EVALUATION DES ERREURS DE CARTES DE VEGETATION AVEC UNE APPROCHE PAR ENSEMBLES FLOUS ET AVEC LA SIMULATION D'IMAGES SATELLITE

Stéphane Couturier

Stéphane Couturier

EVALUATION DES ERREURS DE CARTES DE VEGETATION AVEC UNE APPROCHE PAR ENSEMBLES FLOUS ET AVEC LA SIMULATION D'IMAGES SATELLITE

Directeurs de thèse : Jean-Philippe Gastellu-Etchegorry, UPS
Jean-François Mas, UNAM

THÈSE soutenue le 24 août 2007,
à Mexico, dans l'Université Nationale Autonome du Mexique (UNAM)

Résumé :

Dans les régions de haute biodiversité, caractérisées par des paysages dynamiques, la cartographie détaillée de l'utilisation des sols et du couvert végétal est communément obtenue par la classification d'images satellite. Cependant, les cadres conceptuels d'estimation d'erreurs sur les cartes sont éprouvés pour les zones tempérées et hautement industrialisées.

Une nouvelle méthode est proposée pour l'évaluation de la fiabilité des cartes et une autre méthode pour l'estimation des erreurs de classification par ambiguïtés entre classes sur images satellites. La première méthode comprend un nouveau mode d'échantillonnage et une estimation par ensembles flous des incertitudes positionnelles et thématiques. Elle a été testée sur l'Inventaire Forestier Mexicain de l'an 2000. La deuxième méthode s'appuie sur la simulation d'images satellites avec le modèle de transfert radiatif DART et a été testée sur des images IKONOS de six types de forêts au Mexique, sur terrain plat et en forte pente.

Mots clés : Ensembles flous; erreur de position; incertitude thématique; télédétection; classification de forêt; séparabilité spectrale; transfert radiatif; IKONOS

Discipline : Télédétection de la Biosphère Continentale

Laboratoire de rattachement à Toulouse: Centre d'Etude Spatiale de la BIOsphère, 18 avenue Edouard Belin 31401 TOULOUSE CEDEX 9.

Pour la dryade Adriana Ayodhya, la compagne de ma vie

Pour I'k' te' qui traverse sa huitième vie

Remerciements :

Merci a mes deux directeurs de thèse : Jean-François pour ton mode horizontal de travail, la justesse de tes conseils, ton infaillible disponibilité, et Jean-Philippe pour le généreux partage de ton expertise de physicien et pour être un modèle de rigueur et d'auto-critique. Merci à vous deux ainsi qu'à Alejandro Velázquez, Gerardo Bocco et Jean-Claude Menaut pour avoir su m'accorder votre confiance et votre souplesse, toutes deux très précieuses dans l'exercice délicat d'une cotutelle de thèse. Merci aux rapporteurs et examinateurs de revues scientifiques qui ont accepté de juger ce travail, pour leurs commentaires constructifs et contextualisés.

Merci aux innombrables personnes qui m'ont aidé durant le travail de terrain, pour l'acquisition et le traitement des données. Chaque fois, j'ai essayé de procéder à un échange soit par l'élaboration d'un produit utile à leurs intérêts, soit par l'enseignement de ma discipline, soit par la rémunération, afin de remercier et encourager leurs services envers la communauté scientifique. Une communauté scientifique dont les projets et publications n'ont cessé de se nourrir de ce type de travail de base, et qui pourtant continue honteusement à ne pas prévoir une rétribution appropriée! En particulier je remercie les habitants du village El Cerro, près de Cuitzeo, Michoacán, les techniciens forestiers de la communauté de San Juan Parangaricutiro, près de Tancítaro, Michoacán, les étudiants botanistes du campus CIECO à Morelia, Michoacán, les paysans de l'ejido La Perla près du volcan de San Martin, Veracruz, M. Yakam, le guide de Chan Laguna près de Calakmul, Campeche, Pedro Zamora et ses étudiants biologistes de l'Université Autonome de Campeche.

Merci à l'ensemble des membres de l'équipe DART, qui m'ont infailliblement soutenu en tant qu'utilisateur du simulateur, une interaction sans laquelle l'application que j'ai proposée n'aurait pas pu être mise en oeuvre. Merci notamment aux compagnons scientifiques et aventuriers de la vie le Manu, le Bernat et l'Iskander.

Au personnel administratif de l'Université, qui bien trop souvent supporte encore les frictions avec les étudiants à cause de démarches multiples, redondantes, et inutilement compliquées parce que décidées à huis clôt et séparément par de hauts fonctionnaires de plusieurs administrations. Une révision rigoureuse et une synthèse unificatrice de ces démarches réduiraient aisément la charge nerveuse des étudiants – généralement adultes! - du troisième cycle ainsi que de leurs tuteurs, augmentant d'autant la production scientifique pour l'Université. Pour ceci, une consultation avec les intéressés pour une proposition intelligente, complète et unique de démarches administratives éviterait la sensation de fatalité et médiocrité administrative dans une université aussi prestigieuse que la UNAM…

A Ayodhya, qui aura admirablement résisté à cette période de thèse, initiée peu après notre union, non seulement pour m'avoir accompagné lors des voyages dans mon petit nuage, mais aussi pour m'avoir également emporté dans ton petit nuage de thèses. Au moment de défendre cette thèse devant les jurys, nous aurons créé plusieurs jeux d'idées académiques, construit des liens scientifiques et affectifs avec la nature et de nombreux liens humains, ainsi que fondé un foyer urbain et un foyer campagnard pour la suite.

A I'k'te', de ton enchantement tu n'as cessé d'inspirer nos nuits de travail, puis tu t'es éteinte au moment ou le travail s'est achevé. A Mistlixochitl qui viens de naître et qui guideras le cycle créatif de notre vie qui vient de débuter.

Agradecimientos :

Agradezco a mis dos tutores principales para esta tesis : Jean-François por tu modo horizontal de trabajar, la justeza de tus consejos, tu disponibilidad infallible, y Jean-Philippe por haber generosamente compartido tu especialidad de físico y por ser un modelo de rigor y de auto-crítica. Gracias a ustedes dos así como a Alejandro Velázquez, Gerardo Bocco y Jean-Claude Menaut por haber sabido proporcionarme su confianza y flexibilidad, ambas calidades que me fueron esenciales en el ámbito delicado de una cotutela de tesis. Gracias a los sinodales y revisores de revistas científicas que aceptaron de juzgar este trabajo, gracias por sus comentarios constructivos y contextualizados.

Agradezco a las innumerables personas que me ayudaron para el trabajo de terreno, para la adquisición y el tratamiento de los datos. Cada vez intenté proceder a un intercambio, que sea por la elaboración de un producto útil a sus intereses, o por la enseñanza de mi disciplina, o por una remuneración, con el fín de agradecer y alentar sus servicios hacia la comunidad científica. ¡Comunidad científica cuyos proyectos y publicaciones no han cesado de nutrirse de este tipo de trabajo básico y que vergonzosamente le sigue omitiendo una retribución adecuada! En particular, agradezco a los habitantes del pueblo El Cerro, cerca de Cuitzeo, Michoacán, a los técnicos forestales de la comunidad de San Juan Parangaricutiro, cerca del Tancítaro, Michoacán, a los estudiantes botánicos del CIECO en Morelia, Michoacán, a los campesinos del ejido La Perla, cerca del volcán San Martin, Veracruz, al señor Yakam, el guía de Chan Laguna cerca de Calakmul, Campeche, a Pedro Zamora y sus estudiantes biólogos de la Universidad Autónoma de Campeche.

Agradezco a todos los miembros del equipo DART, quienes infalliblemente me brindaron asistencia como usuario del simulador, interacción sin la cual nunca se habría podido cumplir la aplicación que propuse. Gracias en particular a los compañeros científicos y aventureros de la vida el Manu, el Bernat y el Iskander.

Agradezco al personal administrativo de la Universidad, quien demasiado frecuentemente sigue soportando fricciones con los estudiantes a causa de trámites múltiples, inútilmente complicados y requisitos dedundantes. Decididos al escondito y separadamente por altos funcionarios de distintas administraciones, estos trámites beneficiarían de una revisada rigorosa y una síntesis unificadora, coordinadas por las mismas administraciones. Tal esfuerzo reduciría significativamente la carga nerviosa de los estudiantes – ¡ya adultos! - del posgrado así como de sus tutores, aumentando por lo mismo la producción científica para la Universidad. Con este objetivo, una consulta de los interesados por una propuesta inteligente, completa y única de trámites administrativos evitaría la sensación desgraciadamente muy común de fatalidad y mediocridad administrativa para una universidad tan prestigiosa como la UNAM...

A Ayodhya, quien habrá admirablemente resistido a este periodo de tesis, iniciada poco despues de nuestra unión, no sólo por acompañarme durante los viajes a mi pequeña nube científica, sino también por llevarme a la nube suya de sus tesis. Al momento de defender esta tesis ante el jurado, habremos creado varios sistemas de ideas académicas, habremos construido lazos científicos y afectivos con la naturaleza y otros numerosos lazos humanos, y habremos fundado un hogar urbano y un hogar en el campo para lo que sigue.

A I'k' te', con tus sortilegios no has cesado de inspirar nuestras noches de trabajo, y luego te despediste en el momento en que el trabajo se terminó. A Mistlixochitl recien nacida quien guiará el siguiente ciclo creativo que apenas empieza.

Acknowledgments :

I am grateful to both of my supervisors : Jean-François for your horizontal sense of work, for the fairliness of your comments, for your tireless availability, and Jean-Philippe for generously sharing your expertise in physics and for being a model of rigor and self-criticism. Thanks to both of you as well as to Alejandro Velázquez, Gerardo Bocco and Jean-Claude Menaut for granting me your trust and flexibility, both very precious in the delicate process of a double tutorship. Thanks to the examinators and to the reviewers of scientific journals who have accepted to judge this work, for their constructive and contextual comments.

I am grateful to the numerous persons who helped me during field work, for the adquisition and processing of the data. Every time, I attempted to deal with these persons on an exchange basis, either by generating a product which I knew would serve their interest, or by teaching my speciality, or by loaning their services, in order to thank them for their services to the scientific community and encourage further contributions. Projects and publications of the scientific community are big consumers of field and processing workers, but, rather shamefully, no appropriate retribution is usually contemplated in these projects. For this reason, I am especially grateful to the inhabitants of the village El Cerro, near to Cuitzeo, Michoacán; to the technical foresters of San Juan Parangaricutiro, near to the Tancítaro volcano, in Michoacán; to the students in botany of the CIECO campus, Morelia, Michoacán; to the peasants of the La Perla ejido, near to the San Martin volcano, Veracruz; to Sir Yakam, the guide of Chan Laguna near to the Calakmul reserve in Campeche; and to Pedro Zamora and his biology students of the Campeche Autonomous University.

I am grateful to all members of the DART team, who undefectedly supported me as a user of the model. The application I proposed would never have been carried out without this interaction. Thanks to the scientific fellows and life adventurers the Manu, the Bernat and the Iskander.

I am grateful to the administrative staff of the University, who still too often face frictions with the students due to multiple, redundant and uselessly complicated administrative requisites. Currently, high ranked agents of several administrations decide on these requisites in an unconcerted and authoritative manner. A concerted remodeling of this set of requisites into a unified, unambiguous and simplified list would drastically reduce the negative nervious activity of both the – generally grown up adult! – PhD students and their supervisors. Among the benefits, this would increase the global amount of time dedicated to scientific production rather than to paper handling. As a result, a poll involving all the parties towards the proposal of an intelligent, complete and unique list of requisites might get rid of the currently widespread sensation of fatality and administrative mediocrity in the nonetheless prestigious UNAM university...

To Ayodhya who will have admirably resisted this period of PhD thesis, initiated little after our union, not only for having been my faithful companion in the journeys to my small scientific cloud, but also for having me flown to her own small cloud of scientific reflection. At the time of PhD defense, we will have developed complete sets of academic ideas, constructed a scientific and affective relationship with nature and numerous human ties, and we will have created a household in the city and in the countryside.

To I'k' te', whose spells have never stopped inspiring our nights of work, you went away as the work was completed. To Mistlixochitl, who were just born, you will be the guide for the creative life cycle that has just begun.

EVALUATION DES ERREURS DE CARTES DE VÉGÉTATION AVEC UNE APPROCHE PAR ENSEMBLES FLOUS ET AVEC LA SIMULATION D'IMAGES SATELLITE

Résumé:

Dans les régions de haute biodiversité, la classification d'imagerie satellite est devenue une tâche essentielle pour la cartographie détaillée de l'utilisation des sols et du couvert végétal. Le procédé génère des inexactitudes propres à la complexité des paysages dans ces régions, pour lesquels les cadres conceptuels d'estimation d'erreur sont peu adaptés. Nous proposons deux nouvelles méthodes, l'une pour l'évaluation de la fiabilité des cartes et l'autre pour l'estimation des ambiguïtés de classification sur imagerie satellite.

La première méthode comprend un nouveau mode d'échantillonnage et une estimation par ensembles flous des incertitudes positionnelles et thématiques. Les trois premiers chapitres de cette thèse illustrent différentes phases de construction de cette méthode. La deuxième méthode s'appuie sur la simulation d'images satellites avec le modèle de transfert radiatif DART. La validation et deux applications de cette méthode sont considérées dans les chapitres 4 et 5, et un rapprochement est fait avec le résultat de l'évaluation des cartes du chapitre 2.

Le chapitre 0 situe le contexte de cette recherche. L'historique de la cartographie sur l'utilisation des sols et les couverts de végétation y est rappelé en ce qui concerne la zone d'application (le Mexique), ainsi que les principes de l'évaluation de fiabilité des cartes à échelle régionale. Le premier chapitre décrit les principales étapes de l'évaluation et une nouvelle technique d'échantillonnage y est comparée avec les techniques d'échantillonnage utilisées dans la littérature. Le chapitre 2 est dédié aux résultats thématiques de l'évaluation, sans approche par ensembles flous, de quatre zones d'études qui couvrent une large gamme d'écosystèmes et d'empreinte humaine dans le paysage. Ces résultats sont comparés avec ceux des études publiées pour des zones majoritairement tempérées. La nouvelle approche par ensembles flous est incorporée à la méthode dans le

chapitre 3 et appliquée à une portion de la plus grande masse forestière tropicale au Mexique. La méthode a permis de quantifier les erreurs dues aux incertitudes thématiques et positionnelles des fragments cartographiés de forêts sub-pérennes basses et médianes, peu ou beaucoup modifiées par l'homme.

Les chapitres 4 et 5 sont consacrés à la deuxième méthode pour l'identification de sources majeures d'erreurs sur les cartes. Dans le chapitre 4, le cadre général de la méthode est présenté, avec la prise de mesures optiques et structurelles sur parcelles de forêt. La différence entre la distribution de réflectances simulées et celle de pixels réels a été trouvée inférieure à sa variation à l'intérieur d'un même type de forêt, dans le cas de trois images multi-spectrales IKONOS. Par conséquent, le modèle a été utilisé pour estimer la séparabilité spectrale de trois classes de forêt sur pente d'aspect choisi pour une image IKONOS. Dans le chapitre 5, l'estimation de la séparabilité spectrale a été appliquée à six types de forêt, pour un classificateur conventionnel et pour un classificateur contextuel. Les ambiguïtés entre types de forêt sont confrontées aux estimations de confusions entre classes de l'Inventaire Forestier Mexicain faites dans le chapitre 2.

La conclusion des travaux, ainsi que leurs perspectives, sont enfin présentées dans le chapitre 6 de cette thèse.

EVALUACIÓN DE ERRORES DE CARTAS DE COBERTURA VEGETAL Y USOS DE SUELO CON ENFOQUE DIFUSO Y CON LA SIMULACIÓN DE IMÁGENES DE SATÉLITE

Resumen:

En las regiones de alta bio-diversidad, la clasificación de imágenes de satélite se volvió una tarea esencial para la cartografía detallada del uso de suelo y de la cobertura vegetal. El proceso de clasificación está afectado por errores debidos a la complejidad de los paisajes en tales regiones, para las cuales resultan poco adaptados los cuadros conceptuales existentes de estimación de errores. Dos nuevos métodos están propuestos en esta tesis, uno para evaluar la confiabilidad de cartas en tales regiones, y otro para la estimación de las ambigüedades por clasificación de imágenes de satélite.

El primer método comprende un modo novedoso de muestreo y una estimación, por enfoque difuso, de incertidumbres de posición y de incertidumbres temáticos. Los tres primeros capítulos de esta tesis ilustran diferentes fases de construcción de este método. El secundo método está basado en la simulación de imágenes de satélite con el modelo de transferencia radiativa DART (Discrete Anisotropic Radiative Transfer). Los capítulos 4 y 5 tratan de la validación y de dos aplicaciones de este método. En el capítulo 5 se opera un acercamiento a los resultados de la evaluación de cartas del capítulo 2.

El capítulo 0 sitúa el marco teórico y el contexto de esta investigación. Incluye un historial de la cartografía de uso de suelo y cobertura vegetal en México así como los principios de la evaluación de confiabilidad de cartas a escala regional. Se explica la utilidad de los modelos de transferencia radiativa, y en especial del modelo DART, en la comprensión de las ambigüedades en imágenes de satélite.

En el primer capítulo se encuentran descritas las principales etapas de la evaluación de cartas. El desempeño del modo novedoso de muestreo está comparado con las técnicas de

muestreo utilizadas en la literatura. El capítulo 2 está dedicado a los resultados temáticos de la evaluación, sin enfoque difuso, de la carta del Inventario Forestal Nacional mexicano del año 2000 en cuatro zonas de estudio. Estas zonas cubren una amplia gama de ecosistemas y de modificación antrópica del paisaje en México. Estos resultados están comparados con los resultados de estudios publicados para zonas mayoritariamente templadas. Un enfoque novedoso, por conjuntos difusos, está incorporado al método en el capítulo 3, y aplicado a una porción de la masa forestal tropical más extendida en México. El método permitió cuantificar los errores debidos a incertidumbres temáticos y a incertidumbres de posición entre fragmentos cartografiados de selva baja y mediana sub-perennifolia, poco o mucho modificados por el humano.

Los capítulos 4 y 5 están dedicados al segundo método para la identificación de fuentes mayores de error en cartas. En el capítulo 4 se presenta el cuadro general del método, con la adquisición de medidas ópticas y estructurales en parcelas de bosque. La diferencia entre la distribución simulada de reflectancia y la distribución de reflectancia en píxeles reales se encontró inferior a la variación de reflectancia a dentro de un mismo tipo de bosque, en el caso de tres imágenes multi-espectrales IKONOS. Por consecuencia, el modelo se utilizó para estimar la separabilidad espectral de tres clases de bosque en pendiente, de aspecto seleccionado, en una imagen IKONOS. En el capítulo 5, la estimación de la separabilidad espectral se aplicó a seis tipos de bosque, para un clasificador convencional como para un clasificador contextual. Las ambigüedades entre tipos de bosque fueron confrontadas a las estimaciones de la confusión entre clases del Inventario Forestal Nacional mexicano, realizadas en el capítulo 2.

Las conclusiones de los trabajos, así como sus perspectivas, están reunidas en el capítulo 6 de la tesis.

EVALUATION OF ERRORS IN LAND COVER MAPS WITH A FUZZY SETS APPROACH AND WITH SIMULATED SATELLITE IMAGERY

Abstract:

In highly bio-diverse regions, the classification of remote sensing imagery has become a basic environmental task for detailed land use and land cover mapping. The process is confronted with inaccuracies typical of complex landscapes, and yet, error estimation frameworks have focused on mainly temperate zones. We propose a new method for the accuracy assessment of maps in highly diverse and dynamic environments and a new method for the systematic estimation of ambiguities on satellite imagery.

The first method comprises a novel sampling scheme and a fuzzy sets based framework whereby degrees of positional and thematic tolerance are user defined. The first three chapters of this thesis illustrate different construction stages of this method. The second method is based on the simulation of satellite imagery with the Discrete Anisotropic Radiative Transfer (DART) model. The validation and two applications of this method are treated in chapters 4 and 5.

Chapter 0 situates the background theory upon which this research is based. Background information on Land Use and Land Cover (LULC) mapping in Mexico and basic principles on regional map accuracy assessment are presented. In chapter 1, we describe the general scheme of our regional scale accuracy assessment and the results of our novel

sampling technique are confronted with those of the sampling techniques commonly employed in the literature. Chapter 2 is dedicated to the thematic results of the accuracy assessment, without the fuzzy component, of four study areas which encompass a wide variety of ecosystems and anthropisation of the landscape. These results were compared with those of the published studies which concern mainly temperate zones. The novel fuzzy sets based approach is incorporated to the accuracy assessment in chapter 3 and applied to a portion of the largest tropical forested mass in Mexico. The method allowed us to quantify errors due to either positional or thematic uncertainties of median and low sub-perennial forest patches on the map.

Chapters 4 and 5 were dedicated to the second method for the identification of main sources of misclassification errors on maps. In chapter 4, the general framework is presented, including the optical and structural measurements of forest plots and the image processing steps. Then, the difference between the reflectance distribution in pixels simulated by the model and that of real pixels was found less than the reflectance variation within a forest type in three multi-spectral IKONOS imagery. Consequently, the model was used to estimate the spectral separability between three forest types on terrains with steep slope and varying aspects. In chapter 5, the estimate of spectral separability was applied to six forest types, for a conventional classifier as well as for a contextual classifier. The ambiguities between forest types were confronted with the estimates of confusions between classes of the Mexican National Forest Inventory found in chapter 2.

Finally, chapter 6 synthetizes the results and perspectives of the research.

Table of contents

General Introduction

Geo-strategic and economic interests and, increasingly, environmental issues at various scales, press governments, private corporations, citizens and other non-governmental actors in the quest of spatial information on the continental biosphere. Much of this information is still released through public agencies of large countries or groups of countries, for which a regional scale, intermediate between continental (1:5 000 000) and local (> 1:50 000), is generally adapted to geographical integration strategies. Since the 1990s, classification of satellite imagery had become the standard for Land Use and Land Cover (LULC) mapping programs at regional scale.

For scientific purposes, these cartographic products are useful as input data in many applications, including environmental modelling and land cover change analysis. For these applications, detailed measures on map reliability are necessary. Indeed, maps are affected by misclassification errors, partly due to the intrinsic limitations of the satellite imagery. Misclassification occurs especially when categories of the classification system (classes) are not well distinguished, or ambiguous, in the satellite imagery. Spectral separability, an inverse measure of ambiguity, is a classical remote-sensing index in the evaluation of the performance of automatic classifiers. More generally, it is an indicator of problematic categories for which more cartographic effort is needed.

Assessing the accuracy of LULC maps is a common procedure in geo-science disciplines, as a means, for example, of validating automatic classification methods on a satellite image. For regional scale LULC maps, because of budget constraints and the distribution of many classes over the large extension of the map, the complexity of accuracy assessments is considerably increased. Only relatively recently have comprehensive accuracy assessments, with estimates for each class, been built and applied to regional or continental, detailed LULC maps. In Europe, Büttner and Maucha (2006) reported the accuracy assessment of 44 mapped

classes (including 3 forest classes) of the CORINE Land Cover (CLC) 2000 project. In the United States of America (USA), Laba et al. (2002) assessed the accuracy of 29 LULC classes and Wickham et al. (2004) the accuracy of 21 classes in maps of year 1992 from, respectively, the Gap Analysis Project (GAP) and the National Land Cover Data (NLCD). As a part of the Earth Observation for Sustainable Development (EOSD) program of Canada, Wulder et al. (2006) provide a plan for the future accuracy assessment of the 21 classes in the 2000 Canadian forest cover map.

In this research, we are interested in the evaluation of errors of detailed LULC regional maps in highly bio-diverse regions. These are provided by agencies of countries located in the sub-tropical belt, where no such comprehensive assessment has been done at equivalent taxonomic resolution. Indeed, the accuracy level of these maps is central to issues related with landscape dynamics and bio-diversity conservation in the environment of sub-tropical countries, and in the global environment.

This cartography is characterized by a greater taxonomic diversity (number of classes) than the above cited cartography. For example, in the United States of Mexico (USM), located in a 'mega-diverse' area, the map of the National Forest Inventory (NFI) of year 2000 (Mas et al., 2002) contains 75 LULC classes, including 29 forest cover classes, at the community level of the classification scheme. Taxonomically, the NFI community level in the USM is comparable to the subclass level of the National Vegetation Classification System (NVCS) of the USA (see FGDC, 1997), on which the NLCD map was based.

Higher taxonomic diversity, combined with highly dynamic landscapes, has several implications on the evaluation of errors. First, the numerous sparsely distributed classes represented in the classification scheme pose additional difficulties to the accuracy assessment of the map in terms of representative sampling.

Second, thematic conceptual issues impact the way maps should be assessed, for reasons illustrated in three cases:

- More diversity introduces more physiognomic similarity among taxonomically close classes: for example, cedar forest is an additional conifer forest class in sub-tropical environments, so mixed conifer forest patches are more difficult to classify, and boundaries between conifer forests are more difficult to set. As a result, more uncertainty is introduced in each label of the map as well as in each line of the map.
- Highly dynamic landscapes mean more classes placed along a continuum of vegetation, where some classes are a *temporal transition* to other classes. For example, the sequence of classes 'pasture to secondary forest to primary forest' is characteristic of sub-tropical landscapes. The extremes of such sequence may be spectrally distinct and easily separated, however boundaries between intermediate classes are difficult to interpret.
- More diversity combined with highly dynamic landscapes means more *fragmented landscapes* composed of small patches of different classes. The interpretation of these result in heterogeneous patches difficult to assess.

Third and last, ambiguity between classes on satellite imagery, related to the above situations, becomes more likely. In these conditions, the information on spectral separability could be a systematic tool to prioritise future cartographic efforts.

Confronted with the three implications, we developed two methods based on recent theoretical advances made by the geo-science community. The first method is a novel accuracy assessment framework, adapted to maps of environments with high biodiversity and highly dynamic landscapes. Three phases of construction of the first method were illustrated in chapter 1, 2 and 3 of this thesis. The second method is a model-based systematic framework for the measurement of

ambiguities between forest cover classes on satellite imagery, and this method is illustrated in chapters 4 and 5 of this thesis, with two applications.

In response to the first implication, the first objective of the thesis was to apply a statistically valid accuracy assessment to pilot areas of the Mexican NFI map. An important part of this objective was to document and analyze errors of the NFI map of year 2000 in these areas, and relate them with published error estimates of compatible regional scale cartographic products. To comply with this objective, the first method comprised a sampling design that efficiently controlled the spatial distribution of samples for all classes, including sparsely distributed classes. Previous assessments have relied on two-stage sampling schemes where simple random or stratified by class random sampling was employed in the first stage. We demonstrate that these strategies fail in the context of the Mexican NFI, propose a two-stage hybrid scheme where proportional stratified sampling is employed for sparsely distributed classes, and apply it to four areas in distinct eco-geographical zones of Mexico.

In response to the second implication, the second objective of the thesis was to design a fuzzy sets-based design capable of describing uncertainties due to complex landscapes. Traditionally, it is possible to incorporate a thematic fuzzy component in accuracy assessment designs, but this component, as well as positional uncertainty, are implicitly fixed by the map producer, with no possible change after the design has been applied. Recently, advances in fuzzy classification theory have permitted the comparison of maps incorporating thematic and positional fuzziness. We introduce the notion of degrees of thematic and positional tolerance in accuracy assessments of maps. With the utilization of reference maplets and GIS techniques, this research incorporates thematic fuzziness and positional fuzziness as two parameters in the design, which creates the possibility for a map user to evaluate the map at desired levels of positional and thematic precision. We illustrate the practical usefulness of this possibility in the case of the NFI map, with landscapes composed of intricate tropical forest patches.

As a response to the third implication, the third objective of the thesis was to develop a model-based framework (the second method), by which spectral separability between forest classes is systematically derived from a database on forest structural and biophysical properties. Traditionally, spectral separability is empirically derived from an image, and depends on external parameters such as the sensor type, atmospheric conditions, viewing and illumination configurations. Considering the abundance of forest classes and the complex topography in Mexico, the transferability of empirical measurements was likely to be limited. Conversely, a model-based approach permits the computation of spectral separability in any chosen configuration. This computation is based on the multi-spectral reflectance distribution of a group of pixels on satellite imagery. The quasi totality of radiative transfer models of heterogeneous landscape simulate point-like signals, without the atmospheric impact on the texture of the image. This characteristic prevents their validation on a neighbourhood of pixels. The Discrete Anisotropic Radiative Transfer (DART) model simulates remote sensing images, with atmospheric scattering, of 3D landscapes. First, the method consists in the calibration of the DART model with structural and optical measurements of forest plots, and then in the quantification of spectral separability between forest stands. Very high resolution (4m) imagery was selected for the validation of the approach because of the considerable savings in ground measurements compared to the use of groups of pixels of 20-30m resolution imagery. Validation tests of this model are presented against three IKONOS multi-spectral images, and ambiguity is evaluated among forest stands pertaining to six classes of the NFI map.

In chapter 0, we recall the background theory upon which this thesis is based. Background information on LULC mapping in Mexico and basic principles on regional map accuracy assessment are presented. Traditional methods of fuzzy sets based accuracy assessment, the DART model and its applications are briefly described.

In chapter 1, we describe the general scheme of our regional scale accuracy assessment, and, in particular, our novel sampling design is confronted with previously employed sampling designs for mainly temperate environments. Chapter 1 illustrates all steps of the assessment scheme, including a traditional thematic fuzzy component. This scheme is applied to the NFI map of the closed watershed of the Cuitzeo lake, Michoacán state, an area with very high taxonomic diversity.

In chapter 2, the general scheme of the accuracy assessment, without fuzzy component, is applied to four areas of distinct eco-geographical zones in the Mexican territory. The focus of chapter 2 was the analysis of thematic confusions specific to the NFI map, with respect to previous research in mainly temperate environments; therefore we selected a framework without fuzzy component, comparable to most previously published regional accuracy assessments. The eco-geographical zones include lowland and altiplano environments in Mexico, and contrasted levels of forest cover and anthropisation of the landscape.

In chapter 3, the novel fuzzy sets based design is incorporated in the general scheme of the accuracy assessment and applied to the NFI map in one of the study areas, the watershed of the Candelaria river, Campeche state, where lies part of the largest tropical forested mass in Mexico. Apart from misclassification errors, the method aimed at estimating the positional and thematic fuzziness of median and low sub-perennial tropical forests on the map.

Chapters 4 and 5 were dedicated to the second method, for the identification of main sources of misclassification errors. In chapter 4, the design of the method, based on the radiative transfer model and ground measurements, is presented. Then, the reflectance distribution simulated by the model was tested against real imagery. Next, the method permitted the computation of spectral separability between three forest types for terrains on slopes of varying aspects. Finally, in chapter 5, the evaluation of spectral separability was extended to a contextual classifier. The method was applied to six forest types and the results were

confronted with confusion estimates from the accuracy assessment of the NFI map found in chapter 2.

Chapter 0: Theoretical Framework

1. Land Use and Land Cover mapping in Mexico

1.1 Land cover mapping before year 2000

The production of Land Use and Land Cover (LULC) maps is essential to the understanding of biodiversity dynamics in space and time. LULC maps are also a tool for the measurement of the human presence and social processes in the landscape and for the sustainable use of finite resources on the planet, a growing challenge in our densely populated societies.

As a consequence of its extension over a wide range of physio-graphical, geological and climatic conditions, the Mexican territory is composed of a remarkably large variety of ecosystems and diversity of flora (Rzedowski 1978). The continental territory of Mexico is among the five richest in biological diversity and therefore considered as a mega-diverse area. In turn, this range of environmental conditions predetermined transformations of the landscape by humans in a variety of ways. The intensification of land uses over the last century and the response of the eco-systems to this intensification altogether shaped the complex landscapes in the contemporary Mexico.

In the past three decades, governmental agencies in the North American sub-continent have promoted the production of geographic information at a regional scale, which we define intermediate between continental (1:5 000 000) and local (> 1:50 000). The major historical data set of regional scale (1:250 000) LULC maps in Mexico was developed by the National Institute of Statistics, Geography and Informatics (INEGI). In the nineteen eighties, the first set of 121 LULC maps was published for the entire territory, based on the interpretation of aerial photography collected from 1968 to 1986 (average date 1976) and considerable ground work (INEGI, 1980). This dataset was part of the INEGI first series ('INEGI serie I', in Spanish) cartography. In the mid nineteen nineties, INEGI produced the second series cartography ('INEGI serie II') in a digital and paper format. The LULC maps of INEGI serie II were elaborated using the former series I maps, and visual interpretation of Landsat Thematic Mapper (TM) images acquired in 1993, printed at scale 1:250 000. The INEGI cartography legend included 642 categories to consistently describe LULC in the entire country. For land cover categories, or classes, this detailed

classification scheme was based on physiognomic, floristic and phenologic attributes of plant communities (see table 1) and degrees of anthropic modification.

Table 1 - Classification scheme of the INEGI land use and vegetation cartography (only natural land cover categories are indicated):

Formation	Vegetation Types
Temperate Forest	1. Cedar forest , 2. Fir forest, 3. Pine forest, 4. Conifer scrubland, 5. Douglas fir forest, 6. Pine-oak woodland, 7. Pine-oak forest , 8. Oak-pine forest, 9. Oak forest, 10. Mountain cloud forest, 11. Gallery forest.
Tropical forest	*Perennial & sub-perennial tropical forests:* 12. Tropical evergreen forest, 13. Tropical sub-evergreen forest, 14. Tropical evergreen forest (medium height), 15. Tropical sub-evergreen forest (medium height), 16. Tropical evergreen forest (low height), 17. Tropical sub-evergreen forest (low height) , 18. Gallery forest.
	Deciduous & sub-deciduous forests: 19. Tropical sub-deciduous forest (medium height), 20. Tropical deciduous forest (medium height), 21. Tropical sub-deciduous forest (low height), 22. Tropical deciduous forest (low height), 23. Tropical spiny forest.
Scrubland	24. Sub-montane scrubland, 25. Spiny Tamaulipecan scrubland, 26. Cacti-dominated scrubland 27. Succulent-dominated scrubland, 28. Succulent-cacti-dominated scrubland, 29. Sub-tropical scrubland, 30. Chaparral, 31. Xerophytic scrubland, 32. Succulent-cactus-dominated cloud scrubland,, 33. Rosetophilous scrubland, 34. Desertic xerophytic rosetophilous scrubland, 35. Desertic xerophytic microphilous scrubland,, 36 Propospis spp.-dominated, 37. Acacia spp.-dominated, 38. Vegetation of sandy desertic conditions.
Grassland	39. Natural grassland, 40. Grassland-huizachal, 41. Halophilous grassland, 42. Savannah, 43. Alpine bunchgrassland, 44. Gypsophilous grassland.
Hygrophilous vegetation	45. Mangrove, 46. Popal-Tular (hygrophilous grassland), 47. Riparian vegetation.
Other vegetation Types	48. Coastal dune vegetation, 49. Halophilous vegetation.

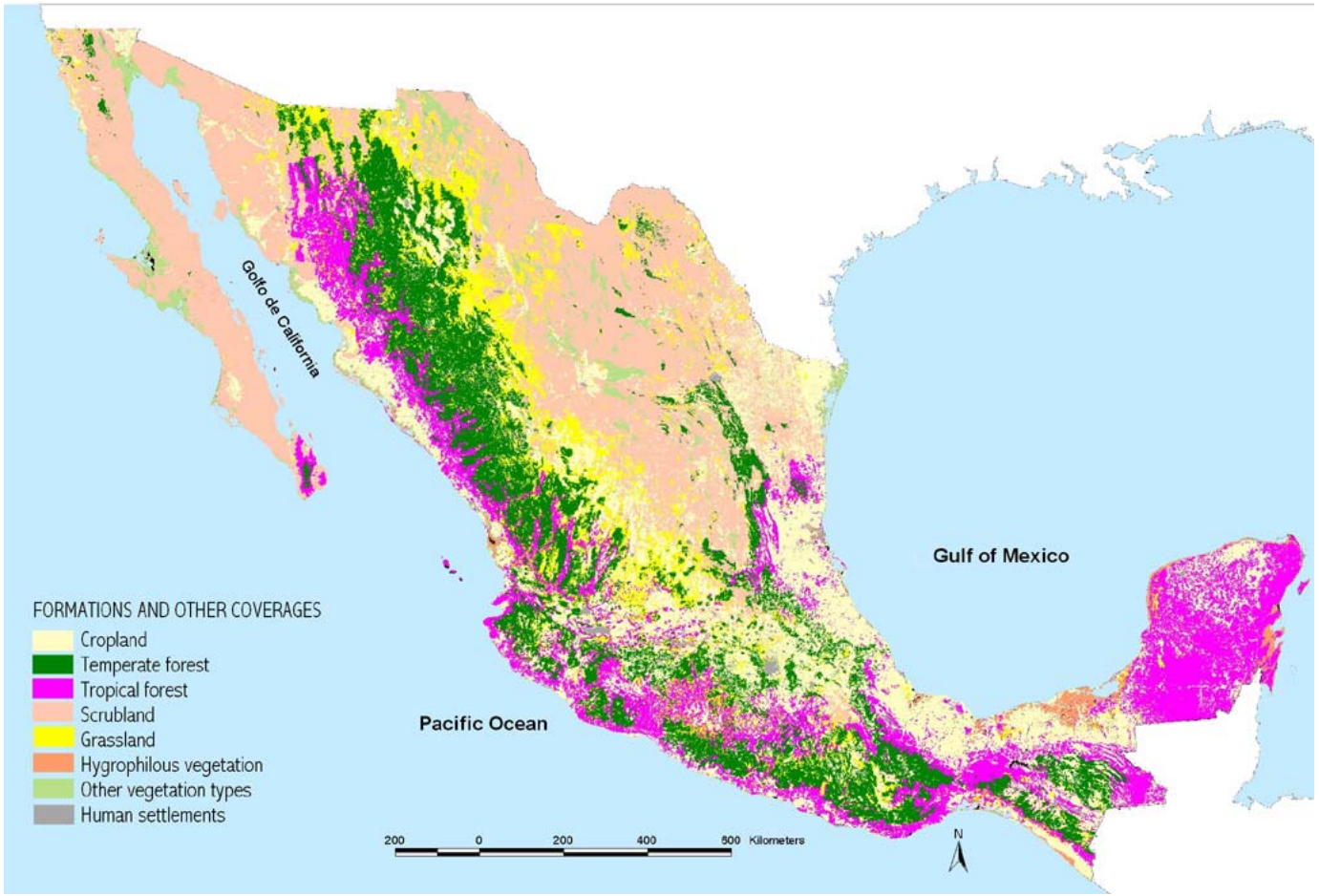
1.2 LULC mapping in year 2000

In the year 2000, the Ministry of the Environment in Mexico (SEMARNAP) attributed the task of updating the LULC map of the country (at scale 1:250 000) to the Institute of Geography of the Universidad Nacional Autónoma de México (UNAM). This task was intended as the first step of a nation-wide detailed forest assessment. In this perspective, the cartographic project was named the National Forest Inventory (NFI) map of year 2000. An important objective of the project was the compatibility with previous cartography in view of LULC change studies. Rapidity (8 months) and low cost of execution were constraints that guided the planning of the project.

Visual interpretation of satellite imagery, with the aid of INEGI previous LULC digital cartography, was selected as the best classification strategy. However, the classification scheme was adjusted to the capacity of the Landsat Enhanced TM plus (ETM+) imagery at discriminating classes, according to previous classification experience in Mexico (e.g. Mas et al., 1996). The 642 categories of the INEGI cartography legend (including 49 vegetation types in table 1) were aggregated into 75 thematic classes (community level, with the inclusion of two levels of human induced modification) and further into three coarser levels of aggregation.

Visual interpretation was done on ETM+ imagery of the drier season, acquired between November 1999 and April 2000. The best option for interpretation was visually selected among various colour composites. The methods and results of the IFN 2000 cartographic project have been published (Mas et al., 2002a; Palacio et al., 2000; Velázquez et al., 2002). Figure 1, taken from Mas et al. (2002a), illustrates the 2000 NFI map at formation level (coarsest level of agregation). The present research focuses on the cartographic product with the finest level of aggregation (community level, with the inclusion of degradation levels) of the NFI map. Indeed, a series of important applications typical of mega-diverse areas rely on their information content: forest degradation and regeneration, biodiversity conservation, environmental services, carbon budget studies, etc. However, the reliability of these taxonomically detailed products is still poorly studied. Since 2001, the National Commission on Forests (CONAFOR) is in charge of the updating of regional scale LULC cartography in Mexico, as a joint task with the more detailed forest inventory of the Mexican territory.

Figure 1: National Forest Inventory map of Mexico in year 2000



Taken from Mas et al. (2002)

2. Regional scale map accuracy assessment

2.1 Present state of map accuracy assessment in Mexico

In many applications still, map reliability and quality are unquestioned, given for granted, just as if each spatial unit on the map perfectly matched the key on the map, which in turn perfectly matched reality. The minimum mapping unit, which defines the scale of the map, is commonly the unique information available about the spatial accuracy of a map. In Mexico, the main databases and cartographic products at regional scale, such as INEGI cartography, and the most recent CONAFOR-managed cartographic products, remain deprived of reliability (accuracy) assessment. This is most unfortunate since the named governmental agency produces controversial statements on recent deforestation rates based on these maps. A spatial error bar is sometimes present aside the legend of INEGI maps and indicates an estimate of errors in the process of map production. However, the procedure leading to this estimate is usually undisclosed, and any objective interpretation of this estimate by the user is thus discouraged (Foody, 2002). Moreover, such error bar indicates a very reduced piece of information with respect to the thematic accuracy of the map.

In the absence of more information, qualitative statements can characterize the reliability of cartographic products, based on a judgment on the quality of the data that was employed in the map production process. For example, the INEGI serie I data are expected to be very reliable in terms of thematic accuracy, because of the quality of the field reference data, but their temporal coherence (accuracy) is low. Conversely, the LULC maps of INEGI serie II are characterized by a high temporal coherence. However, because the visual interpretation of only one colour composite of Landsat imagery (bands 4,3,2) was used to update a map with very high taxonomic precision (INEGI legend of 642 classes), the thematic accuracy of INEGI serie II is likely to be poorer than that of INEGI serie I (Mas et al., 2002b).

2.2 Background on map accuracy assessment

The quantitative information on the accuracy of a cartographic product gives the user an estimate of its agreement with respect to reality. This estimate may serve as a measurement of the risk undertaken by a decision maker using the map. Besides, this information allows error propagation modeling through a Geographical Information Systems or GIS (Burrough, 1994).

Generally, map accuracy is measured by means of reference sites and a classification process more reliable than the one used to generate the map itself. The classified reference sites are then confronted with the map, assuming that the reference site is "the truth". Agreement or disagreement is recorded in error matrices, or confusion matrices (Card, 1982), on the base of which various reliability indices may be derived. For regional scale LULC maps, the abundance and distribution of classes over the large extension of the map, confronted with tight budget constraints, add complexity to accuracy assessments. Only relatively recently have comprehensive accuracy assessments, with estimates for each class, been built and applied to regional or continental LULC maps (e.g. Laba et al., 2002; Stehman et al., 2003; Wickham et al., 2004; Wulder et al., 2006). Because of the high complexity of these products, detailed information on the assessment process itself is needed for the reliability figures to be interpreted properly (Foody, 2002). With this understanding, Stehman and Czaplewski (1998) have proposed a standard structure for accuracy assessment designs, divided into four phases:

1. Representative selection of reference sites (sampling design),
2. Definition and processing of the reference material (evaluation protocol),
3. Classification of the reference sites (labeling protocol),
4. Comparison of the map label with the reference label (synthesis of the evaluation).

Wulder et al. (2006) provide a recent review on issues related to these four steps of the accuracy assessment design. We indicate in the next sub-section the techniques most commonly employed for phase 1 in the literature. Chapter 1 focuses on a novel sampling design. We describe characteristics of phase 2 in subsection 2.4. Phases 2 and 3 were grouped in the same step called 'Verification design' in chapters 1 and 2. Conceptual

issues related to the fuzziness of cartographic objects concern phase 3 and are treated in the third section of this chapter. Chapter 3 of this thesis presents a novel approach for phase 3, based on a detailed analysis of sources of error in the mapping process (background section of chapter 3). Finally, we present the fundamentals of the synthesis of the evaluation (phase 4) in section 2.5.

2.3 Sampling design

The first phase of the accuracy assessment is the sampling design. The selection of the reference sites is a statistical sampling issue (Cochran, 1977), where strategies have varied according to the application and complexity of the spatial distribution. Stehman (2001) defines the probability sampling, where each piece of mapped surface is guaranteed a non-null probability of inclusion in the sample, as being a basic condition for statistical validity. In most local scale applications, reference sites are selected through simple random sampling. Two stage (or double) random sampling means sampling in a set of clusters selected by simple random sampling in a first step. At regional scale, this technique permits much more control over the spatial dispersion of the sample (Zhu et al., 2000), and was adopted for the first regional accuracy assessments in the USA, for LULC maps of 1992 (Laba et al., 2002; Stehman et al., 2003). A random, stratified by class sampling strategy means that reference sites are sampled separately for each mapped class (Congalton, 1988). This strategy is useful if some classes are sparsely represented on the map and, therefore, difficult to sample with simple random sampling. This strategy was adopted at the second stage of their double sampling by Stehman et al. (2003) and Wickham et al. (2004). Systematic sampling refers to the sampling of a partial portion of the mapped territory, where the portion has been designed as sufficiently representative of the total territory. This strategy, adopted as a first stratification step, is attractive for small scale datasets and reference material of difficult access: Wulder et al. (2006) define a systematic stratum for the future (and first) national scale accuracy assessment of the forest cover map in Canada.

In the case of the NFI map of Mexico, a preliminary accuracy assessment was conducted immediately after map production in year 2000. A systematic sampling of the entire country was planned, but the assessment could only take place on a small portion of the planned coverage, in the Northern part of the country (Mas et al., 2002a). The assessment

yielded reliability levels for a few homogeneously distributed classes, and was not designed to attend, in a cost-effective way, the high number of classes of the NFI map and their complex distribution over the entire territory. Indeed, chapter 1 of this thesis shows that existing sampling strategies, previously employed in the USA or planned in Canada are not efficient in the case of the NFI map. The first objective of this research, treated in chapter 1, was therefore to develop a sampling design capable of attending this taxonomic diversity and distribution. This sampling strategy was then extended to four pilot areas of the Mexican territory in chapter 2, where accuracy results and confusion patterns could be discussed and compared with those of published regional scale assessments.

2.4 Evaluation protocol

For regional scale maps, the frame for reference material of phase 2 is typically an aerial photographic coverage (e.g. Zhu et al., 2000), and ground survey is only occasional. Aerial photography is preferred to ground survey in large areas for its cost-effectiveness. Double sampling techniques are effective at controlling the spatial dispersion of the sample among aerial photographs, if photographic frames are taken as the cluster, or Primary Sampling Unit (PSU), for first stage sampling.

Congalton and Green (1993) relate errors of the map to imprecise delineation and/or misclassification. Additionally, the imperfect process of the assessment itself also generates erroneous statements on whether the map represents reality or not. A main topic is the positional error of the aerial photograph with respect to the map. To this respect, a procedure ensuring geometric consistency must be included in the evaluation protocol. For example, the procedure of visually locating sample points on the original satellite imagery, described in Zhu et al. (2000), reduces the inclusion of errors due to geometric inconsistencies. Other sources of fictitious errors occur in phase 3 (labelling protocol), and are related to the thematic and positional uncertainties of maps. This topic is introduced in the next section and fully devised in chapter 3.

2.5 Synthesis of the evaluation

The comparison between the information contained on the map and the information derived from the reference site yields an agreement or a disagreement. Typically, the numbers of agreements and disagreements are recorded and form a confusion matrix.

However, these numbers are reported in the matrix with weights that depend on the probability of inclusion of the reference site in the sample (Stehman, 2001). This probability of inclusion is defined by the sampling design. For example, a simple random selection is associated with a uniform (constant) inclusion probability among all reference sites. For a two stage sampling, the probability of inclusion follows Bayes law: The probability of inclusion p_{2k} of a reference site at the second stage is a multiplicative function of the inclusion probability p_{1k} of the cluster it pertains to, and of the inclusion probability of the reference site, once the cluster has been selected $p_{2|1}$ (conditional inclusion probability):

$$p_{2k}=p_{2|1} * p_{1k} \qquad (1)$$

Accuracy indices per class are derived from these calculations: 'user's accuracy' of class *k* is the account of agreements from all sites of mapped class *k* while the 'producer's accuracy' of class *k* counts agreements from all reference sites labeled as class *k*. The respective disagreements correspond to 'commission errors' and 'omission errors' (Aronoff 1982). The global accuracy index, or proportion correct index, which indicates the accuracy of the map as a whole (all thematic classes), integrates the accuracy level of all classes, weighted by the probability of inclusion specific to each class. In this calculation, weights usually correspond to the relative abundance of the class on the map. Other reliability indices are popular, such as the Kappa index, which takes into account the contribution of chance in the accuracy (Rosenfield and Fitzpatrick-Lins, 1986). However, in regional scale accuracy assessments, the proportion correct indices are preferred, because they are coherent with the interpretation of confusions according to area fractions of the map (Stehman, 2001).

A confidence interval of the accuracy indices can be estimated, although only few accuracy assessments provide this information. A popular estimate of the confidence interval is based on the binomial distribution theory: the confidence interval of the accuracy estimate depends on the sample size and on the reliability value (accuracy estimate) in the following manner (Snedecor and Cochran, 1967, cited by Fitzpatrick-Lins, 1981):

$$d^2 = t^2 \, p \, (1-p) / n \qquad (2)$$

where d is the standard deviation (or half the confidence interval) of the estimate, t is the standard deviate on the Gaussian curve (for example, t = 1.96 for a two-sided probability of error of 0.05), p is the reliability value, and n is the number of sampled points. Although most accuracy assessments refer to it, this binomial distribution formula is only valid for simple random sampling. For more sophisticated sampling designs (e.g. two stage sampling) the confidence interval is influenced by the variance of agreements among clusters. Estimators integrating inter-cluster variance are seldom employed in map accuracy assessments because of their complexity (Stehman et al., 2003). Chapter 1 and 2 only state variance estimates based on the binomial formula. A better estimator, with the integration of an inter-cluster variance term, was used in chapter 3. The estimate is built on stratified by class selection in the second-stage of the sampling (Särndal et al., 1992, section 5.5) and is formulated in chapter 3.

3. Assessing the fuzzy nature of mapped objects

In traditional accuracy assessment, the labeling protocol (phase 3 of the accuracy assessment) consists in attributing one and only one category of the classification scheme to each reference site. However, this procedure assumes that each area in the map can be unambiguously assigned to a single category of the classification scheme (or LULC class). In reality, the mapped area may be related to more than one LULC class because of the characteristics of the landscape in the reference site. This conceptual difficulty is ignored in the traditional (or Boolean) labeling protocol, and may conduce to an under-estimation of map accuracy (Foody, 2002). In particular, this difficulty arises in the following cases:

- The landscape in the reference site has physiognomic similarities with more than one LULC class. For example, a one hectare forest patch containing oak trees and two or three pine trees has physiognomic similarity with forest class ‘oak forest’ and forest class ‘oak-pine forest’. As a result, the map label for this site is affected with uncertainty. The reference site could be in a *transition zone* between an oak forest and a oak-pine forest.
- The landscape in the reference site is a patch within a continuum of vegetation, where the LULC classes represented are a *temporal transition* to other classes. For example, the sequence of classes ‘pasture to secondary forest

to primary forest' is characteristic of some sub-tropical landscapes. As a result, the map label for this site is affected with uncertainty. The extremes of such sequence may be easily identified on the ground, however boundaries between patches of intermediate classes are difficult to set. As a result, lines between mapped objects for this site are affected with uncertainty.

- The landscape in the reference site is a *fragmented landscape*, composed of small patches (below minimum mapping unit) of different land use or land cover. The interpretation of this mixed reference site, because of the scale of the map, must be with a heterogeneous patch. As a result, the map label for this site is affected with uncertainty.

3.1 Fuzzy classification

Due to the above described continuous or fragmented aspects of land use and land cover in a landscape, maps with discrete representation (discrete, or crisp, class assignation) and infinitely small line features (crisp boundaries of objects) necessarily describe reality with a certain margin of uncertainty. In order to take this uncertainty aspect into account, it has been referred to the concept of fuzzy sets (Zadeh, 1965).

In the crisp approach, an element x of the map X belongs totally to a class k of the set C or does not belong to it. A way of representing this is to define a membership function μ, which takes the value '1' if the element x belongs to class k and '0' otherwise. This assignation process can be called Boolean labeling. In a typical case of photo-interpretation for map accuracy assessment, a forest reference site with a crown cover close to 40% may pertain to a transition zone between closed forest (crown cover > 40%) and open forest (crown cover < 40%). If the photo-interpreter characterizes this site as closed forest and the corresponding label on the map is open forest, then this site is interpreted as an error on the map.

In fuzzy sets theory, an element belongs to a set or class with a certain degree of similarity, probability or property, some of these notions being contained in a 'degree of membership', depending on the application. One element x may belong to various classes at a time with different degrees of membership $\mu_k(x)$. For example, quantitative degrees of membership take a value between 0 and 1 to express the partial membership to various

classes of the set. With this approach, the reference site with a tree cover close to 40%, would be characterized for instance by a 0.5 degree of membership in both classes (open and closed forest).

Many authors have rejected the term "fuzzy set theory" to characterize landscape interpretation, in favor of "soft" or "continuous" classification. Critics have noted that the use of a continuous range of membership values does not entail employment of the concepts of fuzzy logic (Haack 1996). Nevertheless, the term "fuzzy classification" will be used here as a compromise, recognizing the heritage of these techniques but emphasizing the classification process over the logic of set theory.

3.2 Fuzzy map assessment

Cartographical models that present a fuzzy classification approach were developed (Fisher and Pathirana 1990, Wang 1990, Foody 1992, Equihua 1990, 1991, Moreno-Sanchez, 2000). These models allow the representation of the landscape features enumerated in the previous subsection. Despite the perspective of a more lawful representation of real landscapes, these models present two limitations:

- The interpretation and manipulation of fuzzy classified maps by users already accustomed to crisp maps is still a pending challenge; each point on the map represents various LULC classes with different degrees of membership. The vast majority of maps, including the existing LULC maps in Mexico, are crisp.

- The coherent production of fuzzy classified maps with quantitative degrees of membership is not possible in all mapping situations. One of the situations where such fuzzy classified map can be easily produced is a binary map of, for example, forest/non-forest where percent crown coverage represents one of the fuzzy labels. A second situation is a map made of ordinal categories, where uncertainty between categories can be modeled by a fuzzy matrix (illustrated in Hagen, 2003). A third situation occurs when automatic processing is constructed so as to generate the quantitative fuzzy labels. A typical example of this third situation is a map of unmixed fractions of LULC classes, extracted from automatic spectral mixing analysis, where the classes

are represented by pure end-member pixels. However, the assignment of quantitative fuzzy labels during visual interpretation, for example, can be affected by subjectivity. This is possibly a reason why quantitative fuzzy labeling has generally not been adopted in mapping situations with visually interpreted material.

In this research, the focus is made on assessing a crisp map with fuzzy classified reference material. As mentioned in the previous section, the typical reference material of regional accuracy assessments is aerial photographs. We were confronted with the subjectivity of interpreters in preliminary attempts at classifying the material with quantitative degrees of membership. For these reasons, we settled for the fuzzy classification technique expressed by linguistic rules, introduced for visual interpretation by Gopal and Woodcock (1994), and commonly employed. This technique of fuzzy classification is described in the verification design of chapter 1.

The use of fuzzy classification techniques in the labeling protocol permits the reduction of fictitious errors in the process of map assessment, fictitious errors being due to the thematic uncertainty of maps. However, as said earlier, maps are also characterized by positional uncertainty. This uncertainty may also affect the accuracy results when the assessed map is compared with the reference material. In current regional accuracy assessments so far, no technique handles positional uncertainties in a systematic way. However, as a result of advances in fuzzy classification theory, much research have focused on the comparison of fuzzy classified maps and on the multi-scale comparison of maps (Pontius and Cheuk, 2006; Remmel and Csillag, 2006; Visser and de Nijs, 2006). In chapter 3, the systematic inclusion of positional uncertainty within regional accuracy assessments is proposed and formalized.

4. Radiative transfer models and spectral separability

4.1 Classification and ambiguities in satellite imagery

Detailed mapping of taxonomically diverse regions is a difficult, yet important task. The majority of the studies dedicated to this task since the 1990s were concerned with the spectral separability of LULC classes on 20-30m resolution imagery (typically from Landsat or SPOT sensors). Numerous classification techniques, applied in confined

geographical areas, reported successful attempts at discriminating a set of classes on this type of imagery (e.g. Tuomisto et al., 1994; Mas and Ramírez, 1996; García and Álvarez, 1994; Haack and Jampoler, 1994; Roy et al., 1991), including image-based automatic procedures (Magnussen et al., 2004; Salovaara et al., 2005; Lu et al., 2003). However, extensive forest maps derived from semi-automatic classification are affected by greater environmental heterogeneity (including steeper topography), and therefore confusion patterns prove quite substantial when comprehensive regional-scale accuracy assessments are made (e.g. Wickham et al., 2004; Remmel et al., 2005). Indeed, authors of case studies admit that spectral separability of forest covers highly depends on topographic conditions and illumination configuration (e.g. Baban and Kamaruzaman, 2001). Additionally, the advent of new sensors broaden possibilities for classifying taxonomically close forests (e.g. Wang et al., 2004) but new empirical studies will be needed to test spectral separability of forest covers with the new sensors.

Case studies on steep terrain that include topographic corrections reported good accuracy results of forest discrimination in temperate zones (e.g. Riaño et al., 2003; Dorren et al., 2003) but fail to report detailed information on the cases where forests remain ambiguous because of the aspect of the slope, because of the illumination configuration or because of atmospheric conditions. A probable reason for this is that classification studies are empirical, image-based, and the transferability of the results is limited. Failed attempts at discriminating forest types in complex environments certainly also exist. However these attempts are seldom published; therefore, for a forest map producer, little information is available on ambiguities between forest types on satellite imagery, although they are a major cause for map inaccuracies.

Conversely, a model-based approach permits the computation of spectral separability with controlled environmental conditions. Considering the abundance of forest types and the complex topography in Mexico, we decided to explore in this thesis a non-empirical approach, based on the recent advances in 3D radiative transfer models.

4.2 Radiative Transfer models and remote sensing imagery of forests

Over the past two decades, the modeling of remotely sensed signals from the earth surface have significantly contributed to the understanding of the potentials and limitations of remote sensing imagery (Pinty & Verstraete, 1991; Strahler, 1996).

Radiative transfer corresponds to the propagation of radiation in a medium. Its equation in the Visible-Near InfraRed (VNIR) wavelengths gives the variation rate of a stationary monochromatic wave of intensity I(r, Ω) (Watts per unit solid angle: $W.sr^{-1}$), at location r, per unit displacement dr and along direction Ω (Hapke, 1993):

$$\frac{dI(r,\Omega)}{dr} = \left[\xi.\frac{d}{dx} + \eta.\frac{d}{dy} + \mu.\frac{d}{dz}\right] I(r,\Omega) = -\alpha(r,\Omega).I(r,\Omega) + \int_{4\pi} \alpha(r,\Omega').\omega(r,\Omega').\frac{P(r,\Omega' \to \Omega)}{4\pi}.I(r,\Omega').d\Omega' \qquad (3)$$

Where:

- ξ, η et μ are the cosine angles of the direction of wave propagation Ω along axes x, y and z.

- $\alpha(r,\Omega)$ is the total extinction coefficient, calculated as the sum of the absorption coefficient $\alpha_a(r,\Omega)$ and the differential scattering coefficient $\alpha_d(r,\Omega' \to \Omega)$, which describes the scattering of photons from direction Ω' to direction Ω.

- $\omega(r,\Omega) = \frac{\alpha_d(r,\Omega \to 4\pi)}{\alpha(r,\Omega)}$ is the scattering albedo from direction Ω.

- $\frac{P(r,\Omega' \to \Omega)}{4\pi} = \frac{\alpha_d(r,\Omega' \to \Omega)}{\alpha_d(r,\Omega')}$, the normalized phase function, is the fraction of radiation along direction Ω' which is intercepted and scattered along direction Ω. The first term of equation (3) expresses the attenuation (absorption plus scattering in all directions) of the wave during propagation, and the second term expresses the scattering towards direction Ω of all incident radiation.

The radiative transfer approach is characterized by two modelling stages: one stage is the mathematical method to solve radiative transfer equation (3), and the other stage is the

representation of the medium of propagation. These two stages are interdependent. With a first group of models, known as 2D models (Idso and de Wit, 1970; Ross, 1981; Verhoef, 1984; Myneni et al., 1989), landscapes are represented with a constant horizontal distribution of absorbing and scattering elements (sheets, branches, etc.).

More sensitive to heterogeneous landscapes, 3D radiative transfer models incorporate three dimensional elements and were able to produce reflectance signals of a vegetation parcel for comparison with low (1km) to medium (25m) resolution satellite imagery (e.g. North, 1996: the Flight model; Govaerts and Verstraete, 1998: the Raytran model; Disney et al., 2006: the Drat model). This approach has been successfully assessed against real imagery of forests in mainly coniferous stands made of repeated individuals of one or two species (Gemmel and Varjo, 1999; North, 2002; Courbaud et al., 2003; Disney et al., 2006). Such models have predominantly been used for the inversion of biophysical parameters on medium resolution imagery for which the average reflectance over a homogeneous stand is adapted.

Spectral separability studies are based on the reflectance distribution of a group of pixels on satellite imagery. The quasi totality of radiative transfer models, such as the ones mentioned above, simulate point-like signals, averaged over the landscape size. Consequently, the impact of atmospheric scattering on the texture of the image, generally non negligible at medium resolution, is impossible to simulate directly. This characteristic prevents their validation on a neighbourhood of pixels on real imagery. Furthermore, the signal cannot be obtained at resolutions approaching crown size, because individual tree characteristics impact the signal more than the average canopy.

4.3 The DART model

The Discrete Anisotropic Radiative Transfer (DART ; Gastellu-Etchegorry et al., 1996; Martin, 2006), developed in CESBIO (Centre d'Etudes Spatiales sur la BIOsphère)-UPS (Université Paul Sabatier) in Toulouse, France, is a 3D radiative transfer model which produces images at satellite sensor. The model integrates the sensor response and the atmospheric signal (Gascon, 2001).

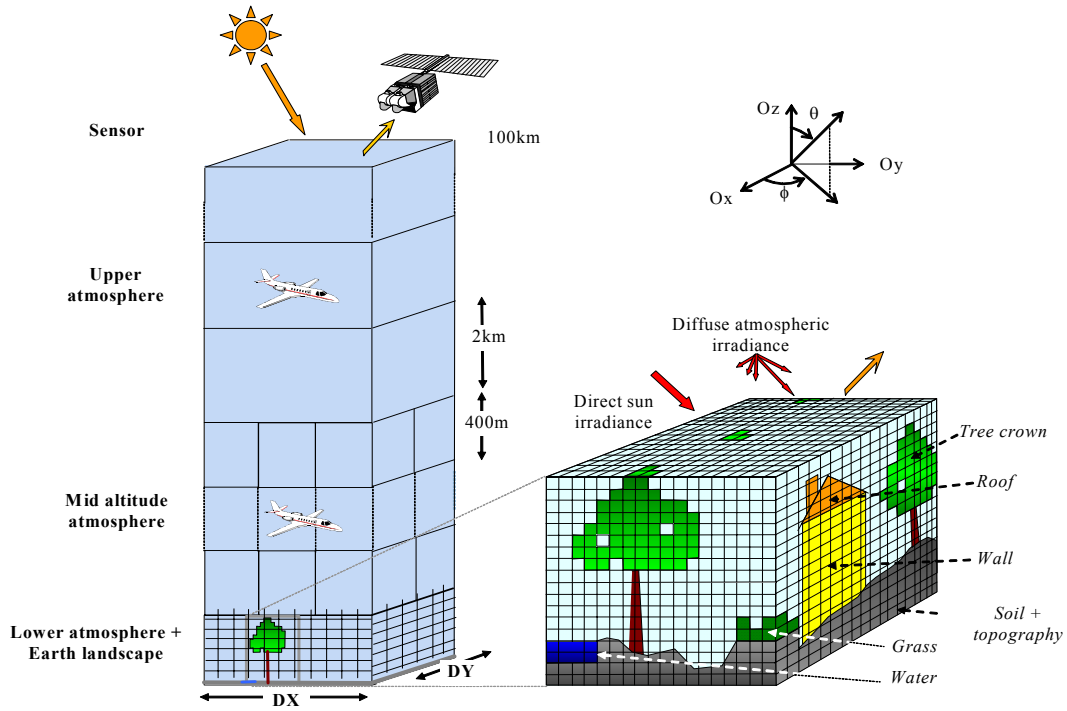
In the method of the discrete ordinates (Kimes and Kirchner, 1982), the angular variable Ω is partitioned into N_{dir} contiguous angular sectors where each sector has a central direction Ω_n ($n \in [1 ; N_{dir}]$) and an angular width $\Delta\Omega_n$. This discrete representation transforms the integral-differential equation (3) into N_{dir} differential equations:

$$\mu_n\left[\frac{d}{dz}+\eta_n\frac{d}{dy}+\xi_n\frac{d}{dx}\right]W(r,\Omega_n)=-\alpha(r,\Omega_n).W(r,\Omega_n)$$

$$+\sum_{m=0}^{N_{dir}}\alpha(r,\Omega_m).\omega(r,\Omega_m).\frac{P(r,\Omega_m\rightarrow\Omega_n)}{4\pi}.W(r,\Omega_m).\Delta\Omega_n \qquad (4)$$

Where $W(r, \Omega_n) = I(r, \Omega_n).\Delta\Omega_n$ ($n \in [1 ; N_{dir}]$) is the power transported at location r, along direction Ω_n and in a cone of angular width $\Delta\Omega_n$, and Ω_0 is the solar incident direction. For precise approximation of equation (3), the angular sectors $\Delta\Omega_n$ must be small, especially in a heterogeneous medium of propagation with anisotropic optical properties such as forest covers.

Figure 2: DART representation of the landscape- atmosphere- sensor system

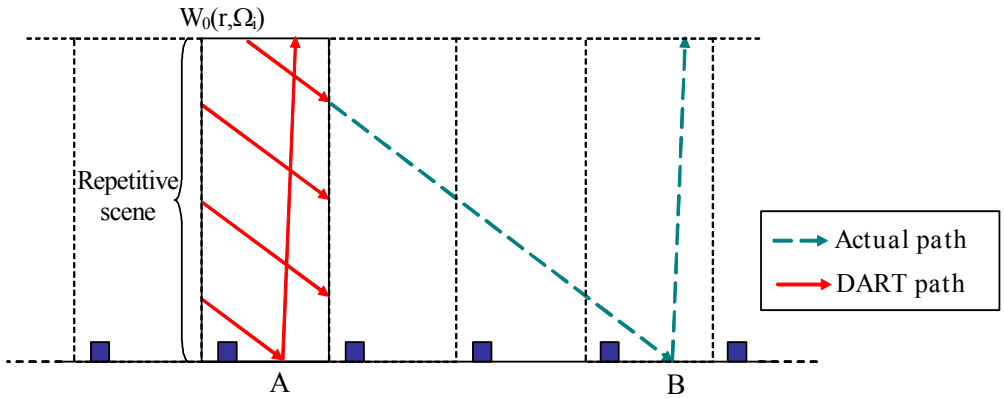


Taken from Martin (2006).

The medium of propagation of the DART model is the 'landscape – atmosphere' system. The computer representation of the medium is a finite tri-dimensional matrix (or scene) of parallelepiped cells of different sizes (see figure 2). The atmosphere system is simulated with 'air' cells, whereas the landscape system is composed of 'air' cells combined with cells associated with 4 types of elements: soil, vegetation, wall, and water. The soil is simulated with or without topography. Scene dimensions along axes Ox, Oy and Oz are DX, DY and DZ. An atmospheric layer of DZ ≤100km is simulated, with DZ= 100km for satellite sensors, because the concentration of major atmospheric components (gas, aerosols) is neglected above 100km. Cell dimensions of the landscape (ΔX, ΔY and ΔZ) are selected as a trade-off between the size of landscape elements and computation time, the latter being governed by the total number of cells in the matrix.

Equation (4) is numerically solved using the ray tracing method. At iteration k=0, a discrete number of solar rays $W_0(r,\Omega_n)$ is simulated above the scene and their propagation is tracked in the scene. Interception of the solar ray occurs upon impinging a non-empty cell of the matrix. The interception is total for opaque cells, such as soil, water and wall cells. In this thesis, interception of rays by the trunk of trees is total as well. Conversely, the interception is partial in vegetation (e.g. crown, grass) cells, for which transmission occurs. During interception, the transmitted power of the ray decreases by absorption and/or scattering, according to the optical and structural properties of the cell. Whenever the ray reaches one side of the scene, the approach of Kimes and Kirchner (1982) is adopted, and consists in the infinite repetition of the scene sideways (see figure 3)

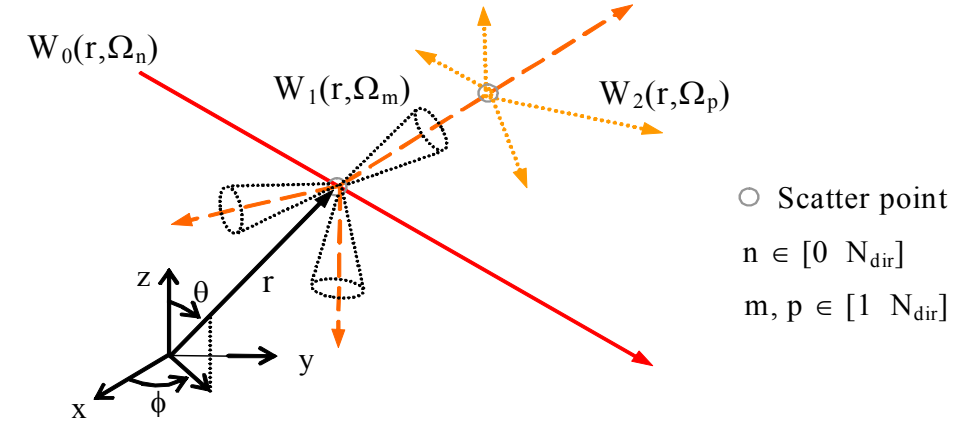
Figure 3: Tracking of rays in an infinite, repetitive landscape



Taken from Martin (2006)

In case of non flat terrain, topography is regularized for the continuous repetition of the soil surface. The scattering of intercepted radiation generates new rays $W_1(r,\Omega_n)$ along the N_{dir} directions. Along their path, these rays generate new scattered rays $W_k(r,\Omega_n)$ ($k \in [2\ \infty[$), as illustrated in figure 4. This process ends if the ray leaves the top of the 3D scene, or whenever the power of the ray becomes smaller than a threshold value.

Figure 4: Multi-scattering mechanism



Taken from Gascon (2001)

4.4 DART studies of the reflectance signal of forests

A major application of the DART model on vegetated landscapes has been the inversion of biophysical parameters (e.g. Leaf Area Index, Crown Coverage) from satellite imagery (Estève, 1998; Gastellu-Etchegorry et al., 2003; Gascon et al., 2004). Typically, the average value of a group of pixels on the image is taken as the output reflectance of the DART model, and inverted into the value of a biophysical parameter. All of these studies operate on pixel values or averaged pixel values in satellite imagery; the texture of the image, or the reflectance distribution (histogram), is not needed. This inversion procedure is therefore possible with the DART model, but also with non-imaging radiative transfer models.

With a previous version of DART (Gastellu-Etchegorry et al., 1996), several studies concerned the reflectance distribution of forests on remote sensing imagery. First, the sensitivity of the Bidirectional Reflectance Factor (BRF) of forests to illumination configurations and major biophysical parameters has been done in a systematic way

(Gastellu-Etchegorry et al., 1999). One aspect of the sensitivity study concerned the standard deviation of the signal at various pixel sizes (1m - 50m). Additionally, a detailed sensitivity study to biophysical parameters was conducted for textural indices (semi-variograms) of the BRF (Bruniquel-Pinel and Gastellu-Etchegorry, 1998). Both studies were applied at the bottom of the atmosphere (BOA).

Next, a comparison of DART- simulated imagery with 20m resolution satellite imagery was done by Trichon et al. (1999). A satisfactory comparison was obtained for average reflectance values between DART simulated BOA reflectance, and the average satellite pixel value after an empirical correction from atmospheric effects was applied. However, the textural indices of the simulated imagery were systematically larger (standard deviation more than two times larger) than the real imagery.

Very interestingly, this result possibly revealed the importance of atmospheric coupling and sensor response in radiative transfer modeling: indeed, the previous version of DART (Gastellu-Etchegorry et al., 1996) only included the landscape system, and was deprived of the coupling with the atmospheric system. In fact, the empirical atmospheric correction permitted to normalize the satellite pixel values to BOA values using a multiplicative coefficient in all the image. However, the texture of the image after such empirical atmospheric correction (a constant multiplicative factor) remained unchanged, and not corrected. In Trichon et al. (1999), the average reflectance value on the satellite imagery, corrected empirically from the atmospheric effect, matched satisfactorily the DART BOA average reflectance, but the textural content of the image could not be compared with the DART BOA simulations.

These results had various implications:

- The sensitivity studies of Bruniquel-Pinel and Gastellu-Etchegorry (1998) were validated at BOA (with airborne data) and therefore their results on texture sensitivity to biophysical parameters should only be applicable to BOA reflectance of forests.
- The sensitivity study on the average reflectance values by Gastellu-Etchegorry et al. (1999), however, may be useful for TOA, because average

values only require a subsequent atmospheric correction for comparison with satellite imagery.

- For future research on reflectance distribution of forests on satellite imagery, the coupling of landscape radiative transfer and atmospheric radiative transfer was probably required. Additionally, non-imaging models possibly cannot yield a correct TOA reflectance distribution of a heterogeneous land cover at 20-30m resolution because they cannot reproduce the result of this atmospheric coupling on image texture.

Since Gascon (2001), the coupling of atmospheric radiative transfer with the BOA imagery in the DART model, as well as the inclusion of sensor impulse response, permit the analysis of reflectance distribution of neighbouring pixels on real satellite imagery. Indeed, the texture attenuation on the satellite imagery due to atmospheric scattering is taken into account.

Another major issue, however, for the validation of simulated reflectance distribution on real imagery, is spatial resolution. Indeed, the validation of a histogram of 50 pixels on 25m resolution imagery (available in the 1990s) represented more than 3 hectares of intensive ground measurements. Moreover, the geo-registration of such satellite imagery with the surveyed parcels was noticed as a problem. For example, Trichon et al. (1999) comment that re-sampling of SPOT (format 1B) imagery may have augmented positional errors between the parcel on the ground and the image pixels. More recently, 1-4m resolution imagery became available, with the perspective of much less ground measurements and visual location of parcels on the imagery. Gascon et al. (2004) and Martin (2006) have explored biophysical parameter inversion using DART on hyper-spatial imagery. Since the major above mentioned issues are nowadays possible to attend, this PhD thesis follows Trichon et al. (1999)'s suggestion that physical modeling should also permit to predict conditions when land cover types are confounded or separable.

4.5 The present contribution to spectral separability studies

In order to attend the variety of eco-systems represented in Mexico, we propose a systematic separability study of forest cover types in view of explaining ambiguities in

remote sensing imagery. This general scheme should ease the prediction of classification errors for an extended and diverse range of eco-systems.

The approach is essentially generic, and examines the utility of a 3D radiative transfer model, with general assumptions on the forest structure - not coupled with a tree architectural model yet - in order to test the correspondence of such general model with the structural information given by 4m multi-spectral imagery, on very different forest types. More specifically, the present work proposes a methodology to efficiently describe the structural and optical properties of the forest canopy and understory at parcel scale, and integrate these elements into the spectral signature of forest stands.

We first tried to test the model on traditional black and white aerial photographs of tropical forests. However, the absence of information on the radiometric response of the camera impeded independent validation procedures. Therefore, chapters 4 and 5 explore the model-based spectral separability study of forest cover types on multi-spectral IKONOS satellite imagery (radiometrically more robust than aerial photographs). The method is validated in chapter 4 and applied on forest stands on slope of major biomes in Mexico. In chapter 5, the method is extended to a contextual classifier and concerns six forest types in the Mexican National Forest Inventory.

Chapter 1: The Sampling Design

A SAMPLING DESIGN FOR THE ASSESSMENT OF LAND COVER MAPS WITH HIGHLY DIVERSE TAXONOMY:

Application to the National Forest Inventory 2000 map on the watershed of the Cuitzeo lake.

The chapter 1 of this thesis shows that existing compatible sampling strategies, previously employed in the USA or planned in Canada are not efficient in the case of the Mexican National Forest Inventory (NFI) map. There is no record so far in the literature of a comprehensive method to assess the accuracy of detailed regional scale Land Use/Land Cover (LULC) maps in the sub-tropical belt. The first objective of this research, treated in chapter 1, was therefore to develop a sampling design capable of attending this taxonomic diversity and distribution.

The elevated bio-diversity and the presence of highly fragmented classes hamper the use of sampling designs commonly employed in previous assessments of mainly temperate zones. A sampling design for assessing the accuracy of the NFI map at community level is presented. A pilot study was conducted on the Cuitzeo Lake watershed region covering 400,000 ha of the 2000 Landsat-derived map. Various sampling designs were tested in order to find a trade-off between operational costs, a good spatial distribution of the sample and the inclusion of all scarcely distributed classes ('rare classes'). A two-stage sampling design where the selection of primary sampling units was done under separate schemes for commonly and scarcely distributed classes, showed best characteristics.

Chapter 1 also illustrates all steps of the assessment scheme employed in this thesis, including a traditional thematic fuzzy component. This scheme is applied to the NFI map of the closed watershed of the Cuitzeo lake, Michoacán state, an area with very high taxonomic diversity.

Sampling design for the assessment of land cover maps with highly diverse taxonomy: application to the National Forest Inventory 2000 map on the watershed of the Cuitzeo lake.[Θ]

Stéphane Couturier

Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)

Centre d'Etudes Spatiale de la Biosphère (CESBIO), CNES/CNRS/Paul Sabatier University

Álvaro Vega

Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)

Jean-François Mas

Instituto de Geografía, Unidad Académica Morelia Universidad Nacional Autónoma de México (UNAM)

Valdemar Tapia

Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)

Erna López

Centro de Investigación en Ecología (CIECO), UNAM

There is no record so far in the literature of a comprehensive method to assess the accuracy of detailed regional scale Land Use/Land Cover (LULC) maps in the sub-tropical belt. The elevated bio-diversity and the presence of highly fragmented classes hamper the use of sampling designs commonly employed in previous assessments of mainly temperate zones. A sampling design for assessing the accuracy of the Mexican National Forest Inventory (NFI) map at community level is presented. A pilot study was conducted on the Cuitzeo Lake watershed region covering 400,000 ha of the 2000 Landsat-derived map. Various sampling designs were tested in order to find a trade-off between operational costs, a good spatial distribution of the sample and the inclusion of all scarcely distributed classes ('rare classes'). A two-stage sampling design where the selection of primary sampling units was done under separate schemes for commonly and scarcely distributed classes, showed best characteristics. A total of 202l secondary sampling units were verified against their NFI map label. Issues regarding the assessment strategy and trends of class confusions are devised.

[Θ] Submitted in Spanish language to the refereed journal 'Boletín de Investigaciones Geográficas' edited by the Institute of Geography, UNAM.

1. Introduction

The accuracy assessment of maps consists in estimating the reliability and making an account of errors. Assessing the accuracy of Land-Use and Land-Cover (LULC) maps is a common procedure in geo-science disciplines, as a means, for example, of testing automatic classification methods on satellite imagery. For regional scale LULC maps, the abundance and distribution of categories (or classes) over the large extension of the map, confronted with tight budget constraints, add complexity to accuracy assessments. Only relatively recently have comprehensive accuracy assessments, with estimates for each class, been built and applied to regional or continental LULC maps; Examples are accuracy assessments of Landsat-derived 1992 LULC maps conducted in the United States of America (USA) by Laba et al. (2002) and Wickham et al. (2004), and the programmed accuracy assessment of the Canadian forest inventory of year 2000 (Wulder et al., 2006).

The National Forest Inventory (NFI) 2000 map of Mexico, sponsored by SEMARNAP, the Ministry of the Environment in Mexico, was conceived and jointly coordinated by the Institute of Geography of UNAM (the National Autonomous University of Mexico) and INEGI, the National Institute of Statistics, Geography and Informatics in Mexico. The cartography was generated in the Geography Institute at UNAM based on visual interpretation at scale 1:250 000 of Landsat imagery of year 2000 (Mas et al. 2002). The aim of the 2000 NFI cartographic project was to update previous LULC maps, especially the 1993 cartography known as the 'serie II' INEGI maps, based on visual interpretation of Landsat images acquired in year 1993. A main trait of the NFI project was to propose a hierarchical classification scheme as a convenient LULC standard in Mexico for future remote-sensing based cartography. Four aggregation levels were considered; the finest taxonomic resolution (denominated community level, including sub-communities) consisted in 75 classes (Palacio-Prieto et al., 2000), a much larger number, for example, than the 21 classes which constitute the legend of fine taxonomic resolution LULC maps (Wickham et al., 2004) in the USA. The taxonomic resolution of the NFI at community level is comparable to the subclass level of the National Vegetation Classification System (NVCS) of the USA (see FGDC, 1997). However, the Mexican NFI map was conceived at

much coarser spatial resolution (1:250 000 scale) than the LULC projects in the USA (spatial resolution close to the Landsat pixel). The differences in taxonomic diversity and spatial resolution, mentioned above, are both likely to have an incidence on the way the accuracy of the NFI map can be assessed.

A preliminary assessment was conducted immediately after the elaboration of the NFI map, based on the interpretation of aerial photographs. However, confidence levels could be reported only for a few classes, abundant in the northern part of the country (Mas et al. 2002). The research described in this paper is motivated by the conception of a full assessment design capable of estimating confidence levels of the total amount of classes and for the entire country.

In the first place, the high number of classes at community level (75) implies the presence of many sparsely distributed classes (rare classes) on the map. This situation complicates a probabilistic sampling scheme (*sensu* Stehman y Czaplewski, 1998) which combines both cost-efficiency and statistical validity. A first objective of this paper was to find a design where the entire set of rare classes could be sampled under given cost constraints; we tried a few sampling strategies which had been successfully applied to the LULC maps in the USA, and compared their performance with novel sampling strategies.

In the second place, the assessment of the fine taxonomic resolution NFI map is confronted with difficulties related with the thematic complexity of Mexican landscapes:

- Presence of tropical and temperate environments
- Parcels of slash and burn agriculture
- Highly fragmented landscapes
- Temporal transition between different vegetation types (e.g. induced pasture to second growth forest)
- Spatial transition between different vegetation types (e.g. eco-tones)
- Vegetation types whose classification is ambiguous.

The occurrence of each of these situations hardens the task of unambiguously labeling a polygon as one class in the process of interpreting Landsat imagery (during map production) as well as in the process of interpreting aerial photographs (during map evaluation). Specifically, the abundance of forest classes in the NFI map augments the risk of ambiguity among various class labels during interpretation of eco-tones. As recommended by Laba et al. (2002), we opted for a fuzzy characterization of landscapes during the process of reference material production for the evaluation of the NFI map. We propose a simplified version of the linguistic fuzzy evaluation introduced by Gopal and Woodcock (1994). This version has the double advantage of being user-friendly for a specialist photo-interpreter, and compatible with automatic procedures in a Geographic Information Systems (GIS), for subsequent computation of accuracy indices.

We tested the sampling designs and the fuzzy evaluation with the NFI map in the watershed of the Cuitzeo lake, state of Michoacán, Mexico (see figure 1). This area is extended on 400,000 hectares, contains 21 classes at community level and includes landscapes with all above mentioned situations, therefore illustrating the complexity of the NFI map.

In the methods section, we present the evaluation strategy, with an emphasis on simulations of various sampling designs. The comparison of their performance, in the results section, allowed us to analyze the rationale behind an optimal design and select one design for the evaluation of the map in the pilot area. Another subsection in the method section describes, with illustrations, the verification design with the fuzzy description of photo-interpreted polygons. The derivation of the confusion matrix and accuracy indices is then detailed in the subsection on the synthesis of the evaluation, taking into account the characteristics of the sampling design. The pattern of confusions registered in the pilot area is then discussed in view of the characteristics of the design. Perspectives of the sampling design are finally devised.

2. Study area and materials

The Cuitzeo lake, second in size in Mexico, is located on the transversal volcanic axis in central western Mexico. The closed watershed of the Cuitzeo lake (area under investigation) is located to the North of the Michoacán state (figure 1.1), and is covered with temperate sub-humid or tropical dry vegetation. The original vegetation was mainly forests of oak, oak-pine (mixed), and pine towards the South-Southwest, and tropical scrubland towards the North-Northeast. Major land uses in year 2000 according to the NFI are annual rain-fed agriculture, irrigated agriculture and induced grassland for pasture (see table 1.1). Apart from the diversity of LULC classes, a key motivation for selecting this pilot area was the availability of detailed reference data exploitable in the process of independently verifying the 2000 Landsat-based NFI. A complete coverage of 244 black and white aerial photographs on paper at scale 1: 37 000 was acquired in 1999 over the Cuitzeo area, therefore quasi synchronous with the Landsat imagery.

Figure 1.1 Location map of the watershed of the Cuitzeo lake

The watershed is the gray shaded area.

Table 1.1 Distribution of classes of the National Forest Inventory map in the Cuitzeo lake watershed.

Class code	Class name	Formation	Area fraction	Area (km2)
100	Irrigated crop	Cropland	0.1411	564.97
110	Hygrophilous crop		0.0048	19.04
200	Perennial crop		0.0021	8.27
210	Annual crop		0.2356	943.14
300	Forest plantation		0.0071	28.24
410	Fir forest	Temperate Forest	0.0037	14.72
420	Pine forest		0.0041	16.32
421	Pine forest & Sec Veg[a]		0.0036	14.31
510	Oak-pine forest		0.0958	383.34
511	Oak-pine forest & Sec Veg[a]		0.0325	130.29
600	Oak forest		0.0232	92.88
601	Oak forest & Sec Veg[a]		0.0553	221.54
700	Cloud forest	Tropical forest	0.0029	11.73
920	Sub-trop scrubland	Scrubland	0.0194	77.58
921	Sub-trop scrubland & Sec Veg[a]		0.0768	307.25
1000	Mezquital		0.0004	1.51
1200	Chaparral[b]			
1330	Induced grassland	Grassland	0.1594	638.04
1410	Hygrophilous grassland	Hygrophilous Veg	0.0209	83.50
1510	Halophilous vegetation		0.0069	27.78
1600	No apparent vegetation cover[b]	Other Cover		
1700	Human settlement		0.0250	100.02
1800	Water	Water	0.0796	318.75
All			1.0000	4003.23

[a] Secondary Vegetation (secondary vegetation, herbaceous and/or shrub-like vegetation, according to NFI classification scheme). [b] This class was not present on the map but appeared as a reference label (see table 1.3).

The following reference data were also gathered and processed:

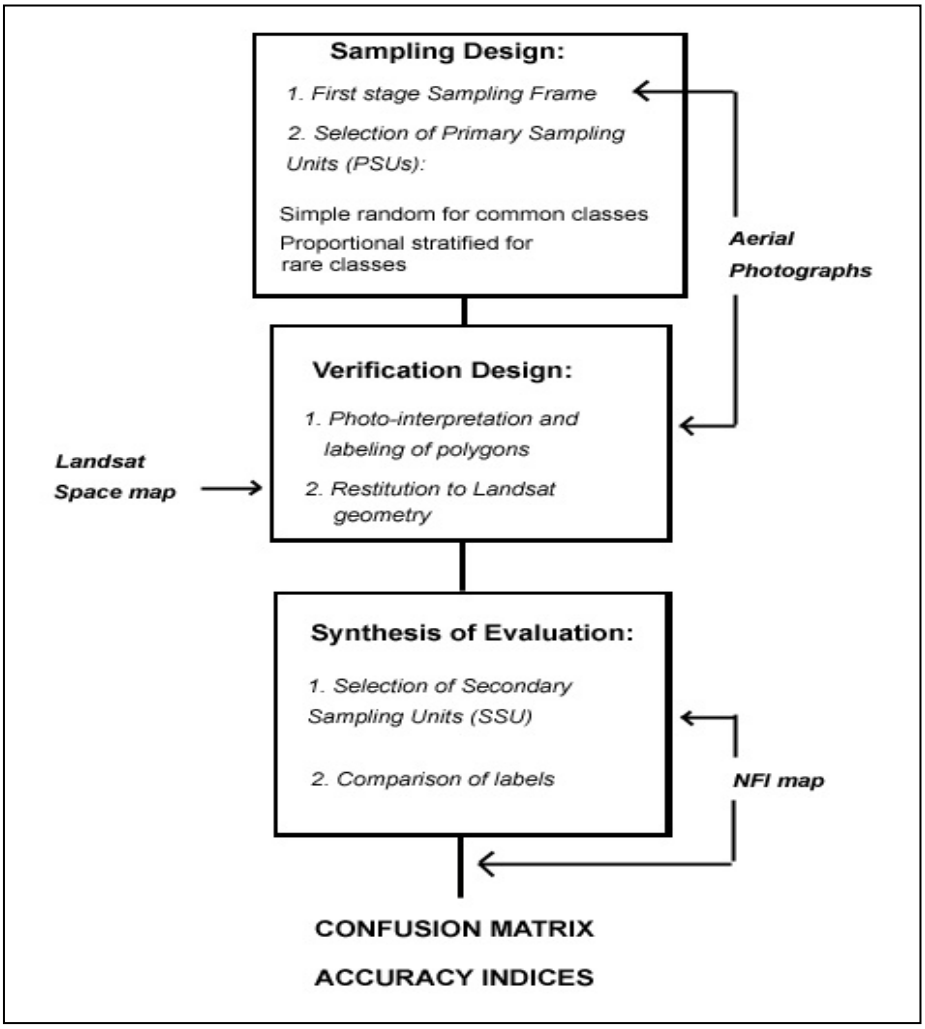
- Twelve topographic INEGI maps at 1:50 000 were scanned.
- One colour composite (bands 541) of a Landsat image mosaic of year 2000 was prepared over the watershed, georectified against the INEGI 1:50 000 topographic maps.

3. Methods

The algorithm adopted for the accuracy assessment, represented in figure 1.2, is organized around three steps:

1. Sampling design: selection of representative reference sites.
2. Verification design: description of a reference site, and labeling of the reference sites by means of interpretation of aerial photograph.
3. Synthesis of the evaluation: Comparison between the information contained in the reference site and the information contained in the map, construction of a confusion matrix and computation of assessment indices.

Figure 1.2 Algorithm for the accuracy assessment

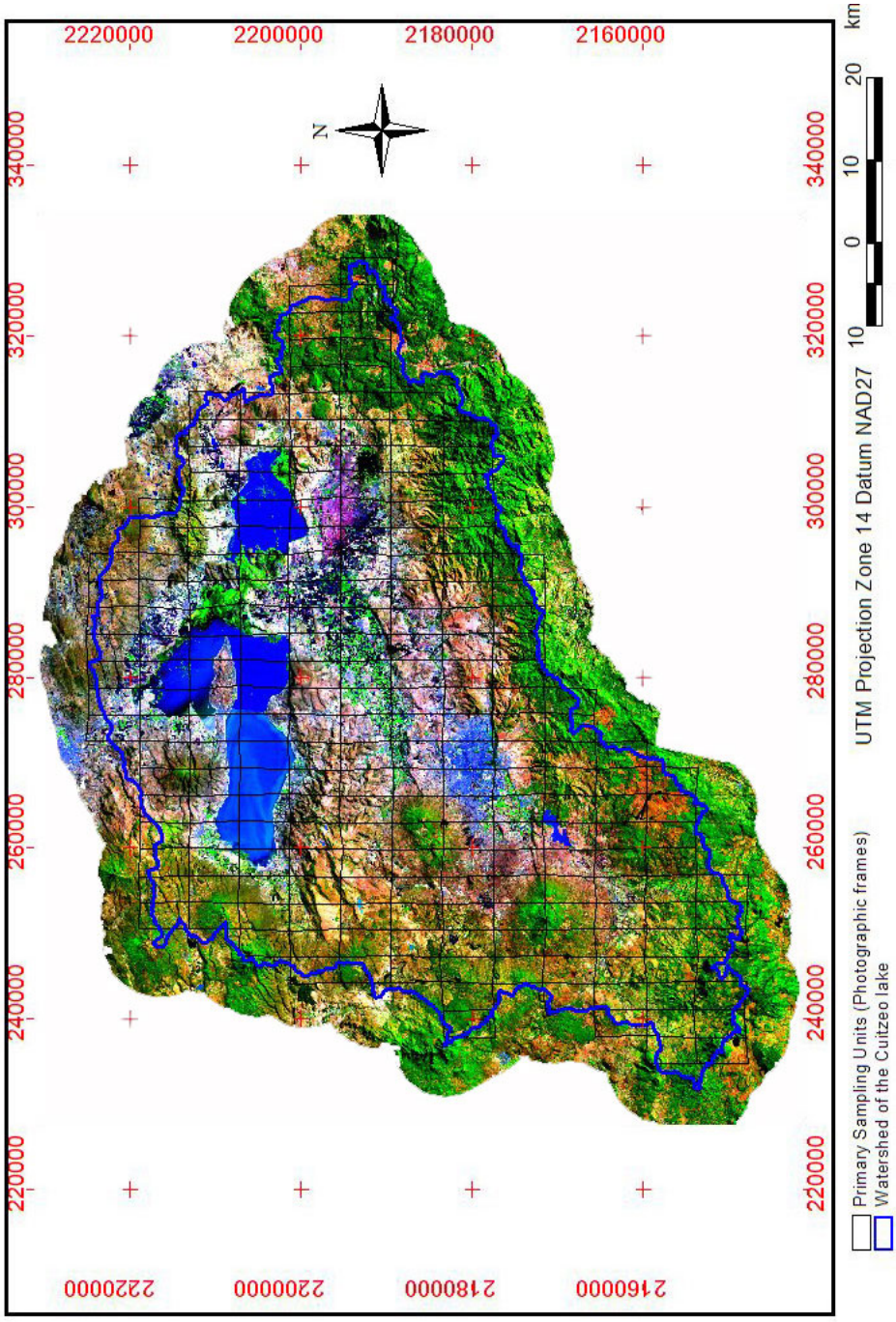


3.1 Sampling designs

The sampling design of a regional accuracy assessment aims at estimating accuracy for each class on the map. The selection of reference sites must be such that a statistically representative number of sites covers each mapped class of table 1.2. Five strategies were contemplated and tested: a first design consisted in a simple random sampling (Simple Sampling, or SS), stratified by class, whereby each class was sampled randomly over the entire territory. Four other designs consisted in two stage, or double, sampling (Double Sampling, or DS). A double sampling design is operated in two steps, first with the selection of Primary Sampling Units (PSUs), and then with the selection of Secondary Sampling Units (SSUs). Conveniently, a PSU corresponds to the effective area of an aerial photograph (photographic frame). The SSUs are dots of a regular grid within the selected PSUs.

Figure 1.3 shows the complete set of PSUs, which was constructed from the grid of all photographic center points, geo-referenced onto the INEGI topographic maps at 1:50,000. The space map represented in figure 1.3 is derived from a mosaic of four Landsat images of year 2000, that were the basis for the NFI map production (Mas et al., 2002). These Landsat images had also been geo-referenced to the INEGI topographic maps at 1:50,000. A photographic frame pertained to the set of PSUs if its center point was included in the watershed. In this way, the random character of the selection was preserved at the edge of the watershed (see Zhu et al., 2000). 218 photographic frames formed the complete set of PSUs (figure 1.3). The four double sampling (DS) designs are distinguished by the selection mode of a PSU, which in turn defines the probability of inclusion of its surface in the sample. For example, a simple random selection is associated with a uniform (constant) inclusion probability among all PSUs of the complete set. The computation of the accuracy index is governed by the probability of inclusion of each sampled point (selected SSU), as detailed in the section on the synthesis of the evaluation.

Figure 1.3 Space map of the watershed of the Cuitzeo lake



The complete set of Primary Sampling Units (photographic frames) is represented. The colour composite is extracted from a mosaic of 4 Landsat images (band composition 541), geo-rectified to INEGI topographic maps at 1:50 000. The grid of photographic frames was processed from the set of geo-referenced photographic center points.

The first double sampling design (DS1) is defined by the simple random selection of the PSUs, as in Laba et al. (2002). DS2 is characterized by the random, stratified by class, selection of PSUs, as in Stehman et al. (2003). DS3 is defined by a proportional, stratified by class, selection of PSUs. For the latter design, not applied in previous published research, the probability of inclusion of a PSU is proportional to the abundance of the class in the PSU. The abundance of a class is its surface fraction, obtainable via easy attribute computation in a GIS. Then, in order to ease accuracy calculations, the probability of inclusion of SSUs at the second stage is defined as being inversely proportional to the abundance of the class in the PSU. Proportional sampling is a known statistical technique (Cochran, 1977) and some characteristics of its application to map accuracy assessment are devised in Stehman et al. (2000). However, DS3 has never been applied in published studies, maybe because it was not necessary for maps with classification systems of mainly temperate countries. Finally, an entirely novel, hybrid design (DS4) includes a simple random selection of PSUs (as in DS1) for common classes (area fraction above 5%, 7 classes in Cuitzeo), and a proportional stratified selection of PSUs (as in DS3) for rare classes (area fraction below 5%, 14 classes in Cuitzeo). After selection of the PSUs, the sample size of SSUs was fixed at 100 per mapped class, a value widely adopted in similar assessments (see Stehman and Czaplewski, 1998).

Under these conditions, it was possible to test the efficiency of the sampling designs. According to the binomial distribution theory, the confidence interval of the accuracy estimate depends on the sample size and on the reliability value (accuracy estimate) in the following manner (Snedecor and Cochran, 1967, cited by Fitzpatrick-Lins, 1981):

$$d^2 = t^2 \, p \, (1-p) / n \qquad (1\text{-}1)$$

Where d is the standard deviation (or half the confidence interval) of the estimate, t is the standard deviate on the Gaussian curve (for example, $t = 1.96$ for a two-sided probability of error of 0.05), p is the reliability value, and n is the number of sampled points. The standard deviation of the accuracy estimate was calculated for each class, as an indicator of the efficiency of the sampling design.

3.2 Verification Design

The verification design is based on the interpretation of aerial photographs according to fuzzy rules. A first photo-interpreter was assigned to delineate homogeneous polygons at scale 1:125 000 (finer resolution than the NFI map, and suitable for photo-interpretation) on transparent slides over the photographic material. This delineation was done on the photographic frames corresponding to the selected PSUs. Next, another interpreter digitized the photo-interpreted polygons over the Landsat geo-rectified space map. This step was meant to minimize the interference of fictitious positional errors between the NFI map and the material of reference (see a similar technique in Zhu et al., 2000). The comparison of numerous repeated trials ensured that a good control over positional errors was achieved. Indeed, the skill of restitution on screen was acquired by the second interpreter from previous research, where photo-interpreted maps had to be digitized on screen over ortho-photographs or scanned topographic maps. This exercise usually serves the purpose of transferring stereo-pair interpretation work (a valuable traditional skill of physical geographers and technical foresters) onto a digital geo-referenced GIS. Figures 1.4 and 1.5 respectively illustrate the process of polygon delineation on a photographic frame, and the process of restitution on screen of the photo-interpreted polygons.

Each sample point (selected SSU) was located on the colour composite of the geo-rectified Landsat mosaic on screen. The location of the point was then visually transferred onto the interpreted transparent slide over the photographic material, with the aid of the digitised polygons on screen.

A linguistic fuzzy qualification (categories of Gopal and Woodcock, 1994) was attributed to each polygon by the first photo-interpreter:

5. Absolutely right: This class corresponds perfectly to the polygon

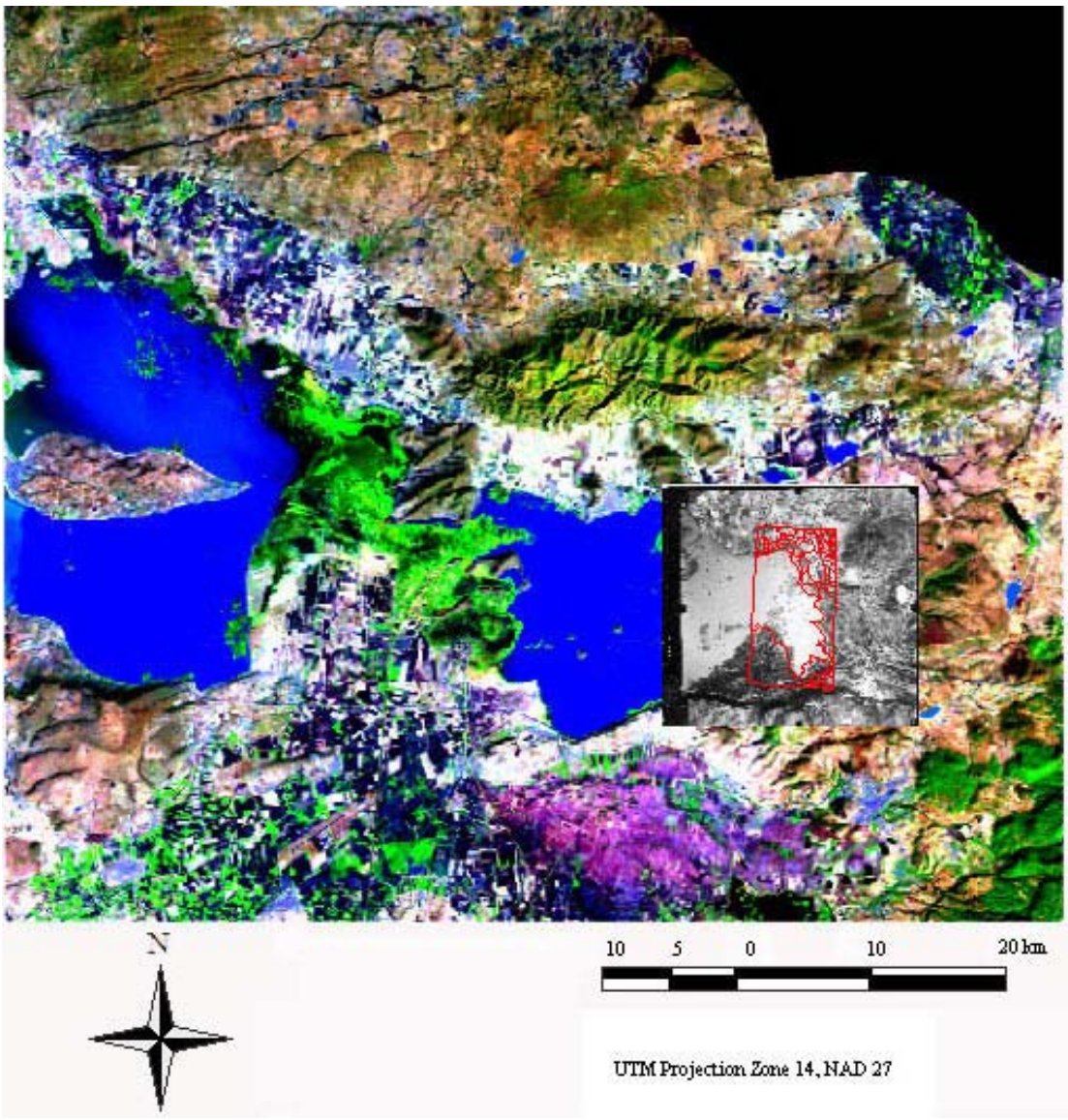
4. Good: This class corresponds well to the polygon

3. Reasonable or acceptable: This class is not the best option but is acceptable for the polygon: Right

2. Understandable, but wrong: This class would be a bad label for the polygon

1. Absolutely wrong: This class has nothing to do with the polygon

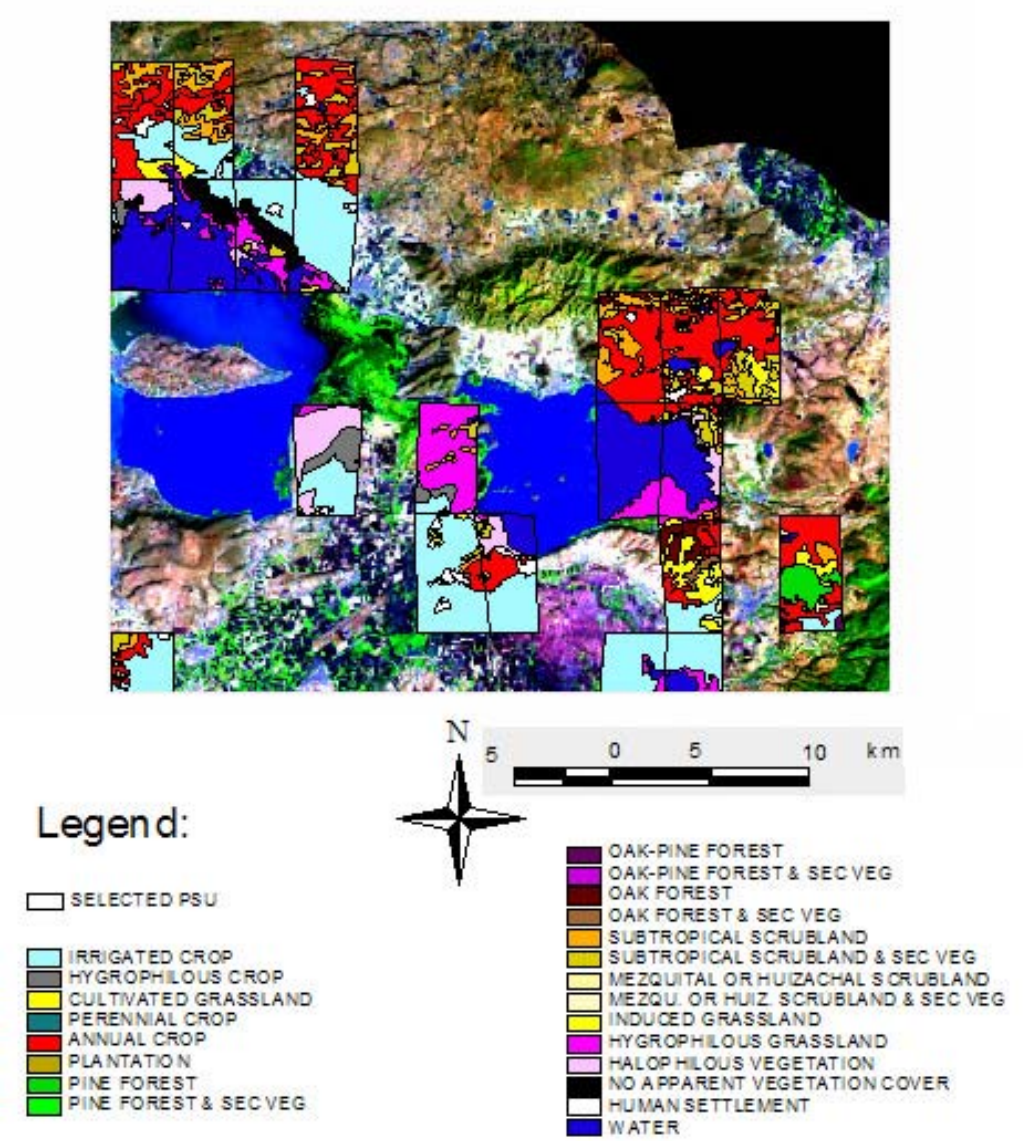
Figure 1.4 Aerial photograph of the side of the Eastern Cuitzeo lagoon, overlaid onto the watershed space map.



The superimposed photographic frame is segmented into reference polygons. The colour composite is extracted from a mosaic of Landsat images (band composition 541), geo-rectified to INEGI topographic maps at scale 1:50 000.

.

Figure 1.5 Primary Sampling Units and watershed space map in the area of the Eastern Cuitzeo lagoon



The selected PSUs contain the reference information obtained from the verification process. The geometric restitution of the photo-interpreted material was done on screen, with the aid of the INEGI topographic maps at scale 1:50 000 and of the Landsat-based space map..

In order to simplify the work of the photo-interpreter in the presence of numerous classes, the qualification was done only for the four classes which most probably described the polygon. This restriction permitted the creation of an attribute table containing qualified class codes (appearing in table 1.2), exploitable by automatic procedures in a GIS.

Finally, a confidence rating was associated with the photo-interpretation:

4. Very high
3. High
2. Average
1. Low

Prior to all the work described, a third interpreter had independently built a LULC map at 1:50,000 of the entire watershed based on the same set of 244 aerial photographs, and with the aid of intensive field work (López et al., 2006). When the confidence rating of the first interpreter was 1 or 2, his interpretation was confirmed or corrected according to the LULC map at 1:50,000.

3.3 Synthesis of the evaluation

In all sample points (selected SSUs), a comparison between fuzzy labels of the reference material and the label of the NFI map was made via GIS, allowing a spatial tolerance of 500m around the SSU in order to take into account the scale of the map (1:250 000). The derivation of the global accuracy index is detailed here for the fuzzy evaluation with sampling design DS4, selected for the present analysis.

For any class k, a double sampling was applied. Following the Bayesian rule, the probability of inclusion of a sample point p_{2k} is a multiplicative function of the inclusion probability p_{1k} of the PSU it pertains to, and of the inclusion probability of the SSU once the PSU has been selected $p_{2|1}$ (conditional inclusion probability):

$$p_{2k}=p_{2|1} * p_{1k} \tag{1.2}$$

In the case of rare classes, the proportional stratified sampling is constructed in such a way that equation (1.2) simplifies into the inclusion probability of a simple stratified sampling (Stehman et al., 2000): $p_k = f_k$, where f_k is the frequency of class k on the map. In the case of common classes, the first stage probability p_{1k} is $p_{1k}= K/N$, where K is the fixed number

of randomly selected PSUs for common classes and N is the total number of PSUs. $p_{2|1}$, the conditional probability, is equal to n_k/N'_k where n_k is the sample size and N'_k is the number of SSUs of class k in the K selected PSUs. Therefore, following equation (1.2), differential weights of $f_k * K/N$ were attributed to common classes when global estimates were computed.

4. Results

4.1 Sampling designs

With the target sample size of 100 per mapped class, we simulated 10 distributions of sample points for every one of the five sampling designs and examined the results. Some characteristics of the distributions are synthesized in table 1.2, in order to evaluate the efficiency of each design.

The strategy of simple sampling (SS) yielded a dispersion of the sample points over more than 85% of the PSUs (see table 1.2). As in other regional accuracy assessments, simple random sampling stratified by class is judged cost-prohibitive (here, in terms of purchase of photography and photo-interpretation labour). Accounting for the costs of operation, the reasonable number of photographic frames was fixed to a maximum of one fourth of the total coverage of the pilot area (one fourth of 218, or 54 photographic frames). Under this condition, the problem of non-statistically representative sample sizes was likely to occur. We required that the confidence interval for any class (equation (1.1)) be less than 15% in order to consider the sample as statistically representative for this class.

The simulated distributions under DS1 left at least three (usually rare) classes with insufficient sample sizes (table 1.2). This design is known for its good capacity of sample dispersion, counterbalancing the clustering effect of the double sampling. For this reason, Stehman et al. (2003) use it for the accuracy assessment of the National Land-Cover Data (NLCD) 1992 en EUA. However, the design was discarded for the pilot study on the Mexican NFI because it cannot handle satisfactorily the high taxonomic diversity (large number of rare classes) in the classification system.

Under DS2, the selection of PSUs is separately operated for each class, and all PSUs where class k is present have the same probability of inclusion for class k. In such a

design, if a PSU only contains one point (SSU) of class k, this PSU has equal probability of being selected than another PSU that contains 50 points (SSUs) of class k. If class k is highly fragmented, a high number of PSUs is therefore necessary in order to statistically cover class k. For this reason, simulations showed that more than 80 PSUs were necessary to cover the 21 classes, a quantity which overrides the admitted 54 photographic frames. Although less problematic than DS1 according to the number of underrepresented classes (table 1.2), the DS2 design remains inadequate in this pilot study of the NFI accuracy assessment, presumably because of the presence of highly fragmented classes. In addition, the computation of inclusion probabilities in the second stage of the design proved complicated to handle via automatic GIS procedures.

Table 1.2: Evaluation of five sampling designs on the National Forest Inventory map in the watershed of the Cuitzeo lake.

	Total number of selected PSUs[a]	Number of classes with non-representative sample size (confidence interval superior to 15%)	Average number of SSUs[b] per PSU[a] among rare classes	Average number of SSUs[b] per PSU[a] among common classes
SS[c]	167-193	0	14-20	9-13
DS[d] 1	54	3-8	21-25	14-21
DS[d] 2	54	2-4	12-23	19-25
DS[d] 3	54	0-4	20-33	26-36
DS[d] 4	54	0	17-23	16-22

[a]Primary Sampling Unit [b]Secundary Sampling Unit
[c]Simple (one-stage) Sampling [d]Double (two-stage) Sampling

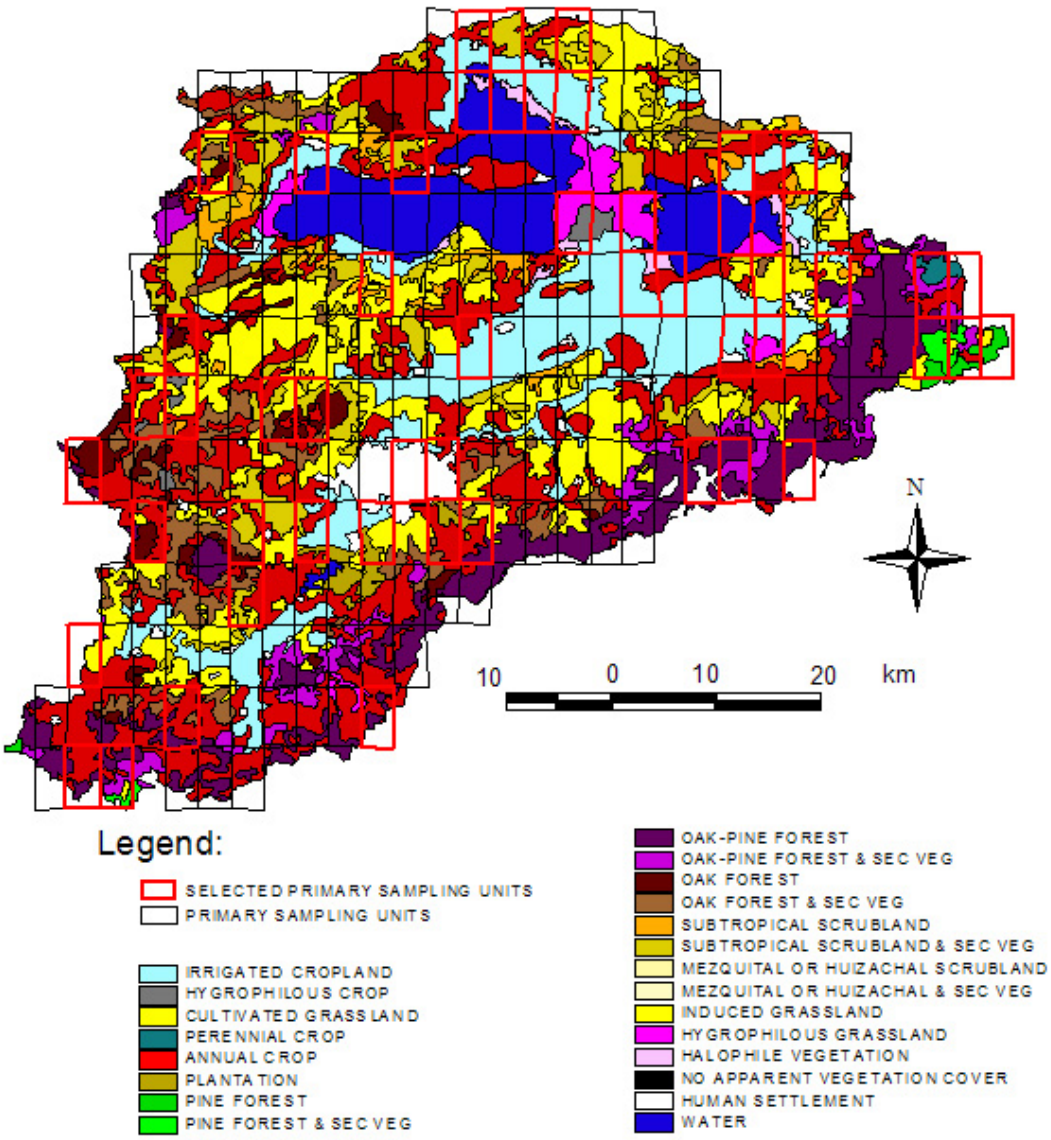
Ten simulations were applied for each of the five sampling designs: SS is a simple random, stratified by class sampling. DS1 consists in a simple random selection of Primary Sampling Units (PSU). DS2 consists in a random, stratified by class selection of PSUs, DS3 is a proportional stratified selection, and DS4 treats separately common classes (to which DS1 is applied) and rare classes (to which DS3 is applied).

By contrast, proportional sampling in DS3 maintains a low level of complexity in the calculations of accuracy indices (i.e. random stratified formulae; see Stehman et al., 2000:604) because all sample points of class k have equal probabilities of inclusion, whatever the PSU they pertain to. However, as in DS2, the selection of PSUs also operates separately for each one of the 21 classes.

With the constraint of 54 photographic frames, preliminary simulations incited us to limit PSUs to three per class. We observed that three PSUs were generally sufficient to sample a majority of rare classes, and approach their real distribution pattern. However, the limitation to three PSUs for common classes generated much spatial clumping of the sample, with respect to the real distribution pattern. These observations are illustrated in the last two columns of table 1.2: the average amount of sample points (SSUs) per PSU raises to 26-36 in comparison with much lower values in simulations of previous designs. In the presence of much spatial clumping, the precision of the estimates is likely to be affected by the spatial autocorrelation of errors.

DS4 was designed with the objective of benefiting from the good spatial dispersion of DS1 (a characteristic which was exploited for the LULC map assessment in the USA) for common classes, and from the advantages of DS3 (inclusion of all classes and easy estimate computation) for rare classes. Effectively, the selection of a group of PSUs for the entire set of 7 common classes permitted a dispersion of the sample comparable to the results of DS1 (see table 1.2). Overall, visual inspection of the results indicated that DS4 yielded the best spatial distribution of samples among double sampling strategies. At the same time, the number of PSUs could be extended to more than three for those rare classes which did not obtain a sufficient sample size in the simulations of DS3, maintaining the total number of PSUs at 54. In this way, all classes were guaranteed a statistically valid sample size. A sequence of ArcView and Visual Basic (for Excel) routines, operating on attributes of the sample points, was developed for proportional sampling in both stages. The figure 1.6 shows the set of selected PSUs as the result of a simulation using DS4. The figure 1.7 shows the dispersion of sample points (selected SSUs) in the area of the Cuitzeo Eastern lagoon. These data were used to compute the accuracy indices.

Figure 1.6 Selected Primary Sampling Units in the watershed of the Cuitzeo lake



The Primary Sampling Units were selected using the hybrid sampling design. The hybrid design is based on simple random sampling of PSUs for common classes and on proportional stratified sampling of PSUs for rare classes. The National Forest Inventory map is laid in the background.

-

4.2 Main confusion patterns of non-forest LULC classes

The global confidence level and confusion patterns among NFI classes are reported in a confusion matrix (table 1.3). The same conventions as matrices and indices employed in Stehman et al. (2003) were retained since these conventions permit a coherent analysis of confusion patterns among classes, based on comparable area fractions. The global confidence level of 81.3% mainly reflects the high accuracy (89%) of the dominant class 'annual crop'. LULC classes such as Water and Human settlement recorded very high confidence levels (more than 94%). Indeed, a high spectral separability with respect to other classes characterises these LULC classes on conventionally-used Landsat colour composites (see figure 1.3). Conversely, very high confusion among aquatic vegetation covers (hygrophilous grassland and halophilous vegetation) is evident from the matrix.

The spectral ambiguity and variability (across inundation phases) of these aquatic vegetation types was one of the key explainers of this observed high confusion. Examining the sites of these errors on the 'serie II' INEGI maps, we observed that the NFI 2000 map was generally reporting false changes with respect to the previous information. Among non-treed vegetation covers, much of the mapped induced grasslands, and some of the mapped 'secondary' oak would be cropland in reality. In turn, much of the mapped cropland should be scrubland, with the subtraction of a small reciprocal confusion pattern.

Altogether, if we compare area fractions of producer's totals to area fractions of user's totals, the presence of scrub (with or without 'secondary' vegetation) is underestimated by 3.5% of the entire region. This result confirms data gathered by Lopez et al. (2006), according to whom the NFI 2000 under-estimates the withdrawal of non-pastoral and pastoral agricultural activity in the region, towards the regeneration of scrub vegetation communities

4.3 Incidence of the fuzzy characterization of forests

The common class oak-pine forest (98%), and rare classes such as sub-tropical scrubland (86%), oak forest (95%), pine forest (81%) and fir forest (84%) registered high accuracy.

Therefore, the 'primary' vegetation classes present in the area are generally characterized by a low level of confusion. Exceptions reside in the sole mapped patches of mezquital and cloud forest in Cuitzeo, which both corresponded to other classes on the ground (0% accuracy, double-verified with ground visits). These NFI errors clearly appeared to be omission on respectively reporting change and rectifying an error from INEGI former maps of the area.

By contrast with the good agreements obtained for 'primary' temperate forest classes, numerous confusions were recorded among classes denominated as temperate forests with 'secondary' vegetation (the NFI classification scheme assimilates forests with 'secondary vegetation, herbaceous, and/or shrub-like vegetation' under this label of 'secondary' forests):

According to the reference photo-interpreted data, 'secondary' pine forest was wrongly mapped as 'secondary' oak, 'secondary' mixed (oak-pine) forest was wrongly mapped as 'secondary' pine, and oak forest was wrongly mapped as 'secondary' mixed forest. These confusions are mainly responsible for the low and very low accuracy levels registered for 'secondary' oak (54%), 'secondary' pine (10%) and 'secondary' mixed (46%) forests. Various interpretations may arise from this observation. If we assume that the reference data reflects reality, much area of these 'secondary' forests lays on eco-tones more difficult to discern than for 'primary' mixed forests. Indeed, the more pronounced presence of human activity augments the class membership value of these areas. Depending on the map user's needs, an upper-threshold on membership values could fix the thematic tolerance of the accuracy assessment. Another interpretation for the low accuracy levels could be related to the positional accuracy of boundaries during polygon delineation. In this sense, the accuracy assessment is also dependent on the width of spatial transitions between clearly distinguished landscapes. This width could correspond to a tolerance level fixed by the user of the map.

Figure 1.7: Selected Secondary Sampling Units around the Eastern Lagoon of the Cuitzeo lake

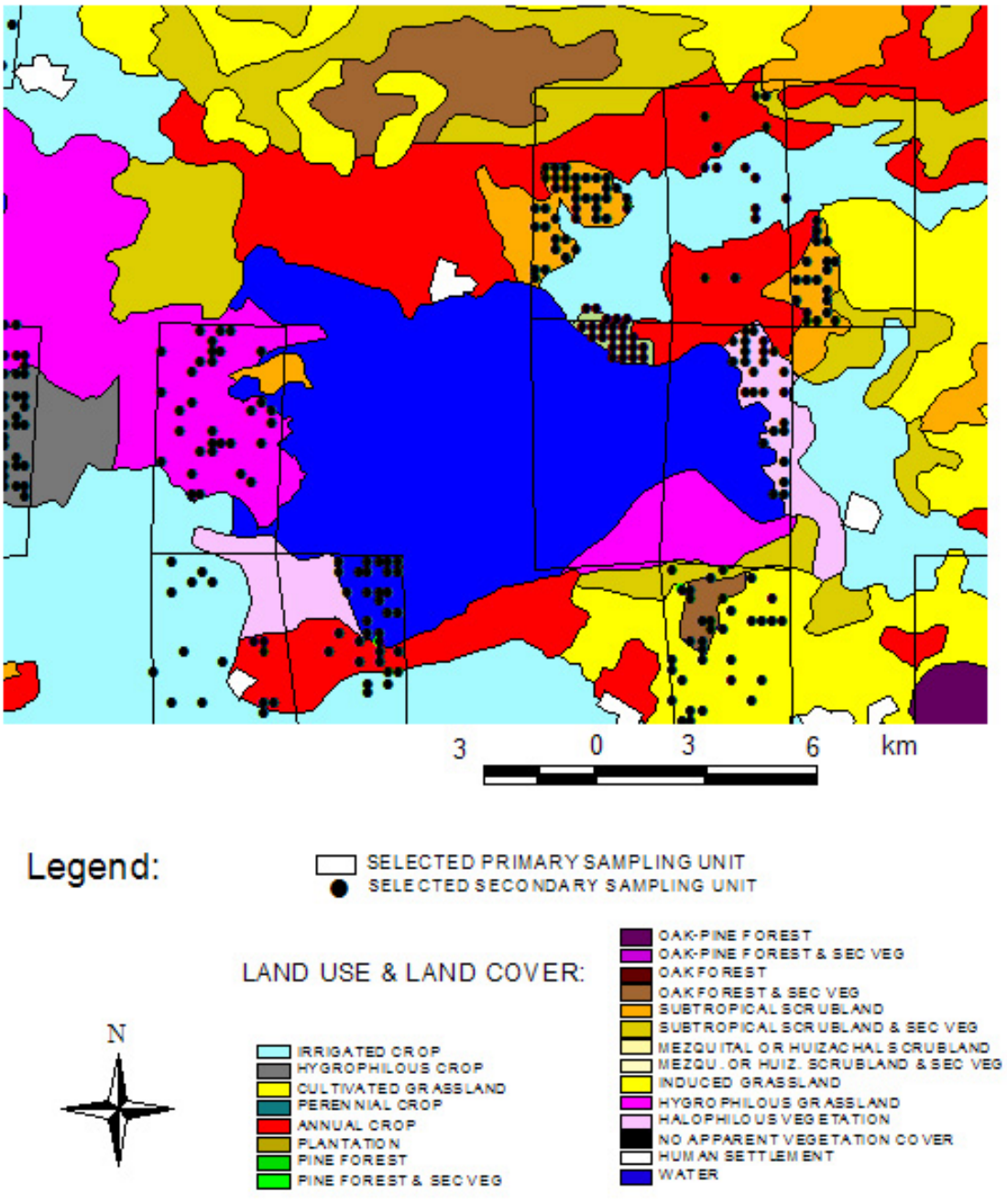


Table 1.3: Confusion matrix for the fuzzy accuracy assessment of the National Forest Inventory map at community level in the watershed of the Cuitzeo lake.

	100	110	200	210	300	410	420	421	510	511	600	601	700	920	921	1000	1200	1330	1410	1510	1600	1700	1800	n^a	Area frac[b]	User[c]	StdE[d]
100	12.4			0.56										0.14	0.14				0.28			0.28	0.28	100	14.1	0.88	0.03
110		0.39		0.04										0.01	0.01					0.01		0.01	0.01	100	0.48	0.82	0.04
200			0.20																			0.01		100	0.21	0.99	0.01
210				21.0							0.24			0,71	1.18							0.47		100	23.6	0.89	0.03
300				0.01	0.59				0.04						0.01		0.06							100	0.70	0.84	0.04
410						0.31			0.05					0.01				0.01						100	0.37	0.84	0.04
420				0.01			0.33		0.03	0.03								0.01						100	0.41	0.81	0.04
421				0.03			0.01	0.04	0.12	0.15								0.01					0.01	100	0.36	0.10	0.03
510								0.10	9.38	0.10														100	9.58	0.98	0.01
511				0.19						1.50	1.20			0.03				0.32						100	3.25	0.46	0.05
600				0.05						0.02	2.20			0.05										100	2.32	0.95	0.02
601				0.50	0.11			0.66	0.17	0.33		2.99		0.50	0.17							0.11		100	5.53	0.54	0.05
700				0.02	0.01		0.22	0.03		0.01			0.00		0.02									100	0.29	0.00	0.00
920				0.23										1.67				0.04						100	1.94	0.86	0.03
921				0.23							0.15			0.08	7.06							0.15		100	7.67	0.92	0.03
1000				0.04											0.01	0.00								23	0.04	0.00	0.00
1200																								0	0.00		
1330	0.32			2.07							1.12	0.16		0.96				11.0				0.32		100	15.9	0.69	0.05
1410	0.10	0.08													0.02				0.98	0.67			0.23	100	2.09	0.47	0.05
1510		0.01		0.01											0.05			0.01	0.17	0.28	0.14		0.03	100	0.69	0.40	0.05
1600																								0	0.00		
1700																						2.50		100	2.50	1.00	0.00
1800				0.24										0.08				0.16					7.48	100	7.96	0.94	0.02
n^a	95	87	99	194	87	84	158	32	163	103	142	55	0	113	120	0	8	88	74	73	20	114	114	2023			
Area frac[b]	12.8	0.48	0.20	25.2	0.71	0.31	0.56	0.82	9.79	2.13	4.91	3.15	0.00	4.22	8.66	0.00	0.06	11.5	1.44	0.95	0.14	3.85	8.04		100.0		
Prod[c]	0.97	0.81	1.00	0.83	0.84	1.00	0.59	0.04	0.96	0.70	0.45	0.95		0.40	0.82		0.00	0.95	0.68	0.29	0.00	0.65	0.93			81.3	
StdE[d]	0.02	0.04	0.00	0.03	0.04	0.00	0.04	0.04	0.02	0.04	0.04	0.03		0.05	0.04		0.00	0.02	0.05	0.05	0.00	0.04	0.02				

The global confidence level was estimated at 81.3%. Entry numbers refer to class codes described in table 1.1.
[a] Sample size. [b] Class area fraction (out of the total Cuitzeo area), sum of the row/column matrix elements. [c] User's (User) and Producer's (Prod) accuracy
[d] Standard Error using binomial theory, with approximation to large sample sizes if applicable.

The results presented here with the linguistic fuzzy rules reflect the confidence levels at one tolerance level. Besides a satisfactory global accuracy, these results revealed the difficulty in classifying secondary vegetation types in the Cuitzeo region, with the type of remote sensing imagery available for map production.

5. Summary and conclusion

No comprehensive method has been published on the accuracy assessment of regional scale LULC maps in sub-tropical countries. We present a hybrid sampling design for the evaluation of LULC maps with high taxonomic diversity and highly fragmented landscapes. This and other four probability sampling designs - some of them previously employed for mainly temperate countries - were tested for the accuracy assessment of the Mexican National Forest Inventory (NFI) map at community level. Each sampling design was evaluated on the watershed of the Cuitzeo lake, a pilot study where a high number of classes is represented. As expected, the double sampling (DS) designs achieved much more control over the spatial distribution of the sample than a simple, random stratified sampling, thereby permitting the control over costs of operation. With fixed operational costs, the only design that systematically provided statistically representative estimates for all classes was the hybrid design (DS4). Additionally, the hybrid design achieved a spatial dispersion of the sample similar to the dispersion achieved by DS1, with simple random selection of Primary Sampling Units (PSUs). DS1 is known for generating a good dispersion of the sample in regional map assessments. For this reason, DS1 was successfully applied in the accuracy assessment of the National Land Cover Data (NLCD) in the USA (Stehman et al., 2003). However, DS1 was discarded in our pilot study because it was not able to handle the high number of rare classes of the NFI, a characteristic which the NFI also has at national scale. Instead, the hybrid design maintains simple random selection of PSUs for common classes, but applies a proportional stratified selection of PSUs for rare classes. This way, DS4 cumulates the advantage of a good sample dispersion for common classes, and the advantages of a sufficient sample size and easy estimate calculation for rare classes.

The entire accuracy assessment strategy, including the hybrid sampling design and a fuzzy verification design, permitted the evaluation of the pilot area within a probability sampling scheme. Therefore, a confusion matrix was built and some confusion patterns could be analysed. The global confidence level of the NFI in the watershed of the Cuitzeo lake is high (81%). High confidence levels also characterize, generally, 'primary' forest and 'other cover types' classes. Conversely, high confusion characterized 'secondary' forest classes (generally scarcely distributed). The way secondary vegetation is included in the classification scheme may have caused thematic ambiguity, and visual polygon delineation may be subject to positional uncertainty in spatial transition zones. Despite thematic uncertainties related to secondary vegetation, a virtue of the NFI classification system is undoubtedly its compatibility with classification schemes of former INEGI maps. We noted that the present accuracy assessment is dependent on specific positional and thematic tolerances for these uncertainties, but the issue of incorporating more choice for this tolerance into the assessment is subject of further research.

For all sampling designs, the dispersion of sample points remained low, due to the relatively small geographic extension of the pilot area, given the coarse scale (1:250 000) of the map and the high number of classes. For this reason, the average number of sample points per PSU remained relatively high for all designs. However, the hybrid sampling design and more generally the entire assessment design presented here is immediately transferable to more extended study areas, with typically available data and skills in Mexico. Specifically, its application to national scale is possible, using additional strata such as a systematic photographic coverage along flight lines (material initially planned in the NFI project) and administrative states.

6. Acknowledgments

This research was done during the PhD thesis of the principal author, who receives a grant from the CONACYT public institution under scholarship number 205000.

7. References

Cochran, W.G., 1977, Sampling Techniques, 3rd ed. John Wiley and Sons, New York, 428p.

FGDC, 1997, FGDC vegetation classification and information standard, June 3, 1996 draft. Reston, VA: Federal Geographic Data Committee, Vegetation Subcommittee (FGDC-VS), FGDC Secretariat.

Fitzpatrick-Lins, K., 1981, Comparison of sampling procedures and data analysis for a land-use and land-cover map, *Photogrammetric Engineering and Remote Sensing*, 47, pp. 343-351.

Gopal S., and C. E. Woodcock, 1994, Accuracy of Thematic Maps using fuzzy sets I: Theory and methods, *Photogrammetric Engineerίng and Remote Sensing* 58, pp. 35-46.

Mas JF, Velázquez A, Palacio-Prieto JL, Bocco G, Peralta A, Prado J (2002) Assessing forest resources in Mexico: Wall-to-wall land use/ cover mapping. *Photogrammetric Engineering and Remote Sensing* 68 (10), pp. 966-969.

Laba M, Gregory SK, Braden J et al. (2002) Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map. *Remote Sensing of Environment* 81, pp. 443-455.

López E, Bocco G, Mendoza M, Velázquez A and Aguirre-Rivera R (2006) Peasant emigration and land-use change at the watershed level: A GIS-based approach in central Mexico, *Agricultural Systems* 90, pp. 62-78.

Palacio-Prieto JL, Bocco G, Velázquez A, Mas JF, Takaki-Takaki F, Victoria A, Luna-González L et al. (2000) La condición actual de los recursos forestales en México: resultados del Inventario Forestal Nacional 2000, *Investigaciones Geográficas* 43, pp. 183-202.

Remmel TK, Csillag F, Mitchell S, Wulder M (2005) Integration of forest inventory and satellite imagery: a Canadian status assessment and research issues. *Forest Ecology and Management* 207, pp. 405-428.

Snedecor, G.W., and W.F. Cochran, 1967, Statistical methods, State University Press, Ames, Iowa, 728 pp.

Stehman SV and Czaplewski RL (1998) Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment* 64, pp. 331-344.

Stehman SV, Wickham JD, Smith JH, and Yang L (2003) Thematic accuracy of the 1992 National Land-Cover Data for the eastern United-States: Statistical methodology and regional results. *Remote Sensing of Environment* 86, pp. 500-516.

Stehman S.V., Wickham J.D., Yang L., and J.H. Smith, 2000, Assessing the Accuracy of Large-Area Land Cover Maps: Experiences from the Multi-Resolution Land-Cover Characteristics (MRLC) Project, *4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Amsterdam, 601-608 pp.

Wickham JD, Stehman SV, Smith JH and Yang L (2004) Thematic accuracy of the 1992 National Land-Cover Data for the western United-States. *Remote Sensing of Environment* 91, pp. 452-468.

Zhu Z., Yang L., Stehman S.V., and R.L. Czaplewski, 2000, Accuracy Assessment for the U.S. Geological Survey Regional Land-Cover Mapping Program: New York and New Jersey Region, *Photogrammetric Engineering & Remote Sensing,* 66, pp. 1425-1435.

Chapter 2: Accuracy Assessment of the NFI Map

ACCURACY ASSESSMENT OF THE MEXICAN NATIONAL FOREST INVENTORY MAP.

A study in four eco-geographical areas:

- The biosphere reserve of Los Tuxtlas and adjacent sugar cane production zone, Veracruz

- The closed watershed of the Cuitzeo Lake, Michoacán

- The biological reserve of the Tancítaro volcano and adjacent avocado production zone, Michoacán

- The Candelaria river watershed, Campeche

In Chapter 1, the entire accuracy assessment strategy, including the hybrid sampling design, permitted the evaluation of the pilot area within a probability sampling scheme.

Therefore, in chapter 2, the general scheme of the accuracy assessment, without fuzzy component, is extended to four areas of distinct eco-geographical zones in the Mexican territory. The focus of chapter 2 was the analysis of thematic confusions specific to the NFI map, with respect to previous research in mainly temperate environments; therefore we selected a framework without fuzzy component, comparable to most previously published regional accuracy assessments.

The eco-geographical zones include lowland and altiplano environments in Mexico, and contrasted levels of forest cover and anthropisation of the landscape. The double sampling strategy was applied to the 2.5 million-ha area under scrutiny. A total of 5955 reference sites were verified against their NFI map label. The availability of detailed quasi-synchronous reference data to the 2000 Landsat-derived NFI and the high diversity of mapped classes allowed a careful thematic analysis on the selected regions, relevant for national extrapolation.

Accuracy Assessment of the Mexican National Forest Inventory map: a study in four eco-geographical areas[Θ]

Stéphane Couturier

Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)

Centre d'Etudes Spatiale de la Biosphère (CESBIO), CNES/CNRS/Paul Sabatier University

Jean-François Mas

Instituto de Geografía, Unidad Académica Morelia Universidad Nacional Autónoma de México (UNAM)

Jorge Benítez

Centro EPOMEX, Universidad Autónoma de CampecheAbstract:

Erna López

Centro de Investigación en Ecología (CIECO), UNAM

The accuracy of the National Forest Inventory (NFI) map is derived in four pilot areas of the Mexican territory using an assessment design tailored for the project. A main challenge of the design was to integrate the high diversity of classes encompassed in the community level (the most detailed level) of the classification scheme, within a cost-controlled statistically sound assessment. A double sampling strategy was applied to the 2.5 million-ha area under scrutiny. A total of 5955 reference sites were verified against their NFI map label. The availability of detailed quasi-synchronous reference data to the 2000 Landsat-derived NFI and the high diversity of mapped classes allowed a careful thematic analysis on the selected regions, relevant for national extrapolation. Global accuracy estimates of 64% to 78% were registered among the four eco-geographical areas (two with mainly temperate climate and the other two with mainly tropical climate), with the lower accuracy levels found in areas more densely covered with forests. According to the estimates, the NFI map tends to underestimate the presence of temperate forest (especially oak) and overestimate the presence of tropical forest in the regions investigated. The analysis of confusions reveals difficulties in unambiguously interpreting or labelling forests with secondary vegetation, herbaceous, and/or shrub-like vegetation as well as distinguishing between aquatic vegetation types. The design proved useful in the perspective of an accuracy assessment at national scale.

[Θ] Submitted to the Singapore Journal of Tropical Geography, Blackwell Publishing 6 June 2006

1. Introduction

The National Inventory Forest (NFI) 2000 map of Mexico, sponsored by SEMARNAP, the Ministry of the Environment in Mexico, was conceived and jointly coordinated by the Institute of Geography of UNAM (the National Autonomous University of Mexico) and INEGI, the National Institute of Statistics Geography and Informatics in Mexico. The cartography was generated at UNAM based on previous INEGI maps and on visual interpretation of Landsat imagery of year 2000 at scale 1:250 000 (Mas et al. 2002). A main trait of the NFI project was to propose a hierarchical classification scheme as a convenient Land-Use and Land-Cover (LULC) standard in Mexico for future remote-sensing based cartography. Four aggregation levels were considered; the finest taxonomic resolution (denominated community level, including sub-communities) consisted in 75 classes. The classification scheme proposed for the NFI is based on a physiognomic hierarchy of vegetation communities, as is, for example, the classification scheme of the United States National Vegetation Classification System (NVCS, see FGDC, 1997). The taxonomic resolution of the NFI at community level is comparable to the subclass level of the NVCS. A growing need for public agencies producing cartography at regional scale is to provide objective information on the reliability of its product. This need has been manifested in the literature by various accuracy assessments of Landsat-derived LULC maps. In Laba et al. (2002) and in Wickham *et al.* (2004), the respective maps of the New York Gap Analysis Project and of the National Land-Cover Data (NLCD) are assessed at regional scale for year 1992 in the United States of America (USA). A comparison between the 2000 Landsat-derived map of the EOSD (Earth Observation for Sustainable Development of Forests) project and the Canadian national forest inventory based on aerial photograph interpretation was reported by Remmel et al. (2005). However, even in these developed countries, the budget devoted to the assessment of a map is tight and poses serious challenges to the designers.

In Mexico, a preliminary accuracy assessment was conducted immediately after the elaboration of the NFI map. However, confidence levels were reported only for a few classes, abundant in the northern part of the country (Mas et al. 2002). At that time, the scope of the project prevented the conception and implementation of a full assessment

design capable of estimating confidence levels of the remaining classes on the national territory.

In 2003, a research project initiated at the Institute of Geography, UNAM, whose proposed tasks were to build academic capacity for assessment of LULC maps in Mexico and to conceive a framework for future accuracy assessments of the NFI.

Such a framework was built in accordance with the typically available verification materials, skills and resources in Mexico. In order to implement the methodology, a pilot study was launched over a set of four distinct eco-geographical areas described in the following section. The 'methods' section provides an overview of the assessment strategy found to meet the following desirable criteria: 1) a probability sampling protocol (*sensu* Stehman and Czaplewski 1998, comprising a sampling design, a verification design and the phase of analysis); 2) an ergonomic and operational design for future NFI map updating missions; 3) a reasonable compromise between operational costs and precision (standard error) of accuracy estimates. Findings arising from the accuracy assessment results of the pilot study are then reported and discussed as the main focus of the article.

Table 2.1: Class distribution in the national forest inventory map (formation aggregation level) over the entire Mexican territory.

Formation	Área km2	%
Cropland	456,870	23.53
Temperate forest	328,507	16.92
Tropical forest	307,349	15.83
Scrubland	554,518	28.55
Grassland	188,474	9.71
Hygrophilous vegetation	20,826	1.07
Other vegetation types	61,986	3.19
Other cover types	23,455	1.21
Total	1,941,984	100.00

Information taken from Palacio-Prieto et al. (2000)

2. Study areas

The set of study areas (located on figure 2.1) considered in this research captures parts of the mega-diversity of the Mexican territory with special focus on some of the main forest

biomes (see table 2.1 and 2.2), and includes contrasted levels of alteration of the original vegetation cover.

Figure 2.1 Location map of all study areas

Accuracy estimates are reported for the four mapping regions shaded in gray.

The first two areas are located on the transversal volcanic chain and contiguous altiplano in central western Mexico. These are the closed watershed of the Cuitzeo Lake, later referred as Cuitzeo, and an area encompassing both the natural reserve of the Tancítaro volcano and the Uruapan avocado production zone, later referred as Tancítaro. Both areas are included in the state of Michoacán and are covered with temperate sub-humid and tropical dry vegetation (table 2.2).

A third area includes the nuclear and buffer zones of the biosphere reserve of Los Tuxtlas, in the state of Veracruz, and the adjacent sugar cane production zone. This area is generally characterised by tropical humid conditions although a temperate humid climate and relief-induced micro-climates hold on two volcanic mounts embedded in the coastal plain.

A fourth area corresponds to the Mexican side of the Candelaria river watershed in the state of Campeche. This area includes a portion of the Calakmul forest reserve and is mainly characterized by tropical sub-humid conditions (table 2.2).

The Candelaria and Tancítaro regions comprise extensive forests (low and high levels of human management, respectively) while most of Cuitzeo and Los Tuxtlas is covered with non forested agriculture land (crop and graze land, respectively). Apart from the informative contrast in LULC, a key motivation for selecting and defining these areas was the availability of reference data exploitable in the process of independently verifying the 2000 Landsat-based NFI. These reference data are detailed in the verification design (next section).

3. Methods:

3.1 Sampling design

Apart from a stratification per eco-geographical area, the sampling design incorporated a two-stage sampling protocol, where aerial photography frames formed the Primary Sampling Units (PSUs). This strategy is employed, for example, by Stehman et al. (2003) for the accuracy assessment of the NLCD in the USA. Figures 2.2 to 2.5 display the entire set of PSUs in each of the four study areas. Instead of Landsat pixels as in Stehman et al. (2003), a regular 500m-spaced two dimensional grid (thereafter referred as the 'second stage grid') formed the set of punctual Secondary Sampling Units (SSUs) of the second stage. Indeed, a scale criterion during map production was to leave out polygons less than 500 meters wide.

In order to preserve equal inclusion probabilities at the second stage, the selection of PSUs was done via proportional stratified random sampling as described in Stehman et al. 2000 (further discussed via personal communication); this mode of selection was appropriate for including all sparsely distributed classes (a frequent occurrence in our case), in the sample while maintaining a low complexity level of statistics (i.e. standard stratified random formulae). As a compromise between operational costs (our budget for the pilot study) and precision of the estimate, the number of PSUs selected was kept below, though approaching, one fourth of the total number of PSUs in each area.

Once the sample PSUs were selected, all points of the second stage grid included within these sample PSUs were assigned to their mapped land cover class. The full second stage sample consisted of 100 points from each class mapped in the area. The selection of this sample was obtained via proportional stratified random sampling from the points contained in the sampled PSUs, this time with a probability inversely proportional to the abundance of the class (see the option of proportional stratified random sampling advocated in Stehman *et al*.(2000:604). A sequence of ArcView and Excel-based Visual Basic simple routines, for easy and fast repeated use on vector attributes of each class, was specifically designed to perform this proportional selection at both stages.

Table 2.2. Distribution of classes (community and formation aggregation levels) of the national forest inventory map, in the four eco-geographical areas

Class	Name	Formation	Cuitzeo		Tancítaro		Tuxtlas		Candelaria		Total area per
			Area frac	Area (km^2)	Area frac[a]	Area (km^2)	Area frac	Area (km^2)	Area frac	Area (km^2)	class (km^2):
100	Irrigated crop	Cropland	0.1411	564.97	0.0106	13.45					578.42
110	Hygrophilous crop		0.0048	19.04							19.04
130	Cultivated grassland						0.6058	1839.75	0.1708	1908.25	3748.01
200	Perennial crop		0.0021	8.27	0.2904	367.70	0.0129	39.16			415.13
210	Annual crop		0.2356	943.14	0.0803	101.69	0.1765	535.84	0.0070	77.85	1658.51
300	Forest plantation		0.0071	28.24							28.24
410	Fir forest	Temperate Forest	0.0037	14.72							14.72
420	Pine forest		0.0041	16.32	0.1658	209.99	0.0011	3.36			229.67
421	Pine forest & Sec Veg [b]		0.0036	14.31	0.0634	80.23	0.0011	3.37			97.90
510	Oak-pine forest		0.0958	383.34	0.1907	241.47					624.82
511	Oak-pine forest & Sec Veg [b]		0.0325	130.29	0.1284	162.54	0.0028	8.48			301.31
600	Oak forest		0.0232	92.88			0.0011	3.44			96.32
601	Oak forest & Sec Veg [b]		0.0553	221.54	0.0017	2.16	0.0041	12.49			236.20
700	Cloud forest	Tropical Forest	0.0029	11.73			0.0035	10.78			22.51
800	Med and high Per Trop forest[c]						0.1213	368.43			368.43
801	Med and high Per Trop For[c] & Sec Veg [b]						0.0292	88.56			88.56
820	Med and high Subper Trop Forest[c]								0.5010	5595.31	5595.31
821	Med and high Subper Trop For & Sec Veg [b]								0.0880	982.82	982.82
830	Low Subper Trop forest [d]								0.1765	1971.60	1971.60
831	Low Subper Trop forest [d] & Sec Veg [b]								0.0025	27.57	27.57
920	Sub-tropical shrubland	Scrubland	0.0194	77.58							77.58
921	Sub-tropical shrubland & Sec Veg [b]		0.0768	307.25							307.25
1000	Mezquital		0.0004	1.51							1.51
1200	Chaparral										0.00
1320	Savannah	Grassland							0.0108	120.13	120.13
1330	Induced grassland		0.1594	638.04	0.0032	4.02	0.0004	1.08	0.0039	43.65	686.80
1400	Mangrove	Hygrophilous vegetation					0.0066	20.15	0.0060	66.93	87.08
1410	Hygrophilous grassland		0.0209	83.50			0.0019	5.86	0.0225	251.24	340.60
1510	Halophilous vegetation	Other vegetation types	0.0069	27.78					0.0039	43.56	71.34
1600	No apparent vegetation cover	Other cover types			0.0390	49.43	0.0007	2.11			51.54
1700	Human settlement		0.0250	100.02	0.0265	33.59	0.0065	19.82	0.0009	10.19	163.62
1800	Water		0.0796	318.75			0.0244	74.01	0.0063	70.11	462.87
	All		1.0000	**4003.23**	1.0000	**1266.28**	1.0000	**3036.69**	1.0000	**11169.21**	**19475.41**

[a] Area percent fraction with respect to total area. [b] Secondary vegetation (secondary vegetation, herbaceous and/or shrub-like vegetation, according to NFI classification scheme)
[c] Median and high sub-perennial tropical forest [d] Low sub-perennial tropical forest,

3.2 Verification Design and Analysis

Black and white aerial photography served as the basic material for the reference LULC classifications in all areas under investigation. Their characteristics are described in table 2.3.

Table 2.3. Characteristics of aerial photography used as the main verification dataset for the accuracy assessment of the national forest inventory map.

Aerial Photography:	*Data type/Interpretation*	*Scale/Resolution*	*Year*	*Number of photographs*
Cuitzeo	Prints/stereoscopic	1:37 000	1999	244
Tancítaro	Prints/stereoscopic	1:24 000	1996	152
Tuxtlas	Digital/on screen	1:75 000 / 1.5m grain	2000 1996	12 14
Candelaria	Prints/stereoscopic	1:75 000	Jan. 2000-March 2002	174

For instance, the first row indicates that 244 aerial photographs at 1: 37 000 covering the Cuitzeo area and acquired in 1999 were used for reference classification via stereoscopic interpretation. The full aerial photographic coverage (152 photographs) acquired in 1996 of the Tancítaro volcano and the Uruapan avocado production defined the Tancítaro area.

Strong ties between UNAM Geography Institute and local institutions in this relatively small area (the least extended study area of the four) allowed supervised field work to document LULC changes between 1996 and 2000 that would otherwise have affected our reference data. All air photography at scale 1:75 000 was taken from the INEGI archive. A set of 26 digital ortho-photographs (12 from year 2000 and 14 from year 1996) covered the Tuxtlas area. A 2003 photo-interpretation based map at 1: 50 000 from the National Reserves Commission (CONANP) served as the basis for reference classification, especially in areas where the 2000 Landsat imagery and 1996 ortho-photographs showed contradictory LULC. Finally, 174 aerial photographs from the INEGI archive from January 2000 to March 2002 covered the Candelaria area and were used for reference classification.

One specialist interpreter was assigned to produce reference maplets at a scale of about 1:125 000 on transparent slides for each photographic frame corresponding to a selected

PSU. The reference maplets were drawn according to the community level of classification of the NFI map; a primary and a secondary (alternate) class were assigned to each polygon of the maplet. The allowance for a secondary label was a standard rule to account for the ambiguity of class labelling, either from mixed classes or transitions between classes (ex. ecotones).

A confidence rating was applied to photograph interpretation, comparable to the one used in Zhu et al. (2000). Wherever the interpreter registered low confidence ratings, his mapping was confirmed or corrected by local information collected in the field. In two areas (Cuitzeo and Candelaria), another specialist ('local interpreter') had previously worked on the same material to generate cartography with equal or more precise spatial and taxonomic resolutions. In such cases, any discrepancy in photo-interpretation was turned to the advantage of the 'local interpreter'.

Geometric consistency between the reference material and the NFI map was ensured in a similar way as in Zhu et al. (2000): Each sample point (selected SSU) was located on a colour composite of the geo-referenced Landsat image (geometric base of the NFI map) on screen. The location of the point was visually transferred onto the transparent slide over the photographic material. The primary and secondary reference labels of the sample point were then digitally registered according to the photo-interpretation maplet. This step was done by another team member than the specialist photo-interpreter, which added objectivity to the labelling process.

Because the SSUs were punctual, we looked for a definition of agreement that would avoid fictitious errors due to the proximity of a polygon edge (and related scale artefacts): We defined agreement between the reference label and the NFI map label when a match occurred between either of the primary or secondary reference labels and any of the NFI map labels within a buffer of 500m of the sample point. This 500m rule was related to scale instructions for map production (see sampling design). A practical inconvenience of the definition was the necessity to manually report NFI map labels which corresponded to the match in the confusion matrix: in case of polygon edge situation, no existing routine permitted us to automatically identify which of the NFI labels within the 500m buffer was the one matching the reference label.

Improvements on the verification design, in aspects such as the automatic treatment of edges (or scale) effects in NFI and reference maplet polygons, and fuzzy representation of classes, are under development and constitute the focus of parallel research. However, preliminary results using these developments suggest that general patterns of thematic errors within and among eco-geographical areas, which are the major focus of this article, apply either with the assessment design described in this paper or an improved version of this design. The main effect of a new design is believed to reside in improved estimates of absolute accuracy levels (less inclusion of mis-registration errors, scale artifacts or errors due to class definitions in reference and NFI maps).

Within a given class, the construction of a proportional random sampling ensured that inclusion probabilities among sample points were equal. The formula for simple random sampling (the binomial standard error formula) was therefore used to compute the estimated accuracy variance for each class. However, no adjustment for cluster design was attempted, as a consequence of poor knowledge of intra-cluster variance. A consequence of this approximation – i.e. of ignoring the effect of intra-cluster variance - is an under-estimation of the confidence interval mentioned for the accuracy indices, common to many existing accuracy assessment studies.

4. Results

4.1 Confusion matrices and global findings

Global accuracy and regional confusion patterns among NFI classes are reported via four confusion matrices (tables 2.4 – 2.7). The same convention as in Stehman et al. (2003) was adopted for the matrix format. Indeed, the normalization of matrix elements to total area fractions allows an inter-class confusion pattern analysis per eco-geographical area. Overall agreement scores ranged from 64% (Candelaria) to 78% (Los Tuxtlas). To assist the reader in the ongoing analysis, this interval was used as quantitative bounds of a practical confusion scale: high confusion or low accuracy will thereafter mean accuracy scores lower than 64%, intermediate confusion or accuracy will designate scores between 64 and 78%, and scores above 78% will be labeled as low confusion or high accuracy. As an arbitrary means of separating LULC classes along levels of spatial distribution, we considered three categories of classes: dominant (area fraction above 20%), common (area

fraction between 5 and 20%), and rare (area fraction below 5%). In order to ease sentences in the analysis, we also introduce the term of 'confusion flux' from class a to class b as being the proportion of sampled mapped class a found to belong to class b in the reference data.

4.2 Analysis per formation

Accuracy scores were lower in forest-dominated Candelaria (64%) and Tancítaro (67%) areas than in non-forest dominated Cuitzeo (75%) and Los Tuxtlas (78%) areas, possibly because of the higher (confusion-prone) diversity of forest classes than non-forest classes in the NFI classification scheme.

For 'other cover types' (no vegetation cover, water and human settlement), a high accuracy (79% and above) was registered, with the only exception of the water class in Candelaria, where water bodies are small, dispersed and often seasonal. Visually, the spectral separability of these land covers within the group and with respect to other groups is indeed among the highest on conventionally-used Landsat color composites. The mangrove class also recorded high accuracy in both Candelaria and Los Tuxtlas areas where mangrove is present. Conversely, very high inter-confusion within aquatic non-treed vegetation covers is evident from the matrices when hygrophilous grassland and halophilous vegetation are both present (Cuitzeo and Candelaria). High confusion fluxes also exist from hygrophilous grassland to induced grassland in Los Tuxtlas and Candelaria. The spectral ambiguity and variability (across inundation phases) of these aquatic vegetation types is probably one of the key explainers of this observed high confusion. Former INEGI maps mostly confirm the reference data in exhibiting such errors of the NFI. Trends registered for forest types and land use categories vary according to the eco-geographical area as described below.

4.3 Analysis per eco-geographical area

Cuitzeo

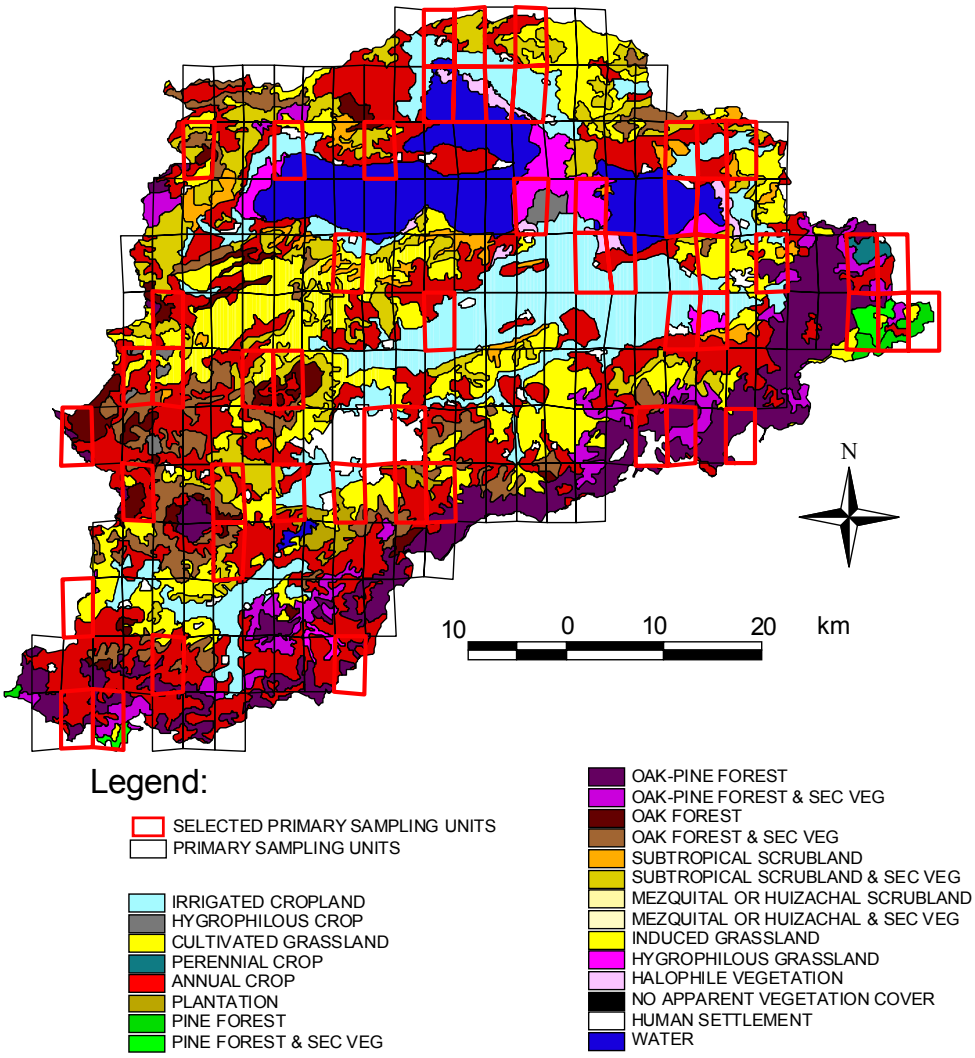
The global estimate of accuracy in Cuitzeo (figure 2.2) primarily reflects the intermediate score (71%) of the dominant class 'annual crop' (table 2.4). The common class oak-pine forest (97%), and rare classes such as sub-tropical scrubland (78%), oak forest (92%), pine

forest (79%) and fir forest (76%) record intermediate-high to high accuracy. Therefore, the semi-arid and temperate 'primary' vegetation classes present in the area are generally characterized by a low level of confusion. Exceptions are the sole mapped patches of mezquital and cloud forest in Cuitzeo, which both corresponded to other classes on the ground (0% accuracy, double-verified with ground visits). These NFI errors clearly appeared to be omission on respectively reporting change and rectifying an error from INEGI former maps of the area.

By contrast with the good agreements obtained for 'primary' temperate forest classes, numerous confusions were recorded among classes denominated as temperate forests with secondary vegetation (more exactly with herbaceous, shrub-like and/or secondary vegetation, as stipulated in the NFI classification scheme, which we will label 'secondary' forests): According to the reference photo-interpreted data, 'secondary' pine forest was wrongly mapped as 'secondary' oak, 'secondary' mixed (oak-pine) forest was wrongly mapped as 'secondary' pine, and oak forest was wrongly mapped as 'secondary' mixed forest.

These confusions are mainly responsible for the low and very low accuracy levels registered for 'secondary' oak (46%), 'secondary' pine (12%) and 'secondary' mixed (45%) forests. Various interpretations may arise from this observation. If we assume that the reference data reflects reality, much area of these 'secondary' forests lays on eco-tones even more difficult to discern than for 'primary' mixed forests. Indeed, the more pronounced presence of human activity further augments the dimension of class membership to these areas.

Figure 2.2. National forest inventory map in the watershed of the Cuitzeo Lake.



The Primary Sampling Units (PSUs) selected by the sampling design are highlighted.

However, doubts may be cast on the actual robustness of the definition of classes denominated 'forest with secondary vegetation'. For example, an old fallow field or an open forest may both exhibit characteristics of several classes in the classification scheme. The decision between a land use class or a forest class will be function of subjective bounds made by interpreters on height and density of vegetation. The decision to attribute a polygon of secondary vegetation the label of one of the 'secondary' forest classes, will be function of its proximity to biotopes of a 'primary' forest type, this proximity also being possibly subjective. This multiplicity of class memberships may call for the aid of a fuzzy representation beyond the option of an alternate class; the fuzzy representation is presently under construction within the same assessment design.

Altogether, some balance results from most of the confusion fluxes generated by these real or fictitious (resulting from the ambiguity on definitions) errors: the fluxes between 'secondary' and 'primary' and among forest types exhibit many feedback loops. Altogether, the producer's totals suggest general trends such as the NFI under-estimation of 'primary' oak forest and over-estimation of 'secondary' oak forest both by 3% of the entire Cuitzeo area.

Among non-treed vegetation covers, much of the mapped induced grasslands, and some of the mapped 'secondary' oak would be cropland in reality. In turn, much of the mapped cropland should be scrubland, with the subtraction of a small reciprocal confusion pattern. Altogether, according to the area fraction numbers on producer's totals, the presence of scrub (with or without 'secondary' vegetation) is underestimated by 6% of the entire region. This could reveal an under-estimation of the withdrawal of non-pastoral and pastoral agricultural activity in the region, towards the regeneration of scrub vegetation communities. This strong under-estimation was confirmed in a fine scale LULC change analysis by López et al. (in press), where scrubland is depicted as covering a much larger area in 2000 than in a former map.

Table 2.4. Confusion matrix of the national forest inventory map in Cuitzeo at community level.

	100	110	200	210	300	410	420	421	510	511	600	601	700	920	921	1000	1200	1330	1410	1510	1600	1700	1800	n[a]	Area frac[b]	User[c]	StdE[d]
100	**12.241**			0.432										0.144	0.144				0.288		0.144	0.432	0.288	98	14.113	0.87	0.03
110	0.098	**0.300**		0.036										0.010	0.010					0.005		0.010	0.005	92	0.476	0.63	0.05
200			**0.204**																			0.002		85	0.207	0.99	0.01
210	0.892			**16.777**	1.071				0.178		0.178			1.785	2.142							0.535		132	23.560	0.71	0.04
300				0.015	**0.585**				0.038					0.008	0.015		0.045							94	0.705	0.83	0.04
410						**0.279**		0.029	0.051					0.004				0.004						100	0.368	0.76	0.04
420				0.022			**0.322**		0.033	0.026								0.004						110	0.408	0.79	0.04
421				0.040			0.003	**0.043**	0.116	0.142								0.007					0.007	108	0.357	0.12	0.03
510								0.176	**9.224**	0.176														109	9.576	0.96	0.02
511				0.206						**1.473**	1.199			0.034				0.343						95	3.255	0.45	0.05
600				0.073						0.024	**2.125**			0.049			0.049							95	2.320	0.92	0.03
601				0.498	0.111			0.664	0.332	0.332	0.277	**2.546**		0.498	0.166							0.111		100	5.534	0.46	0.05
700				0.023	0.003		0.221	0.023		0.003			**0.000**		0.020									90	0.293	0.00	0.00
920				0.268										**1.505**	0.124			0.041						94	1.938	0.78	0.04
921				0.444							0.317			0.063	**6.724**							0.127		121	7.675	0.88	0.03
1000				0.036										0.002	**0.000**									23	0.038	0.00	0.00
1200																								0	0.000		
1330	0.280			2.237							1.258	0.140		0.979	1.258			**9.507**				0.280		114	15.938	0.60	0.05
1410	0.070	0.093													0.023				**0.973**	0.672			0.255	90	2.086	0.47	0.05
1510		0.008		0.016											0.112			0.008	0.175	**0.175**	0.128		0.072	87	0.694	0.25	0.05
1600																								0	0.000		
1700																						**2.499**		91	2.499	1.00	0.00
1800				0.251										0.084				0.503					7.124	95	7.962	0.89	0.03
n^a	114	63	84	212	87	76	156	42	175	103	142	47	0	109	163	0	8	91	66	52	17	106	110	2023			
Area frac[b]	13.581	0.401	0.204	21.374	1.770	0.279	0.547	0.935	9.973	2.177	5.355	2.685	0.000	5.162	10.739	0.000	0.094	10.416	1.437	0.853	0.272	3.996	7.751		**100.00**		
Prod[c]	0.90	0.75	1.00	0.78	0.33	1.00	0.59	0.05	0.92	0.68	0.40	0.95		0.29	0.63		0.00	0.91	0.68	0.21	0.00	0.63	0.92			**74.6**	
StdE[d]	0.03	0.05	0.00	0.03	0.05	0.00	0.04	0.03	0.02	0.05	0.04	0.03		0.04	0.04		0.00	0.03	0.06	0.06	0.00	0.05	0.03				

Global accuracy was estimated at 74.6%

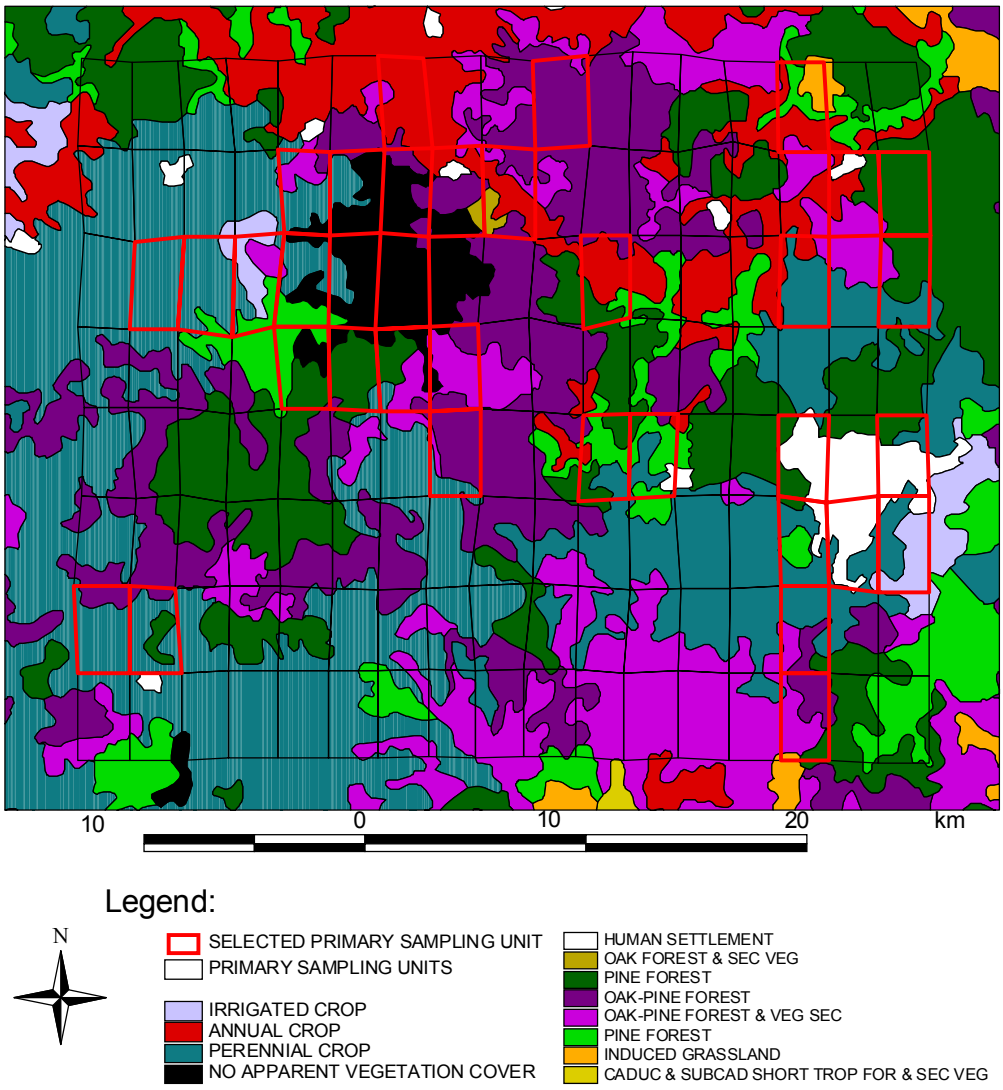
[a] Sample size. [b] Class area fraction (out of the total Cuitzeo area), sum of the row/column matrix elements.

[c] User's (User) and Producer's (Prod) accuracy [d] Standard Error using binomial theory, with approximation to large sample sizes if applicable

Tancítaro

In Tancítaro (figure 2.3), the dominant class, perennial crop (mainly comprising avocado and apricot plantations) suffers only low, mutually compensating confusion, towards mixed and pine forests (Table 2.5). Structural and spectral characteristics of these crops on Landsat imagery usually permit an easy delineation and distinction from native forests. Much confusion occurred from mapped irrigated crop towards perennial crop.

Figure 2.3. National Forest Inventory map in the region of the Tancítaro volcano.



The Primary Sampling Units (PSUs) selected by the sampling design are highlighted.

Table 2.5. Confusion matrix of the national forest inventory map in Tancítaro at community level.

	100	200	210	420	421	510	511	600	601	1330	1600	1700	n^a	Area frac[b]	User[c]	StdE[d]
100	**0.228**	0.411	0.206	0.023		0.126		0.023				0.046	93	1.062	0.22	0.04
200		**25.078**	0.264	1.320		1.848	0.528						110	29.038	0.86	0.03
210	0.592		**7.016**	0.338		0.085							95	8.030	0.87	0.03
420		1.345	0.598	**6.723**		2.839	4.482	0.149			0.149	0.299	111	16.583	0.41	0.05
421		1.068	0.641	1.424	**0.498**	2.136	0.427				0.142		89	6.336	0.08	0.03
510	0.156	0.938	1.250	1.876	0.156	**14.693**							122	19.070	0.77	0.04
511		0.558	1.005	3.460	0.446	0.112	**7.144**	0.112					115	12.836	0.56	0.05
600													0	0.000		
601			0.029	0.059	0.020	0.055			**0.008**				87	0.171	0.05	0.02
1330						0.072	0.130			**0.115**			88	0.318	0.36	0.05
1600		0.439				0.088	0.175				**3.202**		89	3.904	0.82	0.04
1700		0.030								0.060		**2.563**	89	2.652	0.97	0.02
n^a	28	177	147	149	22	213	142	4	4	34	76	92	1088			
Area frac[b]	0.976	29.866	11.008	15.222	1.121	22.052	12.886	0.284	0.008	0.175	3.494	2.907		100.000		
Prod[c]	0.23	0.84	0.64	0.44	0.44	0.67	0.55	0.00	1.00	0.66	0.92	0.88			**67.3**	
StdE[d]	0.08	0.03	0.04	0.04	0.11	0.03	0.04	0.00	0.00	0.08	0.03	0.03				

Global accuracy was estimated at 67.3%.

[a] Sample size. [b] Class area fraction (out of the total Cuitzeo area), sum of the row/column matrix elements.

[c] User's (User) and Producer's (Prod) accuracy [d] Standard Error using binomial theory, with approximation to large sample sizes if applicable

As in Cuitzeo, the 'primary' mixed forest class is attributed a high accuracy. However, all other classes representing non-human made forests ('primary' pine and 'secondary' pine, oak and mixed forests) are perturbed with much internal confusion. Similar observations as in Cuitzeo hold for these temperate forests.

The general trend is an over-estimation of 'secondary' pine by 5% of the Tancítaro area and a under-estimation of mixed forest by 3%. Over the sample, the photo-interpreter described temperate forests richer in oak trees than the NFI map producer did. This may have reflected reality, or to a certain extent a subjective difference of appreciation – so as to define limits between pine and oak-pine forests - between this specialist and other field specialists behind the elaboration of the NFI. No double photo-interpretation was done in Tancítaro to complement views of our specialist photo-interpreter.

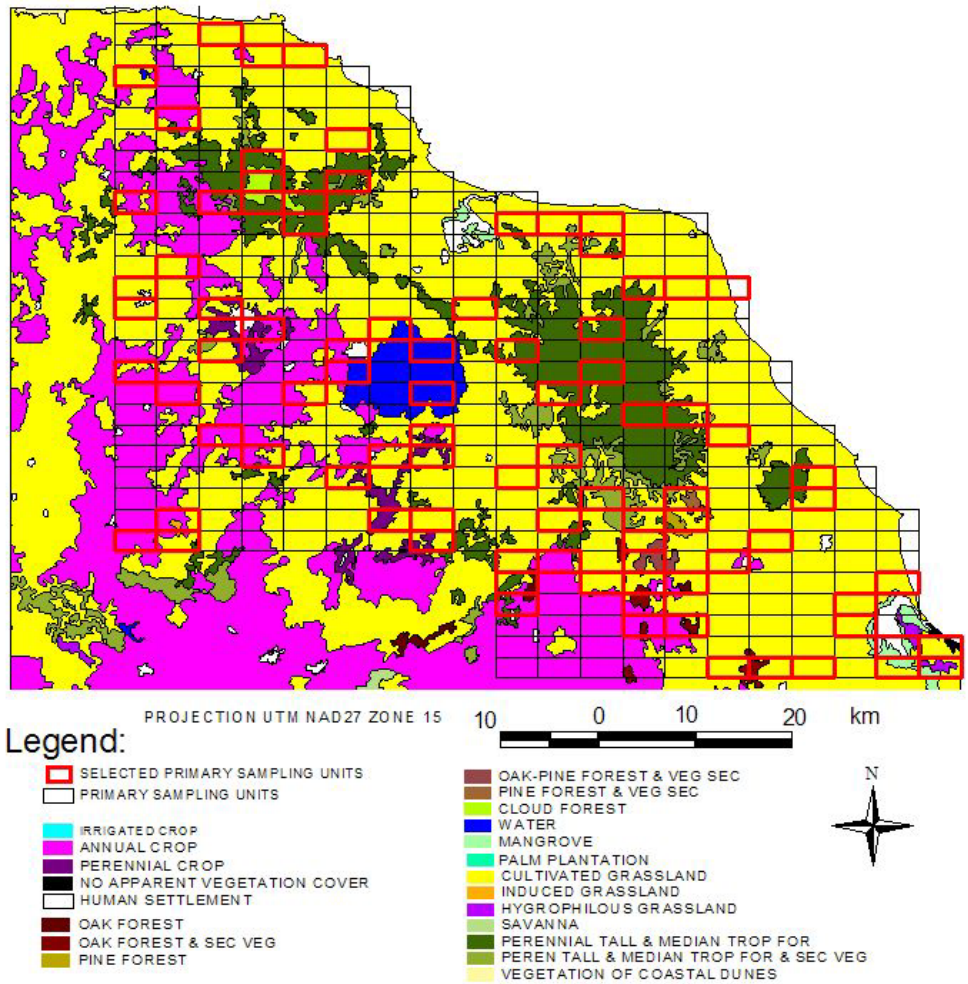
A source of real inter-class confusion in densely forested Tancítaro could be the Landsat interpreter's simplification of polygon edges with respect to scale 1:250 000: Along intricate borders of forest ecosystems, the mapping process may have been conducted according to a 'compensation' strategy, a common practice in visual interpretation of fuzzy limits, whereby digitization of edges is simplified while the overall area of polygons about the edge is kept as close to reality as possible.

Strong confusion fluxes from vegetation classes towards classes representing denser vegetation ('secondary' pine towards 'primary' pine or mixed forest, 'secondary' oak towards 'primary' mixed forest, and induced grassland towards 'secondary' mixed forest) suggest a real under-estimation by the NFI of vegetation density and/or vegetation regeneration processes.

Los Tuxtlas

Most of the global estimate in Los Tuxtlas (figure 2.4) is explained by the score (83% accuracy) of the sole dominant class, cultivated grass for pasture, which according to the map occupies over 60% of the area (table 2.6). To the notable exception of oak forest, 'primary' forest classes including median and high tropical forest (the only common forest class), cloud forest and pine forest recorded low levels of confusion.

Figure 2.4. National Forest Inventory map in the region of Los Tuxtlas.



The Primary Sampling Units (PSUs) selected by the sampling design are highlighted.

As of earlier results, 'secondary' temperate forest classes suffered very high (pine and mixed) to intermediate (oak) confusion. About half of this confusion flux was directed to grassland types and the rest was internal to temperate forests (mainly 'secondary' pine to pine, 'secondary' mixed to pine, and oak to 'secondary' oak). Additionally, strong confusion fluxes existed from oak forest towards cultivated grassland and from 'secondary' median and high tropical forest towards cultivated grassland. The confusion observed for all these rare forest classes suggests a under-estimation of forest cover in reduced class-specific areas of the map. It also reinforces the hypothesis emitted earlier about a 'secondary forest' fuzziness phenomenon, independent of the mapping zone.

Over the whole area however, producer's estimate totalizes 16.9% for 'primary' perennial tropical forest against 12.1% of mapped 'primary' perennial tropical forest, the difference being attributable to misclassified areas of mapped cultivated grassland. These figures suggest that the extent of the forest in the biological reserve is under-estimated in the NFI (over-estimation of the deforested area).

Within cultivated land, much grassland (3.7% of Los Tuxtlas area) was misclassified as annual crop, suggesting a under-estimation of the long-lasting and widespread conversion process from annual production systems to cattle raising in the region of Los Tuxtlas. Many misclassification occurrences for perennial crop, mainly identified as coffee and fruit plantations, are from mapped patches of grassland and of annual crop. In comparison to the area represented by these misclassifications, the area mapped as perennial crop is much (five times) smaller. There might be a sub-representation of 'traditional' treed production systems or market oriented fruit production systems within the mapped areas of cultivated land.

Candelaria

The dominant class in Candelaria (figure 2.5) is the sub-perennial median tropical forest (thereafter median forest), which covers 50% of the area (Table2.7). Other leading common classes are the sub-perennial low tropical forest (thereafter low forest) and cultivated grassland for pasture (both with about 17% coverage). The intermediate-low global accuracy (64.4%) reflects the combination of intermediate confusion levels of median forest and cultivated grassland and high confusion levels of the low forest.

Mutual, nearly counterbalanced, confusion exists between median and low forests. Many associations of tree species compose grades of transition between low and median forests, which could explain such confusion. Conversely, the strong confusion fluxes from the median forest class reported towards 'secondary' forest classes are not counterbalanced. Besides, much cultivated grassland (2.6% of all Candelaria area) was misclassified as median 'secondary' forest. Both results suggest a possibly strong overestimation of median forest by the NFI map.

Table 2. 6. Confusion matrix of the national forest inventory map in Los Tuxtlas at community level.

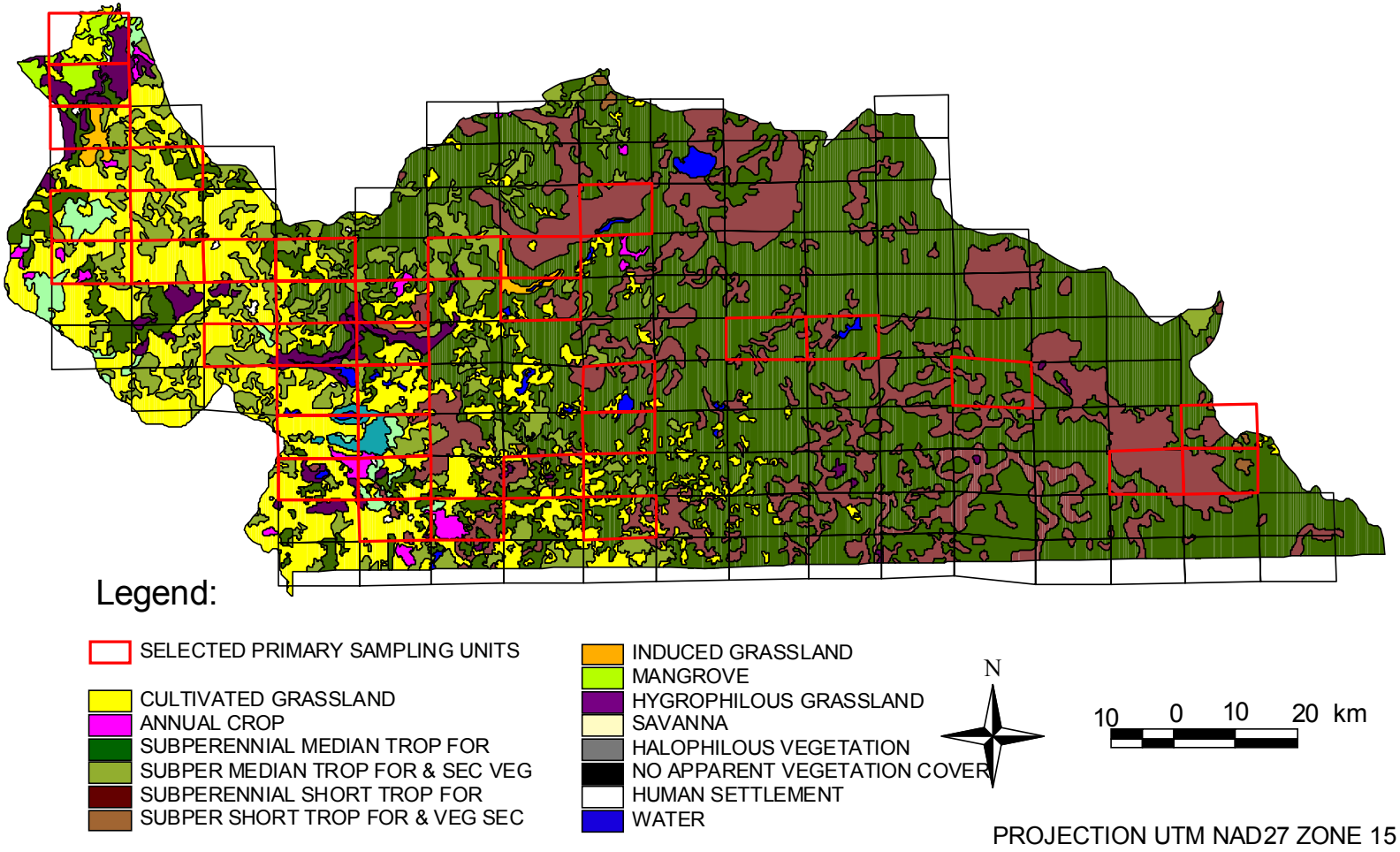
	130	200	210	420	421	511	600	601	700	800	801	1330	1400	1410	1600	1700	1800	n^a	Area frac[b]	User[c]	StdE[d]
130	**50.570**	4.506								5.508								121	60.584	0.83	0.03
200	0.240	**0.735**									0.270					0.045		86	1.290	0.57	0.05
210	3.761	2.603	**9.112**							0.289	1.880							122	17.645	0.52	0.05
420				**0.094**						0.002		0.009				0.005		94	0.111	0.85	0.04
421				0.059	**0.000**							0.052						89	0.111	0.00	0.00
511	0.028			0.152		**0.017**		0.037				0.045						99	0.279	0.06	0.02
600	0.051	0.002				0.004	**0.032**	0.025										96	0.113	0.28	0.05
601	0.018						0.069	**0.288**				0.037						90	0.411	0.70	0.05
700									**0.355**									99	0.355	1.00	0.00
800	0.720									**11.104**	0.308							118	12.133	0.92	0.03
801	0.962	0.120									**1.834**							97	2.916	0.63	0.05
1330			0.001								0.011	**0.024**						67	0.036	0.69	0.03
1400			0.036								0.007		**0.571**				0.050	93	0.664	0.86	0.04
1410										0.006	0.027	0.054	0.004	**0.102**				100	0.193	0.53	0.05
1600											0.003	0.004			**0.060**		0.003	82	0.069	0.87	0.04
1700	0.013	0.013	0.020								0.007					**0.600**		100	0.653	0.92	0.03
1800																	**2.437**	96	2.437	1.00	0.00
n^a	241	84	72	181	0	9	42	97	99	126	134	153	82	53	71	99	106	1649			
Area frac[b]	56.363	7.980	9.168	0.305	0.000	0.020	0.100	0.349	0.355	16.910	4.346	0.226	0.575	0.102	0.060	0.650	2.490		100.000		
Prod[c]	0.90	0.09	0.99	0.31		0.83	0.32	0.82	1.00	0.66	0.42	0.11	0.99	1.00	1.00	0.92	0.98			**77.9**	
StdE[d]	0.02	0.03	0.01	0.03		0.13	0.07	0.04	0.00	0.04	0.04	0.03	0.01	0.00	0.00	0.03	0.01				

Global accuracy was estimated at 77.9%.

[a] Sample size. [b] Class area fraction (out of the total Cuitzeo area), sum of the row/column matrix elements.

[c] User's (User) and Producer's (Prod) accuracy [d] Standard Error using binomial theory, with approximation to large sample sizes if applicable

Figure 2.5. National Forest Inventory map in the watershed of the Candelaria river (Mexican side).



The Primary Sampling Units (PSUs) selected by the sampling design are highlighted.

Table 2.7. Confusion matrix of the national forest inventory map in Candelaria at community level.

	130	210	820	821	830	831	1320	1330	1400	1410	1510	1700	1800	n^a	Area frac[b]	User[c]	StdE[d]
130	**11.852**	0.924		0.154	1.385	2.617				0.154				111	17.085	0.69	0.04
210	0.141	**0.523**								0.017		0.017		84	0.697	0.75	0.05
820	0.380	2.657	**34.915**	4.554	3.036	4.554								132	50.096	0.70	0.04
821	2.611	0.193		**4.835**	0.774	0.387								91	8.799	0.55	0.05
830		1.228	4.298	1.228	**9.210**	0.307		0.307		1.074				115	17.652	0.52	0.05
831			0.056			**0.080**				0.111				93	0.247	0.32	0.05
1320	0.073	0.012			0.367	0.159	**0.232**			0.208		0.012	0.012	88	1.076	0.22	0.04
1330			0.004	0.022	0.075			**0.263**		0.026				89	0.391	0.67	0.05
1400					0.012			0.068	**0.519**					97	0.599	0.87	0.03
1410	0.064		0.107		0.107	0.064		0.300	0.021	**1.585**				105	2.249	0.70	0.04
1510		0.008			0.062	0.004		0.017		0.266	**0.033**			94	0.390	0.09	0.03
1700	0.018											**0.073**		15	0.091	0.80	0.03
1800		0.047	0.024	0.008	0.008	0.008		0.047		0.141	0.047		**0.298**	80	0.628	0.48	0.06
n^a	134	95	150	77	155	83	19	97	85	231	14	15	39	**1194**			
Area frac[b]	15.139	5.592	39.404	10.801	15.036	8.179	0.232	1.002	0.540	3.583	0.080	0.102	0.310		**100.000**		
Prod[c]	0.78	0.09	0.89	0.45	0.61	0.01		0.26	0.96	0.44	0.41	0.72	0.96			**64.4**	
StdE[d]	0.04	0.03	0.03	0.06	0.04	0.01		0.04	0.02	0.03	0.13	0.12	0.03				

Same code conventions as in table 4. Global accuracy was estimated at 64.4%.

[a] Sample size. [b] Class area fraction (out of the total Cuitzeo area), sum of the row/column matrix elements.

[c] User's (User) and Producer's (Prod) accuracy [d] Standard Error using binomial theory, with approximation to large sample sizes if applicable

A local manifestation of the underestimation of vegetation degradation level is the misclassification of hygrophilous vegetation for 'secondary' low forest. The colonization of some hygrophilous grasses is an indicator of high human-induced perturbation of humid habitats. Strong confusion fluxes from the water and savannah classes towards hygrophilous grassland may further suggest a human-induced higher level of degradation than the one depicted by the NFI map. Although much area mapped as savannah corresponded to other LULC according to the reference data, the presence of species pertaining to savannah ecosystems reported in the area of mapped savannah patches (personal communication) confirmed us in the view that human-induced perturbation of former savannah patches was under-estimated by the NFI. Localized under-estimation of vegetation regeneration could also hold: over 3% of Candelaria area of mapped cultivated grassland would in reality be low forests.

Altogether, according to producer's totals, the over-estimation of primary forests would amount to 10% of the entire Candelaria area. Additionally, the total forest cover would be over-estimated by 3% and the hygrophilous grassland would be under-estimated by 1.5% (percentages of the entire Candelaria area).

5. Discussion

5.1 On the assessment methodology

The method described in this research, conceived for an accuracy assessment of the NFI, permitted the evaluation of the pilot areas within a probability sampling scheme. Thereby, the design is ready for operational use at national scale, with typically available data and skills. The two-stage feature achieved much more spatial control over the sample than would have achieved an alternative simple stratified design, thus greatly reducing costs incurred in the purchase of aerial photography and in photo-interpretation labour. Additionally, the stratification per mapped class permitted us to apprehend the taxonomically diverse NFI with a high level of detail.

A limitation of the assessment design in its application to each eco-geographical area was the limited geographic distribution of point samples, given the coarse scale (1:250 000) of the map and the high number of classes. For this reason, some cluster clumping effect did

probably occur in PSUs containing many sample points and the binomial standard errors for these cases under-estimate the real confidence interval of accuracy levels.

As a consequence, our analysis does capture trends of accuracy results and possible interpretations for these trends, although the absolute value of these trends should be taken with caution. To this effect, the trends commented for each eco-geographical area and as a conclusion of the pilot study, are also intended as interpretation keys of matrix confusions to the address of a potential user of the NFI.

The assessment design was conceived to meet requirements of a more extended area than each eco-geographical area taken separately. Indeed, the assessment design would readily be operational for a systematic sample of the entire country stratified by state, for example. For budget reasons so far, the design could only be tested on the mentioned eco-geographical areas. A strength of the assessment design is the relative uniformity of the response design, rendered possible by the vast (across eco-systems) experience of the photo-interpreter specialist in the team.

In most cases of low photo-interpretation confidence, ground visits complemented the reference data although systematic ground visits were ruled out as being cost-prohibitive. Given the budget available for accuracy assessment, likely developments of the design in the near future will focus on the reduction of border and scale effects, on fuzzy representation of landscapes and on the statistical analysis. Applying the assessment to a more extended territory is likely to be a more distant priority.

5.2 On the results of the accuracy assessment

Global accuracy

Globally, the range of accuracy levels found in this pilot study (64-78%) is somewhat above the range (46-66%) reported in Stehman *et al.* (2003) at similar taxonomic resolution (Level II) for Eastern regions (the smaller regions of the NLCD) of the USA. The spatial resolution of the USA map is much finer (Landsat pixel) implying greater difficulties to achieve high accuracy. Although the taxonomic resolution for forests in the Mexican NFI is greater than that of the NLCD Level II, the taxonomic diversity (number of classes) per region is greater in the USA regions (21 classes) than in each separated area

of our pilot study, except in Cuitzeo. This also implies comparatively greater difficulties for the USA map to achieve high accuracy. In Canada, Remmel et al. (2005) report a comparison between the 2000 Landsat-derived map of the EOSD project and the national forest inventory generated via photo-interpretation. The spatial resolution of these maps (Landsat pixel) is also much finer than that of the Mexican NFI. The coincidence reported for the 6 class aggregation level (close to but coarser than the NFI community level) was 64%.

The coarse spatial scale (1:250 000) NFI maps were generated via visual interpretation of Landsat colour prints at scale 1:125 000 (Mas et al., 2002), whereas the above mentioned maps were generated via fine resolution automatic classification of Landsat imagery (unsupervised clustering and visual labelling of classes). Although visual interpretation yielded a coarser map product, the results obtained in this research suggest that this classification method might have produced a map with accuracy levels that are difficult to reach elsewhere at equivalent taxonomic resolution. This preliminary result with regards to taxonomic resolution is worth stating since Mexico mega-diversity is certainly a motive of many IFN applications. However, an accuracy assessment of the entire territory is necessary to yield accuracy levels at national scale.

‘Secondary’ vegetation

In general, ‘primary’ vegetation classes seem relatively well represented (low confusion scores) in the pilot areas. By contrast, the forest classes with herbaceous, shrub and/or secondary vegetation (‘secondary’ forests) are affected by much confusion within their forest biome group. Difficulty in consistently distinguishing forests along levels of tree density has already been observed elsewhere. For example, Remmel *et al*. (2005) report this ambiguity as being the prominent source of confusion at the 20 class aggregation level of their classification scheme. However, in their case, confusion between forests of differing tree densities appeared to be confined to a forest type. Conversely, confusion fluxes from our NFI ‘secondary’ forests are stronger than confusion fluxes from ‘primary’ forests, and are directed not only to the ‘primary’ forest of the same type but to other ‘secondary’ and ‘primary’ forest types.

The secondary forest fuzziness phenomenon, apparently independent of the mapping zone, bears various interpretations. One of these interpretations resides in the intrinsic choice of the classification scheme, where a vegetation cover physiognomically deprived of original vegetation (typically, an abandoned fallow field) is perceived as necessarily pertaining to a vegetation type, the latter being understood as the original vegetation cover before any anthropogenic modification had occurred.

This concept might provoke difficulties in mapping secondary vegetation based on medium resolution (Landsat-type) satellite imagery. Undoubtedly, a strength of the classification scheme is its compatibility with INEGI archive map labels. This fuzziness phenomenon might be one of the necessary trade-offs for the fundamental criteria of compatibility.

6. Summary and conclusion

The objectives of this article were to give a synopsis of the accuracy assessment methodology applied on the Mexican NFI map, and to report results for four mapping areas with contrasting LULC situations. The accuracy assessment met both specified criteria of including a probability sampling protocol and a practical, cost-effective scheme that would render applicable at national scale the evaluation of the most detailed (community) level NFI map. Error matrices were provided and permitted a detailed thematic analysis per class, per eco-geographical area:

Generally, ‘water’ and ‘mangrove’ classes and ‘other cover types’ (no apparent vegetation cover and human settlement) were represented with very high accuracy (mostly above 79%). However, very high inter-confusion was recorded within aquatic non-treed vegetation covers. Most common ‘primary’ forest and land use (cropland) classes were characterized by accuracy levels above 70%. By contrast, patterns of high confusion were registered in all mapping zones for (generally sparsely distributed) ‘secondary’ forest classes, and were possibly due to a combination of ambiguity generated by the inclusion of secondary vegetation in the classification scheme and digitizing practices during visual interpretation.

Altogether, some balance resulted from most of the confusion fluxes generated by this 'secondary forest' fuzziness phenomenon and the accuracy assessment actually served as a useful quantitative estimator of the general trends of these errors.

Next, the following tendencies of misclassification errors were observed. For mostly temperate areas (Cuitzeo and Tancítaro), the level of degradation or anthropisation of some vegetation cover may have been overestimated by the NFI map. For example, the map might under-estimate the presence of oak populations in temperate forests, and classify as crops and induced grasslands much land surface in reality covered by scrubland (in Cuitzeo). For Los Tuxtlas, an opposite tendency is perceived for temperate forests where the NFI map under-estimates the level of degradation. As of croplands, an opposite tendency with respect to temperate zones is perceived as well: the presence of cultivated grasslands for pasture to the expense of other production systems seems to be under-estimated by the NFI. The forest mass (mostly tropical forest) of the biosphere reserve seems under-represented in the NFI (over-estimation of the deforested area). On the contrary, an under-estimation of non-forested or degraded vegetation cover is clearly perceived in Candelaria where much area of the mapped sub-perennial median tropical forest (a piece of the largest tropical forested mass in Mexico) was photo-interpreted as 'secondary' forest types. Some of this perception is also present in the under-estimation of the extent of hygrophilous grasslands, indicator of a human-induced alteration of other ecosystems, and the over-estimation of the extent of other native ecosystems.

In an ecological sense, the area occupied by tropical 'primary' vegetation classes (common or rare) tended to be overestimated in the regions investigated, with the corollary impression that tropical deforestation is underestimated by the NFI, with respect to the INEGI maps of 1976 (series I), the highly reliable cartographic source in Mexico. On the contrary, the presence of native vegetation seemed underestimated in both mostly temperate areas; the tendency reflected there would be the overestimation of temperate deforestation if the NFI were compared with highly reliable 1976 INEGI maps. Among the explained confusions, omissions of changes from former INEGI maps might reflect limitations of the classification process based on visual interpretation of 20-30m resolution satellite imagery, and the harsh time constraint for the NFI map production in year 2000 (interpretation of more than 120 Landsat images in an 8 month process). We suggest our

assessment design be augmented with one stratification level (stratification per administrative state) and be applied on a systematic sampling of the entire country (along flight lines of a systematic aerial photographic coverage).

7. Acknowledgments

This research was funded by CONACYT (the national science and technology board in Mexico) under project number 38965T and under scholarship number 205000. The help of the respective local field and academic teams in all four areas must be duly acknowledged here:

Cuitzeo: Thanks are due to young and elder inhabitants of the El Cerro village for providing information on typical tropical dry LULC in the Cuitzeo area, thanks to Manuel Mendoza for useful comments on landscape dynamics. Thanks to all members of the hydrological Cuitzeo Lake watershed team for field work, and many more in Morelia city.

Tancítaro: Thanks are due to the technical foresters of the San Juan Parangaricutiro community for their knowledge and data on forest distribution and landscape dynamics; Special thanks to Alejandro Velázquez for providing the logistical platforms to the region of Tancítaro and his extensive knowledge on vegetation. Thanks to Arturo Garrido for providing auxiliary maps and thanks to many more.

Los Tuxtlas: Thanks are due to inhabitants of the ejido La Perla San Martin for precious insights on agricultural frontier dynamics; Thanks to Martin Ricker of the UNAM Biological Station for providing interpretation material and easing field work; thanks to Adriana Bermeo, to Lorenzo Arteaga and to the DECOTUX AC team for providing key elements of landscape dynamics. Thanks to Gerardo Bocco and Fernando Rosete from the National Institute of Ecology (INE) for providing auxiliary maps and remote-sensing resources.

Candelaria: Thanks are due to our guide Yam Cayam in the Calakmul area, to Pedro Zamora and all members of the EPOMEX team for helping during field work and for their expertise in sub-perennial forest types.

8. References

FGDC (1997) FGDC vegetation classification and information standard, June 3, 1996 draft. Reston, VA: Federal Geographic Data Committee, Vegetation Subcommittee (FGDC-VS), FGDC Secretariat.

Mas JF, Velázquez A, Palacio-Prieto JL, Bocco G, Peralta A, Prado J (2002) Assessing forest resources in Mexico: Wall-to-wall land use/ cover mapping. *Photogrammetric Engineering and Remote Sensing* 68 (10), pp 966-969.

Laba M, Gregory SK, Braden J et al. (2002) Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map. *Remote Sensing of Environment* 81, pp 443-455.

López E, Bocco G, Mendoza M, Velázquez A and Aguirre-Rivera R (2006) Peasant emigration and land-use change at the watershed level: A GIS-based approach in central Mexico, *Agricultural Systems* 90, pp 62-78.

Palacio-Prieto JL, Bocco G, Velázquez A, Mas JF, Takaki-Takaki F, Victoria A, Luna-González L et al. (2000) La condición actual de los recursos forestales en México: resultados del Inventario Forestal Nacional 2000, *Investigaciones Geográficas* 43, pp 183-202.

Remmel TK, Csillag F, Mitchell S, Wulder M (2005) Integration of forest inventory and satellite imagery: a Canadian status assessment and research issues. *Forest Ecology and Management* 207, pp 405-428.

Stehman SV & Czaplewski RL (1998) Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment* 64, pp 331-344.

Stehman SV, Wickham JD, Smith JH & Yang L (2003) Thematic accuracy of the 1992 National Land-Cover Data for the eastern United-States: Statistical methodology and regional results. *Remote Sensing of Environment* 86, pp 500-516.

Stehman SV, Wickham JD, Yang L & Smith JH (2000) Assessing the accuracy of large-area land cover maps: Experiences from the Multi-resolution Land-Cover Characteristics (MRLC) project. In G. B. M. Heuvelink, & M. J. P. M. Lemmens

(Eds.) *Accuracy 2000: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences,* pp. 601-608. Delft Univ. Press., The Netherlands.

Wickham JD, Stehman SV, Smith JH & Yang L (2004) Thematic accuracy of the 1992 National Land-Cover Data for the western United-States. *Remote Sensing of Environment* 91, pp 452-468.

Zhu Z, Yang L, Stehman SV & Czaplewski RL (2000) Accuracy assessment for the U.S. Geological Survey regional land-cover mapping program: New-York and New Jersey region. *Photogrammetric Engineering and Remote Sensing* 66, pp 1425-1435.

Chapter 3: Thematic and Positional Fuzziness

A FUZZY SETS BASED ACCURACY ASSESSMENT OF LAND COVER MAPS AT USER-DEFINED THEMATIC AND POSITIONAL PRECISION

A methodology for assessing land cover/land use maps with degrees of tolerance in thematic and positional aspects.

In chapter 2, global accuracy estimates of 64% to 78% were registered among the four eco-geographical areas, and the lower accuracy levels were found in areas more densely covered with forests. The analysis of confusions revealed difficulties in unambiguously interpreting or labelling forests with secondary vegetation, herbaceous, and/or shrub-like vegetation as well as distinguishing between aquatic vegetation types.

The method was applied to the Mexican National Forest Inventory (NFI) map in the Candelaria River watershed, covering 1.1 million hectares, where lies part of the largest tropical forested mass in Mexico. A challenging task consists in the identification of thematic and positional errors of the NFI map in terms of the distribution of these forest communities.

In chapter 3, a fuzzy sets based methodology is developed for differentially assessing vector-based maps according to errors related to position, thematic fuzziness and misclassification. Apart from misclassification errors, the method aimed at estimating the positional and thematic fuzziness of median and low sub-perennial tropical forests on the map.

The nature of errors (mainly positional or mainly thematic) for forest classes of the map may be valuable information in the context of environmental modeling over large forest extents. The work frame is applicable at regional scale and provides a means of estimating quantities of positional and thematic errors, at user-defined precision, an asset for the understanding and interpretation of confusions.

A fuzzy sets based accuracy assessment of land cover maps at user-defined thematic and positional precision.[Θ]

Stéphane Couturier
Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)
Centre d'Etudes Spatiale de la Biosphère (CESBIO), CNES/CNRS/Paul Sabatier University
Jean-François Mas, Gabriela Cuevas
Instituto de Geografía, Unidad Académica Morelia Universidad Nacional Autónoma de México (UNAM)
Jorge Benítez
Centro EPOMEX, Universidad Autónoma de CampecheAbstract:
Álvaro Vega, Valdemar Tapia
Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)

Land cover maps with detailed taxonomy constitute a basis for environmental planning in bio-diverse regions. The assessment of these maps faces uncertainty due to high taxonomic diversity and unclear boarders between forest classes. The nature of errors (mainly positional or mainly thematic) for each class of the map may be valuable information in the context of environmental modeling over large forest extents. In this paper, a methodology is developed for differentially assessing vector-based maps according to errors related to position, thematic fuzziness and misclassification. The method was applied to a fragment of the Mexican National Forest Inventory map in the Candelaria River watershed, covering 1.1 million hectares. The work frame is applicable at regional scale and provides a means of estimating quantities of positional and thematic errors, at user-defined precision, an asset for the understanding and interpretation of confusions.

[Θ] In review for publication in *Photogrammetric Engineering and Remote Sensing*

1. Introduction

Land Use and Land Cover (LULC) maps at regional scale in sub-tropical countries serve as input data for continental land use change studies (Achard et al., 2002) and ecological modelling (Houghton et al., 2000; Barbosa et al., 1999; Hely et al., 2006). The quantitative information on the total surface covered by each land cover type is certainly a key variable of such assessments. The information provided by maps on the spatial distribution of vegetation types is also critical in evaluating threats to biodiversity (Hunter, 1996) and understanding the drivers of deforestation (Soares-Filho et al., 2006). For example, the connectivity between neighbouring forest patches is a major parameter in conservation strategies for the implementation of a Mesoamerican Biological Corridor (Miller, 2001). Information on the nature of errors (thematic or positional) in base maps can therefore be most useful. However, accuracy assessments at regional scale are generally tailored to give one measure of error (either global or per class), which synthesizes all types of inaccuracies on the map. In fact, procedures are incorporated in the assessment design to implicitly set a degree of tolerance regarding thematic and/or positional aspects at which the map is evaluated. Throughout this paper, we will refer to the terminology of Stehman & Czaplewski (1998) regarding steps and desirable properties of accuracy assessments (e.g. probability sampling, evaluation and labeling protocols, the synthesis of the evaluation) because they are now widely used and accepted.

Sometimes, two degrees of tolerance regarding a thematic or a positional aspect are included in the assessment design and separate accuracy results (confusion matrices and indices) are provided. For the assessment of the National Land-Cover Data (NLCD) of the USA in 1992, Stehman et al. (2003) operate at two degrees of tolerance regarding position. They fix one degree of tolerance at pixel level (which can be considered a zero positional tolerance), and another one at the mode pixel level within a 3x3 pixel support region of the LULC map. Laba et al. (2002) present separate results at two degrees of tolerance regarding thematic uncertainty. The first assessment is made with a crisp agreement definition, and the second assessment is based on the fuzzy linguistic scale of Gopal & Woodcock (1994). The latter assessment represents more thematic tolerance than the ‘crisp’ evaluation.

For map accuracy assessment, thematic tolerance is generally treated through fuzzy classification, related to fuzzy sets theory. Fuzzy classification is traditionally present under the form of Gopal & Woodcock (1994)'s linguistic metric, rather than a quantitative fuzzy metric, because linguistics are well adapted to visual classification of non-ordinal categories. Positional tolerance, however, is less explicitly treated in map accuracy assessment, or treated in a non-systematical way, perhaps because available reference data is typically obtained punctually in space (e.g. one pixel of Landsat-based maps). Another reason may be that the map producer wants to make the assessment at one particular scale, typically the scale of the map. However, in practice, because of the lack of alternative data, maps from governmental agencies interest a number of users whose working scale is not necessarily the scale of the map. For example, national LULC information is widely available at scale 1:250 000 in Mexico (Mas et al., 2002). For environmental planning, many users could be interested in knowing the reliability of this map below or above map scale.

We reformulate the framework for accuracy assessments at regional scale, with the inclusion of the notion of bi-dimensional degree of tolerance, with a positional and a thematic aspect. To aid this reformulation, we focus on the use of reference maplets, based on aerial photograph interpretation, and we employ the formalism of fuzzy sets theory because of the recent advances in map comparison with positional and thematic aspects by Hagen (2003). We applied the newly formulated accuracy assessment to the Mexican National Forest Inventory (NFI) of year 2000. The NFI is a vector-based LULC map at scale 1:250 000.

In the next section, we identify error types in the context of accuracy assessment of maps. Then we describe the study area, where the distinction between patches of two forests communities in a complex sub-tropical environment (see table 3.1) constitutes a challenge for map assessment. The aim of this case study was the quantification of thematic and position-related errors on the NFI map. In the methods section, we detail each step of the new conceptual framework, including two major steps which are the construction of reference maplets and the fuzzy map comparison. The sampling design and the synthesis of evaluation are also detailed. An accuracy assessment with traditional point-like reference data is also applied on the same area. Part of the results is dedicated to the

quantification of errors among classes of the NFI with the differentiation of error types permitted by the method. We also discuss the differences between the new method and the assessment with traditional point-like reference data.

2. Background

The accuracy assessment of a map aims at quantifying errors regarding its spatial and thematic representation with respect to reality, usually through the elaboration of a confusion matrix. Congalton and Green (1993) relate these errors of the map to imprecise delineation and/or misclassification. However, the imperfect process of the assessment itself generates erroneous statements on whether the map represents reality or not. In such cases, the assessment may report *fictitious* errors on the map (also labelled 'non-error differences' in the confusion matrix by Congalton and Green, 1993), as opposed to the *real* errors of the spatial and thematic representation. After a review of published works describing causes of errors in the confusion matrix (Congalton and Green, 1993; Aspinall and Pearson, 1995; Green and Hartley, 2000; Foody, 2002; Powell et al., 2004), we built a list (see table 3.2) of 7 main error types (A to G) with a series of attributes. For example, the column 'nature of error' refers to whether the error is real or fictitious, and whether the error is positional, thematic or temporal-based.

In accuracy assessment designs, methods that reduce the inclusion of some error types are introduced, as an effort to focus the estimate on real, thematic errors (misclassification errors). For example, the method of visually locating sample points on the original satellite imagery, proposed in Zhu et al. (2000), reduces the inclusion of errors due to geometric inconsistencies between the reference data and the map (a combination of errors A and B). In this paper, we extend Zhu et al. (2000)'s approach and produce two-dimensional reference maplets, co-registered to the original satellite imagery. Even assuming perfect co-registration, another potential source of fictitious positional errors (Error type D) is related to the subjective delineation of boundaries in a heterogeneous or fragmented landscape (Powell et al., 2004). In this paper, we treat this kind of errors with a positional tolerance operator during the fuzzy map comparison process. On the other hand, thematic ambiguities (error types E and F), correlated with the above-mentioned positional uncertainty, are likely to occur in the case of complex forest environments.

Table 3.1 Distribution of classes (community and formation aggregation levels) of the national forest inventory map in the watershed of the Candelaria River.

Class	Community	Formation	Area frac[a]	Area (km2)
1	Cultivated grassland	Cropland	17.08	1908.25
2	Annual crop		0.70	77.85
3	Med. and high sub-per forestb [b]	Tropical Forest	50.10	5595.31
4	Med. and high sub-per forest & sec. veg. [b]		8.80	982.82
5	Low sub-per forestd [d]		17.65	1971.60
6	Low sub-per forest & sec. veg. [c] [d]		0.25	27.57
7	Savannah	Grassland	1.08	120.13
8	Induced grassland		0.39	43.65
9	Mangrove	Hygrophilous vegetation	0.60	66.93
10	Hygrophilous grassland		2.25	251.24
11	Halophilous vegetation	Other vegetation types	0.39	43.56
12	Human settlement	Other cover types	0.09	10.19
13	Water		0.63	70.11
	All		1.00	11169.21

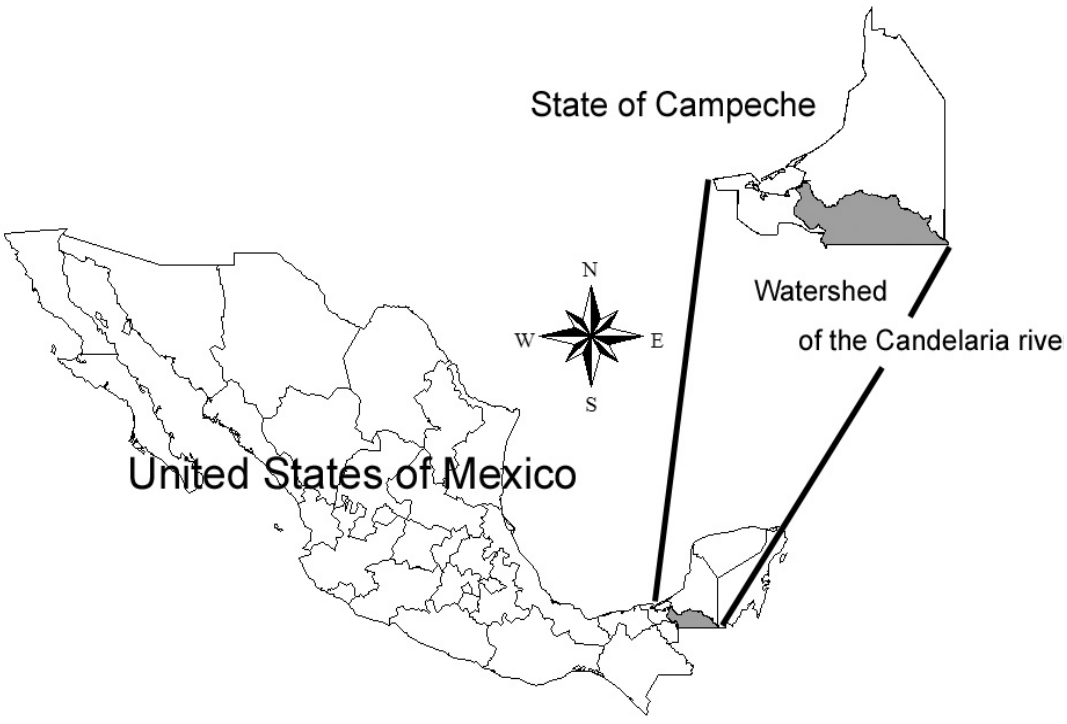
[a] Area percent fraction with respect to total area. [b] Median and high sub-perennial tropical forest
[c] Secondary vegetation (secondary vegetation, herbaceous and/or shrub-like vegetation, according to NFI classification scheme) [d] Low sub-perennial tropical forest,

The assignment of membership values to these sites within a fuzzy linguistic scale has traditionally been used as a means of reducing the inclusion of these fictitious thematic errors (e.g. Laba et al., 2002). In this research, we treat this kind of errors with a thematic tolerance operator; the maximum membership value of a reference site defined the thematic tolerance.

3. Study Area

The watershed of the Candelaria River in the Mexican state of Campeche (see figure 3.1) covers more than 1.1 million hectares. The area includes a portion of the Calakmul Biosphere Reserve, located within the largest tropical forested mass of the country, and contiguous to forested extents in neighbouring Belize and Guatemala. According to the NFI map, median and high sub-perennial tropical forest (thereafter 'median forest') is the most conspicuous physiognomy of the region (with a coverage over 50%, see table 3.1). Another common vegetation community is the low sub-perennial tropical forest (thereafter 'low forest'). These forest types are typical of a tropical sub-humid climate, and the low forest occurs on poorly drained soils. A challenging task of the accuracy assessment consisted in the identification of thematic and positional errors of the NFI map in terms of the distribution of these forest communities. Increasing levels of alteration of the original vegetation cover towards the coast characterize the hydrological basin. A major land use in year 2000 according to the NFI map is cultivated grassland for pasture.

Figure 3.1 Location map.of the Candelaria area



The accuracy assessment was tested on the mapping zone shaded in grey.

Table 3.2 Inventory of error types found in an accuracy assessment

Error type	Nature of error	Possible causes of error	Consequence for user	Improvements in the assessment	Step of Accuracy Assessment
A. Geo-location of the reference data (RD)	Positional Fictitious	1.Ortho-rectification error of RD (ex: positional uncertainty of photo centre) 2. Delineation of sampling units on RD	Distortions may lead to map strips being falsely labelled errors	1.Superposition of RD over RSSM, although then the assessment will not attend real map positional errors (error type E)	Restitution to Landsat geometry
B. Location of the Remote Sensing Support of the Map (RSSM)	Positional Real	1. Geo-referencing error of RSSM 2. Map digitising uncertainty	Distortions may lead to map strips being erroneously mapped	1.Improve quality of map geo-reference	
C. Interpretation of changing landscape	Temporal Fictitious	1. Insufficient temporal coverage of RD	Errors on highly contrasting LULC classes due to rapid changes	1.Superposition of RD over RSSM and track false change 2.More ground knowledge on LULC classes susceptible to change	Reference maplet construction
D. Delineation of heterogeneous landscape	Positional Fictitious	1. Fragments or thin corridors of a class below MMU 2. Fuzzy boarder between LULC classes (transition zone)	Risks of accuracy bias for LULC classes present but non-dominant in mosaics	1. Treatment of intersect polygons below MMU or below user-defined tolerance distance for transition zones	Positional tolerance operator
E. Interpretation of the reference data	Thematic Fictitious	1. Apparent ambiguity in RD and insufficient ground knowledge to distinguish between LULC classes 2. Subjectivity of photo-interpreter 3. Data entry error	Ambiguities or subjective interpretation tend to occur for LULC classes along a continuum	1. More ground visits on uncertain photo-interpreted sites 2. Double check with independent photo-interpretation 3. Adjust the maximum fuzzy membership value	Thematic tolerance operator
F. Interpretation of ambiguous landscape	Thematic Fictitious	1. Mosaic of various LULC classes in a polygon 2. Fuzzy mixture of various LULC classes (eco-tones within a polygon)	Risks of accuracy bias for fragmented LULC classes and LULC classes present in eco-tones	1. Improve the clarity of LULC class definition in the classification system 2. Adjust the maximum fuzzy membership value	Thematic tolerance operator
G. Interpretation of the remote sensing support of the map	Thematic Real	1. Ambiguity in RSSM and insufficient ground knowledge to distinguish between LULC classes 2. Insufficient knowledge or skill of RSSM interpreter 3. Data entry error 4. Deficiency of classification system or mapping conditions	Ambiguities or subjective interpretation tend to occur for LULC classes of similar biomes. Other sources of error may explain confusion between contrasting LULC classes on the map.	1. Study of ambiguity between LULC classes in RSSM to interpret the estimation of errors found in the assessment. 2.Discussion on limitations of the classification system as a result of the analysis of class confusions	Discussion

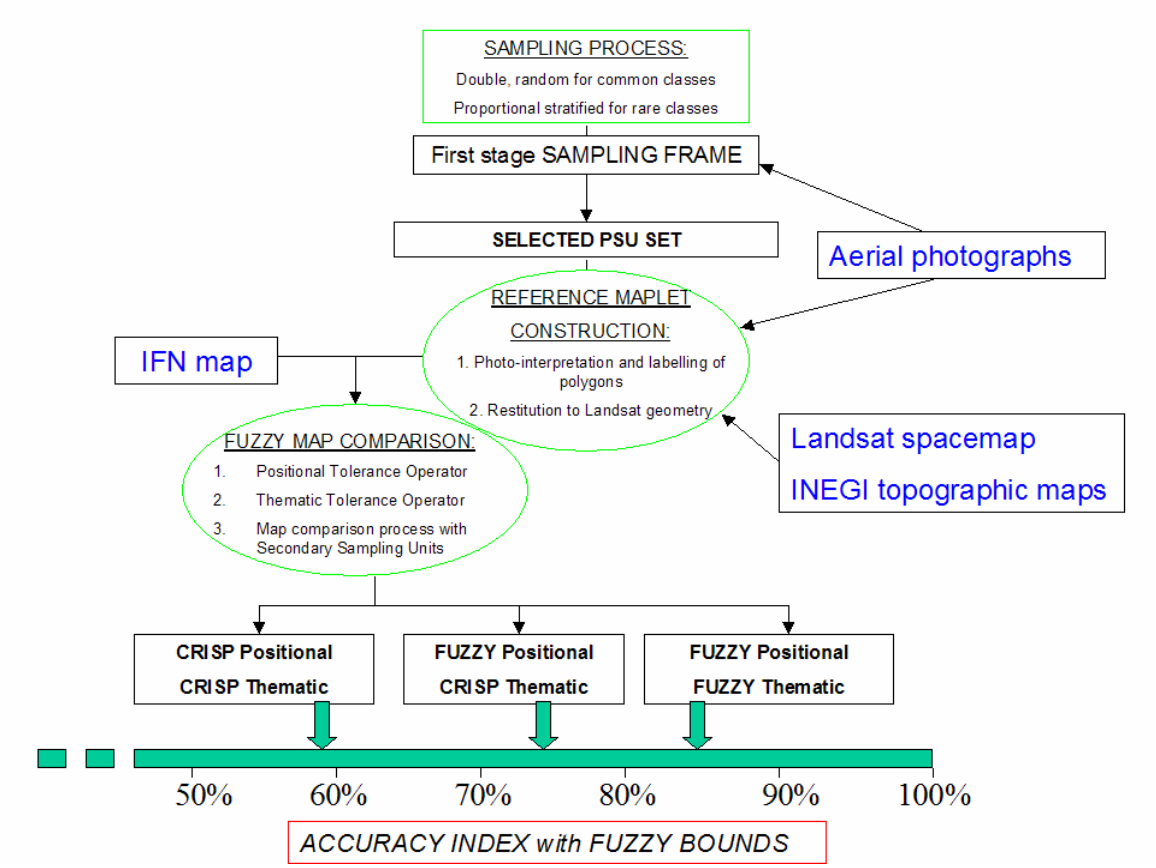
RSSM = Remote Sensing Support of the Map ; RD = Reference Data; MMU: Minimum Mapping Unit; LULC = Land Use/Land Cover; strategies used in this paper are highlighted in bold

Apart from the taxonomic detail and spatial intricacy of LULC classes, illustrative of environmental situations found in many subtropical settings, a key motivation for selecting this area as the case study was the availability of a quasi-synchronous aerial photograph coverage, exploitable in the process of independently verifying the 2000 Landsat-based NFI. These reference data are detailed in the section on photo-interpretation.

4. Accuracy Assessment Algorithm

The accuracy assessment is composed of the sampling process, the reference maplet construction, the fuzzy map comparison and, finally, the synthesis of the evaluation. The flow chart in figure 3.2 illustrates each step of the method.

Figure 3.2 Flow chart of the accuracy assessment method

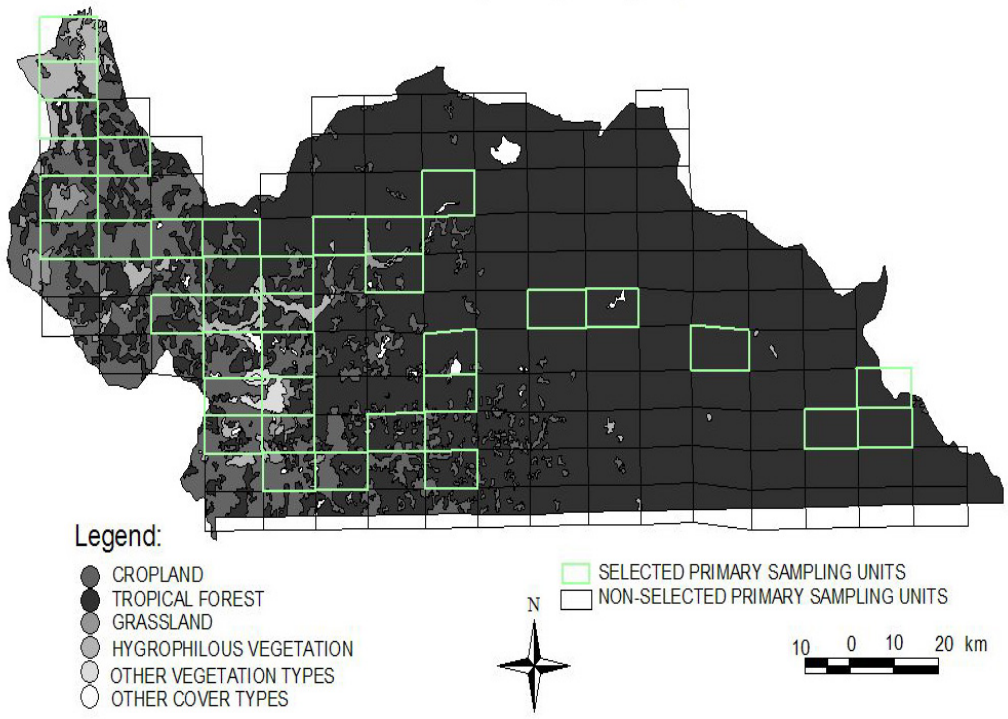


4.1 Sampling Process

The high taxonomic diversity of the map implies the presence of many sparsely distributed classes in the sample. For cost efficiency, the sampling design incorporated a two-stage (double) sampling protocol, where aerial photography frames formed the Primary Sampling Units (PSUs).

The mode of selection for PSUs followed a 'hybrid' scheme: a simple random selection of PSUs (see, for example, Stehman et al., 2003) was employed for common classes (area fraction above 5%) and a proportional stratified selection (as described in Stehman et al. 2000 and further discussed via personal communication) was employed for sparsely distributed or rare, classes (area fraction below 5%). Figure 3.3 illustrates the entire set of PSUs considered in the sampling process, as well as the PSUs which were selected in the sample. A photographic frame was included in the initial set of PSUs if its centre was located in the watershed, an inclusion rule also adopted by Zhu et al. (2000).

Figure 3.3 Selected Primary Sampling Units in the Candelaria area



The National forest inventory map is displayed at formation aggregation level in the watershed of the Candelaria River (Mexican side). The original set of Primary Sampling Units (PSU) included in the sampling design is represented. The PSUs that were selected via the sampling process are highlighted.

A regular 500m-spaced two dimensional grid formed the set of Secondary Sampling Units (SSUs) at the second stage. Indeed, a scale criterion during map production was to leave out polygons less than 500 meters wide. Once the sample PSUs were selected, all points of the second stage grid included within these sample PSUs were assigned their mapped LULC class. The full second stage sample consisted of 100 points from each mapped class.

In order to preserve equal inclusion probabilities at the second stage and maintain a low complexity level of statistics (i.e. standard stratified random formulae to compute estimators of accuracy), the selection of PSUs was done via proportional stratified random sampling (see this option presented in Stehman et al. 2000:604) for rare classes. The selection of the second stage sample was obtained via proportional stratified random sampling from the points contained in the sampled PSUs, this time with a probability inversely proportional to the abundance of the class. Figure 3.4 shows the sample of SSUs in two of the selected PSUs, as a result of the second stage sampling. A high density of SSUs is observed in the polygon labelled 'water', a fragmented and rare class. The four selected PSUs for the 'water' class contained many SSUs, because of the rare occurrence of the class in the entire Candelaria area. By contrast, SSUs are well dispersed on the neighbouring PSU, where common mapped classes 'median forest' and 'low forest' are sampled, according to a random PSU selection.

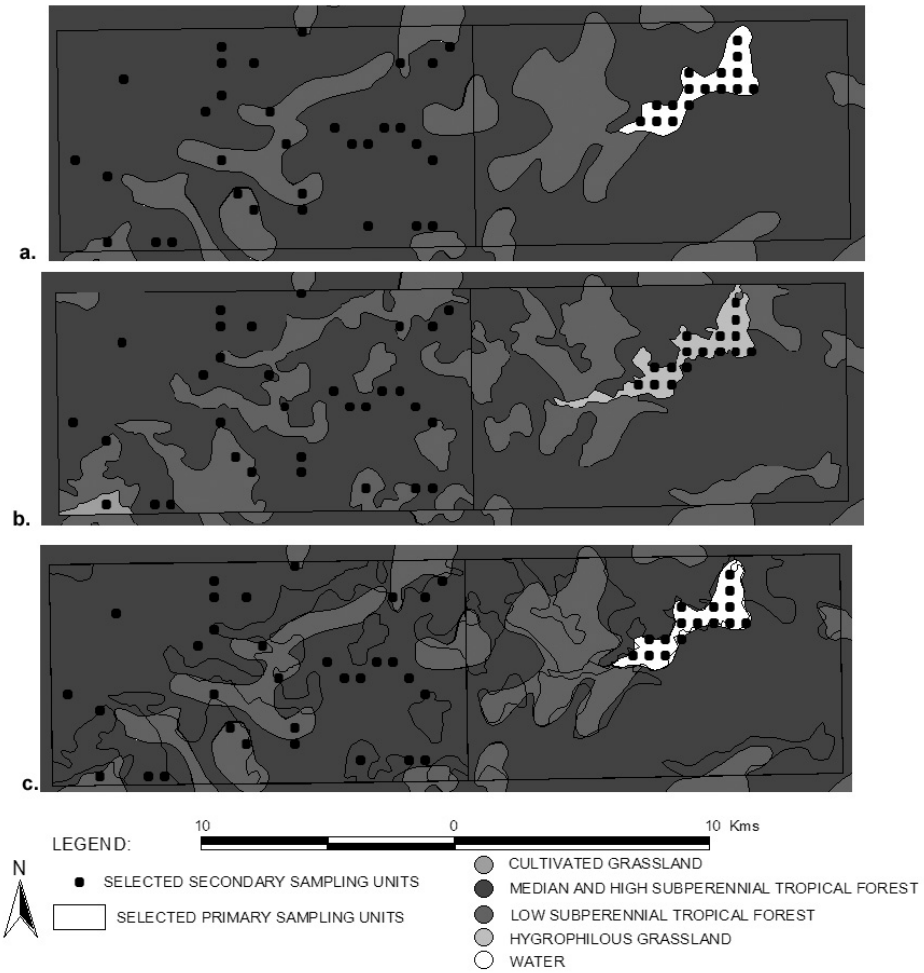
4.2 Reference maplet construction

4.2.1 Photo-interpretation

174 black and white aerial photograph prints at 1: 75 000 from the INEGI archive, acquired from January 2000 to March 2002 over the Candelaria area, were used for reference LULC classification. A first specialist interpreter was assigned to produce, via stereoscopic interpretation, reference maplets at scale 1:125 000 (finer resolution than the NFI maps and suitable for photo-interpretation) on transparent slides for each photographic frame corresponding to a selected PSU. The reference maplets were drawn according to the community aggregation level of the NFI classification scheme; a confidence rating was applied to photograph interpretation, comparable to the one used in Zhu et al. (2000). A team of local specialists had previously worked on the same

photographic material to generate cartography with more precise spatial (1:75 000 scale) and equal taxonomic resolutions, with the help of intensive field work. Both photo-interpretation works were executed in complete independence with respect to each other and with respect to the NFI map production. The photo-interpretation of the first specialist was checked against the more precise cartography. Additionally, wherever the first interpreter registered low confidence ratings, his mapping was confirmed or corrected by information collected in the field by the local team.

Figure 3.4: Comparison of maps in two Primary Sampling Units



Representation of the Secondary Sampling Units in two selected primary sampling units: (a) National Forest Inventory (NFI) map. (b) Reference maplet with same legend as NFI map. Only the class with the highest degree of membership has been represented. (c) NFI map with polygons of reference map overlaid. The reference sites (secondary sampling units) are represented as black dots.

4.2.2 Restitution to Landsat geometry

Geometric consistency between the reference material and the NFI map was ensured through two alternative procedures. The first one, focussed on each sample point, resembled a technique described, for example, in Zhu et al. (2000): Each sample point (selected SSU) was located on the colour composite of the geo-referenced Landsat images (geometric basis of the NFI map) on screen. The location of the point was visually transferred onto the transparent slide over the photographic material. All attributes of the reference labels found on the sample point were then digitally registered according to the photo-interpreted maplet. For this point by point (SSU-based) evaluation, we subsequently defined agreement as a match between the reference label and any of the NFI map labels within a buffer of 500m of the sample point. This 500m rule was related to scale instructions for map production as mentioned in the sampling process. This accuracy assessment method was meant to resemble some existing methods in the literature.

An alternative technique was based on the PSUs rather than on the SSUs and consisted in the digital restitution of the entire reference maplets. This technique permitted the application of the fuzzy map comparison: First, the Landsat imagery was displayed at scale 1:75 000 on a selected rectangular PSU. Second, the transparent slide containing the reference maplet of the corresponding rectangle (PSU) was superimposed over the screen. Digitising of all polygons of the non-georeferenced maplet was then done by visually 'correcting' their shape according to spatial patterns evident on the Landsat image colour composite. Figure 3.4 illustrates such geometrically corrected reference maplets in the case of two selected PSUs. An independent assessment of this visual co-registration was conducted. 200 points were randomly selected on the photographic material. Of these 200 points, 2 points (1%) were found with erroneous land cover on the digital satellite imagery, all within 250m of a polygon with the expected land cover.

4.3. Fuzzy Map Comparison

4.3.1 Fuzzy formalism

The NFI map M_{NFI} , a portion of which is represented in figure 3.4, is a vector-based map where every polygon is represented by a single class and where polygon boundaries have exact location. Therefore, all points of M_{NFI} possess a Crisp attribute Vector V_{crisp} defined as follows (Hagen, 2003):

$$V_{crisp} = \begin{pmatrix} \mu_{crisp,1} \\ \mu_{crisp,2} \\ . \\ . \\ . \\ \mu_{crisp,K} \end{pmatrix} \qquad (3.1)$$

where K is the total number of categories. If i is the class of the observed point, a degree of membership is set for each class according to equation 2:

$$i \rightarrow \mu_{crisp,i} = 5, \mu_{crisp,j} = 1, (i \neq j) \qquad (3.2)$$

Where 1 and 5 respectively represent the lowest and highest scores on the fuzzy linguistic scale of interpretation (Gopal & Woodcock, 1994). On the other hand, a reference maplet M_{REF}, represented in figure 4b, is interpreted with a certain thematic fuzziness. Therefore, the attribute of all points in M_{REF} is a Fuzzy Category Vector V_{cat} defined as follows (Hagen, 2003):

$$V_{cat} = \begin{pmatrix} \mu_{cat,1} \\ \mu_{cat,2} \\ . \\ . \\ . \\ \mu_{cat,K} \end{pmatrix} \qquad (3.3)$$

and a degree of membership is set to each class according to:

$$1 \leq \mu_{cat,j} \leq 5, j = 1,...,K \qquad (3.4)$$

In practice, for each polygon of the reference maplet, the photo-interpreter ranked each class according to their likelihood of presence and attributed linguistic fuzzy scores. The similarity between the two maps was based on the intersection of map M_{NFI} with map M_{REF} which produced a partition of polygons P_m (see figure 3.4c):

$$M_{NFI} \cap M_{REF} = \{ P_m \} \qquad (3.5)$$

in which each polygon P_m was assigned a unique identification number, an attribute vector V_{crisp} from M_{NFI} and an attribute vector V_{cat} from M_{REF}. This partition appears on figure 3.4c, with the NFI map in the background.

4.3.2 The positional tolerance operator

A positional tolerance operator was applied to sufficiently small polygons in the partition (see figure 3.5). The aim of the operator was to modify the reference attribute V_{cat} of any polygon found below a user-defined tolerance band, in function of the reference attributes V_{cat} of its neighbour polygons. We produced a buffer surface from the *intersect* coverage using the tolerance band value τ_p. All polygons of this *intersect* found to be within the buffer surface was highlighted and further treated as an 'edge polygon'. This selection of small polygons with the buffer surface was developed by Mas (2005) to eliminate false change polygons in multi-date map comparisons. In non-edge polygons, the fuzzy representation of the reference material V_{cat} remained unchanged. Conversely, edge polygons were influenced by all their neighbours, and were assigned a Fuzzy Neighbourhood Vector V_{nbh} instead of the original V_{cat} (Hagen, 2003):

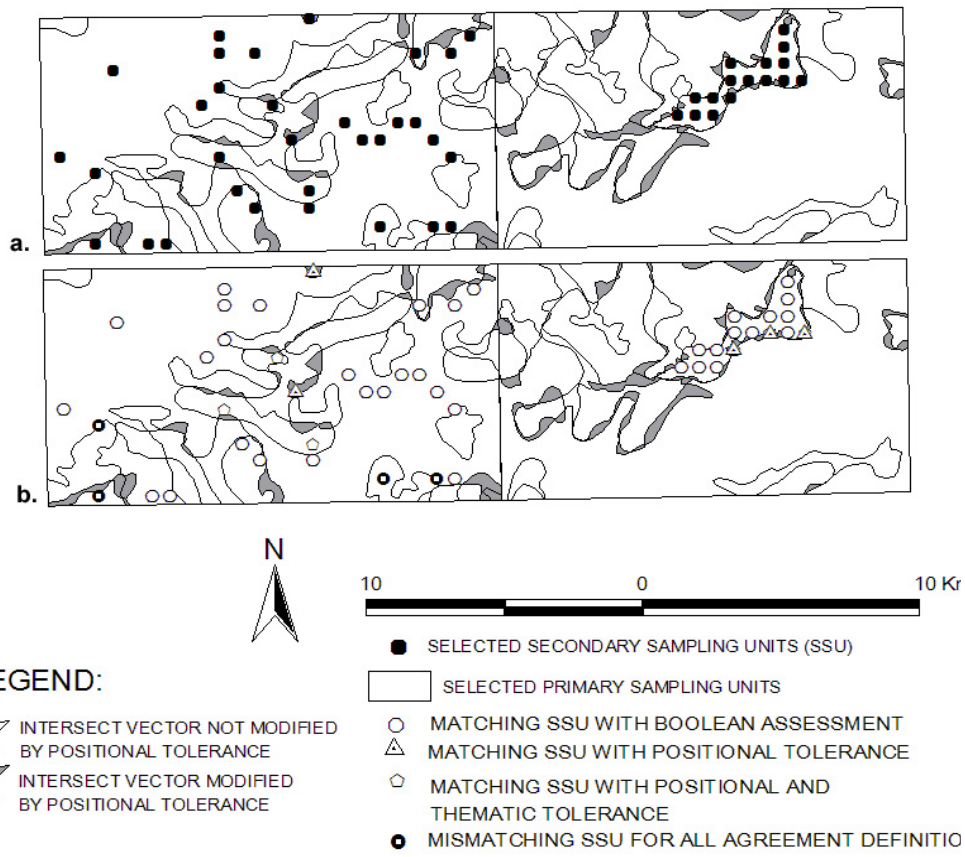
$$V_{nbh} = \begin{pmatrix} \mu_{nbh,1} \\ \mu_{nbh,2} \\ . \\ . \\ . \\ \mu_{nbh,K} \end{pmatrix} \qquad (3.6)$$

Where the different membership functions of the *N-1* neighbouring polygons are combined by calculating their fuzzy union with that of the edge polygon:

$$\mu_{nbh,j} = Max(\mu_{cat,j,1}, \mu_{cat,j,2}, ..., \mu_{cat,j,N}), j = 1, ..., K. \qquad (3.7)$$

Where $\mu_{cat,j,1}$ is the original degree of membership of class j in the edge polygon, and $\mu_{cat,j,n}$, $n = 2,\ldots,N$ are the original degrees of membership of class j in the neighbour polygons. This assignment was done via a procedure of the Arc/Info Macro Language (AML).

Figure 3.5. Edge polygons of the intersect between the NFI and reference maps



(a) Edge (grey) and non-edge (white) polygons of the intersect map; the edge polygons have been selected by the positional tolerance operator with positional tolerance of 500m. Same reference sites as in figure 3.4. (b) is the same as (a) and illustrates results of the fuzzy map comparison: empty circles are SSUs with map agreement at thematic tolerance $\tau_{th}=2$ (traditional Boolean assessment) and zero positional tolerance ($\tau_p=0$); dotted triangles correspond to disagreements at couplet (τ_{th}; τ_p) = (2;0) but agreements at (τ_{th}; τ_p) = (2;500m), these points pertain to edge polygons; dotted pentagons correspond to disagreements at (2;500m) but agreements at (τ_{th}; τ_p) = (4;500m); thick circles correspond to disagreements for all mentioned couplets of tolerance.

The resulting intersect map integrates both thematic and positional fuzziness. This treatment of the intersect map is an extension of Hagen (2003)'s fuzzy formalism to vector-based maps. For the present case study, results were obtained for a positional tolerance band τ_p set to respectively 0m and 500m, in order to capture the positional fuzziness of errors within the range 0-500m (below map scale). Figure 3.5 illustrates the set of edge polygons generated by the positional tolerance operator for τ_p = 500m.

4.3.3 The thematic tolerance operator

For map comparison, a degree of thematic tolerance has to be defined. Stehman et al. (2003) proposed to define the membership value of each reference site as the number of LULC classes for which the linguistic scale value is 3 ('reasonable or acceptable answer') or above.

$$v_{th} = card\left\{ j / \mu_{cat,j} \geq 3, j = 1,...,K \right\} \qquad (3.8)$$

v_{th} expresses a degree of thematic uncertainty. For example, a reference site in a heterogeneous landscape is likely to receive a higher membership value than a site in a homogeneous landscape. Stehman et al. (2003) suggested that the vast majority of their reference sites for the NLCD map of the USA was characterized by membership values of 1 or 2 only. This observation was a rationale for not documenting more than two degrees of membership per sample site in their assessment. Practically, Stehman et al. (2003) set their maximum membership value to 2.

In this research, a preliminary study on fuzzy interpretation of photo-interpreted material in mostly forested areas was realized on a sub-sample of 200 reference sites. This study revealed that 65% of the sub-sample were represented by a range of membership values of 1 to 2, and 33% by a range of membership values of 3 to 4 (higher thematic uncertainty), the rest of the sample (2%) being characterized by membership values of 5 and above. This preliminary study illustrated the thematic difficulties involved in assessing the Mexican LULC map at community level, difficulties also noticed in sub-tropical environments elsewhere (Powell et al., 2004).

Our goal was to estimate the thematic fuzziness inferred by such uncertainties in the accuracy estimate. The thematic tolerance operator consisted in setting a maximum

membership value, τ_{th} , to all reference attribute vectors in the intersect map: reference attribute vectors V_{cat} of non-edge polygons were modified so that:

$$card\left\{j / \mu_{cat,j} \geq 3, j = 1,..., K\right\} \leq \tau_{th} \qquad (3.9)$$

and reference attribute vectors V_{nbh} of edge polygons were modified so that:

$$card\left\{j / \mu_{nbh,j} \geq 3, j = 1,..., K\right\} \leq \tau_{th} \qquad (3..10)$$

In practice, since classes were ranked according to decreasing likelihood of presence in each polygon, the degree of membership of the K-τ_{th} less likely classes were set to 1 in the GIS database. Based on the preliminary analysis on thematic uncertainty in our zone, we produced results for a thematic tolerance τ_{th} set to 2 and 4, in order to capture the thematic fuzziness of errors within the range 3-4.

4.3.4 Map comparison process

Once positional and thematic tolerance values were fixed, the fuzzy map comparison could take place. For non-edge polygon P_m, the similitude between NFI and reference maps is based on the fuzzy set intersection of the two fuzzy vectors (Hagen, 2003):

$$S\ (P_m) = S(V_{cat}, V_{crisp}) = Max(\left\{Min(\mu_{cat,j}, \mu_{crisp,j}), j = 1,..., K\right\}) \qquad (3.11)$$

For edge polygon P_m, the similitude between NFI and reference maps is:

$$S\ (P_m) = S(V_{nbh}, V_{crisp}) = Max(\left\{Min(\mu_{nbh,j}, \mu_{crisp,j}), j = 1,..., K\right\}) \qquad (3.12)$$

Since the fuzzy linguistic RIGHT operator is used, an agreement was obtained on a reference site pertaining to polygon P_m, if $S(P_m)$ is superior or equal to three. Else, there was disagreement between the NFI map and the reference maplet.

4.4 Synthesis of the evaluation

In the case of rare classes, a double proportional stratified sampling was applied and, by construction, the inclusion probabilities of a sample point of class *k* were that of a simple stratified sampling: $p_k = f_k$, f_k being the frequency of class *k*. The formula for stratified random sampling was therefore used for global accuracy estimates. In the case of common classes, the first stage probability of selecting a PSU is expressed by p_{1k}= K/N, where K is

the fixed number of randomly selected PSUs for common classes and N is the total number of PSUs. At the second stage, SSU inclusion probability is p_{2k}=p2|1 * p_{1k}, where p2|1 is the conditional probability; p2|1 should be equal to n_k/N'_k where n_k is the sample size and N'_k is the number of SSUs of class *k* in the K selected PSUs.

If w_{ijk} is the reciprocal of the inclusion probability (weight) for SSU number *j* in PSU number *i*, $w_{ijk} = n_k/N'_k$ * K/N were the weights attributed to common classes when global estimates were computed. Next, we used the formula in the SURVEY- MEANS procedure of the Statistical Analysis Software (SAS, Version 8.3, 2001) in order to calculate the variance of the accuracy estimator (Särndal et al., 1992, section 9.4):

$$\mathrm{var}(\hat{R}) = \sum_{k=1}^{K} \frac{n_k (1 - n_k / N_k)}{(n_k - 1)} \sum_{i=1}^{n_k} (g_{ik.} - \bar{g}_{k..})^2 \qquad (13)$$

where

$$g_{ik.} = \frac{\sum^{m_{ik}} w_{ijk} (y_{ijk} - x_{ijk} \hat{R})}{\sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ik}} w_{ijk} x_{ijk}} \quad \text{and} \quad \bar{g}_{k..} = \left(\sum_{i=1}^{n_k} g_{ik.} \right) \Big/ n_k \qquad (14)$$

in which x_{ijk} and y_{ijk} are the generic indicator variables used in ratios for the calculation of accuracy indices (see SAS program documentation). This estimator takes into account the among-PSU variance of double sampling designs. It assumes PSUs contain SSUs only of one single map class, an assumption not fulfilled by common classes. However, few variance estimators are available in such cases of complex designs.

5. Results

5.1 Confusion matrix and global findings

Global accuracy and confusion patterns among NFI classes are reported in a confusion matrix (table 3.3) for the case of a SSU-based (point by point) geometric restitution, a 500m positional tolerance, and a Boolean (primary and alternate) thematic agreement definition. Compared with the new method we present in this paper, these characteristics are close to the ones employed in the assessment design of Stehman et al. (2003)

Table 3.3 Confusion matrix of the national forest inventory map (community aggregation level) in the watershed of the Candelaria river, with the point-by-point traditional assessment, according to the Boolean agreement definition and with 500m positional tolerance.

	1	2	3	4	5	6	7	8	9	10	11	12	13	*n*	Area frac[a]	User's[b]	Std Err[c]	Binom.[d]
Cultivated grassland	**11.852**	0.924		0.154	1.385	2.617				0.154				111	17.085	0.69	0.066	0.044
Annual crop	0.141	**0.523**								0.017		0.017		84	0.697	0.75	0.117	0.047
Med and high sub-per. forest	0.380	2.657	**34.915**	4.554	3.036	4.554								132	50.096	0.70	0.098	0.040
Med and high sub-per. forest & sec. veg	2.611	0.193		**4.835**	0.774	0.387								91	8.799	0.55	0.146	0.052
Low sub-per. forest		1.228	4.298	1.228	**9.210**	0.307		0.307		1.074				115	17.652	0.52	0.112	0.047
Low sub-per. forest & sec. veg			0.056			**0.080**				0.111				93	0.247	0.32	0.079	0.048
Savanna	0.073	0.012			0.379	0.159	**0.232**			0.196		0.012	0.012	88	1.076	0.22	0.011	0.044
Induced grassland			0.004	0.022	0.075			**0.263**		0.026				89	0.391	0.67	0.055	0.050
Mangrove					0.012			0.068	**0.519**					97	0.599	0.87	0.031	0.035
Hygrophilous grassland	0.064		0.107		0.107	0.064		0.300	0.021	**1.585**				105	2.249	0.70	0.092	0.045
Halophilous vegetation		0.008			0.062	0.004		0.017		0.266	**0.033**			94	0.390	0.09	0.042	0.029
Human settlement	0.018											**0.073**		15	0.091	0.80	0.032	0.030
Water		0.047	0.024	0.008	0.008	0.008		0.047		0.141	0.047		**0.298**	80	0.628	0.48	0.121	0.060
n	134	95	150	77	156	83	19	97	85	230	14	15	39	1194				
Area frac[a]	15.139	5.592	39.404	10.801	15.048	8.179	0.232	1.002	0.540	3.570	0.080	0.102	0.310		100.000			
Prod[e]	0.78	0.09	0.89	0.45	0.61	0.01	1.00	0.26	0.96	0.44	0.41	0.72	0.96			**64.4**		
Std Err.[c]	0.093	0.047	0.097	0.102	0.061	0.019	0.000	0.055	0.006	0.074	0.337	0.108	0.033					
Binom.[d]	0.036	0.030	0.026	0.057	0.039	0.011	0.000	0.045	0.021	0.033	0.132	0.116	0.031					

Global accuracy was estimated at 64.4% (bottom of 'User's' column) and was calculated as is described in the synthesis of the evaluation.

Entry numbers refer to class codes defined in table 3.1. [a]Area percent fraction with respect to total area. These numbers are the sum of matrix columns or rows. [b]User's accuracy.
[c]Standard error, for two-stage sampling, including inter-cluster variance assuming a stratified by class random sampling at second stage.
[d]Binomial estimator of standard error, using the approximation formula for large sample sizes and the full formula for small sample sizes.
[e]Producer's accuracy.

The global accuracy level of 64.4% mainly reflects an intermediate score between the accuracy estimates of 70% and 69% for, respectively, the dominant class 'median and high sub-perennial tropical forest' (hereafter median forest) and for the common class 'cultivated grassland', and the lower accuracy estimate of 52% for the common class 'low sub-perennial tropical forest' (hereafter low forest).

The confusion matrix in table 3.3 reports that mutual, nearly counterbalanced confusion exists between median and low forests. A possible explanation for such mutual confusion could be that the association of some tree species present in median forests may, in some cases, take the physiognomic appearance of a low forest, and vice-versa (taxonomic ambiguity between median and low forest in the classification scheme). Another, alternate, explanation could be the presence of spatially extended eco-tones, or grades of transition, between well-defined patches of median and low forests (uncertainty on the boundary between median and low forest). As a means of approaching the most probable interpretation for the amount of confusion observed, accuracy results were derived from different degrees of tolerance (thematic and positional). Figure 3.5 illustrates agreement counts for four different agreement definitions. White circles represent SSUs where the NFI map agrees with the reference data on applying a thematic tolerance of two and a positional tolerance of zero. The global results for this most conservative degree of tolerance (or Boolean assessment) are reported in table 3.4.

Table 3.4.Confusion matrix of the national forest inventory map (community aggregation level) in the watershed of the Candelaria River, derived from the accuracy assessment with Boolean thematic tolerance (τ_{th} = 2) and zero positional tolerance (τ_p=0).

	1	2	3	4	5	6	7	8	9	10	11	12	13	*n*	Area frac[a]	User's[b]	Std Err[c]	Binom.[d]
Cultivated grassland	**10.422**	1.708		0.683	1.196	2.563		0.342		0.171				100	17.085	0.61	0.074	0.049
Annual crop	0.153	**0.474**			0.007	0.021		0.014		0.014		0.014		100	0.697	0.68	0.134	0.047
Med and high sub-per. forest	0.501	2.505	**31.059**	4.509	6.512	5.010								100	50.096	0.62	0.099	0.049
Med and high sub-per. forest & sec. veg	2.640	0.176	0.264	**4.400**	0.792	0.528								100	8.799	0.50	0.153	0.050
Low sub-per. forest		1.236	7.061	1.236	**5.296**	0.353		0.353		2.118				100	17.652	0.30	0.061	0.046
Low sub-per. forest & sec. veg			0.057		0.027	**0.049**		0.005		0.109				100	0.247	0.20	0.067	0.040
Savanna	0.109	0.011			0.369	0.174	**0.174**			0.217		0.011	0.011	99	1.076	0.16	0.017	0.037
Induced grassland			0.004	0.039	0.102	0.012		**0.203**		0.031				100	0.391	0.52	0.056	0.050
Mangrove					0.018			0.084	**0.497**					100	0.599	0.83	0.024	0.038
Hygrophilous grassland	0.157		0.090		0.090	0.112		0.292	0.067	**1.440**				100	2.249	0.64	0.081	0.048
Halophilous vegetation		0.008			0.066	0.004		0.016		0.257	**0.039**			100	0.390	0.10	0.042	0.030
Human settlement	0.018					0.006						**0.067**		15	0.091	0.73	0.027	0.030
Water		0.047	0.024	0.008	0.008	0.016		0.047		0.141	0.047		**0.290**	80	0.628	0.46	0.127	0.060
n	134	101	136	81	156	84	16	97	86	236	16	14	38	1194				
Area frac[a]	14.000	6.165	38.558	10.874	14.483	8.847	0.174	1.356	0.565	4.498	0.086	0.092	0.301		100.000			
Prod[e]	0.74	0.08	0.81	0.40	0.37	0.01	1.00	0.15	0.88	0.32	0.45	0.73	0.96			**54.4**		
Std Err.[c]	0.109	0.084	0.072	0.103	0.074	0.013	0.000	0.050	0.007	0.072	0.337	0.114	0.034					
Binom.[d]	0.038	0.027	0.034	0.055	0.039	0.008	0.000	0.036	0.035	0.030	0.124	0.119	0.030					

Global accuracy was estimated at 54.4%. Same conventions as in table 3.3.

Entry numbers refer to class codes defined in table 3.1. [a]Area percent fraction with respect to total area. These numbers are the sum of matrix columns or rows. [b]User's accuracy.
[c]Standard error, for two-stage sampling, including inter-cluster variance assuming a stratified by class random sampling at second stage.
[d]Binomial estimator of standard error, using the approximation formula for large sample sizes and the full formula for small sample sizes. [e]Producer's accuracy.

5.2 Positional fuzziness in the forest cover map

Dotted triangles on figure 3.5 represent the case where the most conservative option alone yields a disagreement (counted as an error of the NFI map), but where the addition of a positional tolerance of 500m results in an agreement (i.e. the NFI map is finally found correct). In this case, fictitious errors due to scale and polygon boundary delineation could have biased the accuracy estimate based on the zero positional tolerance. The accuracy estimates with positional tolerance are reported in table 3.5. The comparison of reciprocal confusion amounts of low (code 5) and median (code 3) forests between table 3.5 and table 3.4 gives us an estimate of the contribution of position-related confusions (due to polygon boundary, scale or residues of geometric inconsistency) to the total confusion registered between those forests. Such inspection reveals that the addition of a positional tolerance substantially decreased the confusion between these forest classes: the area of mapped median forest photo-interpreted as low forest decreased from 6.5% of the Candelaria area down to 3% only, this 3.5% decrease in confusion explaining most of the gain in global accuracy for the median forest class. Likewise, the area of mapped low forest photo-interpreted as median forest decreased from 7% of the Candelaria area down to 4.2%, this 2.8% decrease in confusion also explaining most of the accuracy gain of the low forest class. This observation favoured the interpretation that much error occurred because of fuzzy spatial limits between patches of low and median forests rather than because of fuzzy taxonomic limits in class definitions. Residues of geometric inconsistency between photographs and Landsat imagery, both georeferenced to INEGI topographic maps, may have contributed to this fuzziness of spatial limits: each georeferencing task in the context of poorly drained and densely forested Campeche plains was challenging; the double positional precaution proposed in this research (geometric consistency step through the restitution of reference maplets and positional tolerance operator) may have been necessary to efficiently exclude fictitious, non-thematic error types for forest classes in this challenging context.

Table 3.5. Confusion matrix of the national forest inventory map (community aggregation level) in the watershed of the Candelaria River, derived from accuracy assessment with degrees of tolerance (τ_{th}; τ_p) = (2;500m).

	1	2	3	4	5	6	7	8	9	10	11	12	13	*n*	Area frac[a]	User's[b]	Std Err[c]	Binom.[d]
Cultivated grassland	**11.789**	1.025		0.171	1.367	2.563				0.171				100	17.085	0.69	0.056	0.046
Annual crop	0.153	**0.516**								0.014		0.014		100	0.697	0.74	0.119	0.044
Med and high sub-per. forest	0.501	2.505	**35.067**	4.008	3.006	5.010								100	50.096	0.70	0.094	0.046
Med and high sub-per. forest & sec. veg	2.640	0.176		**4.664**	0.792	0.528								100	8.799	0.53	0.156	0.050
Low sub-per. forest		1.236	4.237	1.236	**9.003**	0.353		0.353		1.236				100	17.652	0.51	0.113	0.050
Low sub-per. forest & sec. veg			0.057			**0.081**				0.109				100	0.247	0.33	0.079	0.047
Savanna	0.087	0.011			0.326	0.174	**0.239**			0.217		0.011	0.011	99	1.076	0.22	0.010	0.042
Induced grassland			0.004	0.027	0.074			**0.254**		0.031				100	0.391	0.65	0.055	0.048
Mangrove					0.012			0.072	**0.515**					100	0.599	0.86	0.029	0.035
Hygrophilous grassland	0.067		0.090		0.090	0.067		0.292	0.022	**1.620**				100	2.249	0.72	0.094	0.045
Halophilous vegetation		0.008			0.066	0.004		0.016		0.257	**0.039**			100	0.390	0.10	0.042	0.030
Human settlement	0.018											**0.073**		15	0.091	0.80	0.032	0.030
Water		0.047	0.024	0.008	0.008	0.008		0.047		0.141	0.047		**0.298**	80	0.628	0.48	0.125	0.060
n	139	102	125	77	146	86	0	102	87	239	16	15	39	1194				
Area frac[a]	15.255	5.523	39.478	10.113	14.743	8.788	0.239	1.034	0.538	3.796	0.086	0.098	0.309		100.000			
Prod[e]	0.77	0.09	0.89	0.46	0.61	0.01	1.00	0.25	0.96	0.43	0.45	0.75	0.96			**64.2**		
Std Err.[c]	0.098	0.057	0.094	0.100	0.063	0.018	0.000	0.051	0.006	0.077	0.337	0.104	0.033					
Binom.[d]	0.036	0.029	0.028	0.057	0.040	0.010	0.000	0.043	0.021	0.032	0.124	0.112	0.029					

Same conventions as in table 3.3. Global accuracy was estimated at 64.2%.

Entry numbers refer to class codes defined in table 3.1. [a]Area percent fraction with respect to total area. These numbers are the sum of matrix columns or rows. [b]User's accuracy. [c]Standard error, for two-stage sampling, including inter-cluster variance assuming a stratified by class random sampling at second stage. [d]Binomial estimator of standard error, using the approximation formula for large sample sizes and the full formula for small sample sizes. [e]Producer's accuracy.

Another source of error apparent on table3.3 is that much area of mapped median forest was photo-interpreted as 'secondary' forest classes (here, 'secondary' forest stands for forest with 'secondary vegetation, herbaceous, and/or shrub-like vegetation'), a confusion which is not counterbalanced by reciprocal confusion. Unlike for the confusion with low forest, the inspection of table 3.5 reveals that levels of confusion with 'secondary' forest types do not decrease when geometric tolerance is applied to the agreement definition. The modification of degrees of thematic tolerance may provide elements for the interpretation of these confusions.

5.3 Thematic fuzziness in the forest cover map

The dotted pentagons on figure 3.5 represent the case where the low degree of thematic tolerance (maximum membership value of two, typically employed as a 'Boolean' assessment), even with a 500m positional tolerance, yield a disagreement, but where the application of a thematic tolerance of four yields an agreement. In this case, the difficulty of interpreting remotely sensed material (possibly because of the high taxonomic diversity of the classification scheme) could be a source of uncertainty for the accuracy estimate. The global estimates for the increased thematic tolerance, without and with positional tolerance are reported respectively in table 3.6 and in table 3.7. Finally, the dotted black circles on figure 3.5 represent disagreement whatever degree of tolerance is applied in the agreement definition (reported as errors on table 3.7), therefore being most probably real misclassification errors of the NFI map. The comparison of confusion amounts between table 3.4 (Boolean assessment) and table 3.6 (assessment with increased thematic tolerance) may give an estimate of the contribution of thematic-related confusions (uncertainty of interpretation, fuzziness of taxonomic boundaries) to the total confusion recorded between forest classes. This thematic fuzziness is represented by a light grey bar for per class accuracy indices in figure 3.6a (user's accuracy) and 3.6b (producer's accuracy). The inspection of these results reveals that the passage from a Boolean assessment to a thematic tolerance of 4 substantially decreased the confusion of the median forest class (code 3) with secondary forest classes (codes 4 and 6): the area of mapped median forest photo-interpreted as median 'secondary' forest decreased from 4.5% of the Candelaria area down to 0.5% only (a 4% decrease).

Table 3.6 Confusion matrix of the national forest inventory map (community aggregation level) in the watershed of the Candelaria River, derived from accuracy assessment with degrees of tolerance (τ_{th}; τ_p) = (4;0m).

	1	2	3	4	5	6	7	8	9	10	11	12	13	*n*	Area frac[a]	User's[b]	Std Err[c]	Binom.[d]
Cultivated grassland	**11.447**	0.854		0.513	1.196	2.563		0.342		0.171				100	17.085	0.67	0.071	0.047
Annual crop		**0.627**			0.007	0.021		0.014		0.014		0.014		100	0.697	0.90	0.019	0.030
Med and high sub-per. forest	0.501	2.505	**37.071**	0.501	6.012	3.507								100	50.096	0.74	0.064	0.044
Med and high sub-per. forest & sec. veg	2.640	0.176	0.264	**5.280**	0.440									100	8.799	0.60	0.142	0.049
Low sub-per. forest		1.236	6.531	1.236	**6.178**			0.353		2.118				100	17.652	0.35	0.044	0.048
Low sub-per. forest & sec. veg			0.057		0.027	**0.067**		0.005		0.091				100	0.247	0.27	0.073	0.044
Savanna	0.109	0.011			0.348	0.174	**0.217**			0.196		0.011	0.011	99	1.076	0.20	0.010	0.040
Induced grassland			0.004	0.039	0.102	0.012		**0.203**		0.031				100	0.391	0.52	0.056	0.050
Mangrove					0.018			0.084	**0.497**					100	0.599	0.83	0.024	0.038
Hygrophilous grassland	0.090		0.090		0.090	0.112			0.067	**1.800**				100	2.249	0.80	0.023	0.040
Halophilous vegetation		0.008			0.066	0.004		0.016		0.257	**0.039**			100	0.390	0.10	0.042	0.030
Human settlement	0.018					0.006						**0.067**		15	0.091	0.73	0.027	0.030
Water		0.047	0.024	0.008	0.008	0.016				0.141	0.047		**0.337**	80	0.628	0.54	0.121	0.060
n	115	118	145	82	154	80	20	78	86	242	16	14	44	1194				
Area frac[a]	14.805	5.464	44.040	7.576	14.491	6.481	0.217	1.016	0.565	4.819	0.086	0.092	0.348		100.000			
Prod[e]	0.77	0.11	0.84	0.70	0.43	0.01	1.00	0.20	0.88	0.37	0.45	0.73	0.97			**63.8**		
Std Err.[c]	0.104	0.055	0.082	0.051	0.048	0.027	0.000	0.060	0.007	0.070	0.337	0.114	0.028					
Binom.[d]	0.039	0.029	0.030	0.051	0.040	0.011	0.000	0.045	0.035	0.031	0.124	0.119	0.026					

Same conventions as in table 3.3. Global accuracy was estimated at 63.8%.

Entry numbers refer to class codes defined in table 3.1. [a]Area percent fraction with respect to total area. These numbers are the sum of matrix columns or rows. [b]User's accuracy. [c]Standard error, for two-stage sampling, including inter-cluster variance assuming a stratified by class random sampling at second stage. [d]Binomial estimator of standard error, using the approximation formula for large sample sizes and the full formula for small sample sizes. [e]Producer's accuracy.

Table 3.7 Confusion matrix of the national forest inventory map (community aggregation level) in the watershed of the Candelaria River, derived from accuracy assessment with degrees of tolerance (τ_{th}; τ_p) = (4;500m).

	1	2	3	4	5	6	7	8	9	10	11	12	13	*n*	Area frac[a]	User's[b]	Std Err[c]	Binom.[d]
Cultivated grassland	**15.718**				0.683	0.513				0.171				100	17.085	0.92	0.031	0.027
Annual crop		**0.683**								0.014				100	0.697	0.98	0.011	0.014
Med and high sub-per. forest	0.501	1.503	**42.081**		2.505	3.507								100	50.096	0.84	0.046	0.037
Med and high sub-per. forest & sec. veg	2.288	0.176		**6.336**										100	8.799	0.72	0.116	0.045
Low sub-per. forest		1.236	3.707	1.236	**9.885**			0.353		1.236				100	17.652	0.56	0.094	0.050
Low sub-per. forest & sec. veg			0.057			**0.116**				0.074				100	0.247	0.47	0.091	0.050
Savanna	0.087	0.011			0.304	0.152	**0.315**			0.185		0.011	0.011	99	1.076	0.29	0.020	0.046
Induced grassland			0.004	0.027	0.020			**0.309**		0.031				100	0.391	0.79	0.076	0.041
Mangrove					0.012			0.066	**0.521**					100	0.599	0.87	0.034	0.034
Hygrophilous grassland			0.090		0.090	0.067			0.022	**1.980**				100	2.249	0.88	0.020	0.032
Halophilous vegetation		0.008			0.066	0.004		0.016		0.257	**0.039**			100	0.390	0.10	0.042	0.030
Human settlement												**0.091**		15	0.091	1.00	0.000	0.000
Water		0.047			0.008	0.008				0.141	0.047		**0.377**	80	0.628	0.60	0.132	0.060
n	127	119	133	86	122	76	29	96	88	237	16	16	49	1194				
Area frac[a]	18.594	3.663	45.938	7.599	13.573	4.367	0.315	0.743	0.544	4.089	0.086	0.102	0.388		100.000			
Prod[e]	0.85	0.19	0.92	0.83	0.73	0.03	1.00	0.42	0.96	0.48	0.45	0.89	0.97			**78.4**		
Std Err.[c]	0.085	0.054	0.071	0.060	0.047	0.064	0.000	0.044	0.006	0.076	0.337	0.083	0.024					
Binom.[d]	0.032	0.036	0.024	0.040	0.040	0.018	0.000	0.050	0.021	0.032	0.124	0.077	0.024					

Global accuracy was estimated at 78.4%.

Entry numbers refer to class codes defined in table 3.1. [a]Area percent fraction with respect to total area. These numbers are the sum of matrix columns or rows. [b]User's accuracy. [c]Standard error, for two-stage sampling, including inter-cluster variance assuming a stratified by class random sampling at second stage. [d]Binomial estimator of standard error, using the approximation formula for large sample sizes and the full formula for small sample sizes. [e]Producer's accuracy.

Likewise, the area of mapped median forest photo-interpreted as low 'secondary' forest decreased from 5% of the Candelaria area down to 3.5% (a 1.5% decrease). This 5.5% total decrease in confusion explained most of the accuracy gain of the median forest class using the higher degree of thematic tolerance (see figure 3.6).

By contrast, almost no decrease in these confusions was recorded with the increase in positional tolerance (between table 3.4 and table 3.5, dark grey bar in figures 3.6 and 3.7). This observation favoured the interpretation that much error occurred because of the difficulty of unambiguously labelling maps with the presence or not of 'secondary' vegetation within tropical forests (fuzziness of taxonomic boundaries). A strength of the classification scheme is undoubtedly its compatibility with INEGI archive map labels. The thematic fuzziness related to the difficult inclusion of secondary vegetation in the classification scheme might be one of the necessary trade-offs for the fundamental criterion of compatibility.

Indeed, low accuracy scores were recorded for classes denominated as tropical forests with 'secondary' vegetation. For example, the increase in thematic tolerance (from table 3.4 to table 3.6) yielded an accuracy improvement of 0.9% of the total Candelaria for the median 'secondary' forest whereas the increase in positional tolerance (from table 3.4 to table 3.5) improved accuracy by only 0.3%.

Globally, the increase in thematic tolerance improved the accuracy of the map from 54.4% (table 3.4) to 63.8% (table 3.6), about the same amount as with the passage from zero to 500m positional tolerance (64.2%, table 3.5). This result confirmed the relevance of maximum membership values of up to four in the assessment of thematic fuzziness in highly diverse taxonomy maps such as the NFI map. The increase in both thematic and positional tolerance cumulated to a global accuracy of 78.4% (table 3.7). Besides, much cultivated grassland (2.6% of all Candelaria area, see table 3.3) was misclassified as median 'secondary' forest, an amount that did not diminish with the increase in tolerance degree in both dimensions (positional or thematic). The above mentioned confusions towards less densely vegetated classes altogether suggest a possibly of a strong sub-evaluation by the NFI map of the deforestation in the median and high sub-perennial tropical forest.

Figure 3.6a. User's accuracy index with fuzzy bounds for each class of the map.

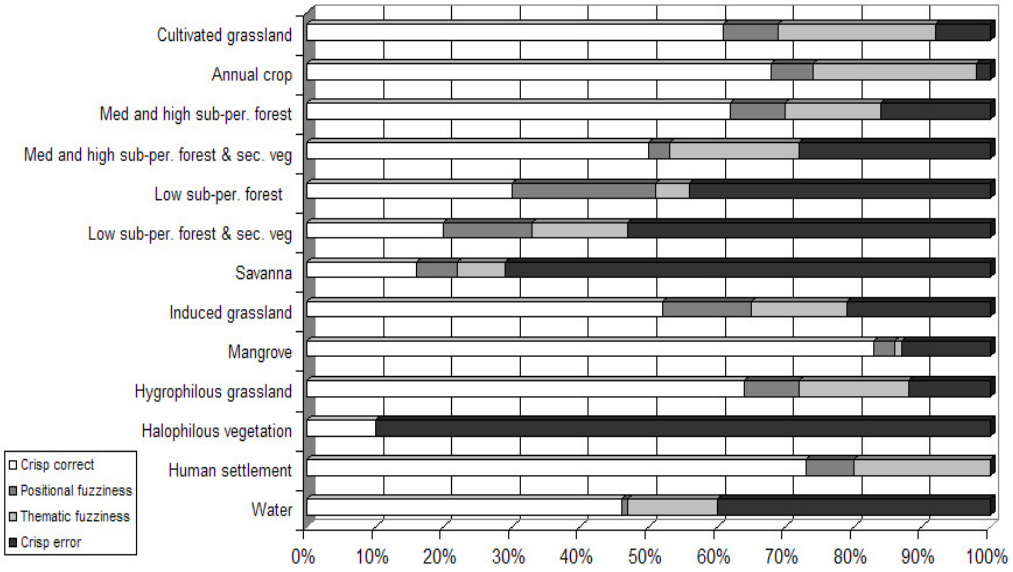
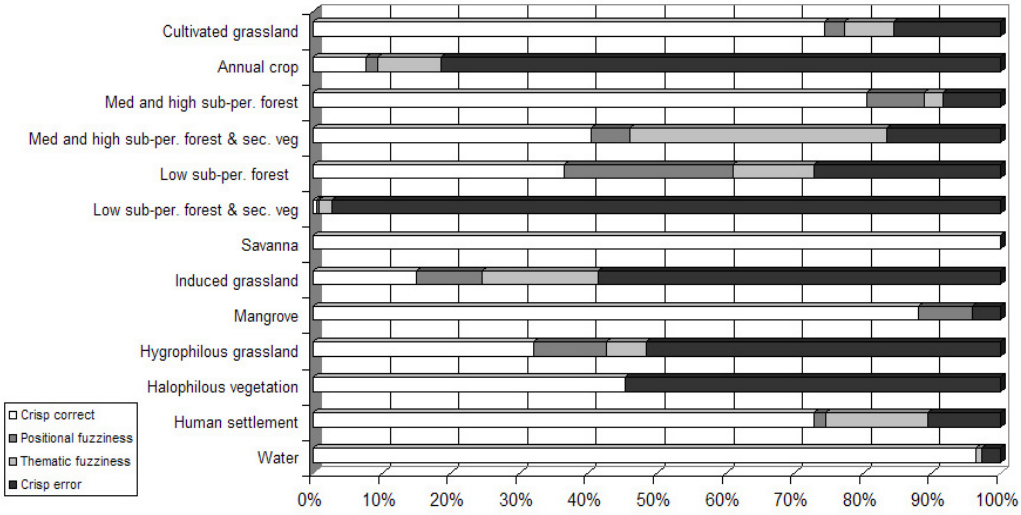


Figure 3.6b Producer's accuracy index with fuzzy bounds for each class of the map.



The black bar portion (crisp error) is the amount of disagreement with highest degree of tolerance ((τ_{th}; τ_p) = (4;500m)); the white bar portion (crisp correct) is the amount of agreement with lowest degree of tolerance ((τ_{th}; τ_p) = (2;0)); the dark grey bar portion is the accuracy difference between assessments at two degrees of positional tolerance (positional fuzziness τ_p = 0 - 500m); the light grey bar portion is the accuracy difference between assessments at two degrees of thematic tolerance (thematic fuzziness for τ_{th}= 2 - 4);

6. Discussion on the method

Although many assessment strategies have favoured point-like reference datasets (Stehman et al., 2003; Powell et al., 2004; Wulder et al., 2006) derived from aerial photograph interpretation, we focus on the use of reference maplets, also based on aerial photograph interpretation. Indeed, with this concept, the labeling protocol can be considered as a comparison of two maps: the map to be assessed and a reference maplet. This statement is meaningful since the concept of positional tolerance can then benefit from recent theoretical advances on the formalization of spatial fuzzy objects and topological relationships (e.g. Cheng et al., 2001; Tang et al., 2006), fuzzy similarity of maps (Hagen, 2003; Hagen et al., 2005) or advances regarding the multi-scale comparison of thematic maps (Pontius and Cheuk, 2006; Power et al., 2001; Remmel and Csillag, 2006; Visser and de Nijs, 2006). The estimates with low and high degrees of thematic tolerance could be regarded respectively as pessimistic and optimistic bounds for the accuracy assessment with reference material representing the real thematic conditions of the map. Likewise, the positional operator based on a user-defined tolerance band may also provide means of defining 'optimistic' (upper threshold) and 'pessimistic' (lower threshold) bounds of the accuracy level at a certain scale of application.

The necessity of labour intensive photo-interpretation and restitution needed for the construction of reference maplets can appear as a drawback of this method. However, in this study, because of the existing skills, the amount of working time for maplet construction was comparable with the amount of working time for equivalent point-like data collection.

The evaluation protocol based on point like data collection does not permit a flexible analysis on positional fuzziness. Although implicit in the discussions on accuracy assessment methods (e.g. Stehman et al., 2003), the notion of degree of tolerance has not been brought forward as a standard concept. The task of comparing assessments of map products around the world would certainly benefit from standardization. This method, with the same formalism, can accommodate quantitative fuzzy instead of linguistic fuzzy

interpretation of landscapes, as well as recent advances in raster-based multi-scale map comparisons. Finally, it can be extended to the accuracy assessment of fuzzy maps.

7. Summary and conclusion

This paper reports an accuracy assessment method for regional scale, taxonomically diverse Land Use and Land Cover (LULC) maps. For every LULC class of the map, this method reports the accuracy level and quantifies thematic and positional fuzziness associated with uncertainties in the interpretation of complex landscapes. The evaluation protocol consists in the construction of reference maplets. The labelling protocol consisted in a fuzzy map comparison for which degrees of positional and thematic tolerance can be user-defined. Taking from recent advances in fuzzy comparison of maps, we propose the formalism for the accuracy assessment of vector-based maps. The geometric restitution of reference maplets and the positional tolerance operator constitute innovations in terms of assessment design, still subject to further improvements.

The method was applied to the NFI map in the densely forested watershed of the Candelaria River in Southeast Mexico. Substantial mutual confusion exists between median and low sub-perennial forests. The method suggested that part of these confusions be essentially related to the fuzziness of boundaries between patches of median and low forests, and not to a taxonomic ambiguity in the classification scheme. Conversely, a substantial part of the confusion registered for ‘secondary’ forest classes were essentially related to thematic uncertainty of interpretation and possibly due to the way secondary vegetation is included in the classification scheme.

The method permits the standardization of the notion of degrees of tolerance in accuracy assessments, in both thematic and positional dimensions. More explicitly than before, a user could, with this method, take the GIS-based data generated by the accuracy assessment, modify degrees of positional or thematic tolerance, and generate an accuracy estimate at his scale of interest. We are conscious that the notion of positional tolerance for map accuracy assessment is delicate since the accuracy assessment primarily makes sense at the scale of the map, for which it has been initially produced. However, this paper provides a universal conceptual framework for analyzing and comparing accuracy assessment strategies. Altogether, the accuracy assessment method is proposed as a useful

estimator of general trends for fictitious and real errors on regional scale LULC maps of bio-diverse regions.

8.References

Achard, F., H.D. Eva, H.J. Stibig, P. Mayaux, J. Gallego, T. Richards and J.P. Malingreau, 2002. Determination of deforestation rates of the world's humid tropical forests, *Science* 297 (5583): 999-1002.

Aspinall, R., and D.M. Pearson, 1995. Describing and managing uncertainty of categorical maps in GIS, *Innovations in GIS 2* (P. Fisher, editors), Taylor & Francis, London, pp. 71-83.

Barbosa, P.M., D. Stroppiana, J.M. Gregoire, and J.M.C. Pereira, 1999. An assessment of vegetation fire in Africa (1981-1991): Burned areas, burned biomass, and atmospheric emissions, *Global Biogeochemical cycles* 13(4): 933-950.

Cheng, T., Molenaar M., and H. Lin, 2001. Formalizing fuzzy objects from uncertain classification results, *International Journal of Geographical Information Science* 15 (1): 27-42.

Congalton, R.G., and K. Green, 1993. A practical look at the sources of confusion in error matrix generation, *Photogrammetric Engineering and Remote Sensing* 59: 641-644.

Foody, G.M., 2002. Status of land cover classification accuracy assessment, *Remote Sensing of Environment* 80:185-201.

Gopal, S., and C. Woodcock, 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets, *Photogrammetric Engineering and Remote Sensing* 60: 181-188.

Green, D.R., and W. Hartley, 2000. Integrating photo-interpretation and GIS for vegetation mapping: some issues of error, *Vegetation Mapping from Patch to Planet* (Alexander, R. and A.C. Millington, editors), John Wiley & Sons Ltd., pp. 103-134.

Hagen, A., Fuzzy set approach to assessing similarity of categorical maps, 2003. *International Journal of Geographical Information Science*, 17 (3): 235-249,

Hagen-Zanker, A., Straatman B., and I. Uljee, 2005. Further developments of a fuzzy set map comparison approach, *International Journal of Geographical Information Science* 19 (7): 769-785

Hely, C., L. Bremond, S. Alleaume, B. Smith, M.T. Sykes, and J. Guiot, 2006. Sensitivity of African biomes to changes in the precipitation regime, *Global Ecology and Biogeography*, 15(3): 258-270.

Houghton, R.A., D.L. Skole, C.A. Nobre, J.L. Hackler, K.T. Lawrence, and W.H. Chomentowski, 2000. Annual fluxes or carbon from deforestation and regrowth in the Brazilian Amazon, *Nature*, 403 (6767): 301-304.

Hunter, M.L., Jr., 1996. *Fundamentals of conservation biology*, Blackwell Scientific, Cambridge 482 p.

Laba, M., S.K. Gregory, J. Braden, D. Ogurcak, E. Hill, E. Fegraus, J. Fiore and S.D. DeGloria, 2002. Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map, *Remote Sensing of Environment* 81: 443-455.

Mas, JF., 2005. Change estimates by map comparison: a method to reduce erroneous changes due to positional error, *Transactions in GIS* 9(4): 619-629.

Mas, JF., A. Velázquez, J.L. Palacio-Prieto, G. Bocco, A. Peralta, and J. Prado, 2002. Assessing forest resources in Mexico: Wall-to-wall land use/ cover mapping, *Photogrammetric Engineering and Remote Sensing* 68(10): 966-969.

Miller, K.E., E. Chang, and N. Johnson, 2001. *Defining common ground for the Mesoamerican Biological Corridor*, World Resources Institute, Washington DC 45 p.

Pontius R.G., and M.L. Cheuk, 2006. A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions, *International Journal of Geographical Information Science*: 20 (1), 1-30

Powell, R.L., Matzke N., de Souza C., Clark M., Numata I., Hess L.L., and D.A. Roberts, 2004, Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon, *Remote Sensing of Environment* 90: 221-234.

Power, C., Simms A., and R. White, 2001. Hierarchical fuzzy pattern matching for the regional comparison of land use maps, *International Journal of Geographical Information Science*: 15 (1), 77-100

Remmel T.K., and F. Csillag, 2006. Mutual information spectra for comparing categorical maps, *International Journal of Remote Sensing*, 27 (7): 1425-1452

Särndal, C.E., Swensson V., and J. Wretman, 1992, *Model-assisted survey sampling*, New-York: Springer-Verlag.

Soares-Filho, B.S., D.C. Nepstad, L.M. Curran, G.C. Cerqueira, R.A. Garcia, C.A. Ramos, E. Voll, A. McDonald, P. Lefebvre and P. Schlesinger, 2005. Modelling conservation in the Amazon basin, *Nature* 440:520-523.

Stehman, S.V., and R.L. Czaplewski, 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment* 64: 331-344.

Stehman, S.V., J.D. Wickham, J.H. Smith, and L. Yang, 2003. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United-States: Statistical methodology and regional results, *Remote Sensing of Environment* 86: 500-516.

Stehman, S.V., J.D. Wickham, L. Yang, and J.H. Smith, 2000. Assessing the accuracy of large-area land cover maps: Experiences from the Multi-resolution Land-Cover Characteristics (MRLC) project, *Accuracy 2000: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences* (Delft Univ. Press., The Netherlands), pp. 601-608.

Tang, X., Fang Y., and W. Kainz, 2006. Fuzzy topological relations between fuzzy spatial objects, *Lecture Notes in Computer Sciences*, LNAI 4223, 97, Springer-Verlag, Germany 324-333

Visser, H., T. de Nijs, 2006. The map comparison kit, *Environmental Modelling & Software*: 21, 346-358

Wulder, M.A., S.E. Franklin, J.C. White, Linke J., and S. Magnussen, 2006. An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data, *International Journal of Remote Sensing* 27(4): 663-683.

Zhu, Z., L. Yang, S.V. Stehman and R.L. Czaplewski, 2000. Accuracy assessment for the U.S. Geological Survey regional land-cover mapping program: New-York and New Jersey region, *Photogrammetric Engineering and Remote Sensing* 66: 1425-1435.

Chapter 4: Spectral Separability of Forests on slope

Spectral separability of forest stands on slope in real and simulated IKONOS imagery

In chapter 1 to 3, the research effort was focused on the evaluation and description of existing errors on LULC maps of regions with high biodiversity and highly dynamic landscapes. Another facet of the evaluation of errors is the identification of their cause, for their correct interpretation and for future cartographic efforts. A major cause of errors is the intrinsic limitation of the satellite imagery from which these maps are derived. Misclassification occurs especially when categories of the classification system (classes) are not well distinguished, or ambiguous, in the satellite imagery. The study of these ambiguities permits, for example, the adaptation of the classification system to the mapping capacity of remote sensing imagers.

In chapter 4, spectral separability, a classical remote-sensing index, is used as an inverse measure of this ambiguity. The relevance of such indicator is to detect these problematic classes for which more cartographic effort is needed. Traditionally, spectral separability is empirically derived from an image, and depends on external parameters such as the sensor type, atmospheric conditions, viewing and illumination configurations. Considering the abundance of forest classes and the complex topography in Mexico, the transferability of measurements of previous image-based studies on confined geographic regions was likely to be limited. The second method developed in this thesis is a model-based systematic framework for the measurement of ambiguities between forest cover classes on satellite imagery. The method is illustrated in chapters 4 and 5. First, the method consists in the calibration of the DART model with structural and optical measurements of forest plots, and then in the quantification of spectral separability between forest stands. Very high resolution (4m) imagery was selected for the validation of the approach because of the considerable savings in ground measurements compared to the use of groups of pixels of 20-30m resolution imagery. In chapter 4, validation tests of this model are presented against three IKONOS multi-spectral images.

The presence of topography is known to complicate the correct distinction between forest types, and few studies, still, successfully separate forests on steep terrain.. Indeed, the shadowed area and the reflectance of the understory may drastically affect the spectral separability of forests with the aspect of the slope. This phenomenon is not well understood because most studies are empirical, image-based. In chapter 4, we apply the model-based method to study ambiguity patterns of forest stands on steep terrain.

Spectral separability of forest stands on slope in real and simulated IKONOS imagery[Θ]

Stéphane Couturier

Centre d'Etudes Spatiale de la Biosphère (CESBIO), CNES/CNRS/Paul Sabatier University

Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)

Jean-Philippe Gastellu-Etchegorry and Emmanuel Martin

Centre d'Etudes Spatiale de la Biosphère (CESBIO), CNES/CNRS/Paul Sabatier University

Pavka Patiño

Centro de Investigación en Ecología (CIECO), UNAM

Heterogeneity in the reflectance signal of forests, especially on steep terrain, makes difficult their classification from remote sensing imagery. The ambiguity between classes is not well predicted because most studies on classification of forests on steep terrain are empirical, image-based. We propose a model-assisted method to evaluate spectral separability between forest stands on existing sites of flat and steep terrain, allowing for the aspect to vary. Multi-spectral IKONOS images were acquired on areas where three types of major forest biomes (oak, pine and high tropical forest) are represented in Mexico. A detailed structural record, optically-based Leaf Area Index (LAI) estimates and bi-directional reflectance measurements of the ground were performed for two forest plots (on flat and on steep terrains) of each class. The potential of the Discrete Anisotropic Radiative Transfer (DART) to approximate the signal of real, very high resolution imagery was first tested on the existing slope aspects. The Euclidian distance, employed in commonly used unsupervised classifiers, was not sufficient to distinguish forest types in most aspect configurations. Tropical forest on flat terrain was spectrally undistinguishable from tropical forest on the existing slope aspect, although it could be distinguishable in other aspect scenarios. Finally, separability between pine and tropical forests depended on the slope aspect, presumably because of the presence of pronounced shadows. This work stressed the usefulness of 3-D radiative transfer models for providing quantitative information about the forest structure that appears on real imagery.

[Θ] In review for publication in the refereed journal *International Journal of Remote Sensing*, Taylor & Francis

1. Introduction

The classification of remote sensing images for forest mapping at fine taxonomic resolution is a difficult challenge. At high (*circa* 25m) spatial resolution, recent case studies based on automatic classification methods for temperate forests (e.g. Magnussen et al., 2004) and for subtropical forests (Salovaara et al., 2005, Lu et al., 2003) do reach satisfactory results. However, extensive forest maps derived from automatic classification are affected by greater environmental heterogeneity, and therefore confusion patterns prove quite substantial when comprehensive accuracy assessments are made (Wickham et al., 2004; Remmel et al., 2005).

At very high (1-5m) resolution, obtained by an increasing group of imagers, case studies of automatic classification already proved encouraging in distinguishing forest types (Wang et al., 2004, Thenkabail et al., 2004), and even individual tree species (Clark et al., 2005) when focussed at crown scale. Nonetheless, most case studies beyond crown scale, such as those mentioned above, deal with forest cover on relatively flat terrain.

The presence of pronounced topography is an environmental factor that is likely to affect classification accuracy at landscape scale. The understory reflectance and the shadowed area of a forest on steep slope, and consequently its overall spectral appearance, may be significantly different from those of a forest on flat terrain, and its spectral separability with respect to other forest types may as well be modified. Some strategies of classification have been to consider sub-classes of forests corresponding to two illumination aspects (shadowed or sunlit) of the same forest (Helmer et al., 2000, Baban and Kamaruzaman, 2001). However, as Baban and Kamaruzaman (2001) suggest, the success of the strategy relies on an *a priori* high separability among forest classes (in their case, rubber, mangrove, and inland forests) on the image at hand, information that a user might want to know from a systematic, non-empirical way.

Major mountainous formations cross the Mexican territory, a country where national thematic cartography has deliberately relied on visual interpretation (Mas et al., 2002), rather than on unsupervised clustering adopted elsewhere as a first classification step. A key question of this research is whether a model-oriented approach may serve as a guide for landscape-scale forest automatic classification in rugged-terrain scenarios. Simulations

of remote sensing images allow the exploration of a wide range of environmental settings. The Discrete Anisotropic Radiative Transfer (DART) model (Gastellu-Etchegorry et al., 2004) simulates remote-sensing images of forest plots in many experimental configurations (Gastellu-Etchegorry et al., 1999) such as view and sun directions, atmosphere, topography, multiple tree species, etc.

A variety of metrics and associated spectral distances is used as a quantitative means of evaluating the level of separability between two classes, depending on the classification method; The Bhattacharyya distance (BD), suitable to measure separability over an entire spectral range, and valued for its correlation with the probability of correct classification (Landgrebe, 2000), was used between classes represented by the set of pixels in a forest plot. We also evaluated the separability between classes according to the Euclidian distance (ED) between mean pixel values.

A first step of the research was to evaluate the potential of the DART model to represent the heterogeneous signal of real, very high resolution imagery. To this end, we conducted a detailed survey of 6 forest plots (3 on flat terrain and 3 on slope) appearing on 3 IKONOS images. A detailed structural record, optically-based Leaf Area Index (LAI) estimates and bi-directional reflectance measurements of the ground were performed. These three types of measurements, and environmental parameters matching conditions of the IKONOS overpasses, served as inputs to the DART model for the simulation of forest plot images. The simulated images of plots on flat terrain were calibrated to the IKONOS image for atmospheric parameters and leaf optical properties, as described in the section on simulated imagery.

Within a given forest type, the variability of the measurements was also assessed on the ground. This variability was used when simulating spectral images of plots of the same forest type. A set of distances was computed from all pairs of these images and we called their maximum the spectral separation within a forest type, or ‘intra-class separation’. Then, distances between the DART simulated image of each forest plot and its corresponding IKONOS sub-image were compared with the intra-class separation, in the flat and steep terrain cases.

Finally, we conducted a study aimed at quantifying separability between the 3 forest plots on slope, simulated in various aspects with respect to sun illumination. This study was

intended as a possibly valuable contribution of DART for the prediction of confusions when automatic classification are launched on forest types in rugged terrain scenarios.

We successively detail the methods for ground measurements, preparation of satellite imagery and calibration of simulated imagery in sections 3, 4 and 5. We then describe the methods for DART validation, and then separability assessments in real and simulated settings in section 6.

2. Study Areas

We focused on three sites in Mexico (see figure 4.1) where major forest biomes are represented. In each site, one forest type was selected and two plots (one on flat terrain and one on slope) were surveyed (see table 4.1). The 'Tancítaro' site is located in the communal lands of San Juan Parangaricutiro, Michoacán state, near to the natural reserve of the Tancítaro volcano peak. This site is about 2200m above sea level and the climate is close to temperate sub-humid. Two plots, dominated by species of pine with very close botanic characteristics, namely *Pinus Montezumae* (flat terrain) and *Pinus Pseudostrobus* (steep terrain), were selected in Tancítaro.

The 'Cuitzeo' site is located in the ejidal lands of El Cerro, Michoacán state, close to the Cuitzeo lake. This site is about 2050m above sea level and the climate is intermediate between tropical dry and temperate sub-humid. Two plots of oak forests were selected, both dominated by *Quercus Deserticola*.

Finally, the 'Los Tuxtlas' site is located in the communal lands of La Perla de San Martin ejido, Veracruz state, within the biosphere reserve of Los Tuxtlas. The village is situated near to the coast, on the flanks of the San Martin volcano at about 900m above sea level.

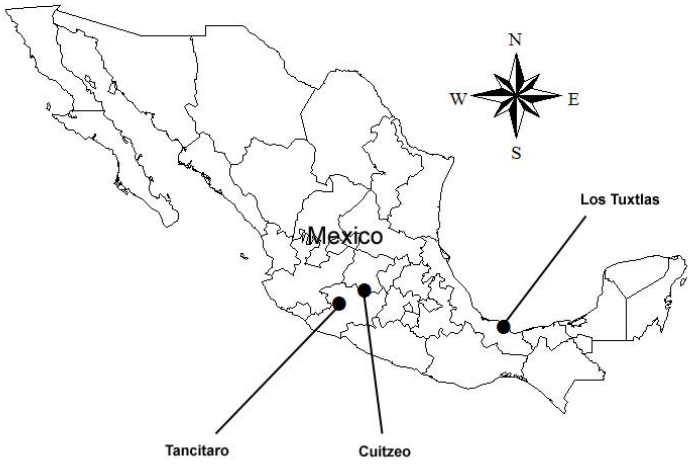
Table 4.1. Geometric characteristics of forest plots.

Forest plot	Code	Site	Plot size		Inclination	Exposition
			on the ground (m)	on imagery (pix)[a]		
Oak Flat	OF	Cuitzeo	50 x 50	156	Flat	
Oak Slope	OS	Cuitzeo	20 x 40	47	18°	North-East
Pine Flat	PF	Tancítaro	50 x 50	156	Flat	
Pine Slope	PS	Tancítaro	25 x 50	70	27°	South
Tropical Flat	TF	Tuxtlas	40 x 40	100	Flat	
Tropical Slope	TS	Tuxtlas	50 x 50	137	30°	East

[a]Plot size on IKONOS multi-spectral scenes are expressed in pixels ('pix')

The climate is mainly characterized by tropical humid conditions. Two plots of well-preserved old-growth high tropical forest were selected. In the latter site, relief-induced transitions in ecosystems explain the presence of conifer as well as oak-dominated broadleaf forests within kilometres of the tropical humid forest. This situation, not uncommon in Mexico and Sub-tropical America, is the rationale for studying spectral separability between the mentioned types of forests. Thereafter, we denominated 'forest stands' as being portions of the forest that include the forest plot and whose orientation is similar as that of the forest plot.

Figure 4.1. Location of forest sites.



3. Ground Measurements

The ground survey of forest plots consisted in the detailed record of the canopy structure, LAI measurements and soil reflectance evaluation, as detailed below. Additionally, for each forest plot, five nearby parcels with the same size as the plot, in the same forest stand, were selected and used for the variability study. In these parcels (thereafter the 'reference parcels'), crown coverage estimation, LAI measurements and soil reflectance evaluation were conducted, as indicated below.

3.1 Forest structure

A key criterion for the selection of the plots was that their crown coverage (CC) would be representative of the surrounding forest stand. This criterion was first submitted to expert local foresters guiding the research team in the field. Then, it was verified against both visual inspection of the panchromatic IKONOS scenes and the estimation of CC in the reference parcels. This CC estimation was done using two crossed metric tapes on the ground: the portion of the tapes that intersected the projected crown cover on the ground was measured and divided by the total length of the tapes.

Table 4.2: Structural parameters of forest plots and stands

1	2	3	4	5	6	7
Forest code	Number of species	Crown Cover in plot	Crown Cover in forest stand	LAI in plot	LAI in forest stand	
					Min	Max
OF	12	64%	54-72%	2.5	2.0	3.2
OS	4	80%	75-88%	2.3	1.8	2.6
PF	3	87%	75-94%	3.5	2.9	4.1
PS	9	64%	48-75%	5.1	3.8	7.0
TF	31	97%	92-100%	7.9	7.0	8.9
TS	46	92%	84-96%	9.5	8.9	10.5

The forest plots were rectangular, marked with metric tapes to record relative positions of trees. Rugged terrain forest plots were selected on a relatively steady slope whose average was measured with a clinometer. First, a botanic inventory was performed for trees larger than 10cm DBH (old-growth tropical forest) or than 5cm DBH (other forest types). The number of species inventoried appears on table 4.2. No association of a particular species was noticed in tropical forest plots.

Second, the measurement of crown dimensions was conducted for each individual inventoried tree. Trees outside the plot but whose crown infringed in the parcel were measured as well. The envelope of each crown was approximated by an ellipsoid whose axes were aligned to the orientation of the rectangular parcel, and whose extreme horizontal coordinates were visually evaluated on the ground, using compasses and metric tapes. Vertical dimensions were evaluated with the assessment of the local expert foresters and assisted with a clinometer. Sometimes, more than one ellipsoid was needed to approximate the geometry of the crown. The 3D structure of the canopy, as well as the CC displayed in table 4.2, were derived from these measurements.

3.2 Leaf Area Index

Effective LAI (term employed by Chen and Cihlar, 1996, but thereafter referred to as simply 'LAI') estimations were made from measurements taken with the SUNSCAN Canopy Analysis System (Delta-T Devices Ltd, Cambridge, UK). A linear probe containing 64 photodiodes measures the radiation intercepted by the canopy while at the same time, shaded and unshaded sensors placed in a clearing read direct and diffuse sun radiations. Assumptions and principles underlying the derivation of LAI from these measurements are detailed in Potter et al. (1996).

For each plot, a set of 16 LAI estimates was done along two 20m-long, perpendicular transects within its perimeter. An additional 10 LAI estimates were made in a similar way and averaged within each reference parcel (a total of 50 LAI measurements). Some of these 50 measurements that were associated to large LAI values could be identified in the panchromatic IKONOS image. They corresponded to 8mx8m parcels on flat terrain with lit crowns, 100% CC, and high signal with respect to the surroundings in the co-registered near infrared (NIR) IKONOS scene. The canopy structure (average height and thickness of crown) of the trees was estimated in this 8mx8m area. These 'high LAI values' were selected for calibration of the simulated imagery, as explained in the section on simulated imagery.

With respect to LAI derivations using hemispherical photography (unavailable for our campaign), our method certainly suffered from the dependence of gap fraction measurements on the zenith angle of direct illumination (Jonckheere et al., 2004:31). Our

emphasis, given tight constraints on available material and time for ground campaigns, was to obtain LAI measurements as coherent as possible between plots/forest types, rather than aim accurate LAI estimates in an absolute sense. Therefore, in order to attenuate the bias due to variable sun zenith angles among plots, the measurement campaign occurred at the same period of year (March-April 2005, the season corresponding to the Cuitzeo overpass, the only site where LAI of forests is strongly seasonally-dependent) and at a time of the day when the sun was at 35-40^{o} zenith angle, markedly away from noon but within the domain of LAI retrieval by the SUNSCAN software. For example, the measurements occurred at around 10:30 am local time in the Los Tuxtlas site. The LAI estimates in forest plots and in reference parcels are indicated in, respectively, column 5 and 6 of table 4.2.

3.3 Soil Reflectance

Bare soil, rock and soil covered with different types of litter characterized the ground conditions in all forest plots, except for the 'pine on slope' plot where the soil was almost entirely covered with grass. In the presence of forest plots with sparse canopy or low LAI, an optical characterization of the understory was deemed necessary. Pinty et al. (1998) suggest that in case of low LAI, strong biases may occur on top of canopy reflectance if soil reflectance is assumed Lambertian. Therefore, we undertook to estimate the bi-directional optical behaviour of the understory along seven viewing angles (nadir; 30 and 60 degrees off nadir in the principal plane away from the sun; 60 and 75 degrees off nadir in the principal plane towards the sun; 30 and 60 degrees off nadir in the perpendicular plane), using a 512-channel hand-held spectro-radiometer (Geophysical and Environmental Research Group [GER] 1500) attached to a clinometer. The bi-directional reflectance is approximated with the SOILSPECT model (Jacquemoud et al., 1992); i.e., an analytic function with 6 parameters (single scattering albedo, soil roughness, and 4 phase function parameters). Ground measurements along the seven viewing directions were considered to derive these 6 parameters with an inversion procedure based on a genetic algorithm (Carroll D., University of Illinois). These 6 parameters are then used by the DART model (Gastellu-Etchegorry, 2006).

After a few attempts, we realized that unsteady (direct and diffuse) illumination conditions and poor control over the projected field of view of our hand-held instrument on spatially

heterogeneous understories impeded a direct appraisal of optical properties in the field. Instead, we proceeded to a visual description of the understory in terms of quantitative fractions of bio-physical constituents present on the ground over the whole forest plot. The bio-physical constituents encountered in the field were then visually related to a reduced set of seven most frequent, conspicuous constituents: a type of humus-rich bare soil, a type of basaltic rock, litter of yellow dried grass, orange-brown leaf litter, dark brown leaf litter, grey (decomposed) leaf and branch litter, and brownish needle litter. These seven constituents were collected in the field and then arranged in the laboratory as an opaque surface to reproduce spatially homogeneous understory scenarios on which optical measurements were performed with illumination zenith angle similar to the ones of the IKONOS overpasses.

Five replicate measurements of the ground's upwelling radiance were taken using the radio-spectrometer held 1m above the centre of the targeted area, at all above-mentioned viewing angles, with a 4^{o} field of view. The instrument integrated exposure times of 5msec at a spectral resolution of 3nm FWHM over the range of wavelengths from 300nm to 1050nm. Mean reflectance spectra of homogeneous understories were computed after a normalization to the radiance response of a Labsphere Teflon reference panel, which diffusely reflects incident radiation over all angles. The bi-directional reflectance of each 'real' understory was then obtained from a linear combination, according to the fractions observed in the field, of the spectra obtained in homogeneous conditions. This operation was tied to the assumption that the overall reflectance response of the understory, a mosaic of bio-physical constituents, would be equivalent to the linear spectral mixture of responses of homogeneous understories (i.e. understories with a single bio-physical constituent) put aside. This assumption could be verified in field measurements of understories made of conspicuous patches of two constituents.

SOILSPECT parameters were then derived (table 3.3) from these constructed bi-directional spectra once they had been convolved to the IKONOS response in every multispectral band of the sensor (NIR, Red, Green and Blue). A preliminary study revealed that the spectral independence of phase function parameters, experimentally observed by Jacquemoud et al. (1992), was verified for laboratory-derived understory reflectance whereas this independence was not verified for some understory reflectance

values measured directly in the field. This observation strengthened our view that measurements taken directly in the field suffered from target heterogeneity.

In order to estimate the variability of soil reflectance in the surrounding forest, we characterized the understory of the 5 reference parcels in each forest stand. In the same way as above, we derived their SOILSPECT parameters.

Table 4.3: Soil optical properties of forest plots

Forest code	Dominant understory component	Albedo single scattering				Roughness *h*	Phase function parameters			
		NIR	Red	Green	Blue		*b1*	*b2*	*c1*	*c2*
OF	Bare soil, humus-rich	0.43	0.40	0.24	0.22	0.56	-0.12	0.20	0.62	-0.36
OS	Grey decomposed litter	0.55	0.42	0.26	0.23	0.36	0.23	0.16	0.19	0.28
PF	Needle litter	0.52	0.42	0.33	0.31	0.16	-0.15	-0.25	0.38	0.12
PS	Bare soil, humus-rich (+grass)	0.45	0.42	0.23	0.21	0.49	0.94	-0.37	0.30	-0.23
TF	Dark brown leaf litter	0.52	0.39	0.32	0.28	0.43	0.28	-0.76	0.25	-0.09
TS	Dark brown leaf litter	0.52	0.39	0.32	0.28	0.43	0.28	-0.76	0.25	-0.09

4. Satellite imagery

Three panchromatic and multi-spectral IKONOS scenes, each covering one site, were acquired. The acquisition characteristics are described in table 4.4.

Table 4.4: Acquisition parameters of IKONOS imagery.

Site	Geometric Coordinates Latitude (°)	Longitude (°)	Date	Local Time[a]	θ_s (°)	φ_s (°)	θ_v (°)	φ_v (°)
Cuitzeo	20.0	-101.1	04/03/2003	10:36	33.2	141.5	14.7	358.9
Tancítaro	19.4	-102.2	20/11/2002	10:47	40.9	162.9	21.9	290.4
LosTuxtlas	18.5	-95.1	29/11/2004	10:04	43.1	157.6	21.8	33.3

[a]*Local time is GMT-7:00.*

The scenes were provided in 'Standard Geometrically Corrected' mode by Space Imaging. Nonetheless, geometric distortions on the imagery were likely to occur due to local topography. Six sub-images, each containing a forest plot, were created from the multi-spectral IKONOS scenes. A fraction of the corresponding panchromatic sub-images is shown in figures 4.2 - 4.3 ((a) and (c)) and figure 4.4 ((a) and (e)).

Figure 4.2: Oak forest plots on panchromatic imagery

The plot on flat terrain on IKONOS (a) and DART (b) imagery. The plot on steep terrain on IKONOS (c) and DART (d) imagery. DART images result from the fusion of DART multi-spectral images.

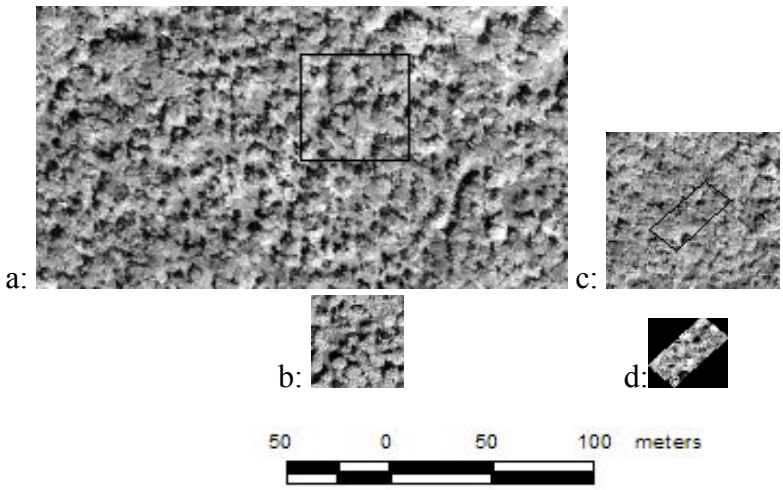
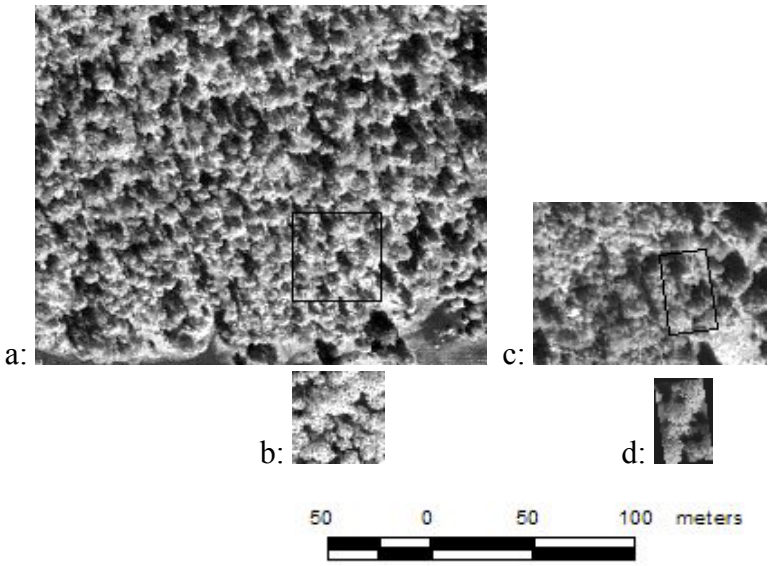


Figure 4.3. Pine forest plots on panchromatic imagery

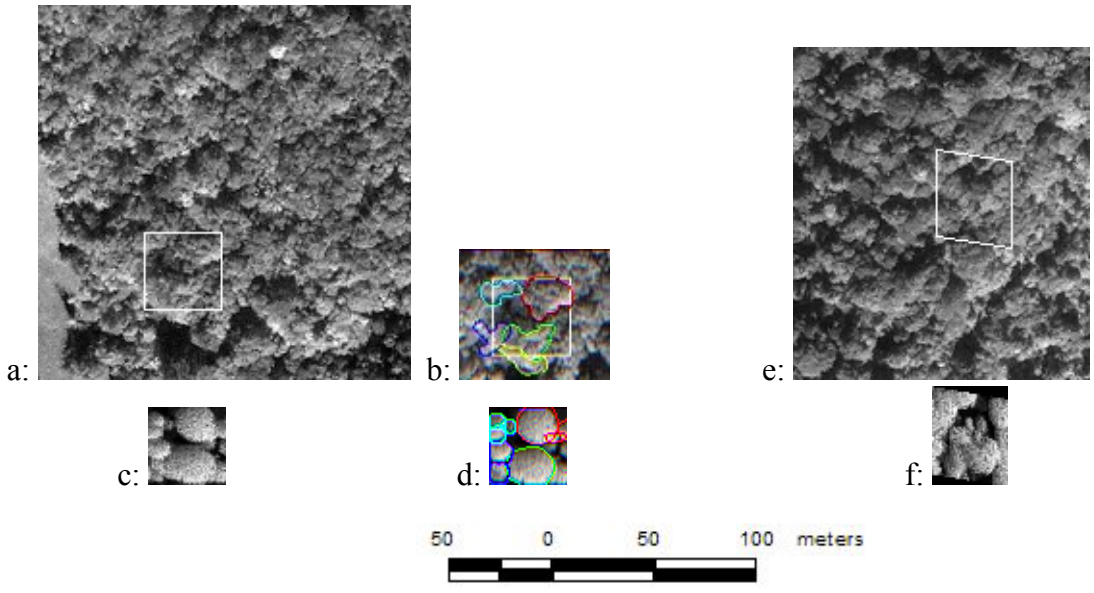
The plot on flat terrain on IKONOS (a) and DART (b) imagery. The plot on steep terrain on IKONOS (c) and DART (d) imagery. DART images result from the fusion of DART multi-spectral images.



The geometry of each sub-image was independently corrected using the corresponding panchromatic sub-image as a reference, once the panchromatic sub-image was corrected via identifiable landmarks: readings from a GARMIN® GPS unit (WAAS mode) formed tracks around crowns of isolated trees in clearings, and along clearing borders/access trails in the vicinity of the forest plots. 2nd degree polynomial transformations were sufficient in all cases to make the panchromatic image match these tracks at less than 4m accuracy (sub-pixel for multi-spectral imagery). Forest plots were then located in the corrected sub-image mostly through visual identification of pre-eminent trees or gaps in the panchromatic image, and alternatively with repeated GPS readings inside the plot, with less than 3m accuracy, using climbing equipment on trees when necessary.

Figure 4.4: Tropical humid forest plots on panchromatic imagery

The plot on flat terrain on IKONOS (a, b) and DART (c, d) imagery. The plot on steep terrain on IKONOS (e) and DART (f) imagery. DART images result from the fusion of DART multispectral images. Emergent crowns are highlighted on (b, d).



The multi-spectral IKONOS raw digital numbers were converted to top-of-atmosphere reflectance using the standard solar irradiance formula provided by Space Imaging. We did not attempt to work with atmospherically corrected bottom-of-atmosphere radiometry. Indeed, for hyper-spatial imagery, such inversion process only derives accurately mean radiance values for patches of homogeneous surfaces and fails to recover the original texture of the image, attenuated in a complex way through atmospheric multi-scattering. Instead, the ability to distinguish forest types on remote sensing support was naturally tested on at-sensor reflectance, and all simulations were therefore carried on using the atmospheric radiative transfer of the DART model (Gastellu-Etchegorry et al., 2004).

5. Simulated imagery

5.1 General Model Assumptions

In the context of this research, the DART model was tested in its forward mode, as a proxy for IKONOS forest scenes. Four systems of parameters (sun with polar position (θ_s, φ_s), atmosphere, landscape, and sensor, with polar position (θ_v, φ_v)) are considered as inputs of each simulation S. No atmosphere measurement was available at IKONOS overpass. Besides, the limited equipment available for this study and the great diversity of leaves/needles present in the forest plots hindered the measurement of their optical properties, not reported in existing databases. Instead, measurements on homogeneous surfaces and in reference parcels of each IKONOS scene allowed us to invert from the imagery atmospheric parameters and leaf optical parameters of forests on flat terrain, using auxiliary simulations as explained below.

Optical properties from the ASTER database (ASTER, 2003) were employed for the grass constituting the understory of the pine forest on slope. Besides, for all forest plots, the shape of tree crowns was assumed ellipsoidal and the leaf angle distribution was assumed spherical.

5.2 Atmospheric Calibration

Bi-directional measurements were acquired with the GER-1500 spectro-radiometer in at least two recognizable reference surfaces per IKONOS scene: large crossings of asphalted and dry dirt roads in Los Tuxtlas, gravel and asphalt roads in Tancítaro, crossing of asphalted roads and a small quarry site in Cuitzeo. The TOA reflectance of these surfaces was simulated assuming a tropical rural atmospheric model with 23km visibility (Berk et al., 1989), but leaving the aerosol optical depth (AOD) as a variable. The AOD was inverted in the blue and green bands, from the simulations that gave the reflectance closest to the IKONOS signals. Then, AOD and total transmittance of the vapour column (T_{H2O}) were considered as variables in the remaining red and NIR bands, since the absorption of water is a significant parameter in these spectral regions.

We tried two strategies to derive AOD in these bands from AOD found in the blue and green bands: AOD (τ_A) is dependent on wavelength according to a power law $\tau_A \sim \lambda^{\alpha}$. A first strategy was to derive α from the quotient of AODs in the blue and green regions, and then compute AODs in the red and NIR. A second strategy was to employ an α value typical of a tropical rural aerosol model. Once AOD known in red and NIR, simulations of the reference surfaces was used to derive T_{H2O} in these bands. The first strategy yielded unrealistically low values of T_{H2O} in the Tancítaro site; an extrapolation of the power law derived from spectral regions as close as blue (480nm) and green (550nm) may indeed generate significant errors in the red and NIR bands. We therefore settled for the second strategy and obtained values for atmospheric parameters as shown in table 4.5.

Table 4. 5: Atmospheric parameters and leaf optical properties (reflectance and transmittance)

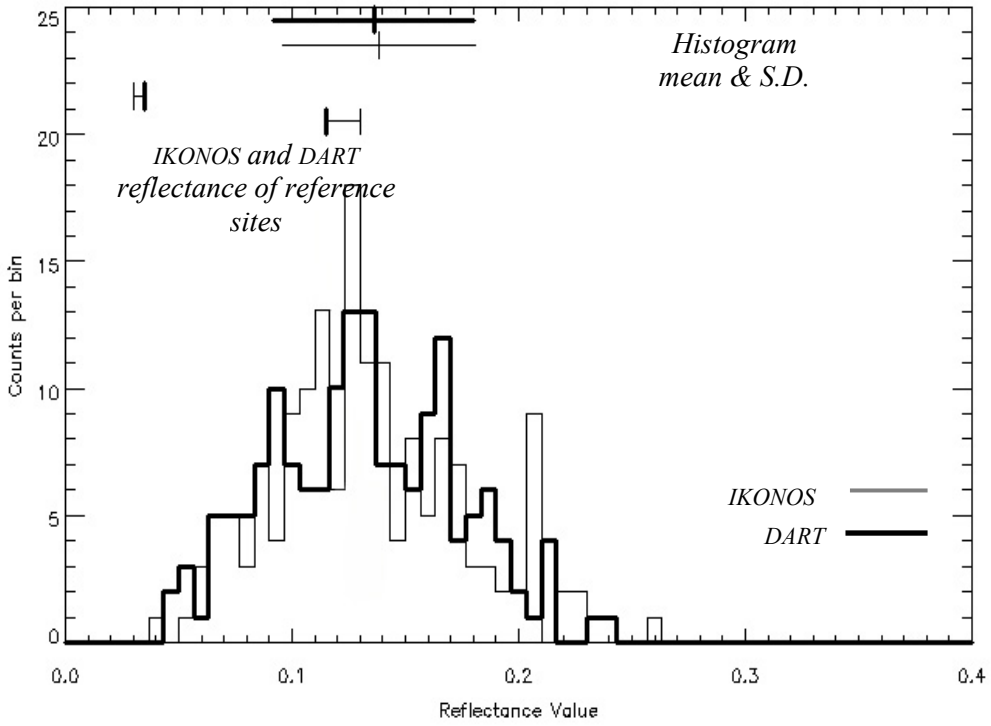
Site	Aerosol Optical Thickness				Water transmittance		Forest type	Leaf optical properties			
	Blue	Green	Red	NIR	Red	NIR		Blue	Green	Red	NIR
Cuitzeo	0.372	0.325	0.250	0.199	0.999	0.910	Oak	0.049	0.087	0.047	0.462
Tancítaro	0.358	0.302	0.241	0.191	0.990	0.850	Pine	0.048	0.082	0.044	0.467
Los Tuxtlas	0.362	0.313	0.243	0.194	0.995	0.860	Tropical	0.052	0.088	0.044	0.470

Figure 4.5: Distribution of pixel reflectance in the NIR band

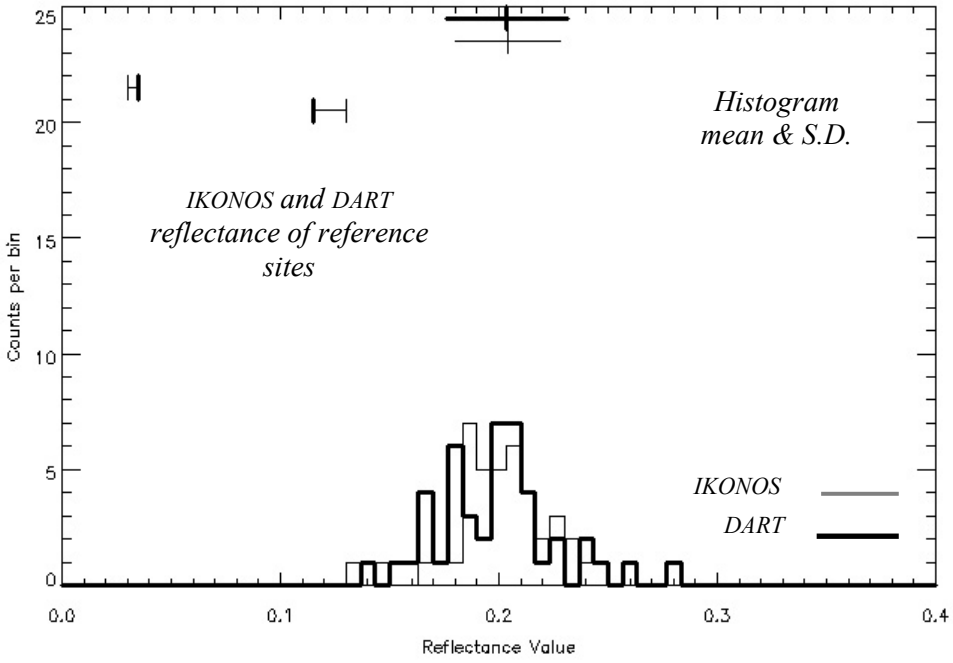
Near infrared signal of IKONOS (thin line) and DART-simulated (thick lines) imagery. Oak (a), pine (c) and tropical (e) forests on flat terrain. Oak (b), pine (d) and tropical (f) forests on steep terrain. Histogram mean and standard deviation are shown in the top part of the graph. IKONOS (thin bound) and DART simulated (thick bound) reflectance values of the 2 reference sites per image (selected for atmospheric calibration) are shown right above the histograms.

5a. Oak forest

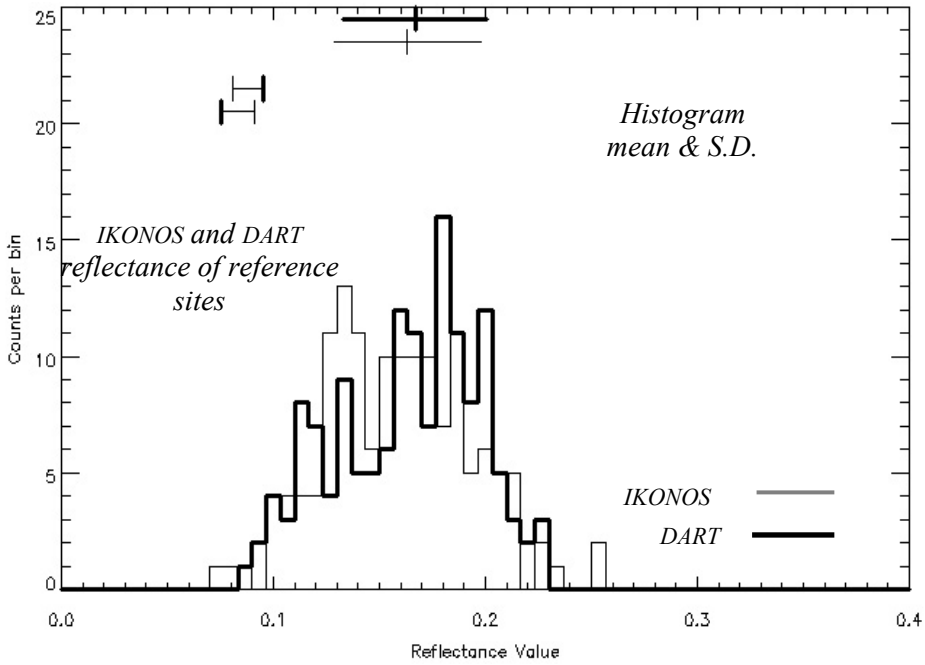
Flat terrain



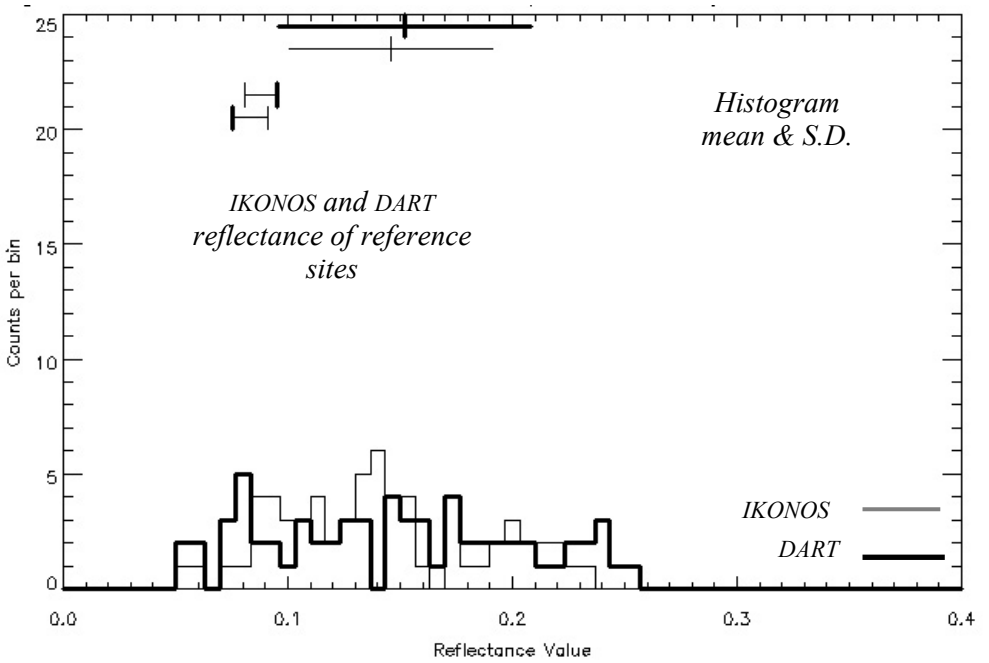
5b. Oak forest
Steep terrain



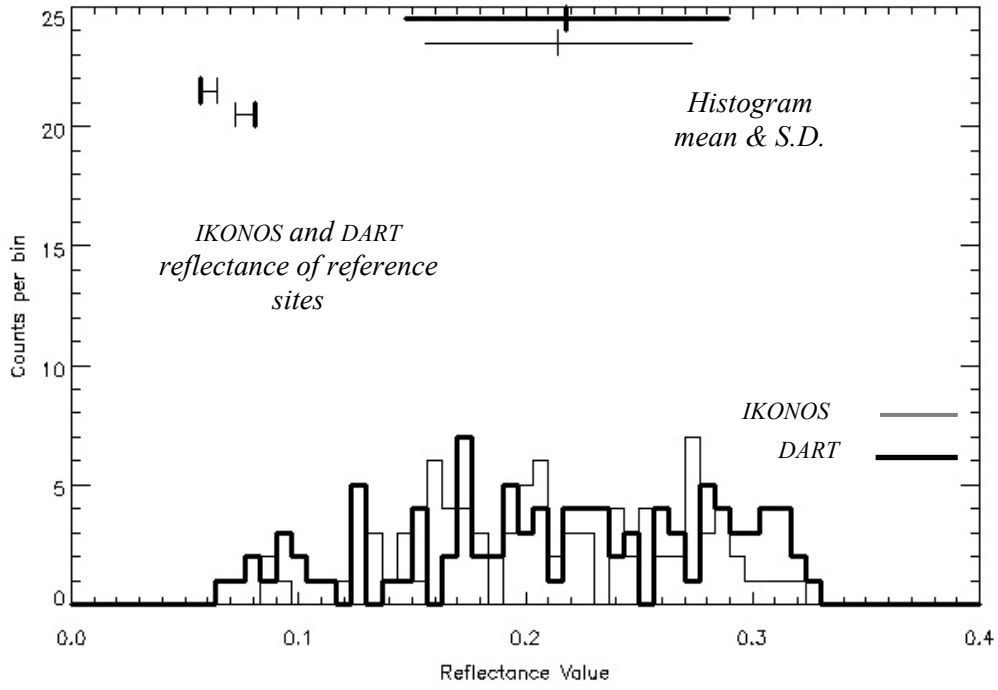
5c. Pine forest
Flat terrain



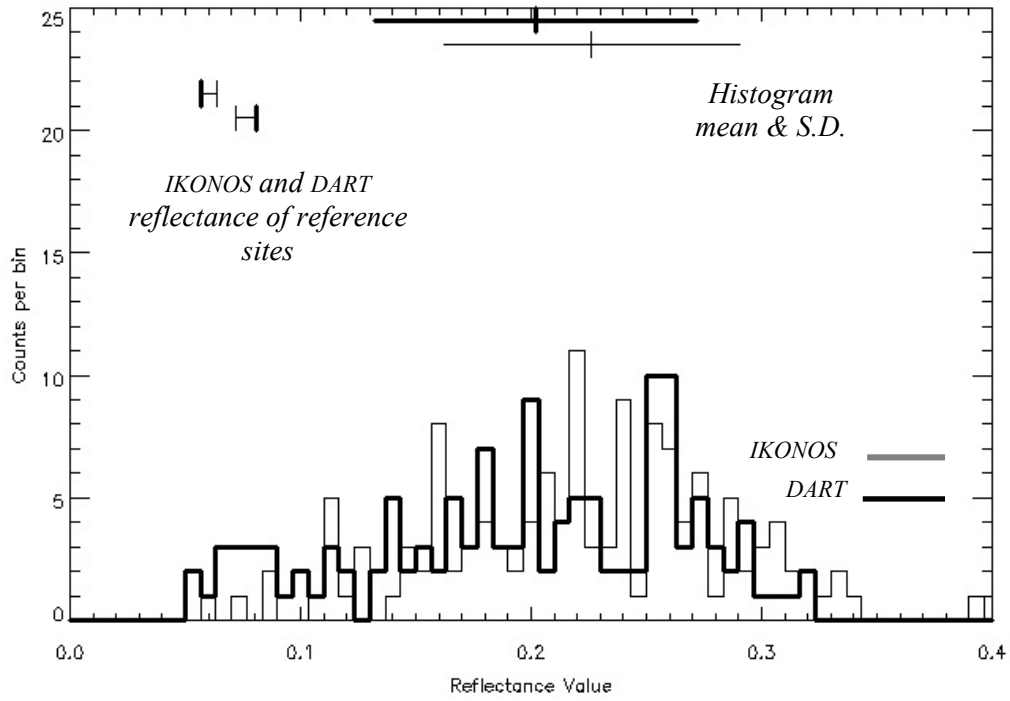
5d. Pine forest
Steep terrain



5e Tropical forest
Flat terrain



5f. Tropical forest
Steep terrain



5.3 Calibration of Leaf Optical Properties

In order to calibrate the Leaf Optical Properties (LOP) of each forest plot on flat terrain, we used the ground measurements (LAI, structure) made in the 8mx8m area corresponding to lit crowns and 100% CC; The LOP value was inverted from such a simulated 8mx8m scene, i.e. when the simulated TOA reflectance value from the simulated lit crowns fitted the observed IKONOS reflectance. Leaf reflectance (both sides) and transmittance were assumed equal and LOP was given a unique spectral value for all trees in the forest plot. Indeed, while using this double assumption, Gascon et al. (2004) obtained robust results for biophysical parameter inversion from very high resolution imagery.

Although in the tropical case many tree species are present (see table 4.2), the assumption of a unique LOP for the entire forest plot may actually make sense since there is a strong

contribution - through atmospheric multi-scattering - of neighbouring pixels (and thus neighbouring crowns) to the TOA reflectance of the 8mx8m area where LOP is calibrated. In this sense, the derived LOP value should be viewed as the average optical property of a group of tree crowns, representing the optical property of crowns in the forest stand, rather than the LOP of a specific tree crown. The LOP values obtained via this calibration process are displayed in table 4.5.

Figures 4.2-4.3 ((b), (d)) and figure 4.4 ((b), (d), (f)) illustrate DART simulations of the forest plots seen on the panchromatic IKONOS scenes. These DART-derived images resulted from the fusion of multi-spectral images simulated at 1m resolution, to approximate the panchromatic signal in the IKONOS sensor. However, they were used only for illustration purposes because panchromatic imagery is of limited radiometric reliability.

6. Method for separability assessments

The Bhattacharyya distance BD_{ij} between forest plots *i* and *j* seen on multi-spectral IKONOS imagery was calculated according to (Landgrebe, 2000):

$$BD_{ij} = \frac{1}{8}\left[\mu_i - \mu_j\right]^T \left[\frac{\Sigma_i + \Sigma_j}{2}\right]^{-1} \left[\mu_i - \mu_j\right] + \frac{1}{2}\ln\frac{\left|\frac{1}{2}\left[\Sigma_i + \Sigma_j\right]\right|}{\sqrt{\left|\Sigma_i\right|\left|\Sigma_j\right|}} \qquad (4..1)$$

where μ_i is a vector of *B* elements (*B* is the number of spectral bands) containing the mean spectral value of all pixels in forest plot *i*, and Σ_i is the covariance matrix of the vectors corresponding to the pixel values in forest plot *i*.

Besides, the Euclidian distance $ED_{ij} = |\mu_i - \mu_j|$ measured the separation between classes *i* and *j*. Because of the high variability of the signal in hyper-spatial imagery of forests (e.g. see Gascon et al., 2004), mean values strongly depend on the resolution considered and ED discriminates classes with greater success at lower resolutions. Consequently, mean values taken for ED distances were extracted from a 25mx25m subplot (which corresponds to 39 pixels, less than the size of our smallest plot).

BD and *ED* were first computed between DART-simulated and IKONOS images of the six real forest plots, yielding the 'DART-IKONOS distances'. Next, in the six forest settings, modifications were applied to three input parameters of the simulations, namely CC, LAI, and ρ_u (understory reflectance), according to the variation measured in the reference parcels: for each forest setting, the minimum and maximum CC, LAI (see table 4.2 for both parameters) and ρ_u measurements were selected. Minimum and maximum ρ_u refer to the understory characterized by, respectively, the lowest and the highest single scattering albedo from all reference parcels. As a result, a set of eight simulations S(CC, LAI, ρ_u) per forest setting was computed, with these three parameters equal to either the minimum or the maximum value. For example, S(CC_{min}, LAI_{max}, $\rho_{u\ min}$) is one of these simulations. Distances were computed between all pairs of these eight simulations, which led to a total of $C_8^2 = 28$ distances for each pair of forest plots. The maximum of these distances expressed the 'intra-class separation' of a forest setting. We assumed that this maximum distance reasonably approached the maximum of all possible distances within a forest type if the entire space of variables (CC, LAI, ρ_u) had been explored.

Confronting the DART-IKONOS distances with the intra-class separation was a means of testing DART as a proxy for the real, very high resolution IKONOS imagery. The steep terrain scenario was also expected to test the validity of DART against IKONOS, since calibration for leaf optical properties was done only on flat terrain settings. Distances between forests on flat and steep terrains of real IKONOS imagery were also computed and compared to intra-class separations.

As a second step, six aspect configurations were simulated for the three forest plots on slope, setting viewing orientation, sun position, and atmospheric conditions to the values found in the Tancítaro area. Expositions of 0, 60, 120, 180, 240, and 270 degrees off of South were selected. Distances between pairs of forest types as well as intra-class separations were calculated for each aspect, the comparison of which was aimed at assessing the separability of forest classes with variable terrain topography (i.e. variable sun and/or view directions).

7. Results

7.1 Validation study of forest plots on real imagery

Table 4.6 reports the computed ED and BD distances for the DART-IKONOS comparison, for the DART-derived intra-class study and for the IKONOS-IKONOS flat to slope separation, in all forest settings. The distances were computed in the whole spectral range (column denominated 'all bands') and for each mono-spectral band. ED scores in all bands, based on the comparison of mean reflectance values, helped us apprehend the signal observed. In all forests, the DART-IKONOS ED was lower than intra-class ED separation, indicating a good agreement between IKONOS imagery and DART-derived imagery, even in steep terrain settings where leaf optical properties of DART simulations had not been calibrated to the vegetation signal of the IKONOS imagery. For example, the DART-IKONOS distance was 0.3, 0.4, and 0.8 percent reflectance in, respectively, oak, pine, and tropical forests on slope, whereas the intra-class distances were 1.1, 1.0, and 2.1, respectively.

Beyond the comparison of mean values, the similarity of pixel value distributions in DART imagery with respect to IKONOS imagery can be appreciated in figure 4.5 (a-f), where histograms of corresponding scenes in the NIR band (the spectral band with the strongest dynamic range) are displayed. BD expresses a distance between multidimensional distributions, assuming a Gaussian shape (Kailath, 1967). DART-IKONOS BD scores in the 'All bands' column of table 4.6 are lower than intra-class BD, indicating the DART-IKONOS similarity of pixel value distributions in the four dimensional space.

Flat to slope distances, when compared to intra-class variation, indicate whether or not the forest plot on steep terrain is separable from the one on flat terrain. For example, the ED distance between the oak plot on slope and the oak plot on flat terrain (2.0 percent reflectance) is larger than intra-class variations of both oak on flat (1.3%) and slope (1.1%) terrains. In this sense, considering our DART-based definition of intra-class variation, the oak forest plot on flat terrain is separable from the oak plot on steep terrain.

Table 4.6: Distances between and among DART-simulated and real scenes of six forest plots appearing on IKONOS imagery.

		OAK FLAT					OAK SLOPE				
		All bands	Blue	Green	Red	NIR	**All bands**	Blue	Green	Red	NIR
Euclidian Distance[a]	DART vs IKONOS	**0.3**	0.4	0.4	0.1	0.2	**0.3**	0.7	0.0	0.3	0.1
	Intra-class[b]	**1.3**	1.4	1.0	1.0	3.0	**1.1**	0.9	1.4	1.5	1.1
	Flat vs Slope[c]	**2.0**	0.2	0.8	0.4	6.4	**2.0**	0.2	0.8	0.4	6.4
Bhattacharyya Distance	DART vs IKONOS	**1.04E-01**	3.40E-05	1.10E-05	3.80E-06	1.70E-04	**1.26E-01**	4.60E-04	6.90E-05	1.50E-03	3.10E-06
	Intra-class[b]	**3.28E-01**	4.70E-03	4.20E-03	1.30E-04	3.40E-05	**2.54E-01**	9.00E-04	4.50E-03	3.30E-03	1.50E-03
	Flat vs Slope[c]	**1.38E+00**	1.00E-03	2.00E-03	2.50E-04	3.30E-02	**1.38E+00**	1.00E-03	2.00E-03	2.50E-04	3.30E-02
				PINE FLAT					**PINE SLOPE**		
Euclidian Distance[a]	DART vs IKONOS	**0.3**	0.5	0.5	0.0	0.3	**0.4**	0.4	0.4	0.5	0.4
	Intra-class[b]	**1.5**	1.2	2.8	0.8	2.4	**1.0**	0.8	2.0	0.4	0.7
	Flat vs Slope[c]	**0.6**	0.1	0.3	0.7	1.5	**0.6**	0.1	0.3	0.7	1.5
Bhattacharyya Distance	DART vs IKONOS	**9.70E-02**	5.30E-03	1.20E-04	3.80E-05	5.70E-05	**1.57E-01**	1.30E-04	1.40E-03	1.10E-04	2.80E-05
	Intra-class[b]	**1.32E-01**	9.20E-03	4.30E-04	2.30E-04	6.40E-04	**1.60E-01**	3.50E-03	2.30E-04	3.30E-05	7.80E-06
	Flat vs Slope[c]	**2.43E-01**	3.20E-05	3.70E-04	1.50E-03	2.50E-04	**2.43E-01**	3.20E-05	3.70E-04	1.50E-03	2.50E-04
				TROPICAL FLAT					**TROPICAL SLOPE**		
Euclidian Distance[a]	DART vs IKONOS	**0.5**	0.8	0.6	0.1	0.3	**0.8**	0.2	0.2	0.3	2.4
	Intra-class[b]	**1.6**	1.0	1.7	0.5	4.2	**2.1**	1.3	2.2	1.3	4.9
	Flat vs Slope[c]	**1.0**	1.6	1.1	0.4	1.1	**1.0**	1.6	1.1	0.4	1.1
Bhattacharyya Distance	DART vs IKONOS	**2.23E-01**	1.40E-04	4.30E-04	4.00E-05	5.20E-05	**2.06E-01**	7.20E-04	1.20E-04	2.90E-04	1.20E-04
	Intra-class[b]	**2.59E-01**	1.30E-03	7.60E-03	1.30E-03	7.80E-06	**3.03E-01**	5.40E-04	9.00E-04	3.30E-03	1.10E-02
	Flat vs Slope[c]	**6.60E-02**	6.00E-05	2.90E-06	4.00E-05	4.10E-05	**6.60E-02**	6.00E-05	2.90E-06	4.00E-05	4.10E-05

[a] *Euclidian distances are displayed in percent reflectance.*

[b] *'Intra-class' refers to the maximum distance among the 8 DART images of the forest plot simulated with minimum and maximum parameter (CC, LAI, ρ_w) values measured within the forest class.*

[c] *'Flat vs Slope' refers to distances among forest plots in IKONOS imagery.*

A close look at individual mono-spectral distances shows that generally, no such separability is achievable considering any of the blue, green or red band alone. BD and ED applied to the NIR band alone, conversely, separated the oak forest plots on flat versus steep terrain. The understory strongly impacted the signals of low LAI oak forests. The difference in terrain orientation, as well as strong intrinsic differences of understory single scattering albedos (table 4.3), may both account for the dissimilarity of oak forest on flat versus steep terrains in the NIR. By contrast, NIR distances of flat versus steep terrains are much lower for pine and tropical forests, presumably because their higher LAI reduces the role of the understory.

Table 4.7: Mean reflectance values (%) of forest plots simulated by DART for various aspects, with environmental parameters set to the Tancítaro area.

	Oak forest				Pine forest				Tropical forest			
	Blue	Green	Red	NIR	Blue	Green	Red	NIR	Blue	Green	Red	NIR
Flat[a]	8.2	6.7	5.7	12.7	7.1	4.9	4.3	16.7	7.4	6.1	3.9	18.9
South	8.6	7.1	6.1	18.7	7.7	6.2	4.4	15.7	8.0	6.7	4.3	19.5
South East	9.4	7.8	6.8	19.8	7.7	6.3	4.6	15.6	7.6	6.6	4.2	19.4
South West	8.5	7.0	6.0	17.5	7.7	6.3	4.6	14.3	7.6	6.5	4.1	17.8
North East	8.3	6.9	5.8	17.1	7.5	6.0	4.2	14.0	7.6	6.5	4.2	16.3
North	8.3	6.7	5.7	15.0	7.6	6.1	4.4	12.5	8.0	6.8	4.4	15.4
North West	9.3	7.7	6.2	16.0	7.6	6.1	4.4	12.8	7.6	6.3	4.1	15.0

[a] *Flat terrain condition.*

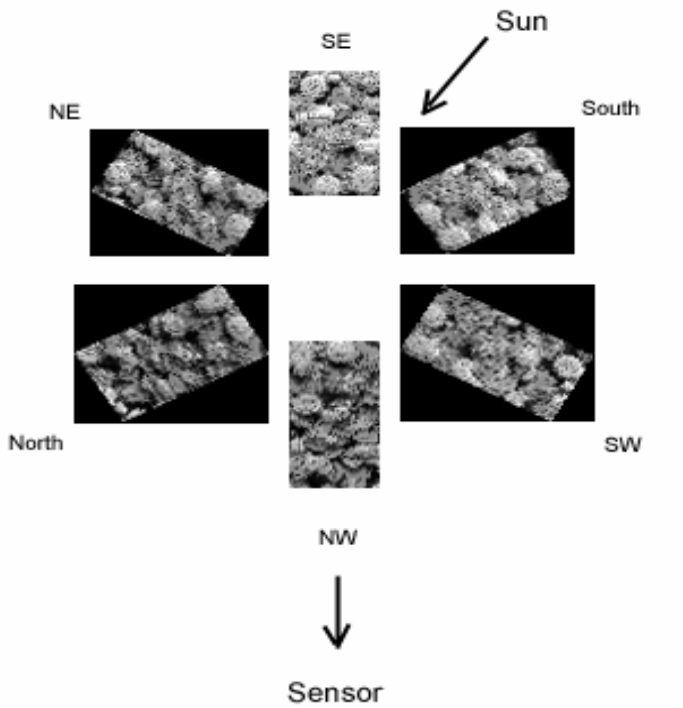
Besides, according to table 4.6, plots of the pine forest on flat versus steep terrain are separable with BD (flat versus slope BD is greater than intra-class BD) but are not separable with ED (flat versus slope ED is lower than intra-class ED). Finally, neither ED nor BD separated the tropical forest plots on flat versus steep terrains. The apparent signal of the tropical forest plots on the IKONOS imagery (see figure 4.4 (a) and (e)) is indeed very similar, compared to signals within other forest types (figures 4.2 and 3). The difference in reflectance from the earth surface due to topography can be attenuated by the forest cover and this phenomenon hampers uniform topographic corrections of satellite imagery on surfaces with varying land cover. In this case, the structure and optical properties of tropical forest stands on flat and steep terrains were found more similar than those within other forest types (tables 4.2 and 4.3).

7.2 Separability study of forest types on varying terrain condition

Once the relative similarity of DART simulations with IKONOS real imagery was established, the next step was to investigate the separability between forest types on a given image, in different terrain conditions. To this end, figures 4.6 - 4.8 illustrate the set of simulations of each forest type, on slopes with comparable aspects.

We also investigated if the aspect of the slope (exposition towards the East) could make the tropical forest plot appear similar to the one on flat terrain, as found above, whereas the aspect of the oak (exposition towards the North-East) and pine (exposition towards the South) forest plots on slope made them appear quite different from the plots on flat terrain. Simulations of forest plots on flat terrain were therefore carried out in the same 'Tancítaro' setting as for forests on slope.

Figure 4.6 – Near Infrared imagery of the oak forest plot on slope



Images of the oak forest plot on steep terrain (18 degrees slope) simulated in the NIR band at 1 meter resolution, sensor plane for 6 expositions (0, 60, 120, 180, 240, and 270 degrees off of South). Sun position and view zenith angle are those of the Tancitaro IKONOS imagery. The same radiometric enhancement was applied to each image.

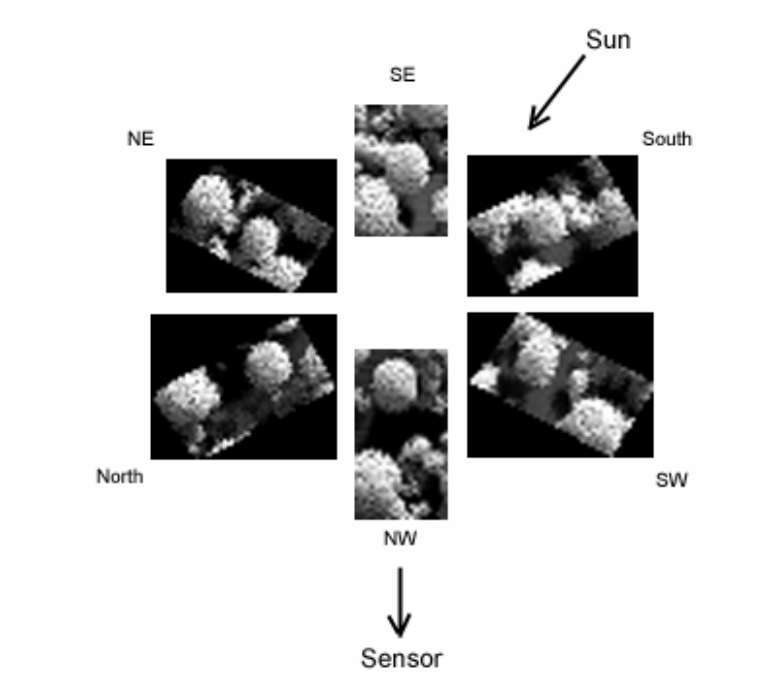
Table 4.7 shows mean reflectance values for all simulations. Note that differences in structure and in understory reflectance between forests on flat terrain and forests on slope may explain some great differences observed in values on the first row (flat terrain conditions) compared to other rows. Unlike in the other spectral bands, a general tendency is clearly perceived for NIR reflectance variation in function of aspect: forests exposed towards the sun (South-Southeast) reflected more NIR light towards the IKONOS sensor, than forests exposed away (North-Northwest) from sunlight direction. NIR mean reflectance varies by a maximum of 4.8, 3.2, and 4.5% for, respectively, oak, pine and tropical forests from sunlit to shadow prone configurations. This variation is not negligible with respect to intra-class variations registered in table 4.6. Thus, it is therefore likely that slope aspect has an effect on separability between classes.

Table 4.8: Inter and intra-class distances between forests on flat terrain and on slope for six illumination aspects. Environmental parameters were set to the Tancítaro area.

	Inter-class:	**Flat**	**South**	**South East**	**South West**	**North East**	**North**	**North West**
Euclidian Distance[a]	Trop.[b] slope vs flat		0.5	0.4	0.4	0.9	1.3	1.1
	Pine vs Oak	*2.1**	1.6	*2.4**	*1.6**	*1.6**	1.3	*2.1**
	Oak vs Tropical	*2.4**	0.9	1.5	0.9	0.9	0.5	*1.6**
	Pine vs Tropical	1.0	1.2	1.1	1.1	0.7	1.0	0.7
Bhattacharyya Distance	Trop.[b] slope vs flat		0.314	0.208	*0.262**	*0.362**	*0.662**	*0.573**
	Pine vs Oak	*1.707**	*1.684**	*3.616**	*1.446**	*2.034**	*2.887**	*3.568**
	Oak vs Tropical	*0.786**	0.369	*1.404**	*0.848**	*0.848**	*0.486**	*1.226**
	Pine vs Tropical	*0.255**	0.237	0.276	*0.334**	*0.366**	0.170	0.095
	Intra-class :	**Flat**	**South**	**South East**	**South West**	**North East**	**North**	**North West**
Euclidian Distance[a]	Oak	1.3	1.8	1.6	1.5	1.2	1.2	1.3
	Pine	1.5	0.9	1.2	0.9	1.0	1.4	0.8
	Tropical	2.1	2.2	2.4	1.7	1.8	1.6	1.6
Bhattacharyya Distance	Oak	0.273	0.376	0.354	0.342	0.281	0.132	0.201
	Pine	0.132	0.146	0.284	0.283	0.256	0.295	0.103
	Tropical	0.209	0.320	0.328	0.169	0.177	0.133	0.177

* Case where classes are separable (inter-class distance is greater than intra-class distances).
[a]Euclidian distances are displayed in percent reflectance.
[b]Tropical. 'vs' stands for 'versus'.

Figure 4.7: Near Infrared imagery of the pine forest plot on slope

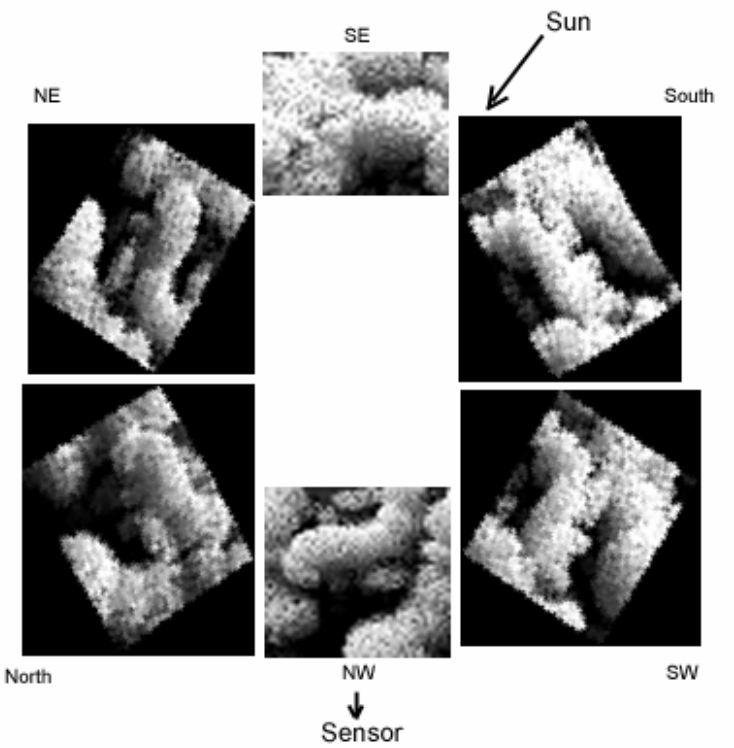


Images of the pine forest plot on steep terrain (27 degrees slope) simulated in the NIR band, at 1 meter resolution, sensor plane for 6 expositions (0, 60, 120, 180, 240, and 270 degrees off of South). Sun position and view zenith angle are those of the Tancitaro IKONOS imagery. The same radiometric enhancement was applied to each image.

Table 4.8 displays inter and intra-class ED and BD scores between forests on slope. Starred figures indicate cases where classes are separable (i.e. when inter-class distance is found larger than intra-class distances). Only few cases resulted separable using the ED distance. The Bhattacharyya distance, since it takes account of pixel value distribution across bands and not only of mean values independently in each band, better discriminates classes than the simpler minimum distance strategy; for example, we observed that mean reflectance in the NIR easily vary by several percents (large ED, see figure 4.5 (c-f)) within a forest class (see table 4.6), while the histogram roughly keeps the same shape (small BD). This explained why our study based on ED was much less discriminative than BD for forest type classification.

However, pine and tropical forests, visually relatively distinct (figures 4.7 and 4.8), were not separated either by BD in four aspects out of 6. A possible explanation is that the presence of shadows (e.g. pine and tropical forests on slope, in figures 4.7 and 4.8) creates a secondary peak (mode) in the histogram, whose shape tends to diverge from the Gaussian bell shape. In such case, the statistics expressed in BD are less robust at describing the characteristic pixel value distribution of the forest type, and intra-class distances are artificially high. The Gaussian assumption is a known weakness of the Bhattacharyya distance, since it was developed as a statistical measure between two normal distributions (Kailath, 1967).

Figure 4.8 Near Infrared imagery of tropical forest plot on slope



Images of the tropical forest plot on steep terrain (30 degrees slope) simulated in the NIR band, at 1m resolution, sensor plane for 6 expositions (0, 60, 120, 180, 240, and 270 degrees off of South). Sun position and view zenith angle are those of the Tancitaro IKONOS imagery. The same radiometric enhancement was applied to each image.

An alternative distance (e.g., histogram-based distance, such as the distance between Grey-Level based density functions used in Xu et al., 2003) might be recommendable for the distinction of pine and tropical forest plots with the presence of pronounced shadows, such as in our case. By contrast, the oak forest type could generally be separated from the other forest types, presumably because its structure contained a limited shadowed area in all aspects compared to other forest types.

Finally, according to table 4.8, the tropical forest exposed South and Southeast was inseparable from tropical forest on flat terrain, an observation which converges with the result observed in table 4.6 for the real IKONOS case (exposition to the East). However, BD could discriminate tropical forests on slope from tropical forest on flat terrain for all of the other aspects studied, confirming that tropical forest does not appear similar in all slope aspects, and that East exposition (our real case) is only a particular case where image similarity was high.

8. Discussion

The phenomenon of apparent dissimilarity among forests because of changing topography incites classification strategies on topographically corrected imagery or where each forest type is divided into topography-induced sub-classes. In the absence of Digital Elevation Model (DEM) at the resolution of hyper-spatial imagery (our case), this approach provides a means of predicting confusions of automatic classification results. In case the IKONOS imagery were topographically corrected, this method of prediction is also useful and directly applicable: a topographic correction normalizes reflectance values of pixels on rugged terrain, with a given aspect and slope, to pixel reflectance values on flat terrain. Consequently, the reflectance of all pixels on a given aspect and slope (the case of our simulated forest plots) is modified by the same multiplicative coefficient. Hence, the above separability assessment could be repeated with the displayed reflectance distributions, augmented by a common multiplicative factor dependent on the correction type.

Topographic corrections can be efficient with 20-30m resolution imagery (e.g. Riaño et al., 2003; Dorren et al., 2003) for the classification of forests that are spectrally

distinguishable by the sensor in flat terrain conditions. Fewer studies exist at hyper-spatial resolution (e.g. Nichol et al., 2006; Svoray & Carmel, 2005), and are successful for vegetation types physiognomically more distant than in our study. Further research should consist in comparing spectral separability derived from the corrected reflectance distributions and results from such classifications.

The poor discrimination power of ED was predictable, since studies on forest classification have not relied on minimum distance between mean pixel values, but instead on maximum likelihood (e.g. Wang et al., 2004) or more sophisticated strategies. Besides, the types of forest at hand were quite distinct from one another. Types of forests which are spectrally closer should be tested with the same methodology. In case Bhattacharyya distance fails to distinguish them, distances of more sophisticated (contextual) classifiers could be tested. In this sense, further developments of this methodology could include the use of optimal neighbourhoods rather than the full set of pixels in the plot. The scope of this study could be broadened through the control over other parameters such as optical properties of woody elements and leaves, which would require extensive laboratory measurements, and pixel resolution, which would require surveys over larger plots with lower resolutions.

Our simplifying assumptions on forest canopy structure (e.g. crowns modelled by ellipsoid shapes) limited the resemblance of DART-simulated imagery with real IKONOS imagery at 1m resolution, especially for tropical forest (see differences in the appearance of emergent crowns in figure 4.4 (b) and (d)). In spite of this limitation, this study proved that DART simulation could approach the 4m real signal (multi spectral IKONOS image) more accurately than intra-class signal variation, even for tropical forest stands. Developments towards the modelling of more realistic and diverse tree elements are currently under construction in CESBIO, and could serve the purpose of approaching the 1m-resolution texture of IKONOS imagery, also used in forest type classification (Wang et al., 2004).

9. Summary and conclusion

Most forest classifications report results on smooth topography. Effects such as the shadowed area or the reflectance of the understory are likely to make forests on slope

appear drastically different from forests of the same type on flat terrain, rendering image-based classification tasks rather puzzling and unpredictable. This research sets a model-based methodology for studying spectral similarity between forests of the same type on flat terrain and on slope, and among forests on slopes of the same aspect. The methodology was applied to IKONOS imagery on three cases pertaining to common forest biomes of Mexico. First, the DART model was tested against the real IKONOS image on each forest plot on slope, where DART simulation was not calibrated on IKONOS for leaf optical properties. The separation of DART images from IKONOS images was found within intra-class variability in the three forest types. Consequently, DART was used as a proxy for IKONOS imagery for evaluating separability of forest classes, simulated in the same conditions of slope aspect.

As expected, very little separability was observed in general with Euclidian distance at 25m resolution since the minimum distance strategy, confronted with very high intra-class variations, is not used for forest classification but rather for more general land cover classification. Instead, in most cases, the very distinct forest types were found separable by the Bhattacharyya distance.

Low LAI oak forests appeared very dissimilar on flat versus on slope scenarios, owing to the strong difference in understory appearance. Unlike plots of other forest types, the plots of high tropical forest on steep slope and on flat terrain appeared spectrally inseparable on real IKONOS imagery. However, when changing the aspect of the plot on slope, the same plots became spectrally distinct for the Bhattacharyya distance. Conversely, pine and tropical forests, spectrally separable on flat terrain, were inseparable on four of the 6 aspects studied. The presence of pronounced shadows on imagery of both forest types on slope, as well as high intra-class structural variability, may partly explain the incapacity of the Bhattacharyya distance to distinguish them, revealing a possible weakness of the maximum likelihood classifier in rugged terrain scenarios. The oak forest type showed spectrally distinct from other forests in most cases, probably because the type of oak forest in the study was structurally very different, less prone to topography-induced shadows, from the other two forests.

Considering the above results, the same methodology could be used for predicting confusions in forest maps of mountainous areas, in function of terrain conditions and

classifying method. The prediction is directly applicable to satellite imagery corrected from topography, since the shape of the reflectance distribution is conserved through the correction. It should be confronted with real classification results of topographically corrected imagery on specific aspects and slopes in the future. This effort should be of interest to bio-diversity estimates since the major deforestation schemes in the sub-tropical belt have already occurred on easily accessible land, characterized by smooth topography. Further developments of DART, some under way, are needed to meet very high resolution requirements for texture rendering of forest structure.

10. Acknowledgments

This research and the acquisition of IKONOS imagery were funded by CONACYT (the national science and technology board in Mexico) under project number 38965T. Field work was carried out with a local team composed of biologists, botanists and technical foresters in the community of San Juan Parangaricutiro, Tancítaro, in the ejido village of La Perla de San Martin, Los Tuxtlas, and in the El Cerro village near to the Cuitzeo lake. For the completion of this work, special thanks are due to Isela Zermeño Hernández, Moisés Méndez Toribio, Rafael Pompa Vargas and Alfredo Patiño Siciliano. We are indebted to the entire support team involved in the development of the DART model, as well as in the development of its SOILSPECT inversion procedure.

11. References

ASTER, Spectral database, NASA/JPL, http://speclib.jpg.nasa.gov, 2003

Baban, S.M.J. and W.Y. Kamaruzaman, Mapping land use/cover distribution on a mountainous tropical island using remote sensing and GIS, *International Journal of Remote Sensing*, 22 (10), 1909-1918, 2001.

Berk A., Bernstein L.S., and D.C. Robertson, MODTRAN: A moderate resolution model for LOWTRAN7, GL-TR-89-0122, Geophysical Laboratory, Bedford, MA, USA, 38pp., 1989.

Chen, J.M., and Cihlar, J., Retrieving leaf area index of boreal conifer forests using Landsat TM images, *Remote Sensing of Environment*, 55, 153-162, 1996.

Clark M.L., Roberts D.A., and D.B. Clark, Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales, *Remote Sensing of Environment*, 96, 375-398, 2005.

Dorren, L.K.A., Maier B., and A.C. Seijmonsbergen, Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification, *Forest Ecology and Management*, 183, 31-46, 2003.

Gascon, F., J.P. Gastellu-Etchegorry, Lefèvre-Fonollosa M.J., and E. Dufrêne, Retrieval of forest biophysical variables by inverting a 3-D radiative transfer model and using high and very high resolution imagery, *International Journal of Remote Sensing*, 25 (12), 5601-5616, 2004.

Gastellu-Etchegorry, J.P., Guillevic P., Zagolski F., Demarez V., Trichon V., Deering D., and M. Leroy, Modelling BRF and radiation regime of boreal and tropical forests: I. BRF, *Remote Sensing of Environment*, 68, 281-316, 1999.

Gastellu-Etchegorry, J.P., Martin E., and F. Gascon, DART: a 3-D model for simulating satellite images and surface radiation budget, *International Journal of Remote Sensing*, 25(1), 75-96, 2004.

Gastellu-Etchegorry, J.P., Hapke handbook, CESBIO, Toulouse, France (http://www.cesbio .ups-tlse.fr), 15pp., 2006.

Helmer, E.H., Brown S., and W.B. Cohen, Mapping montane tropical forest successional stage and land use with multi-date Landsat imagery, *International Journal of Remote Sensing*, 21 (11), 2163-2183, 2000.

Jacquemoud, S., Baret F., and J.F. Hanocq, Modelling spectral and bi-directional soil reflectance, *Remote Sensing of Environment*, 41, 123-132, 1992.

Jonckheere, I., Fleck S., Nackaerts K., Muys B., Coppin P., Weiss M., and F. Baret, Review of methods for in situ leaf area index determination, Part I. Theories, sensors and hemispherical photography, *Agricultural and Forest Meteorology*, 121, 19-35, 2004.

Kailath, T., The divergence and Bhattacharyya distance measures in signal selection, *IEEE Transactions in Communication Theory*, 15, 152-160, 1967.

Landgrebe, D.A., Information extraction principles and methods for multispectral and hyperspectral image data, in *Information Processing for Remote Sensing, Chap. 1*, edited by C.H. Chen, World Scientific Publishing, River Edge, New Jersey, 2000.

Lu, D., Moran E., and M. Batistella, Linear mixture model applied to Amazonian vegetation classification, *Remote Sensing of Environment*, 87, 456-469, 2003.

Magnussen, S., Boudewyn P., and M. Wulder, Contextual classification of Landsat TM images to forest inventory cover types, *International Journal of Remote Sensing*, 25 (12), 2421-2440, 2004.

Mas, J.F., Velázquez A., Palacio-Prieto J.L., Bocco G., Peralta A., Prado J., Assessing forest resources in México: Wall-to-wall land use/ cover mapping, *Photogrammetric Engineering and Remote Sensing*, 68 (10), 966-969, 2002.

Nichol, J., Hang L.K., and W.M. Sing, Empirical correction of low Sun angle images in steeply sloping terrain: a slope-matching technique, *International Journal of Remote Sensing*, 27 (3-4), 629-635, 2006.

Pinty, B., Verstraete M.M., and N. Gobron, The effect of soil anisotropy on the radiance field emerging from vegetation canopies, *Geophysical Research Letters*, 25 (6), 797-800, 1998.

Potter E., Wood J., and C. Nicholl, Sunscan Canopy Analysis System, User Manual v1.05, Delta-T Devices Ltd, Cambridge, UK (http://www.delta-t.co.uk/manual.html), 1996.

Remmel, T.K., Csillag F., Mitchell S., and M. Wulder, Integration of forest inventory and satellite imagery: a Canadian status assessment and research issues, *Forest Ecology and Management* 207, 405-428, 2005.

Riaño, D., Chuvieco E., Salas J., and I. Aguado, Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types, *IEEE Transactions on Geoscience and Remote Sensing*, 41 (5), 1056-1061, 2003.

Salovaara, K.J., Thessler S., Malik R.N., and H. Tuomisto, Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery, *Remote Sensing of Environment*, 97, 39-51, 2005.

Svoray, T., and Y. Carmel, Empirical method for topographic correction in aerial photographs, *IEEE Geoscience and Remote Sensing Letters*, 2 (2), 2005.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S., Legg, C., and M.J. De Dieu, Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests, *Remote Sensing of Environment*, 90, 23-43, 2004.

Wang, L., Sousa W.P., Gong P., and G.S. Biging, Comparison of IKONOS and Quickbird images for mapping mangrove species on the Caribbean coast of Panama, *Remote Sensing of Environment*, 91, 432-440, 2004.

Wickham, J.D., Stehman S.V., Smith J.H., and L. Yang, Thematic accuracy of the 1992 National Land-Cover Data for the western United-States, *Remote Sensing of Environment*, 91, 452-468, 2004.

Xu, B., Gong P., Seto E., and R. Spear, Comparison of gray-level reduction and different texture spectrum encoding methods for land-use classification using a panchromatic IKONOS image, *Photogrammetric Engineering and Remote Sensing*, 69 (5), 529-536, 2003.

Chapter 5: Performance of Classifiers

A model-based performance test for forest classifiers on remote sensing imagery

Ambiguity between forest types on remote-sensing imagery is a major cause of errors found in accuracy assessments of forest inventory maps. In chapter 4, spectral separability was derived between pairs of stands of three taxonomically very distinct forest types, and the focus was to assess the difficult distinction of forests on steep terrain. In chapter 5, the method presented in chapter 4 is extended to six forest types representing community classes of the National Forest Inventory (NFI) map.

The spectral separability index is also used in the evaluation of the performance of sophisticated, parametric, automatic classifiers, which are sometimes presented as improving the results obtained by more conventional classifiers. Chapter 5 presents the method, based on forest plot inventory, ground measurements and simulated imagery, for systematically quantifying ambiguities in the sense of the minimum distance (MD), maximum likelihood (ML), and frequency-based (FB) classifiers.

In chapter 5, the method is tested with multi-spectral IKONOS images acquired on areas containing six major communities (oak, pine, fir, primary and secondary high tropical forests, and avocado plantation) of the NFI map in Mexico. As in chapter 4, Intra-class signal variation was modelled using the Discrete Anisotropic Radiative Transfer (DART) simulator of remote sensing images. Atmospheric conditions were inferred from ground measurements on reference surfaces and leaf optical properties of each forest type were derived from the IKONOS forest signal.

Next, all forest types were simulated, using a common environmental configuration, in order to quantify similarity among all forest types, according to MD, ML and FB classifiers. Classes were considered ambiguous when their dissimilarity was smaller than intra-class signal variation.

A model-based performance test for forest classifiers on remote sensing imagery[Θ]

Stéphane Couturier
Centre d'Etudes Spatiale de la Biosphère (CESBIO), CNES/CNRS/Paul Sabatier University
Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM)
Jean-Philippe Gastellu-Etchegorry and Emmanuel Martin
Centre d'Etudes Spatiale de la Biosphère (CESBIO), CNES/CNRS/Paul Sabatier University
Pavka Patiño
Centro de Investigación en Ecología (CIECO), UNAM

Ambiguity between forest types on remote-sensing imagery is a major cause of errors found in accuracy assessments of forest inventory maps. This paper presents a methodology, based on forest plot inventory, ground measurements and simulated imagery, for systematically quantifying these ambiguities in the sense of the minimum distance (MD), maximum likelihood (ML), and frequency-based (FB) classifiers. The method is tested with multi-spectral IKONOS images acquired on areas containing six major communities (oak, pine, fir, primary and secondary high tropical forests, and avocado plantation) of the National Forest Inventory (NFI) map in Mexico. Intra-class signal variation was modelled using the Discrete Anisotropic Radiative Transfer (DART) simulator of remote sensing images. Classes were considered ambiguous when their dissimilarity was smaller than intra-class signal variation The oak forest and the secondary tropical forest were both distinguishable from all other classes using an MD classifier at 25m resolution, whereas pine and primary tropical forests were ambiguous with three other classes using MD. By contrast, only two pairs of classes were found ambiguous for the ML classifier and only one for the FB classifier. The avocado plantation was confounded with the primary tropical forest for all classifiers, presumably because the reflectance of both types of forest is governed by a deep canopy and a similar shadow area. We confronted the results of this study with the confusion matrix from the accuracy assessment of the NFI map.

[Θ] Submitted to the refereed journal *Forest Ecology and Management*, Elsevier Science

1. Introduction

The classification of remote sensing images for forest cartography is essential to regional biodiversity mapping. Yet, because of the heterogeneity of forest settings, the distinction between forest types remains a difficult challenge. The application of common automatic classifiers (e.g. ISODATA, K-means) and visual labelling of the resulting unsupervised clusters still seem, at least until recently, a widely used strategy in operational forest mapping programs at regional or national scales (i.e. Benjamin et al., 1996; Vogelmann et al., 2001; Wulder et al., 2003). Comprehensive or partial accuracy assessments of these maps were achieved (Laba et al., 2002; Wickham et al., 2004; Remmel et al., 2005). At high taxonomic resolution, confidence levels reported for many forest classes were lower than those reported in other studies on the same type of imagery with more sophisticated classifiers (e.g. Magnussen et al., 2004; Salovaara et al., 2005; Lu et al., 2003). However, the latter studies occurred in more confined, more homogeneous geographical areas and concerned a smaller number of classes. In some of these case studies, classifiers with ancillary information on forest stand structure yield better accuracy than the one obtained with classifiers that do not incorporate such information. For example, Spectral Mixture Analysis (SMA) uses spectral end-members (shadow or sunlit fractions) that contain sub-pixel structural information, and SMA yielded accuracy improvements for the classification of successional stages of the Amazonian tropical forest (Lu et al., 2003). Moreover, Lu (2005) found that a more detailed structural information (the measurement of tree height distribution on the ground), was highly correlated with a combination of Landsat TM bands; a linear regression of these quantities further improved the accuracy of the classification. A drawback of these approaches, however, is again their limited transferability to broader scales. Indeed, the relationships that linked forest stand structural information to the reflectance data were empirically derived and only valid for the environment close to the image at hand (Lu, 2005). From these observations, a set of questions may be posed by the forest map producer: Would any more sophisticated classifier perform better than the common automatic classifiers at regional or national scales, for a wider set of classes? Or: For which set of classes is it worthwhile to look for better, more sophisticated, algorithms instead of settling for the commonly used (or available) classifiers?

Considering the above-mentioned improvements in accuracy, one way to address these questions could be to refer to an expert system, based on the description of the structure of forest stands, capable of estimating *a priori* ambiguities among classes for a given classifier. Indeed, high ambiguity between two forest types for a few conventional classifiers would encourage the forest map producer to look for a more sophisticated classifier. Such result could even pose the question of whether it is suitable to classify these forest types via automatic classification or if it is preferable to use visual classification instead, a strategy employed for example in the case of the National Forest Inventory (NFI) in Mexico (Mas et al., 2002; Mas and Ramírez, 1996). Conversely, low ambiguity between two forest types for a given classification method would ensure the appropriateness of the classification method and stop the quest for a better classifier (e.g. Baban and Kamaruzaman, 2001). It is recently argued indeed (Wilkinson, 2005) that more research efforts should be dedicated to improving other areas of the map production process (Fassnacht et al., 2006) rather than mainly focusing on better classification algorithms.

In this research, we propose a framework based on forest plot inventory, ground measurements and simulated imagery, in order to test a few classifiers' *a priori* performance on pairs of classes among a given set of forest types and for any given remote sensing platform. This framework is applied to the case of 6 forest types pertaining to classes at community level (Palacio et al., 2000) of the Mexican NFI map (see Mas et al., 2002), namely oak, pine, fir, primary and secondary high tropical forests, and avocado plantation.

2. Background

3D-modeling of forest plots offers various advantages to foresters. For example, it allows tree mapping based on plot inventory, which enables the precise estimation of key structural indices of the forest such as canopy gap distribution (Silbernagel and Moeur, 2001). Additionally, radiative transfer coupled with 3D modeling has permitted the evaluation of radiation budgets of forest parcels and intercepted radiation by tree crowns (e.g. Gastellu-Etchegorry & Trichon, 1998; Courbaud et al., 2003). In terms of bi-directional reflectance, the comparison between simulated and real imagery of forests is a difficult task. At scales well above tree crown, the average signal is dominated by the

macroscopic properties of illuminated and shadowed crown and ground components. In this case, the average reflectance of homogeneous stands is simulated with forward models of canopy scattering (e.g. Goel & Thompson, 2000; Pinty & Verstraete, 1991; Strahler, 1996; North, 1996). This approach has been successfully assessed against real imagery in mainly coniferous stands made of repeated individuals of one or two species (Gemmel and Varjo, 1999; Courbaud et al., 2003; Disney et al., 2006). No such exercise was achieved for scales approaching crown size and with a flexible scheme integrating a greater diversity of tree species and heterogeneous understory spectral signatures.

The 3D Discrete Anisotropic Radiative Transfer (DART*) ray-tracing model (Gastellu-Etchegorry et al., 2004) simulates remote-sensing images of heterogeneous natural and urban landscapes, using 3D generic representations of these landscapes. This simulator has been successfully tested against reflectance results of other radiative transfer models (Pinty et al., 2004). Gastellu-Etchegorry and Trichon (1998) stress the difficulty in accurately positioning a simulated plot on a 25 m resolution imagery. Simulating remote sensing imagery and not only average reflectance allows to work at resolutions close to crown scales. At very high (1-5m) resolution, more recently available, forest plots are more easily identifiable on the image, and the amount of ground measurements necessary for statistically meaningful comparisons between simulated and real pixel sets is at least $52 = 25$ times less extensive than for 25m resolution imagery, in which a one hectare plot contains only 16 pixels. Multi-spectral IKONOS imagery (4m resolution) has recently been used for the classification of taxonomically close forests (e.g. Wang et al., 2004, Thenkabail et al., 2004).

We used the DART model to simulate the inherent heterogeneity of forest stands and the associated IKONOS multi-spectral remote sensing images. This paper describes an approach that is essentially generic, and examines the utility of a 3D structural model with general assumptions on the canopy, in order to test the correspondence of such general model with the structural information given by 4m multi-spectral imagery of a large variety of forest types. Our goal was to study ambiguities among these forest types on

* DART is a patented model that is freely available for scientific applications at www.cesbio.cnes.fr.

remote sensing imagery in a systematic manner. For this purpose, a set of parameters, measurable in the field and in the laboratory, was handled as variables, while environmental heterogeneity exterior to the forest setting (viewing and illumination configurations, sensor response and atmospheric scattering) was controlled and fixed.

Clusters from commonly available unsupervised classifiers such as ISODATA or K-means are organized around the minimum distance (MD) principle. Another conventional parametric method is the maximum likelihood (ML) classifier. MD and ML principles were selected for the ambiguity estimate concerning relatively common classifiers. Among more sophisticated methods, the frequency-based (FB) classifiers have been successfully used for contextual image classification of very high resolution imagery (Lira and Maletti, 2002; Xu et al., 2003). The first order FB classifier is built on the quantitative comparison of histograms - or density functions - of pixel neighbourhoods (Maletti et al., 2002).

A classifier was considered to perform well at distinguishing a pair of classes if the spectral dissimilarity between images of simulated plots of these classes (or 'inter-class variation') was higher than the spectral dissimilarity among simulated plots within a given class (or 'intra-class variation'). In the opposite case, i.e. if inter-class variation was lower than intra-class variation, the classes were considered ambiguous for this classifier. The spectral distance characterizing each classifier is defined in the section on classifiers' dissimilarity measures. For example, the Euclidian distance between mean spectral values defines spectral dissimilarity for the MD criterion.

In this work, a botanic inventory, including a structural record of the canopy, was performed on 6 forest plots (one per forest type) appearing on 3 IKONOS scenes (multi-spectral and panchromatic). Optically-based measurements were conducted in order to estimate the Leaf Area Index (LAI) and the reflectance of the understory in each plot. These measurements and environmental parameters matching conditions of the IKONOS overpasses served as inputs to the DART model for the simulation of the forest plots. The comparison of each simulated image against its corresponding IKONOS scene allowed us to account for atmospheric effects and to compute the leaf optical properties of each forest type through an imagery inversion process, as indicated in the section on satellite imagery.

In order to test the potential of DART for simulating the 4m-resolution real IKONOS imagery of the forest plots, the histograms of simulated and real imagery were compared.

In a second step, for ambiguity assessments, all forest plots were simulated under acquisition and environmental conditions of one IKONOS scene. The variability of measurements within a given forest type was assessed on the ground and images of plots of the same forest type were generated from simulations incorporating this variability. Inter-class and intra-class variations were derived and pair-wise comparisons were made for all combinations of the 6 forest types. The aim of this exercise was to enlist, for each of the three classifying methods, pairs of classes that are *a priori* ambiguous and pairs of classes that are *a priori* separable. Finally, in the last part of the paper, we discuss the results of this study in the context of the confusions recorded among the same classes, in a previous accuracy assessment of the NFI map.

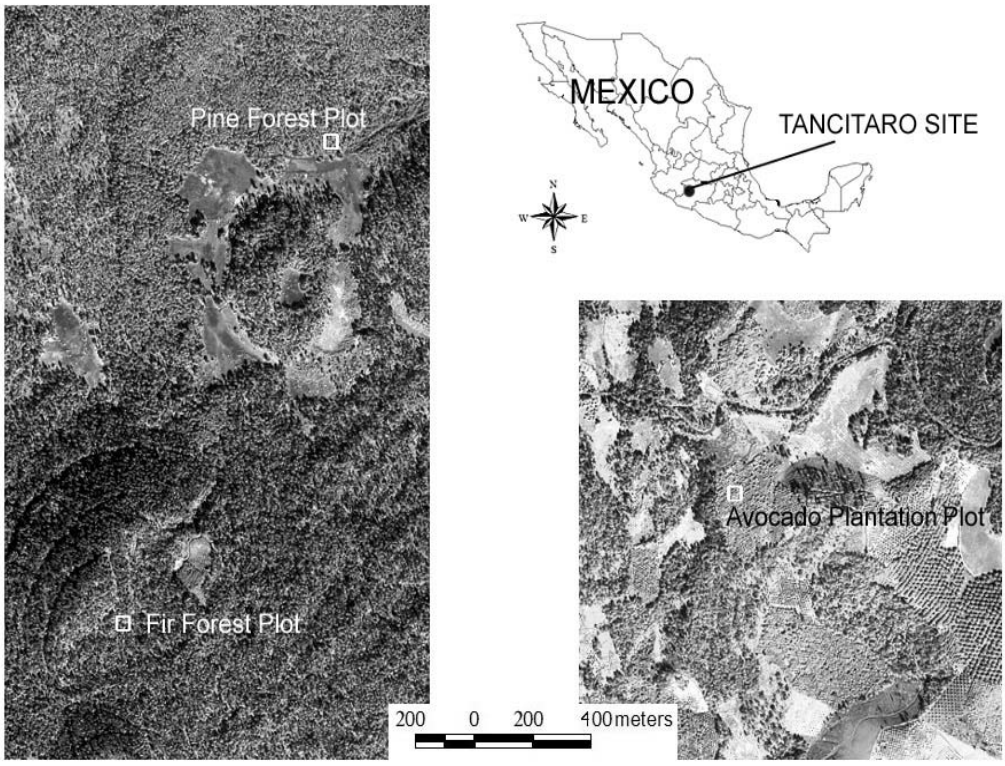
3. Study Areas

An extensive coverage of both temperate and tropical types of forests is reported by the NFI map of mega-diverse Mexico (Mas et al., 2002). Much spatial intricacy occurs among polygons of forest classes at community level, even between parches of temperate and tropical forests (see for example the NFI map of the state of Chiapas), the distinction of which is an important task for biodiversity mapping. We focused on three sites that contain a variety of forest types included in the classification system of the NFI and where IKONOS imagery was available. Two sites, labelled 'Tancítaro' and 'Cuitzeo', are located on the transversal volcanic chain and contiguous altiplano in central western Mexico (see figures 5.1 and 5.2). The 'Tancítaro' site is included in the communal lands of San Juan Parangaricutiro, Michoacán state, near to the natural reserve of the Tancítaro volcano peak. This site is about 2200m above sea level and the climate is close to temperate sub-humid. The 'Cuitzeo' site is located in the ejidal lands of El Cerro, Michoacán state, close to the Cuitzeo lake. This site is about 2050m above sea level and the climate is intermediate between tropical dry and temperate sub-humid. Another site, labelled 'Los Tuxtlas', is situated near to the coast of the Gulf of Mexico (see figure 5.2), on the flanks

of the San Martin volcano at about 900m above sea level. This site is included in the communal lands of La Perla de San Martin, Veracruz state, within the biosphere reserve of Los Tuxtlas. The climate is mainly characterized by tropical humid conditions.

We focused on six forest plots (listed in table 5.1) among these three sites. One plot pertains to an oak forest dominated by *Quercus Deserticola* in the Cuitzeo site; one plot of pine forest dominated by *Pinus Montezumae*, one plot of fir forest dominated by *Abies Religiosa* and one plot of avocado plantation (*Persea Americana*) were selected in the managed forests of Tancítaro; one plot of mature, well-preserved high tropical forest ('primary' tropical forest) and one plot of secondary tropical forest were selected in the Los Tuxtlas site. The secondary forest was regenerating vegetation on a site where

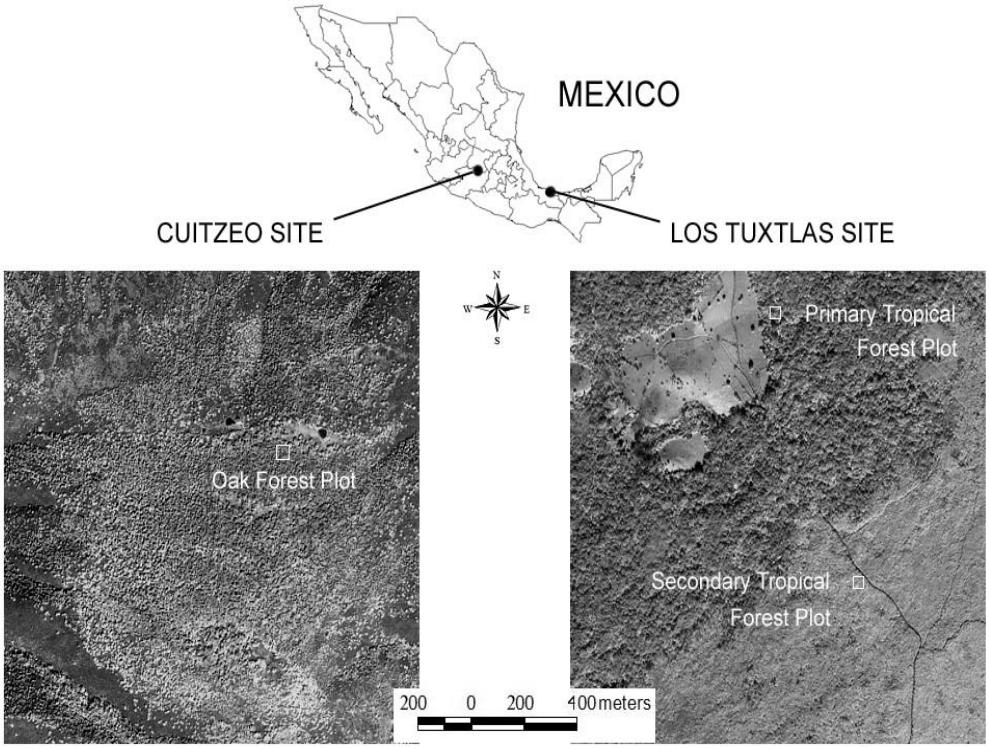
Figure 5.1 Location of forest plots on multi-spectral IKONOS imagery in the Tancítaro site.



The black and white scenes are derived from the RGB composite of bands (red,NIR,green).

primary forest had been depleted by a 1998 fire episode. All plots lay on relatively flat terrain. Thereafter, we denominate 'forest stand' as being a larger portion of forest of the same type, on relatively flat terrain, that includes the forest plot.

Figure 5.2 Location of forest plots in Cuitzeo and Los Tuxtlas sites on multi-spectral IKONOS imagery.



The black and white scenes are derived from the RGB composite of bands (red,NIR,green).

4. Ground Measurements

For each forest class, we proceeded to the survey of one plot and five reference parcels within the same forest stand. The data collected in the reference parcels were used to model the intra-class variation of the signal. A structural record of the canopy was measured in the forest plot. Crown coverage (CC) estimation, LAI measurements and soil reflectance evaluation (see optical measurements) were performed in the plot as well as in the reference parcels.

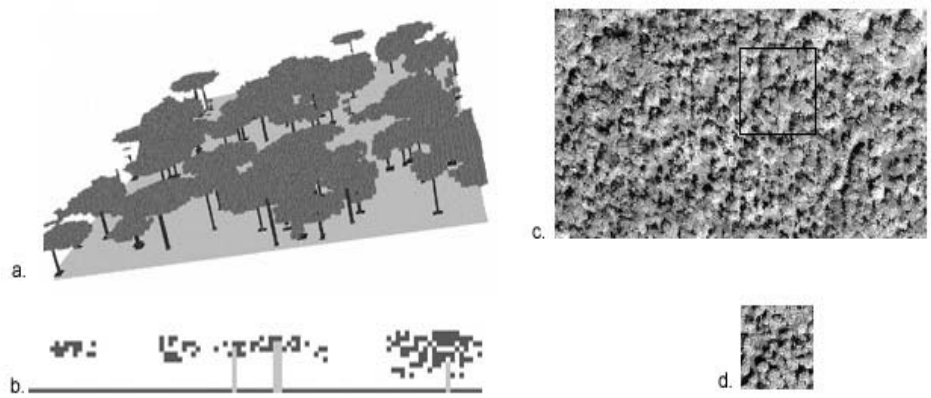
4.1 Forest stand structure

The square-shaped forest plots (see dimensions on table 5.1) were marked with metric tapes to record relative positions of trees. All trees larger than 10cm diameter at breast height (DBH) were inventoried in the plot of primary tropical forest, and larger than 5cm DBH in all other plots. Species were identified (the number of species is mentioned in table 4.1) and crown dimensions were measured assuming an ellipsoidal envelope, the axes of which are aligned to the orientation of the parcel: the extreme horizontal coordinates of the ellipsoidal envelope were visually evaluated on the metric tape using compasses. Vertical dimensions were assessed by the local foresters guiding the team in the field, using reference trees whose height was estimated with a clinometer. In some cases, more than one ellipsoid was required to approximate the crown. Trees with stems outside the plot but whose crown infringed in the plot were also measured and included in the simulation. With the aid of two metric tapes placed perpendicularly across the entire plot (or reference parcel), CC was measured (see table 5.1) as the fraction of the tapes that intersected the projected crowns on the ground.

Table 5.1 Structural parameters of forest plots and stands

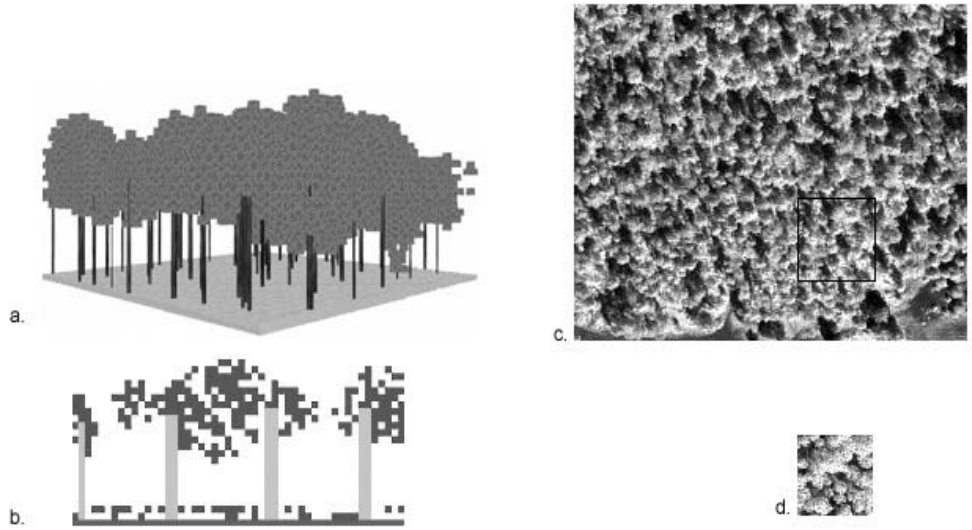
Forest stand	Site	Number of species	Plot size (m)	Crown Cover in plot	Crown Cover in forest stand	LAI in plot	LAI in forest stand Min	Max
Oak	Cuitzeo	12	50 x 50	64%	54-72%	2.5	2.0	3.2
Pine	Tancítaro	3	50 x 50	87%	75-94%	3.5	2.9	4.1
Fir	Tancítaro	6	50 x 50	75%	59-86%	3.9	2.6	5.8
Avocado	Tancítaro	1	50 x 50	77%	72-85%	6.1	5.6	6.4
Trop. Second.	Tuxtlas	21	40 x 40	96%	92-98%	5.1	4.5	6.1
Trop. high	Tuxtlas	31	40 x 40	97%	92-100%	7.9	7.0	8.9

Figure 5.3 Oak forest plot in the Cuitzeo site.



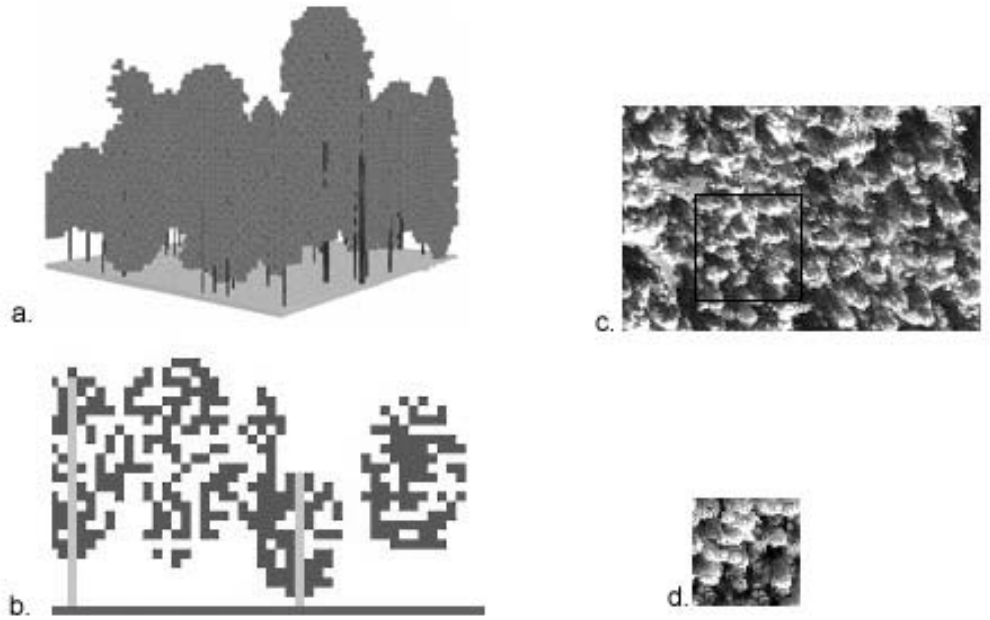
Simulated 3D structure (a) and vertical 2D profile at 50cm resolution (b). Panchromatic IKONOS (c) and DART (d) images of the forest plot. The DART panchromatic image is derived from a fusion of DART multi-spectral images.

Figure 5.4 Pine forest plot in the Tancítaro site.



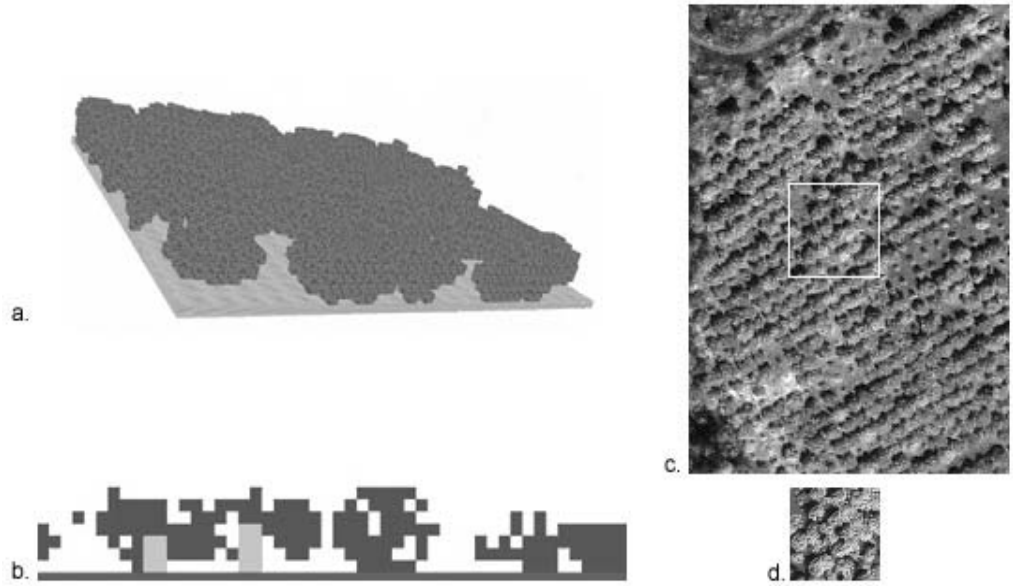
Simulated 3D structure (a) and vertical 2D profile at 1m resolution (b). Panchromatic IKONOS (c) and DART (d) images of the forest plot. The DART panchromatic image is derived from a fusion of DART multi-spectral images.

Figure 5.5 Fir forest plot in the Tancítaro site



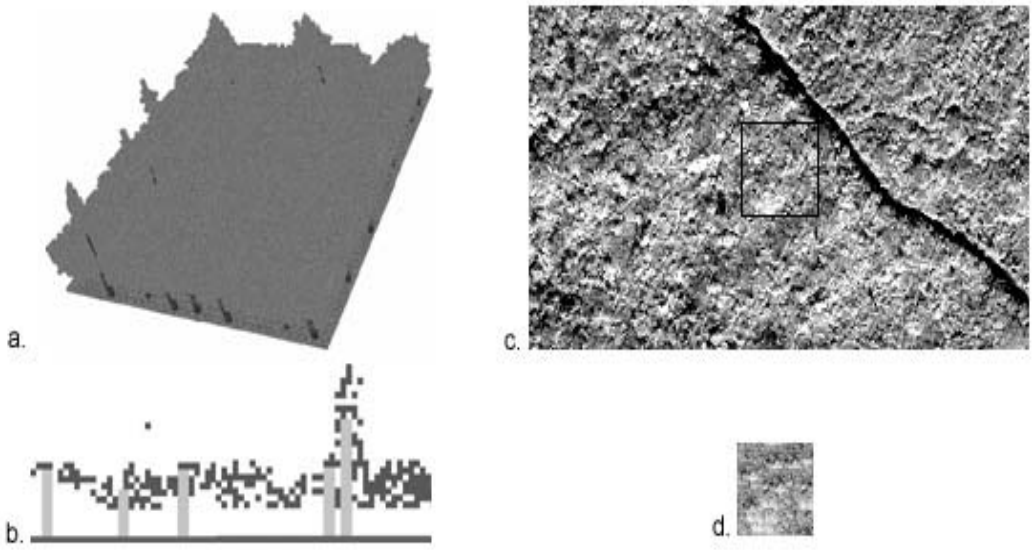
Simulated 3D structure (a) and vertical 2D profile at 1m resolution (b). Panchromatic IKONOS (c) and DART (d) images of the forest plot. The DART panchromatic image is derived from a fusion of DART multi-spectral images.

Figure 5.6 Avocado plantation plot in the Tancítaro site.



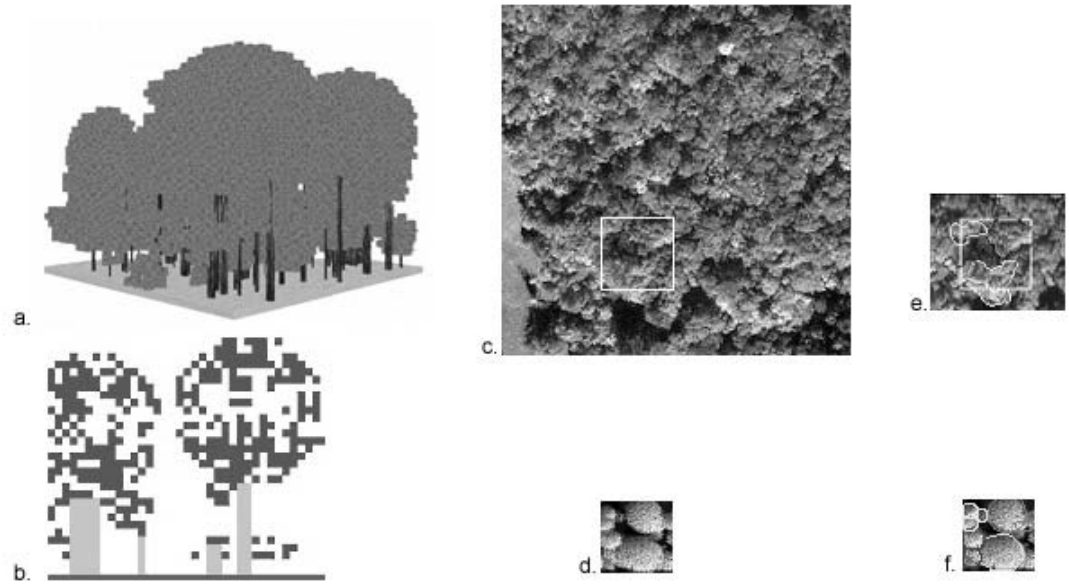
Simulated 3D structure (a) and vertical 2D profile at 1m resolution (b). Panchromatic IKONOS (c) and DART (d) images of the plantation plot. The DART panchromatic image is derived from a fusion of DART multi-spectral images.

Figure 5.7 Plot of the secondary tropical forest in the Los Tuxtlas site.



Simulated 3D structure (a) and vertical 2D profile at 50cm resolution (b). Panchromatic IKONOS (c) and DART (d) images of the forest plot. The DART panchromatic image is derived from a fusion of DART multi-spectral images.

Figure 5.8 Plot of the primary tropical forest in the Los Tuxtlas site.



Simulated 3D structure (a) and vertical 2D profile at 1m resolution (b). Panchromatic IKONOS (c, e) and DART (d, f) images of the forest plot. The DART panchromatic image is derived from a fusion of DART multi-spectral images. Emergent crowns are highlighted on (e, f).

These measurements were synthesized into an ASCII file which was converted by the DART model into a 3D representation based on turbid or empty cubic cells for the canopy, and opaque surfaces for tree trunks (Gastellu-Etchegorry et al., 2004). The 3D structure and vertical profile of the canopy are illustrated with the JAVA 3D interface of the DART model in figures 5.3- 5.8 (a,b).

The satellite imagery was simulated by the DART model, using 3D representations that comprised all measured trees (i.e. an area larger than the plot itself). However, data processing was conducted on the sub-images corresponding to the forest plots. This procedure was adopted to avoid any artificial shadow created by the absence of tree crowns at the border of the plots.

4.2 Optical Measurements

We conducted a series of 16 LAI estimates in the plot and 10 LAI estimates in each reference parcel, using the SUNSCAN Canopy Analysis System (Delta-T Devices Ltd, Cambridge, UK). A linear probe containing 64 photodiodes measures the radiation intercepted by the canopy while at the same time, sensors with and without shade placed in a clearing read diffuse and direct sun radiations. The effective LAI is automatically derived according to the principles and assumptions detailed in Potter et al. (1996). The measurements with the apparatus were made along two 20m-long perpendicular transects within the plot and reference parcels. They were subsequently averaged. These average values are displayed in columns 7 and 8 of table 5.1, respectively.

The measurement campaign occurred in March-April 2005, the season corresponding to the satellite overpass on Cuitzeo, the only site where forest LAI is strongly seasonally-dependent. Since hemispherical photography (a technique which suffers less dependence on sun zenith angle for gap fraction estimation: Jonckheere et al., 2004, p31) was unavailable for our campaign, the emphasis of the study, rather than aim precise absolute LAI values, was to measure coherent LAI values among plots and sites. Therefore, we always realized measurements at a time of the day when the sun was at 35-40^{o} zenith angle, markedly away from noon but within the domain of LAI retrieval by the SUNSCAN software. The LAI estimates in forest plots and in reference parcels (extreme values) are shown in columns 7 and 8 of table 5.1, respectively.

Bare soil, rock and soil covered with different types of litter represented the ground conditions of all forest plots. In the presence of forest plots with sparse canopy or low LAI, an optical characterization of the understory was necessary (Pinty et al., 1998). The direct appraisal of the optical properties in the field was hindered by the heterogeneous ground conditions. Therefore, we first proceeded to a visual description of the understory in terms of quantitative fractions of bio-physical constituents present on the ground over the whole forest plot. The bio-physical constituents encountered in the field were then spectrally related to a reduced set of seven most frequent, conspicuous constituents: a type of humus-rich bare soil, a type of basaltic rock, litter of yellow dried grass, orange-brown leaf litter, dark brown leaf litter, grey (decomposed) leaf and branch litter, and brownish needle litter. These seven constituents were collected in the field and then arranged in the laboratory as opaque surfaces to reproduce spatially homogeneous understory scenarios.

Optical properties (i.e., directional reflectance) of homogeneous understory scenarios were measured in laboratory with illumination zenith angle similar to the ones of the IKONOS overpasses. For this purpose, a series of bi-directional measurements of the ground's upwelling radiance was taken using a 512-channel hand-held spectro-radiometer (Geophysical and Environmental Research Group [GER] 1500) operating at 3nm FWHM over the range of wavelengths from 300nm to 1050nm. The instrument was held 1m above the targeted area. Mean reflectance spectra of homogeneous understories were computed after a normalization to the radiance response of a Labsphere Teflon reference panel, which diffusely reflects incident radiation over all angles.

In a second step, the reflectance of each 'real' understory was computed as a linear combination of the homogeneous understory spectra, with weights being the fractions observed in the field. This computation relies on a linear simplifying hypothesis that was successfully tested with field measurements of understories made of conspicuous patches of two constituents. In order to estimate the variability of soil reflectance in the surrounding forest, we similarly characterized the understory of the 5 reference parcels. The directional reflectance values, convolved to each IKONOS multi-spectral band, were inserted in the Hapke module (Gastellu-Etchegorry, 2006) of the DART model in order to obtain the coefficients of the Hapke (1993) soil model that fits them. These coefficients were used as descriptors of the understory optical properties in the simulation.

5. Satellite imagery

5.1 Image preparation

We acquired panchromatic and multi-spectral IKONOS scenes for each of the three sites (Figures 5.1 and 5.2; Table 5.2).The scenes were provided in 'Standard Geometrically Corrected' mode by Space Imaging. They were subsequently geometrically corrected using GPS measurements.

Table 5.2 Acquisition parameters of IKONOS imagery.

Site	Geometric Coordinates Latitude (°)	Longitude (°)	Date	Local Time	θ_s (°)	φ_s (°)	θ_v (°)	φ_v (°)
Cuitzeo	20.0	-101.1	04/03/2003	10:36	33.2	141.5	14.7	358.9
Tancítaro	19.4	-102.2	20/11/2002	10:47	40.9	162.9	21.9	290.4
LosTuxtlas	18.5	-95.1	29/11/2004	10:04	43.1	157.6	21.8	33.3

Local time is GMT-7:00.

In a first step, the panchromatic image was corrected via identifiable landmarks. We used a GARMIN® GPS unit (WAAS mode) to record tracks around crowns of isolated trees in clearings and along clearing borders / access trails in the vicinity of the forest plots. We then matched the IKONOS imagery to the GPS positions at less than 4m RMS (1 pixel of the multi-spectral image), using a 2nd degree polynomial transformation. Forest plots were then located in the image through visual identification of pre-eminent trees or gaps in the panchromatic image. Alternatively, we used the average of repeated GPS readings inside the plot, using climbing equipment on trees when necessary, until getting less than 3m location accuracy. Figures 5.3 – 5.8 (c) illustrate portions of the panchromatic IKONOS image where the forest plots are located. In a second step, the geometry of the multi-spectral images was corrected using the corresponding panchromatic image as a reference.

The multi-spectral IKONOS raw digital numbers were converted to top-of-atmosphere (TOA) reflectance using the standard solar irradiance formula provided by Space Imaging. No atmosphere measurement was available at IKONOS overpass. Besides, the limited equipment available for this study and the great diversity of leaves/needles present in the forest plots hindered the measurement of their optical properties. In the absence of

available reflectance databases, we used measurements on homogeneous surfaces and in reference parcels of each IKONOS scene in order to infer atmospheric and leaf optical parameters from auxiliary simulations, as explained below.

5.2 Image calibration

Bi-directional measurements were acquired with the GER-1500 spectro-radiometer in at least two recognizable reference surfaces per IKONOS scene: large crossings of asphalted and dry dirt roads in Los Tuxtlas, gravel and asphalt roads in Tancítaro, crossing of asphalt roads and a small quarry site in Cuitzeo. The aerosol optical depth (AOD), and total transmittance of the vapour column (T_{H2O}), two major atmospheric parameters in the spectral bands of interest, were computed with the help of the simulated TOA reflectance of the two reference surfaces, assuming a tropical rural atmospheric model with 23km visibility (Berk et al., 1989).

In order to calibrate the Leaf Optical Properties (LOP) of each forest plot, we considered IKONOS pixels associated with large LAI values in the reference parcels. These pixels were identified in the panchromatic IKONOS image as 8mx8m parcels with lit crowns, 100% CC, and high signal with respect to the surroundings in the co-registered near infrared (NIR) IKONOS scene. The crown structure (average height and thickness of crown) of the trees was estimated in this 8mx8m area. An inversion procedure led to LOP values such that the TOA reflectance value of the simulated lit crowns was equal to the observed IKONOS reflectance values. In this approach, leaf reflectance (both sides) and transmittance values were assumed equal. Moreover, LOP was given a unique spectral value for all trees in the forest plot. These two simplifying assumptions were adopted because they already led to a robust approach for biophysical parameter inversion from very high resolution imagery (Gascon et al., 2004). The shape of tree crowns was assumed ellipsoidal and leaf angle distribution was assumed spherical. Optical properties of tree trunks were taken from the ASTER database (ASTER, 2003).

6. Methods

6.1 Classifiers' dissimilarity measures

Three classifiers were selected for the performance test, namely the Minimum Distance (MD), the Maximum Likelihood (ML) and a first order Frequency-Based (FB) classifier. A key algorithmic step among these parametric classifiers is the minimization of the spectral distance.

The Euclidian distance $ED_{ij} = |\mu_i - \mu_j|$, where μ_i is a vector containing the mean spectral value of pixels in forest plot *i*, measures the spectral separation between classes *i* and *j* for the MD classifier. Because of the high variability of the signal in hyper-spatial imagery of forests (Gascon et al., 2004), we observed that MD discriminated classes with greater success at lower resolutions. Consequently, mean values taken for ED distance calculations were extracted from a 25mx25m subplot (selected as the high resolution reference value of typically available sensors) of the forest plot. This means that MD was tested for an equivalent 25m-resolution sensor rather than for the original 4m-resolution IKONOS imagery.

The Bhattacharyya distance expresses the dissimilarity between two clusters of pixels for the ML classifier (Landgrebe, 2000). Here, The Bhattacharyya distance BD_{ij} between forest plots *i* and *j* seen on multi-spectral IKONOS imagery was calculated according to (Landgrebe, 2000):

$$BD_{ij} = \frac{1}{8}\left[\mu_i - \mu_j\right]^T \left[\frac{\Sigma_i + \Sigma_j}{2}\right]^{-1} \left[\mu_i - \mu_j\right] + \frac{1}{2}\ln\frac{\left|\frac{1}{2}\left[\Sigma_i + \Sigma_j\right]\right|}{\sqrt{\left|\Sigma_i\right|\left|\Sigma_j\right|}} \qquad (5.1)$$

where μ_i contains the mean spectral value of all pixels in forest plot *i*, and Σ_i is the covariance matrix of the vectors corresponding to all pixel values in forest plot *i*.

Finally, the Frequency-Based Distance (FBD) is the dissimilarity measure for the FB classifier. FBD operates, as well as BD, on all individual multi-spectral pixel values inside a forest plot. These multi-spectral values were first grouped in bins to form a multi-spectral histogram. The bin size was set as the sum of the IKONOS and DART noise

levels, in each spectral band. The DART noise level was assessed as the standard deviation of the differences of reflectance values simulated with DART and with a Monte Carlo model, included in the DART model package. The Monte Carlo model is assumed to be exact, but very computer extensive which explains why it is not commonly used. Martin (2006) showed that DART relative accuracy with respect to the Monte Carlo model is better than 2%. FBD_{ij} between forest plots *i* and *j* seen on multi-spectral IKONOS imagery was calculated according to Maletti et al. (2002: p702):

$$FBD_{ij} = \frac{1}{2B}\sum_{b=1}^{B}\sum_{n=1}^{Q}\left|h_i[b,n] - h_j[b,n]\right| \qquad (5.2)$$

where *B* is the number of spectral bands, h_i is the per band normalized histogram of the pixels in forest plot *i*, and *Q* is the number of quantization bins.

6.2 Intra-class and inter-class studies

ED, BD and *FBD* were first computed between DART-simulated and IKONOS images of the six forest plots, yielding the 'DART-IKONOS dissimilarities'. Next, the six forest plots were simulated with varying CC, LAI, and ρ_u (understory reflectance) values in order to account for the variation measured in the reference parcels. For this purpose, we selected for each forest setting the minimum and maximum CC, LAI and ρ_u values that were measured in the associated set of reference parcels. As a result, a set of eight simulations S(CC, LAI, ρ_u) per forest setting was computed, with these three parameters equal to either the minimum or the maximum measured value. For example, S(CC_{min}, LAI_{max}, $\rho_{u\ min}$) is one of these simulations. Distances were computed between all pairs of these eight simulations. The maximum of these distances expressed the 'intra-class dissimilarity' of a forest setting. This definition of intra-class dissimilarity was viewed as a trade off. Indeed, the calculated distances tend to overestimate intra class variability because the impact of the CC, LAI and ρ_u extreme values on intra class variability may be overestimated. On the other hand, they underestimate intra class variability because the variation of other parameters such as leaf optical properties, that could not be measured, were not taken into account. Confronting the DART-IKONOS dissimilarity value with the intra-class dissimilarity was a means of testing DART as a proxy for pixel value distribution of real, very high resolution IKONOS imagery.

As a second step, the same forest plots were simulated, setting viewing orientation, sun position, and atmospheric conditions to the values found in the Tancítaro area. Dissimilarities between pairs of classes, as well as intra-class dissimilarities, were calculated for each forest type, the comparison of which was aimed at determining whether forest classes were ambiguous, or separable.

7. Results

7.1 Intra-class study with simulated imagery

Figures 5.3-5.8 (d) illustrate results of 1m-resolution DART simulations of all forest plots. Table 5.3 reports the computed dissimilarities corresponding to the MD, ML and FB classifiers, for the DART-IKONOS comparison of every forest type, as well as for the intra-class studies. The dissimilarities were computed for all four bands (line denominated 'all') and for each mono-spectral band. In all forests, the DART-IKONOS dissimilarity according to MD is lower than intra-class dissimilarity, indicating a good agreement (within intra-class variation) of mean values between IKONOS imagery and DART-derived imagery. Intra-class variations of mean reflectance are uneven across spectral bands: The NIR band and, to a lesser extent, the green band, generally record larger variations within a forest type than the red and blue band. This observation falls in accordance with the magnitude of absolute reflectance value of green vegetation in those spectral domains.

The similarity of pixel value distributions in DART imagery with respect to IKONOS imagery can be appreciated in figure 5.9 (a-f), where histograms of corresponding scenes in the NIR band (the spectral band with the strongest dynamic range) are displayed. The DART-IKONOS dissimilarity in terms of pixel value distribution in the four dimensional space (all four bands) was also found less than intra-class variability, as attested by the results of ML and FB classifiers in table 5..3

Table 5.3 Spectral dissimilarity $\Delta_{D,I}$ between DART and IKONOS signals, and DART intra class variability σ_D.

		Oak		Pine		Fir		Avocado		Tropical Secondary		Tropical Primary	
	Band	$\Delta_{D,I}$	σ_D	$\Delta_{D,I}$	σ_D	$\Delta_{D,I}$	σ_D	$\Delta_{D,I}$	σ_D	$\Delta_{D,I}$	σ_D	$\Delta_{D,I}$	σ_D
MD	All bands	**0.3**	**1.3**	**0.3**	**1.5**	**0.4**	**1.3**	**0.2**	**1.6**	**0.4**	**1.8**	**0.5**	**1.6**
	NIR	0.2	3.0	0.3	2.4	0.4	1.8	0.0	2.2	0.3	3.2	0.3	4.2
	Red	0.1	1.0	0.0	0.8	0.4	0.8	0.4	0.8	0.3	0.5	0.1	0.5
	Green	0.4	1.0	0.5	2.8	0.1	1.4	0.4	2.0	0.7	2.7	0.6	1.7
	Blue	0.4	1.4	0.5	1.2	0.8	1.1	0.1	1.2	0.2	0.9	0.8	1.0
ML	All bands	**0.104**	**0.328**	**0.097**	**0.132**	**0.058**	**0.138**	**0.048**	**0.081**	**0.129**	**0.257**	**0.223**	**0.259**
	NIR	1.7E-04	3.4E-05	5.7E-05	6.4E-04	2.3E-05	9.1E-04	3.6E-05	2.2E-05	2.5E-04	2.0E-04	5.2E-05	7.8E-06
	Red	3.8E-06	1.3E-04	3.8E-05	2.3E-04	1.2E-04	6.0E-04	1.0E-04	6.6E-05	1.0E-05	4.2E-04	4.0E-05	1.3E-03
	Green	1.1E-05	4.2E-03	1.2E-04	4.3E-04	5.8E-06	1.8E-03	1.5E-06	2.0E-03	1.2E-03	8.4E-03	4.3E-04	7.6E-03
	Blue	3.4E-05	4.7E-03	5.3E-03	9.2E-03	1.1E-03	1.7E-03	5.5E-05	4.1E-04	2.6E-04	5.6E-04	1.4E-04	1.3E-03
FB	All bands	**0.26**	**0.41**	**0.31**	**0.44**	**0.27**	**0.40**	**0.23**	**0.45**	**0.17**	**0.33**	**0.37**	**0.44**
	NIR	0.35	0.33	0.25	0.37	0.28	0.38	0.35	0.34	0.35	0.31	0.44	0.40
	Red	0.28	0.36	0.39	0.42	0.20	0.29	0.23	0.39	0.04	0.29	0.41	0.47
	Green	0.27	0.55	0.31	0.42	0.15	0.55	0.13	0.65	0.23	0.57	0.25	0.62
	Blue	0.14	0.60	0.27	0.56	0.44	0.58	0.21	0.55	0.07	0.23	0.37	0.42

Values are for each forest class and for 3 classifiers: "Minimum Distance (MD)", with Euclidian distance in percent reflectance, "Maximum Likelihood (ML)" with Bhattacharyya distance, and "Frequency-Based (FB)", with Density function-based distance.

Figure 9: NIR reflectance distribution of real IKONOS and DART-simulated imagery

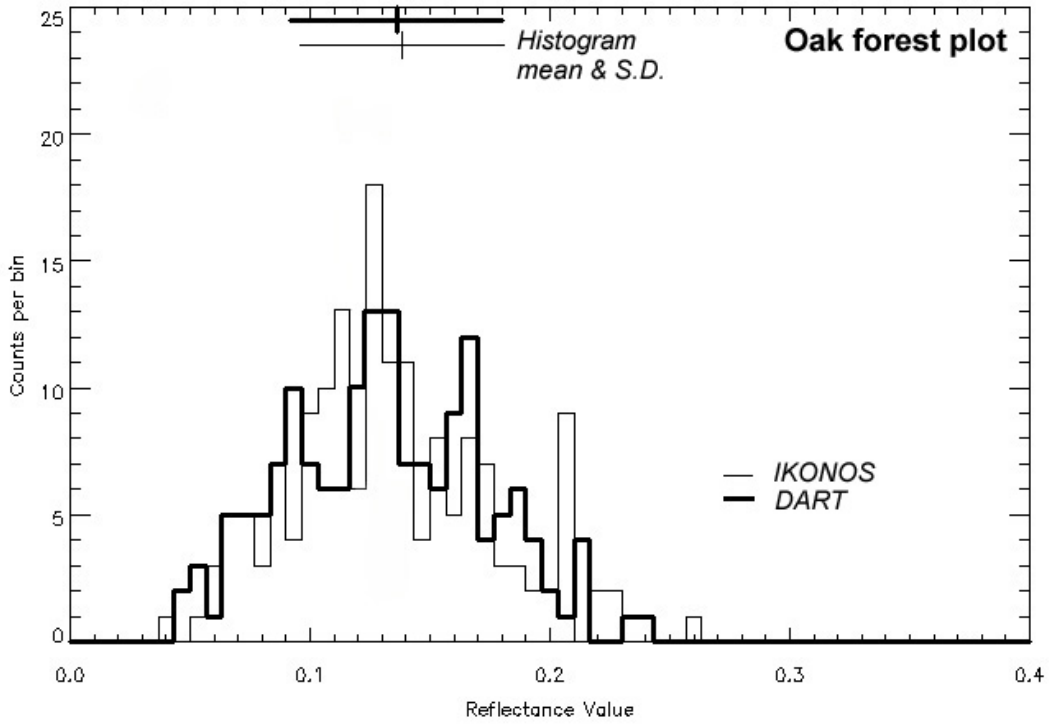


Figure 9a: Oak forest

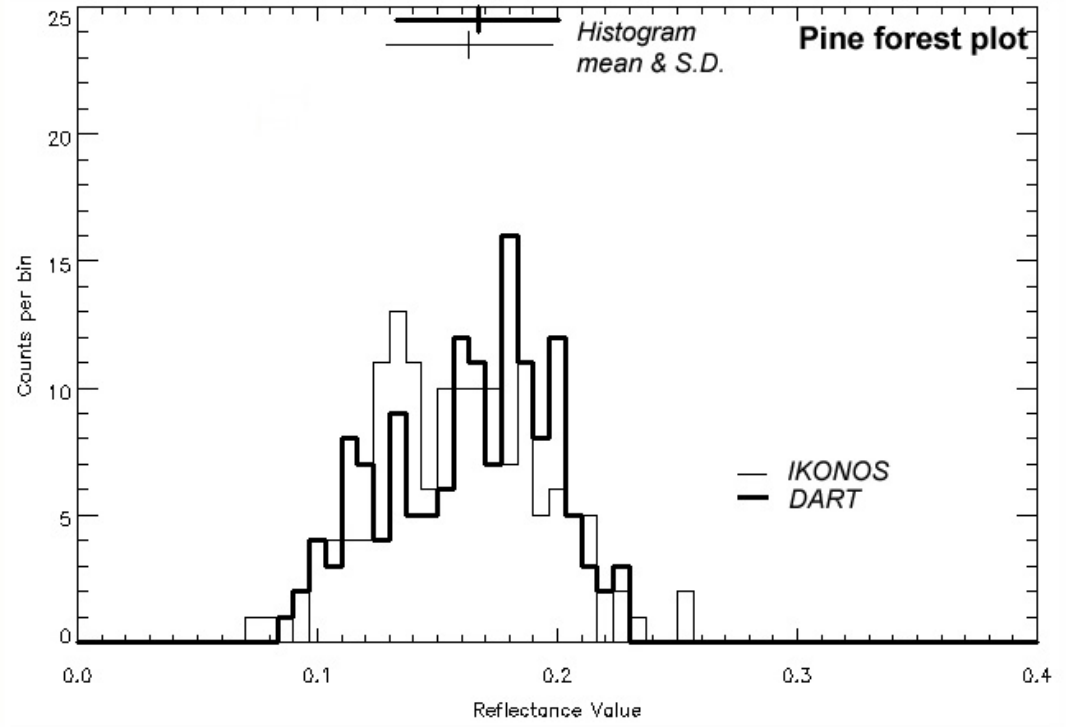


Figure 9b: Pine forest

Figure 9: NIR reflectance distribution of real IKONOS and DART-simulated imagery

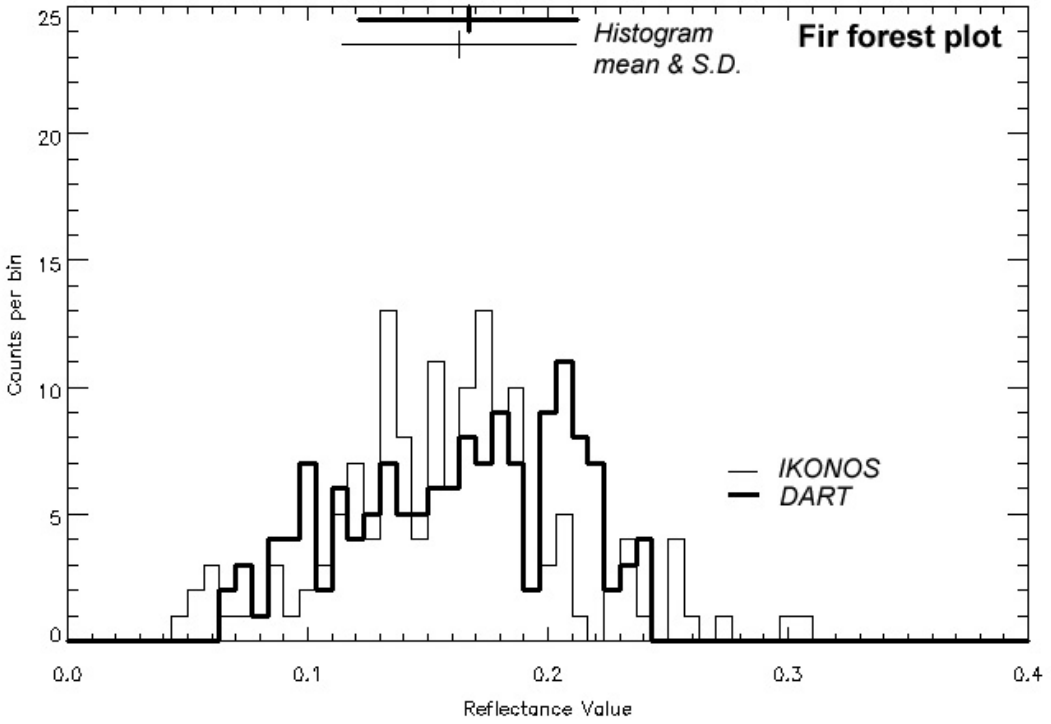


Figure 9c: Fir forest

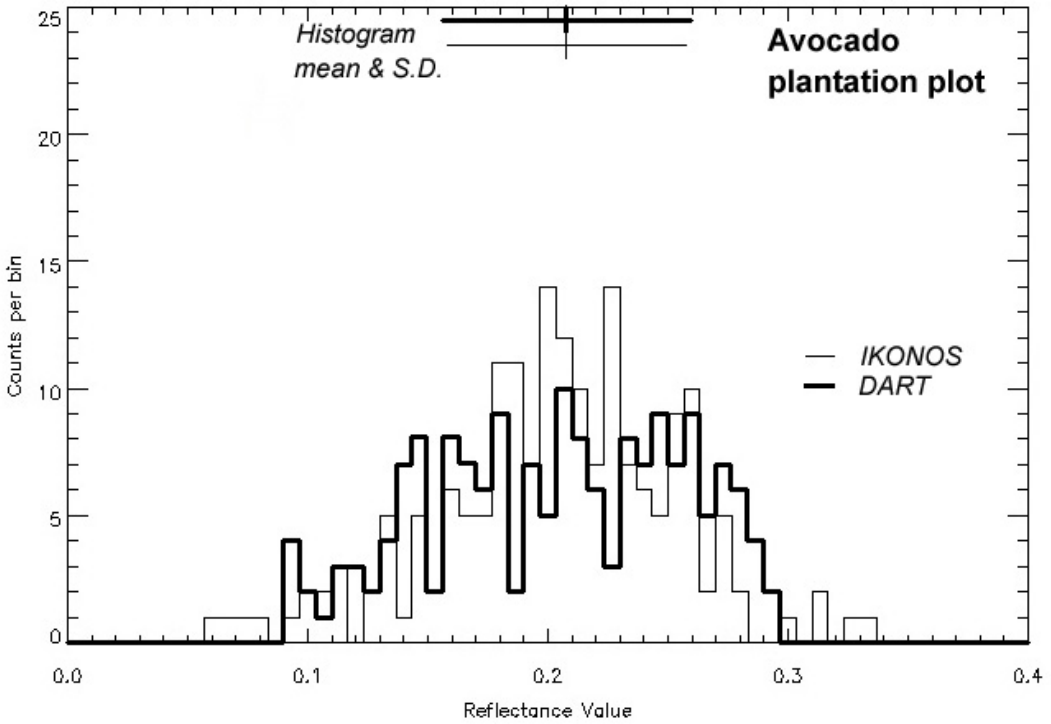


Figure 9d: Avocado plantation

Figure 9: NIR reflectance distribution of real IKONOS and DART-simulated imagery

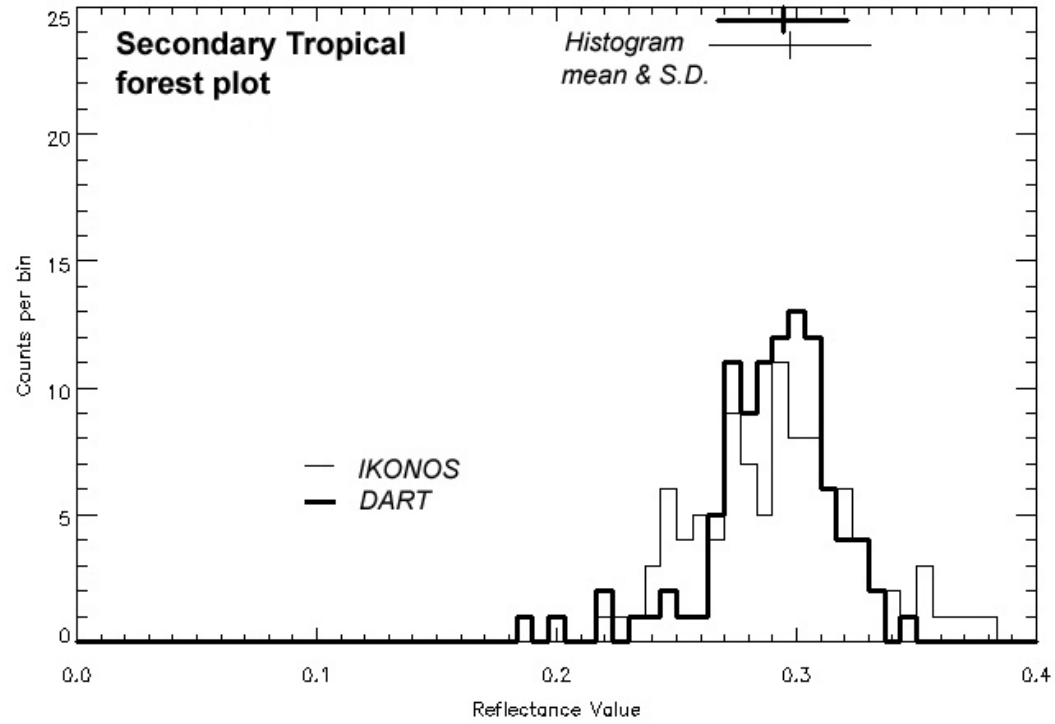


Figure 9e: Secondary Tropical forest

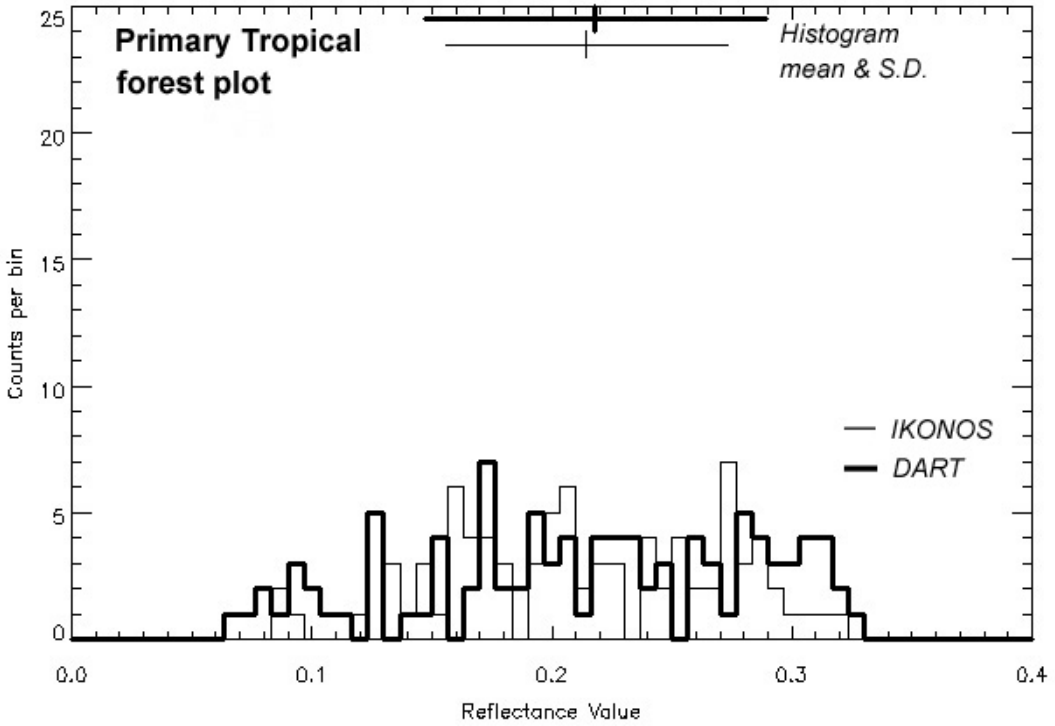


Figure 9f: Primary Tropical forest

7.2 Ambiguity pattern in real IKONOS imagery

We further tested DART-simulated signals against real IKONOS signals by investigating ambiguity patterns among the three forest classes present in the Tancítaro setting. These patterns were studied using an 'ambiguity matrix', displayed in table 5.4. Diagonal data (within outlined rectangles) represent the intra-class dissimilarity for each forest type (information contained in table 5.3). Off-diagonal data give, for all classifiers, dissimilarity values between pairs of classes, computed from the real IKONOS imagery (upper right part of the matrix) or from the DART-simulated imagery (bottom left part of the matrix). Dissimilarity values from DART-derived information (bottom left part) were found highly correlated with dissimilarity values computed from the real IKONOS imagery (upper right part), with the highest correlations recorded when all bands are considered ($r^2 = 0.98$ for MD, $r^2 = 0.99$ for ML and $r^2 = 0.94$ for FB).

Table 5.4: Ambiguity matrix for the pine, fir and avocado classes in the IKONOS Tancítaro image.

		Pine			Fir			Avocado		
Forest:	Band:	MD[a]	ML[b]	FB[c]	MD[a]	ML[b]	FB[c]	MD[a]	ML[b]	FB[c]
Pine	All bands	**1.5**	**0.132**	**0.44**	**0.6**	**0.074**	**0.48**	**1.5**	**0.565**	**0.67**
	NIR	2.4	6.4E-04	0.37	0.1	9.2E-04	0.48	4.0	7.9E-03	0.67
	Red	0.8	2.3E-04	0.42	1.0	2.6E-03	0.60	0.6	4.3E-03	0.88
	Green	2.8	4.3E-04	0.42	0.8	2.2E-03	0.57	0.8	8.8E-04	0.45
	Blue	1.2	9.2E-03	0.56	0.4	1.2E-03	0.26	0.7	4.2E-03	0.69
Fir	All bands	**0.3**	**0.121**	**0.51**	**1.3**	**0.138**	**0.40**	**2.3**	**1.294**	**0.77**
	NIR	0.4	2.4E-03	0.58	1.8	9.1E-04	0.38	4.5	2.5E-03	0.67
	Red	0.6	4.9E-03	0.60	0.8	6.0E-04	0.29	2.0	3.2E-03	0.70
	Green	0.3	2.3E-03	0.58	1.4	1.8E-03	0.55	1.6	4.6E-03	0.80
	Blue	0.1	1.3E-03	0.29	1.1	1.7E-03	0.58	1.1	8.2E-03	0.91
Avocado	All bands	**1.6**	**0.607**	**0.62**	**1.9**	**1.868**	**0.81**	**1.6**	**0.081**	**0.45**
	NIR	4.1	2.3E-03	0.67	4.1	2.7E-03	0.65	2.2	2.2E-05	0.34
	Red	0.6	2.6E-03	0.68	1.2	2.6E-03	0.74	0.8	6.6E-05	0.39
	Green	1.6	3.7E-03	0.82	1.9	1.1E-02	0.97	2.0	2.0E-03	0.65
	Blue	0.2	7.3E-04	0.32	0.4	6.0E-03	0.87	1.2	4.1E-04	0.55

The a priori performance of the three classifiers (MD, ML, FB), mentioned in Table 5.3, is represented for real and DART-simulated IKONOS imagery. The diagonal of the matrix (outlined rectangles) gives intra-class distances. The upper-right matrix gives dissimilarities between classes in the actual IKONOS imagery. The bottom left matrix gives dissimilarities between forest classes in the corresponding DART-simulated imagery.

Off-diagonal (inter-class) dissimilarities, when compared to intra-class dissimilarities on the diagonal of table 5.4, provide the information of whether the forest types are ambiguous or separable. For example, pine was found ambiguous with fir and avocado forests using MD on all bands, whereas only pine and fir forests were ambiguous with ML. Indeed, MD and ML dissimilarities (all bands) for 'pine-fir' are 0.3 and 0.121 (bottom left part of the matrix), or 0.6 and 0.074 (upper right part of the matrix), superior to fir intra-class dissimilarities 1.3 (MD) and 0.138 (ML), respectively. Conversely, the fir and pine forests were separable (as well as the other classes) for the FB classifier, since off-diagonal figures 0.51 or 0.48 are larger than diagonal figures 0.44 (pine intra-class) and 0.40 (fir intra-class). BD expresses a distance between multidimensional distributions assuming a Gaussian shape (Kailath, 1967) while FBD, based on the density function, relies on the actual values of the distribution, thereby reflecting dissimilarities with more robustness. Altogether, the same ambiguity pattern was found, using all bands, among the three classes present on the Tancítaro image, whether using real IKONOS or DART-derived dissimilarity values. This finding, together with the high correlation registered between IKONOS-based and DART-based dissimilarities, provided some confidence for extending the ambiguity study among all six forest classes on a common, simulated environment

7.3 Ambiguities among all forest types

DART simulations of all forest plots were carried out in the 'Tancítaro' setting (the natural setting of the pine, fir and avocado classes). Table 5.5 shows mean reflectance values for all simulations, and one 'ambiguity matrix' per classifier. These matrices are constructed the same way as table 5.4, but they contain dissimilarities calculated for all four bands only. Figures in shadowed boxes indicate cases where classes are ambiguous (i.e., inter-class dissimilarity is smaller than intra-class dissimilarity).

The oak forest and the secondary tropical forest were separable from the other forest types by all classifiers. In both cases, the stand structure makes them appear clearly distinct from other forests. The signal of the low LAI oak forest is largely influenced by the understory condition, as illustrated by the mean visible and NIR reflectance values in table 5, relatively distinct from values of other forest types. Besides, a smooth canopy (see figure 7a) and very high crown cover (table 1) characterize the secondary tropical forest. The

very reduced shadow area inferred by these characteristics causes a bright signal compared to other forest canopies, as illustrated, for example, by the elevated NIR mean reflectance (table 5.5). Most pairs involving other classes (five pairs in total), however, are ambiguous for MD. The limit to MD discrimination power at 25m resolution in the four spectral bands considered appears for all classes except oak and secondary tropical forests. For example, pine and primary tropical forests are confounded with three other classes each.

Table 5.5 Matrix of classifier's theoretical performance for all forest classes.

		Oak	Pine	Fir	Avocado	Tropical Second.	Tropical Primary
Mean Reflectance	NIR	12.7	16.7	16.3	20.8	29.3	18.9
	Red	5.7	4.3	3.7	4.9	4.9	3.9
	Green	6.7	4.9	4.7	6.5	6.4	6.1
	Blue	8.2	7.1	7.0	7.3	8.7	7.4
		Oak	Pine	Fir	Avocado	Tropical Second.	Tropical Primary
Minimum Distance Classifier	Oak	**1.3**					
	Pine	2.1	**1.5**				
	Fir	2.2	*0.3*	**1.3**			
	Avocado	2.5	*1.6*	1.9	**1.6**		
	Trop. Second.	4.6	4.1	4.4	2.5	**1.8**	
	Trop. Primary	2.4	*1.0*	*1.2*	*0.9*	3.2	**2.1**
Maximum Likelihood Classifier	Oak	**0.273**					
	Pine	1.707	**0.132**				
	Fir	3.523	*0.121*	**0.138**			
	Avocado	0.914	0.607	1.868	**0.081**		
	Trop. Second.	0.645	1.872	3.100	0.378	**0.267**	
	Trop. Primary	0.786	0.255	2.609	*0.137*	0.303	**0.209**
Frequency Based Classifier	Oak	**0.42**					
	Pine	0.79	**0.44**				
	Fir	0.78	0.51	**0.40**			
	Avocado	0.73	0.62	0.81	**0.45**		
	Trop. Second.	0.64	0.70	0.91	0.56	**0.33**	
	Trop. Primary	0.60	0.58	0.75	*0.44*	0.51	**0.45**

This matrix is derived from DART images simulated with environmental parameters of the Tancítaro site. The matrix diagonal (bold) gives intra-class distances. The matrix bottom left gives dissimilarities between DART and IKONOS forest classes, for the Tancítaro environmental conditions. Shadowed boxes represent cases where classes are ambiguous (i.e. where inter-class dissimilarity value is smaller than intra-class dissimilarity value).

The poor discrimination power of MD on quite distinct forests is not surprising. Indeed, most studies focussed on forest classification do not rely on minimum distance between mean pixel values, but instead on maximum likelihood (e.g. Wang et al., 2004) or more sophisticated strategies (e.g. Magnussen et al., 2004).

Neither MD nor ML can separate fir from pine forest. Besides being both conifer forests for which no successful attempt of classification was found in the literature, the present pine and fir forest stands possess close structural characteristics (relatively similar CC and LAI; see table 5.1). Moreover, fir and pine forest stands were characterized by low to moderate LAI, which implies that the understory, partially covered with brownish needle litter in both cases, has a strong influence on the canopy reflectance. However, FB, due to its better sensitivity to forest stand structure, could separate fir from pine forest.

Finally, the avocado plantation is confounded with the primary tropical forest for all classifiers, presumably because the reflectance of both types of forest is governed by a deep canopy and a similar shadow area.

7.4 Relating ambiguities to confusions in the NFI map assessment

With the purpose of understanding the role of an ambiguity study in the context of real map production, we confronted the ambiguity pattern with confusion estimates of the NFI map from a previous study. Table 5.6 contains a confusion matrix extracted from the accuracy assessment of the NFI map at community level over three large areas (8306 km^2 in total, see Couturier et al., submitted) containing the 3 sites of this paper. Each forest stand in our ambiguity study is included in a class of that matrix (for example, the avocado plantation is included in the 'Perennial Crop' class). The classification strategy for NFI map production relied on visual interpretation of Landsat colour composites printed at scale 1:125,000 (see Mas et al., 2002). Its accuracy assessment was based on aerial photograph interpretation. The accuracy assessment was achieved with a conservative Boolean agreement definition, which leads to confusions that tend to be superior to those of a fuzzy assessment. The visual classification, because of synoptic abilities of the human eye, is certainly more successful than MD applied to the four IKONOS bands, which correspond to the Blue, Green, Red and NIR Landsat bands. However, the ambiguity pattern on those bands using MD might give trends of the expected success of visual

classification. Additionally, the measurement of ambiguities using ML and FB on finer 4m-resolution imagery may suggest alternative classification strategies in case of pronounced confusions.

According to table 5.6, major confusions did occur between pine and perennial crop: the confused area summed up to 47.3 km^2 versus 100.9 km^2 of pine forests effectively mapped as pine. The confusion was entirely observed in the Tancítaro area where 90% of perennial crop is avocado plantation. This confusion actually corresponds to an ambiguity between pine and avocado plantation for MD (table 5.5). However, table 5.5 also indicates that this pair of classes is separable using ML and FB on 4m-IKONOS imagery. This result could constitute an incentive towards diminishing NFI map confusions using finer resolution, eventually with ML or FB classifiers.

According to table 5.6, fir forests were somewhat confounded with pine forests (1.2 km^2 confused area versus 11.2 km^2 of fir forest mapped as fir). The two classes were also reported ambiguous using MD. The ambiguity is probably more pronounced than the previous case since this time, only the FB classifier (and not ML) could separate the pair of class at 4m-resolution. The latter information might be interesting in the perspective of improving the accuracy of the map on the reduced area where fir forests occur.

Table 5.6 NFI confusion matrix of six forest classes over three zones containing the study sites.

		Oak	Pine	Fir	Perennial Crop	Tropical Second.	Tropical Primary
	Oak	86.04					
NFI	Pine	1.89	100.88				
Confusion	Fir	0	1.16	11.17			
Matrix	Perennial Crop	0.06	47.27	0	348.04		
	Trop. Second.	0	0	0	0	55.69	
	Trop. Primary	0	0.06	0	11.84	9.35	337.19

This matrix is derived from a confusion matrix (Couturier et al., submitted) computed for an area (8306 km^2) that includes the study sites and contains 26 classes. Matrix elements are given in km^2. Diagonal figures a_{ii} show the total area of correctly mapped class i. Off-diagonal figures a_{ij} correspond to the sum of areas where class i was wrongly mapped as class j and where class j was wrongly mapped as class i.

Next, 9.4 km^2 of land was subject to a confusion between secondary and primary tropical forest, against 55.7 km^2 of correctly mapped secondary tropical forest. However, this pair of classes seemed to cause no separability problem on the IKONOS data (no ambiguity for any of the classifiers). The structure of the secondary forest stand studied in this paper has age-specific, homogeneous characteristics compared to the variety of common secondary forests patches, the most mature of which are certainly more difficult to distinguish from primary forest. A limitation of this method is therefore the correct representation of the diversity within a class. For example, forest plots of various age successions would be necessary to calculate secondary forest intra-class variation. Besides, no confusion was recorded between secondary tropical forests and other classes, coinciding with the null ambiguity in our study, although anyways little interface existed between secondary forest and those classes on the large area where the map had been assessed.

Only a small confusion pattern was detected (Table 5.6) between perennial crop and primary tropical forest (11.8 km^2 of confused area versus 337.2 km^2 of correctly mapped primary tropical forest), whereas avocado plantation was systematically confused with primary tropical forest in our experiment. Indeed, the NFI confusion occurred in the region of Los Tuxtlas and concerned shaded coffee or mango plantations and not avocado, so the results of our experiment cannot be applied to the Los Tuxtlas region, and further work should concern shaded coffee or mango plantations.

A small confusion pattern was also detected between oak and pine forests (1.9 km^2 confusion area versus 86.0 km^2 of oak forests) whereas pine and oak forests are clearly separated in the ambiguity study. In reality, oak forests, dominated by *quercus deserticola* , may occur close to pine forests in the fringe between sub-humid temperate and tropical dry climate zones. However, the interface between oak and pine forests is certainly more abundant within the temperate zone of the state of Michoacán, where the structure of oak forests (taller and rounder crowns) and species dominance are very distinct to the ones of the forest studied in this work, and probably more prone to ambiguities with pine forests. Thus, the non-ambiguity results from this study should not be generalized to the entire 'oak' forest type in the NFI classification system; instead, this experiment ought to be extended to oak stands situated in sub-humid temperate conditions before concluding on *a priori* classification ambiguities of the forest inventory in that eco-zone. The oak forest

had very little interface with the remaining set of classes on the map, resulting in null or almost null confusion with these other classes.

A step further into the exploration of ambiguity patterns, also possible with this methodology, would consist in modelling mixed forests, characterized by the physiognomic dominance of both conifers and broadleaf species, because mixed forest is also a major forest class in the sub-humid temperate eco-zone.

8. Discussion

8.1 The database on forest ambiguities

This work aims at contributing to the following research question: given a set of forest classes, which remote sensor and automatic classification technique are optimal to generate a reasonably accurate map? It is argued that a database containing structural properties of forests at plot scale could yield a systematic *a priori* estimation of ambiguities between pairs of classes. This database could constitute a useful management tool for forest mapping/monitoring. The usual approach for forest mapping with remote sensing imagery is to directly apprehend inter-class and intra-class signals from an image of the sensor at hand, and test various classifying techniques using a set of reference field data. A drawback of this approach is that any updating of the map based on a new image involves repeating the tests and updating the field data for verification, because the new image is characterized by new environmental conditions, or was taken by a new sensor. Instead, once validated, the modelling approach offers predictive capabilities on new remote sensing supports and environmental conditions.

8.2 Field data collection and resolution issues

Due to the highly heterogeneous nature of forests, the 3D-modelling approach we propose demands at first a wealth of ground measurements, but then no repetition of such field data should be needed to update results to future conditions. Typically, high resolution (20-30m) resolution imagery is widely used for forest mapping. However, at such a resolution, an enormous amount of crown-scale field data would have been needed, per forest type, to directly compare simulated to real imagery in a statistically meaningful way (for example, 50 Landsat pixels would require the detailed survey of 4.5 ha!). Taking advantage of the

availability of very high resolution imagery (1-5 meters resolution), this work proposes to model reflectance distribution of forests with field data gathered from plot sizes typical of forest inventory surveys (40mx40m or 50mx50m). Additionally, in the estimation of intra-class variation, the modelling approach is useful to minimize the collection of field data. For example, only CC and LAI were measured on the reference parcels and no additional detailed inventory was needed.

Although the texture of 1m-resolution imagery is sometimes used as an auxiliary band to improve discrimination of forests or forest stand structures (Wang et al., 2004; Colombo et al., 2003), only the 4m-resolution IKONOS data was retained in this study. Indeed, the simplifying assumptions used in the DART model for forest canopy (e.g. crowns modelled by ellipsoids) limited the resemblance of DART-simulated imagery with real IKONOS imagery at 1m resolution; this phenomenon is especially visible for primary tropical forest (see differences in the appearance of emergent crowns between figures 5.8 (e) and 5.8(f)). In fact, preliminary tests demonstrated notable differences between first order texture channels of simulated versus real imagery, which limited the usefulness of our modelling approach at 1m resolution. This problem originates from the DART module that simulates 3D forest covers, and not from the radiative transfer module that can work with any 3D environment. In order to solve this problem, CESBIO is developing an interface between the DART radiative transfer module and a 3D model dedicated to the simulation of realistic forest covers. In spite of this limitation, a result of this study is that DART simulation could approach the 4m real signal (multi spectral IKONOS image) more accurately than intra-class signal variation, even for tropical forest stands (table 5.3).

8.3 Structural and optical properties of the forest

Inter-class dissimilarities were measured on both simulated and real imagery and both quantities compared favourably (table 5.4). Intra-class variation was modelled for forest stand structure on relatively flat terrain, which implies that the ambiguity results of this study only apply to forest classes on relatively flat terrain. The intra-class variation also compared favourably to the intra-class variation that we observed on the reference parcels (also on flat terrain) of the real imagery. However, classification exercises on large areas with pronounced relief are affected by more reflectance variability than measured in this

study (greater intra-class variation), and subject to more ambiguities among classes. Further research is under way to model and study the same forest types on steep terrain, since much of the forest outside the plots investigated occurred on micro relief or pronounced slopes, in which case the forest stand structure was distinct. This additional work may open the way to *a priori* ambiguity measures on the entire forest mass of the IKONOS image. In this context, the DART model is useful because it can operate on 3D landscapes with topography.

Optical properties of leaves, not available on existing databases, were inferred from the imagery. We did not account for the fact that leaf optical properties display spatial variations, even within individual tree crowns. Moreover, we did not account for horizontal variations of leaf volume density within individual tree crowns. The radiative transfer module of DART can account for these 3D variations. However, we could not measure them and, thus, use this information in the DART simulations.

8.4 Type of classifier

In spite of forest stands taxonomically very distinct, MD performed poorly at distinguishing various class pairs at 25m resolution. At higher resolution, we attested that MD performed worse due to a very large intra-class signal (forest heterogeneity); however, at lower resolution, intra-class mean value variation is smaller and MD would probably have performed equal or better. Indeed, many forest classification studies recommend to low pass filter the high resolution (20-30m) imagery prior to classification; for example, in recent studies, Salovaara et al. (2005) smoothened the Landsat image with a 5x5 pixel window (*i.e.*, 150mx150m), and Arroyo-Mora et al. (2005) smoothened it with a 3x3 pixel window (*i.e.*, 90mx90m). Reaching these resolutions with our approach is possible, but requires the survey of larger plots, and the integration of scenarios with uneven topography.

Only two cases were ambiguous using the ML classifier versus five cases using MD. This study confirms ML, which takes into account the pixel value distribution across bands, as a reasonably good classifier of forests on multi-spectral IKONOS imagery. This was illustrated by Wang et al. (2004) on taxonomically close forests. When the reflectance

distribution deviates too much from normality, FB classifier might reduce ambiguities found using ML. An example in this study may have been the fir forest, whose NIR reflectance distribution tends to deviate from normality (figure 5.9c), and was confounded with pine forest using ML whereas the pair of classes was separable using FB.

8.5 Scaling up to confusions in the NFI map

Some ambiguity results may have given insights on possible causes for confusions in the NFI map, as reported in the previous section. These insights, however, have to be placed in their context since this study presents cases of automatic classification applied to a reduced set of spectral bands, as compared to those of the Landsat sensor, support of the NFI map production. Another limitation to this comparison exercise is the necessary information on other possible sources of errors in the map such as geometric or scale-induced.

Besides, as evident from the confrontation of ambiguity and confusion matrices, many classes of this pilot study are not comparable to the classes at community level of the NFI map, in the sense that they are contained in taxonomically larger classes. Surveys need to be done in all forest stands representative of the forest class on the map, before the ambiguity study may apply to the entire class. In this sense, a limitation of this approach, apart from shortcomings of the model itself in approximating real forest imagery, is that a certain amount of sample forest stands is needed before the data base can serve as a performance test for an entire classification system.

However, the exercise can virtually be extended to any user-defined forest class (whether physiognomic, taxonomic or crown coverage-defined). An advantage of this method is that the database benefits cumulatively from any additional input of classes. Additional input in the database improves the accuracy of the estimates and/or the extent of its applicability to new forest classes, whereas the sum of previous empirically-based classification methods hardly tells about the outcome of a classification in a new setting with a new remote sensor.

9. Conclusion

This research sets a model-based methodology for studying ambiguity patterns among forest types on satellite imagery. The methodology was applied to multi-spectral IKONOS imagery on six forest stands pertaining to community level classes of the Mexican National Forest Inventory. First, the DART model was successfully tested against the real IKONOS image on each forest plot in the sense that the dissimilarity between DART and IKONOS images was found within intra-class variation in the six forest types. Consequently, DART was used as a proxy for IKONOS imagery in order to estimate ambiguity among all forest classes.

As expected, a poor discrimination power was usually observed with MD at 25m resolution. This is coherent with the fact that the minimum distance strategy, confronted with very high intra-class variations, is not used for forest classification but rather for more general land cover classification. Instead, in most cases, the distinct forest types were found separable by the ML and FB classifiers.

Low LAI oak forests appeared very dissimilar to other forests owing to the strong understory influence on the reflectance. The secondary tropical forest stand appeared very bright with respect to other forests because of its smooth canopy and absence of gaps. The plots of pine and fir forests were ambiguous for MD and ML on real IKONOS imagery. However, the FB classifier, probably because it better accounts for the reflectance value distribution, could separate both forests. Conversely, avocado plantation and primary tropical forest were ambiguous for all classifiers, presumably because their reflectance is governed by a deep canopy and a similar shadow area.

The confrontation of these results with the accuracy assessment of the NFI map stressed the need to extend the approach to more forest stands in the perspective of testing the classification scheme of the NFI. Some practical limitations of the approach are related to the difficulty in modelling very high resolution imagery and in incorporating the complexity of optical properties. However, improvements are technically feasible and under development.

Considering the above, this methodology is to be extended to other forest stands, and the results are to be compared with classification accuracy results on satellite imagery. The

perspective of this research is to predict confusions in forest maps, in function of the type of classifier and the taxonomic resolution of the map. In this sense, 3D-modeling could be a valuable management tool for better assessing the expected accuracy of automatic classification of forest types of interest, as a function of the characteristics of the remote sensing system and of the experimental configuration of the image acquisition. Consequently, it could be of help in the decision process related to buying satellite imagery and/or classification software.

10. Acknowledgments

This research and the acquisition of IKONOS imagery were funded by CONACYT (the national science and technology board in Mexico) under project number 38965T. Field work was carried out with a local team composed of biologists, botanists and technical foresters in the community of San Juan Parangaricutiro, Tancítaro, in the ejido village of La Perla de San Martin, Los Tuxtlas, and in the El Cerro village near to the Cuitzeo lake. For the completion of this work, special thanks are due to Isela Zermeño Hernández, Moisés Méndez Toribio, Rafael Pompa Vargas and Alfredo Patiño Siciliano. We are indebted to the entire team that develops the DART model.

11. References

Arroyo-Mora, J.P., Sánchez-Azofeifa G.A., Kalacska M.E.R., Rivard B., Calvo-Alvarado J.C., and D.H. Janzen, Secondary forest detection in a neotropical dry forest landscape using Landsat 7 ETM+ and IKONOS imagery, *Biotropica*, 37(4), 497-507, 2005.

ASTER, Spectral database, NASA/JPL, http://speclib.jpg.nasa.gov, 2003.

Baban, S.M.J. and W.Y. Kamaruzaman, Mapping land use/cover distribution on a mountainous tropical island using remote sensing and GIS, *International Journal of Remote Sensing*, 22 (10), 1909-1918, 2001.

Benjamin, S., White J.M., Argiro D., and K. Lowell, Gap analysis – a landscape approach to biodiversity planning (Eds H.M. Scott, T.H. Tear, F.W. Davis), American Society for Photogrammetry and Remote Sensing, Bethesda, MD, pp. 279-288, 1996.

Berk A., Bernstein L.S., and D.C. Robertson, MODTRAN: A moderate resolution model for LOWTRAN7, GL-TR-89-0122, Geophysical Laboratory, Bedford, MA, USA, 38pp., 1989.

Colombo, R., Bellingeri D., Fasolini D., and C.M. Marino, Retrieval of leaf area index in different vegetation types using high resolution satellite data, *Remote Sensing of Environment*, 86, 120-131, 2003.

Courbaud, B., de Coligny F., and T. Cordonnier, Simulating radiation distribution in a heterogeneous Norway spruce forest on a slope, *Agricultural and Forest Meteorology*, 116 (1), 1-18, 2003.

Couturier, S., Mas J.F., López E., Benítez J., Tapia V., and A. Vega, Accuracy Assessment of the Mexican National Forest Inventory map: a study in four eco-geographical areas, *Singapore Journal of Tropical Geography*, submitted.

Disney, M., Lewis P., and P. Saich, 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains, *Remote Sensing of Environment*, 100, 114-132, 2006.

Fassnacht, K.S., Cohen W. B., and T.A. Spies, Key issues in making and using satellite-based maps in ecology: a primer, *Forest Ecology and Management*, 222, 167-181, 2006.

Gascon, F., J.P. Gastellu-Etchegorry, Lefèvre-Fonollosa M.J., and E. Dufrêne, Retrieval of forest biophysical variables by inverting a 3-D radiative transfer model and using high and very high resolution imagery, *International Journal of Remote Sensing*, 25 (12), 5601-5616, 2004.

Gastellu-Etchegorry, J.P., Martin E., and F. Gascon, DART: a 3-D model for simulating satellite images and surface radiation budget, *International Journal of Remote Sensing*, 25(1), 75-96, 2004.

Gastellu-Etchegorry J.P., and V. Trichon, A modelling approach of PAR environment in a tropical rain forest in Sumatra: application to remote sensing, *Ecological Modelling*, 108, 237-264, 1998.

Gastellu-Etchegorry, J.P., Hapke handbook, CESBIO, Toulouse, France (http://www.cesbio .ups-tlse.fr), 15pp., 2006.

Gemmel, F., and J. Varjo, Utility of reflectance model inversion versus two spectral indices for estimating biophysical characteristics in a boreal forest test site, *Remote Sensing of Environment*, 68, 95-111, 1999.

Goel, N.S., and R.L. Thompson, A snapshot of canopy reflectance models, and a universal model for the radiation regime, *Remote Sensing Reviews*, 18, 197-225, 2000.

Hapke B., Theory of Reflectance and Emittance Spectroscopy, Cambridge, 455 pp., 1993.

Jonckheere, I., Fleck S., Nackaerts K., Muys B., Coppin P., Weiss M., and F. Baret, Review of methods for in situ leaf area index determination, Part I. Theories, sensors and hemispherical photography, *Agricultural and Forest Meteorology*, 121, 19-35, 2004.

Kailath, T., The divergence and Bhattacharyya distance measures in signal selection, *IEEE Transactions in Communication Theory*, 15, 152-160, 1967.

Laba, M., S.K. Gregory, J. Braden, D. Ogurcak, E. Hill, E. Fegraus, J. Fiore and S.D. DeGloria, Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map, *Remote Sensing of Environment*, 81, 443-455, 2002.

Lira, J., and Maletti G., A supervised contextual classifier based on a region-growth algorithm, *Computers & Geosciences*, 28, 951-959, 2002.

Lu, D., Moran E., and M. Batistella, Linear mixture model applied to Amazonian vegetation classification, *Remote Sensing of Environment*, 87, 456-469, 2003.

Lu, D., Integration of vegetation inventory data and Landsat image for vegetation classification in the western Brazilian Amazon, *Forest Ecology and Management*, 213, 369-383, 2005.

Magnussen, S., Boudewyn P., and M. Wulder, Contextual classification of Landsat TM images to forest inventory cover types, *International Journal of Remote Sensing*, 25 (12), 2421-2440, 2004.

Maletti, G., Ersbøll B., and K. Conradsen, A contextual classifier that only requires one prototype pixel for each class, *IEEE Transactions on Nuclear Science*, 49 (3), 700-706, 2002.

Martin E., DART: modélisation d'images optiques de télédétection. *PhD thesis, Paul Sabatier University*, Toulouse, France, 2006.

Mas J.F. and I. Ramírez, Comparison of land use classifications obtained by visual interpretation and digital processing, *ITC Journal*, 3(4), 278-283, 1996.

Mas J.F., Velázquez A., Palacio-Prieto J.L., Bocco G., Peralta A., Prado J., Assessing forest resources in Mexico: Wall-to-wall land use/ cover mapping. *Photogrammetric Engineering and Remote Sensing* 68 (10), pp 966-969, 2002.

North, P.R.J., Three-dimensional forest light interaction model using a Monte Carlo method, *IEEE Transactions on Geoscience and Remote Sensing*, 34 (4), 946-956, 1996.

Palacio-Prieto J.L., Bocco G., Velázquez A., Mas J.F., Takaki-Takaki F., Victoria A., Luna-González L. et al., La condición actual de los recursos forestales en México: resultados del Inventario Forestal Nacional 2000, *Investigaciones Geográficas* 43, 183-202, 2000.

Pinty, B., & M.M. Verstraete, Extracting information on surface properties from bi-directional reflectance measurements, *Journal of Geophysical Research*, 96, 2865-2874, 1991.

Pinty, B., Gobron N., Widlowski, J.-L., Gerstl, S.A.W., Verstraete, M.M., Antunes M., et al., RAMI II: The second radiative transfer model intercomparison exercise, *Journal of Geophysical Research*, 109 (D6), D06210, 2004.

Potter E., Wood J., and C. Nicholl, Sunscan Canopy Analysis System, User Manual v1.05, Delta-T Devices Ltd, Cambridge, UK (http://www.delta-t.co.uk/manual.html), 1996.

Remmel, T.K., Csillag F., Mitchell S., and M. Wulder, Integration of forest inventory and satellite imagery: a Canadian status assessment and research issues, *Forest Ecology and Management* 207, 405-428, 2005.

Salovaara, K.J., Thessler S., Malik R.N., and H. Tuomisto, Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery, *Remote Sensing of Environment*, 97, 39-51, 2005.

Silbernagel, J. and M. Moeur, Modeling canopy openness and understory gap patterns based on image analysis and mapped tree data, *Forest Ecology and Management* 149, 217-233, 2001.

Strahler, A.H., Vegetation canopy reflectance modelling: Recent developments and remote sensing perspectives, *Remote Sensing Reviews*, 15, 179-194, 1996.

Thenkabail, P.S., Enclona E.A., Ashton M.S., Legg C., and M.J. De Dieu, Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests, *Remote Sensing of Environment*, 90, 23-43, 2004.

Vogelmann, J.E., Howard S.M., Yang L., Larson C.R., Wylie, B.K., and N. Van Driel, Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat thematic mapper data and ancillary data sources, *Photogrammetric Engineering and Remote Sensing*, 67, 650-662, 2001.

Wang, L., Sousa W.P., Gong P., and G.S. Biging, Comparison of IKONOS and Quickbird images for mapping mangrove species on the Caribbean coast of Panama, *Remote Sensing of Environment*, 91, 432-440, 2004.

Wickham, J.D., Stehman S.V., Smith J.H., and L. Yang, Thematic accuracy of the 1992 National Land-Cover Data for the western United-States, *Remote Sensing of Environment*, 91, 452-468, 2004.

Wilkinson, G. G., Results and implications of a study of fifteen years of satellite image classification experiments, *IEEE Transactions on Geoscience and remote sensing*, 43 (3), 433-440, 2005.

Wulder, M.A., Deshka J.A., Gillis, M.A., Luther, J.E., Hall, R.J., Beaudouin, A., and S.E. Franklin, Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program, *Forest Chronicles* 79, 1075-1083, 2003.

Xu, B., Gong P., Seto E., and R. Spear, Comparison of gray-level reduction and different texture spectrum encoding methods for land-use classification using a panchromatic IKONOS image, *Photogrammetric Engineering and Remote Sensing*, 69 (5), 529-536, 2003.

Chapter 6: General Conclusion

The study of sources of errors in the production and assessment of land use and land-cover maps is a particularly important issue in areas with highly dynamic landscapes and high bio-diversity. Yet, most efforts on accuracy assessment of regional scale maps have focused so far on mainly temperate zones. Major characteristics of regional maps of sub-tropical environments are a high taxonomic diversity (number of classes) and highly dynamic landscapes. The objectives of this thesis were three-fold:

- Evaluate the accuracy of the NFI map of year 2000 using a sampling design with a sufficiently high sample size for all classes, including the numerous sparsely distributed (rare) classes,
- Propose an accuracy assessment design where errors due to the positional and thematic uncertainties of maps are quantifiable,
- Propose a framework for the measurement of ambiguities among classes on satellite imagery, applicable in a wide range of environmental settings.

To fulfill the objectives, we developed two methods, based on recent theoretical advances in geo-science disciplines. The first technique is an accuracy assessment design based on fuzzy classification theory. The second technique is based on the radiative transfer modeling of 3D landscapes. First, the significant contribution and results of the methods are described, and then the perspectives of this research are discussed.

1. Main Results

1.1 Sampling design

The sampling part of the first objective was answered in chapter 1: for LULC maps of high taxonomic diversity, we propose a strategy based on a double sampling, with separate schemes for common and for rare classes. The first stage sampling for common classes follows a simple random selection, while the first stage sampling for rare classes follows a proportional stratified selection. We demonstrated the efficiency of this hybrid sampling design with respect to other sampling designs previously employed in regional accuracy assessments. For operational costs fixed relatively high (one fourth of the total photographic coverage), the concurrent sampling designs were unable to provide sufficient

sample sizes to all classes. Similar unpublished results were obtained for varying operational costs, as well as on the other three study areas (chapter 2) of this thesis.

Given the relatively small geographic extent of our well documented study areas, the coarse scale (1:250 000) of the map and the high number of classes, the approach was affected by the limited distribution of sample points. For this reason, the average number of sample points per PSU remained relatively high for all designs (see chapter 1). Some cluster clumping effect did occur in PSUs containing many sample points and the binomial standard errors for these cases under-estimate the confidence interval of accuracy levels. Complex designs, typical of regional scale accuracy assessments, imply complex variance estimators. Better estimators of variance, taking into account the intra-cluster covariance, were available at an advanced phase of our method (chapter 3), and our published confidence intervals were in the range of the confidence intervals in Stehman et al. (2003).

1.2 Results of the accuracy assessment

It was possible to evaluate the accuracy of the NFI map of year 2000 in four distinct eco-geographical areas (chapter 2), namely the watershed of the Cuitzeo lake, the region of the Tancítaro volcano and adjacent avocado production zone, the biosphere of Los Tuxtlas and adjacent sugar cane production zone, and the watershed of the Candelaria River. Global accuracy estimates of 64% to 78% were registered, and the lower accuracy levels were found in areas with denser forest cover. This range of estimates was compared with results from compatible assessment designs of other Landsat-based regional scale maps at similar taxonomic resolution (chapter 2). It is somewhat above the range (46-66%) reported in Stehman et al. (2003) of LULC maps in the four smallest of the 10 environmental regions administrated in the USA, and is above 64%, an accuracy estimate reported by Remmel et al. (2005) for the Canadian forest cover map. The coarse scale NFI map was generated via visual interpretation of Landsat imagery (Mas et al., 2002), whereas the above mentioned maps were generated via semi-automatic classification of Landsat imagery. Although visual interpretation yielded a map product at coarser resolution, the results obtained in this research suggest that this classification method may have produced accuracy levels difficult to reach elsewhere at equivalent taxonomic resolution. An accuracy assessment has to be conducted on the entire territory, though, to

investigate if the trend of accuracy levels reported in this study is confirmed at a national scale.

Trends of errors per class could be quantified (chapter 2) by the assessment. For example, the presence of temperate forests tends to be underestimated whereas the presence of tropical forest tends to be overestimated in the regions investigated. In other words, for these areas, calculations of deforestation rates before 2000 using the presumably highly accurate INEGI series I data of *circa* 1976 and the NFI 2000 map (Mas et al., 2002b) could be overly optimistic for tropical forests and overly pessimistic for temperate forests. The estimation of more precise rates, with their margin of error, can be derived from the results of this thesis. The analysis of confusion patterns unleashed issues that are intrinsic to the NFI map and the selection of the classification scheme. For example, the way secondary vegetation is included in the classification scheme (always associated with a forest type) may have caused thematic uncertainty for these classes in the classification scheme. However, this thematic uncertainty was possibly a necessary trade-off for the compatibility of the NFI 2000 map with classification schemes of former INEGI maps. Another possible interpretation for these errors was related to the positional uncertainty of the boundary of polygons on the map, certainly affected by the width of spatial transitions between clearly distinguished land cover types.

1. 3 The fuzzy sets based method

The evaluation protocol of the traditional accuracy assessment employed in chapters 1 and 2 fails to answer whether these errors were mainly due to thematic uncertainty, positional uncertainty, or real, misclassification error. Since this question is useful for environmental modeling based on this map, the full, fuzzy sets based method was incorporated to the accuracy assessment design in chapter 3. The evaluation protocol consisted in the construction of reference maplets, and the labelling protocol consisted in a fuzzy map comparison for which degrees of positional and thematic tolerance are user-defined. An account of errors was made in the case of complex forested landscapes of Southeastern Mexico. The method suggested that a substantial part of the confusion registered for ‘secondary’ forest classes was effectively related to thematic uncertainty, whereas high level of confusions between median and low sub-perennial tropical forests was essentially related to the fuzziness of boundaries between patches of these forests, and not to a

taxonomic ambiguity in the classification scheme. Therefore, the accuracy indices with fuzziness bounds provided by the model added useful and quantitative insights in the analysis and discussion of confusions. Conceptually, the method permitted the standardization of the notion of degrees of tolerance in accuracy assessments of maps, in both thematic and positional dimensions.

1.4 The model-based spectral separability study

From the analysis of results of chapter 2, we observed that very high accuracy was generally obtained by classes characterized by high visual spectral separability with respect to other classes on satellite imagery (e.g. water and 'other cover types'). Conversely, the highest confusion levels were observed for classes that are possibly ambiguous with physiognomically close classes (e.g. aquatic vegetation types, 'secondary' forests). The accuracy levels of 'primary' forest types were generally intermediate and variable. Although the NFI map was the result of visual interpretation of Landsat imagery, other contemporaneous regional scale LULC maps generally rely on semi-automatic classification of satellite imagery.

In this research, a model-based spectral separability study of forest covers (chapters 4 and 5) was tested as a guide in the selection of the best option for the classification of forests. The model-based method was validated on IKONOS images in three of the eco-geographical zones. The DART 3D model could satisfactorily approach the reflectance distribution of pixels of real 4m multi-spectral images for flat and rugged-terrain scenarios (chapter 4). Consequently, DART was used as a proxy for IKONOS imagery for evaluating separability of forest classes, simulated in the same conditions of slope aspect. Although spectral separability was found very sensitive to the aspect of the slope, previous studies do not usually explore the accuracy of classification results in function of slope aspect. The study demonstrates that physiognomically distinct forest types like pine and perennial tropical forest, both occurring in the Los Tuxtlas region, for example, may appear inseparable on some slope aspects for this type of imagery. In the presence of complex topography, results such as this one permit the evaluation of classification options for future land cover mapping in Mexico.

Additionally, this thesis contributed to demonstrate the usefulness of a 3D radiative transfer model for classification studies at resolutions near to crown scale. Indeed, the quasi totality of radiative transfer models, which simulate a point-like signal, are not adapted to this resolution. Resolution near to crown scale permitted a design with ground measurements on reasonable forest plot sizes (40mx40m or 50mx50m), typical of forest inventory surveys. A similar study on Landsat-like imagery, would have required intensive surveys of at least 4.5 hectares (!) per forest type, which is probably one of the reasons why a similar systematic study had not been done in the past with radiative transfer models.

Spectral separability depends on the classifying technique employed for the classification of satellite imagery. The Minimum Distance (MD), the Maximum Likehood (ML) and the Frequency Based (FB) classifiers were incorporated in the study on six forest types in chapter 5. Five, two and one cases of ambiguity were reported for respectively the MD, ML and FB classifiers. Avocado plantation and primary forest were ambiguous for all classifiers, presumably because their reflectance is governed by a deep canopy and a similar shadow area. The confrontation of these results, in chapter 5, with the accuracy assessment of the NFI map of chapter 2 stressed the need to extend the approach to more forest stands in the perspective of testing the classification scheme of the NFI.

2. Perspectives of the research

2.1 National scale accuracy assessment

For budget reasons, the design could only be tested on four pilot areas. However, the hybrid sampling design and more generally the entire assessment design presented here is entirely transferable to more extended study areas, with typically available data and skills in Mexico. The dispersion is better on larger areas because of a larger total number of PSUs in the initial set. We suggest the application of the accuracy assessment scheme presented in this thesis at national scale, augmenting the scheme with the systematic sampling along a photographic coverage, and with the stratification by administrative state. The cost-efficiency of such an accuracy assessment is ensured by the hybrid design, therefore the existing national photographic coverage of year 2000, unused so far for cost reasons, could serve this important purpose. Additionally, we strongly recommend the

application, by an independent institution, of this, or a concurrent, accuracy assessment design on current CONAFOR-managed maps with more recent, or future, photographic coverage. As long as this is not done, the public figures for recent rates of LULC change must only be viewed by the citizen as unreliable estimates, controlled by governmental agents and therefore subject to justified suspicion.

2.2 The user-defined thematic and positional aspects

In accuracy assessments, it is admitted that no single approach can realistically encompass all possible situations in answering the question: *How accurate is a map?* (Stehman, 1997; Foody, 2002). Rather, accuracy assessment designs develop from consideration of the following question: *Is the map sufficiently accurate for a specific application?*. However, in the context of large area land cover products, the needs of all possible future applications cannot be foreseen or accommodated (Wulder et al., 2006). Therefore, the initial purpose of the map usually guides criteria for the design of regional scale accuracy assessments. In the proposed research, user defined positional and thematic tolerance bounds permit the modification of some criteria decided for measuring accuracy, in function of the application. In this sense, the use of the fuzzy sets based design may contribute to overcome this conceptual difficulty in future accuracy assessments.

2.3 The database on spectral separability

In theory, pairs of classes confounded on satellite imagery could be detected by the model-based spectral separability study. However, as evident from the confrontation of ambiguity results and confusion matrices of the NFI in chapter 5, many classes of the pilot study were not comparable to the classes at community level of the NFI map, in the sense that they are contained in taxonomically larger classes. Surveys need to be done on a set of forest stands, representative of the forest class on the map, before the ambiguity study may apply to the entire class. In this sense, a limitation of this approach is that a certain amount of sample forest stands is needed before the data base can serve as a performance test for an entire classification system. An advantage of this method, however, is that the database benefits cumulatively from any additional input of classes. Additional input in the database improves the accuracy of the estimates and/or the extent of its applicability to new forest classes, whereas the sum of previous empirically-based classification methods

hardly tells about the outcome of a classification in a new setting with a new remote sensing imager.

The case of forests on slope, for example, can be extended to more forest class pairs. This effort should be of interest to bio-diversity estimates since the major deforestation schemes in the sub-tropical belt, including in Mexico, have already occurred on easily accessible land, characterized by smooth topography, and less modification occurs on steep terrain. Spectral separability of forests on slope is easily applicable to satellite imagery corrected from topography, since the shape of the modelled reflectance distribution is conserved through topographic normalization. In fact, studies on topographically corrected classification of forest have treated so far temperate forests easily distinguished on flat terrain, with 20-30m resolution imagery. A next step of this research is to confront our ambiguity results with real classification results of topographically corrected imagery on specific aspects.

Some practical limitations of the model-based approach are related to the difficulty in modelling very high resolution imagery and in incorporating the complexity of optical properties. For example, the simplifying assumptions used in the DART model for forest canopy (e.g. crowns modelled by ellipsoids) limited the resemblance of DART-simulated imagery with real IKONOS imagery at 1m resolution. Improvements should be made with the future development of an interface between the DART radiative transfer module and a 3D model dedicated to the simulation of more realistic forest architecture.

The perspective of this method is to predict confusions in forest maps, in function of the type of classifier and the taxonomic resolution of the map. In this sense, 3D-modeling could be a valuable management tool for better assessing the expected accuracy of automatic classifications of forest types. Refinements of class definitions, more adapted to the functionalities of remote sensing imagery, can also be envisaged. The application of the method to non forest land-cover classes is a possible extension of this research. Together with the accuracy assessment information, the method is expected to generate a spatial knowledge of main land-cover ambiguity issues, on the base of which recommendation lines can be drawn for classification strategies with future sensors and class definitions.

Global References

1. Achard, F., H.D. Eva, H.J. Stibig, P. Mayaux, J. Gallego, T. Richards and J.P. Malingreau, Determination of deforestation rates of the world's humid tropical forests, *Science* 297 (5583): 999-1002, 2002.

2. Aronoff, S., Classification Accuracy: A user approach, Photogrammetric Engineering and Remote Sensing, 48 (8), 1299-1307, 1982.

3. Arroyo-Mora, J.P., Sánchez-Azofeifa G.A., Kalacska M.E.R., Rivard B., Calvo-Alvarado J.C., and D.H. Janzen, Secondary forest detection in a neotropical dry forest landscape using Landsat 7 ETM+ and IKONOS imagery, *Biotropica*, 37(4), 497-507, 2005.

4. Aspinall, R., and D.M. Pearson. Describing and managing uncertainty of categorical maps in GIS, *Innovations in GIS 2* (P. Fisher, editors), Taylor & Francis, London, pp. 71-83, 1995.

5. ASTER, Spectral database, NASA/JPL, http://speclib.jpg.nasa.gov, 2003.

6. Baban, S.M.J. and W.Y. Kamaruzaman, Mapping land use/cover distribution on a mountainous tropical island using remote sensing and GIS, *International Journal of Remote Sensing*, 22 (10), 1909-1918, 2001.

7. Barbosa, P.M., D. Stroppiana, J.M. Gregoire, and J.M.C. Pereira, An assessment of vegetation fire in Africa (1981-1991): Burned areas, burned biomass, and atmospheric emissions, *Global Biogeochemical cycles* 13(4), 933-950, 1999.

8. Benjamin, S., White J.M., Argiro D., and K. Lowell, Gap analysis – a landscape approach to biodiversity planning (Eds H.M. Scott, T.H. Tear, F.W. Davis), American Society for Photogrammetry and Remote Sensing, Bethesda, MD, 279-288, 1996.

9. Bruniquel-Pinel V. and J.P. Gastellu-Etchegorry, Sensitivity of texture of high resolution images of forest to biophysical and acquisition parameters, Remote Sensing of Environment, 65, 61-85, 1998.

10. Burrough, P.A., Accuracy and error in GIS, In Green, D.R. y D Rix (ed.), The AGI Sourcebook for Geographic Information Systems 1995, AGI, London, 87-91, 1994.

11. Büttner, G, and G. Maucha, The thematic accuracy of CORINE Land Cover 2000: Assessment using using LUCAS, EEA Technical Report/No7/2006, http://reports.eea.europa.eu/ , accessed 04/2007, 2006.

12. Berk A., Bernstein L.S., and D.C. Robertson, MODTRAN: A moderate resolution model for LOWTRAN7, GL-TR-89-0122, Geophysical Laboratory, Bedford, MA, USA, 38p., 1989.

13. Card, A., Using known map category marginal frequenties to improve estimates of thematic map accuracy, *Photogrammetric Engineering and Remote Sensing*, 48(3), 431-439, 1982.

14. Chen, J.M., and Cihlar, J., Retrieving leaf area index of boreal conifer forests using Landsat TM images, *Remote Sensing of Environment*, 55, 153-162, 1996.

15. Cheng, T., Molenaar M., and H. Lin, Formalizing fuzzy objects from uncertain classification results, *International Journal of Geographical Information Science* 15 (1), 27-42, 2001.

16. Clark M.L., Roberts D.A., and D.B. Clark, Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales, *Remote Sensing of Environment*, 96, 375-398, 2005.

17. Cochran, W.G., Sampling Techniques, 3rd ed. John Wiley and Sons, New York, 428p., 1977.

18. Colombo, R., Bellingeri D., Fasolini D., and C.M. Marino, Retrieval of leaf area index in different vegetation types using high resolution satellite data, *Remote Sensing of Environment*, 86, 120-131, 2003.

19. Congalton, R.G., Comparison of sampling scheme use in generating error matrices for assessing the accuracy of maps generated from remotely sensed data, *Photogrammetric Engineering and Remote Sensing*, 54 (5), 593-600, 1988.

20. Congalton, R.G., and K. Green, A practical look at the sources of confusion in error matrix generation, *Photogrammetric Engineering and Remote Sensing*, 59, 641-644, 1993.

21. Courbaud, B., de Coligny F., and T. Cordonnier, Simulating radiation distribution in a heterogeneous Norway spruce forest on a slope, *Agricultural and Forest Meteorology*, 116 (1), 1-18, 2003.

22. Couturier, S., Mas J.F., López E., Benítez J., Tapia V., and A. Vega, Accuracy Assessment of the Mexican National Forest Inventory map: a study in four eco-geographical areas, *Singapore Journal of Tropical Geography*, submitted.

23. Disney, M., Lewis P., and P. Saich, 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains, *Remote Sensing of Environment*, 100, 114-132, 2006.

24. Dorren, L.K.A., Maier B., and A.C. Seijmonsbergen, Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification, *Forest Ecology and Management*, 183, 31-46, 2003.

25. Equihua, M., Fuzzy clustering of ecological data, *Journal of Ecology*, 78, 519-534, 1990.

26. Equihua, M., Análisis de la vegetación empleando la teoría de conjuntos difusos como base conceptual, *Acta Botánica Mexicana*, 15, 1-16, 1991.

27. Estève F., Traitement d'Images et Simulation d'Observations Satellitaires, end of degree report in : École Nationale Supérieure d'Ingénieurs de Constructions Aéronautiques (ENSICA), Toulouse, France, 1999.

28. Fassnacht, K.S., Cohen W. B., and T.A. Spies, Key issues in making and using satellite-based maps in ecology: a primer, *Forest Ecology and Management*, 222, 167-181, 2006.

29. FGDC, FGDC vegetation classification and information standard, June 3, 1996 draft. Reston, VA: Federal Geographic Data Committee, Vegetation Subcommittee (FGDC-VS), FGDC Secretariat, 1997.

30. Fisher, P. y S. Pathirana, The evaluation of fuzzy membership of land cover classes in the suburban zone, *Remote Sensing of Environment*, 34, 121-132, 1990.

31. Fitzpatrick-Lins, K., Comparison of sampling procedures and data analysis for a land-use and land-cover map, *Photogrammetric Engineering and Remote Sensing*, 47, 343-351, 1981.

32. Foody, G.M., A fuzzy sets approach to the representation of vegetation continuum from remotely sensed imagery: an example from lowland heath, *Photogrammetric Engineering and Remote Sensing*, 55, 221-225, 1992.

33. Foody, G.M., Status of land cover classification accuracy assessment, *Remote Sensing of Environment* 80, 185-201, 2002.

34. García, M.C. and R. Álvarez, TM digital processing of a tropical forest region in south-eastern Mexico, *International Journal of Remote Sensing*, 15 (8), 1611-1632, 1994.

35. Gascon, F., Modélisation physique et traitement d'images de télédétection optiques. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, 169p., 2001.

36. Gascon, F., J.P. Gastellu-Etchegorry, Lefèvre-Fonollosa M.J., and E. Dufrêne, Retrieval of forest biophysical variables by inverting a 3-D radiative transfer model and using high and very high resolution imagery, *International Journal of Remote Sensing*, 25 (12), 5601-5616, 2004.

37. Gastellu-Etchegorry J.P., Demarez V., Pinel V., Zagolski F., Modeling radiative transfer in heterogeneous 3-D vegetation canopies, *Remote Sensing of Environment*, 58, 131-156, 1996.

38. Gastellu-Etchegorry, J.P., Guillevic P., Zagolski F., Demarez V., Trichon V., Deering D., and M. Leroy, Modelling BRF and radiation regime of boreal and tropical forests: I. BRF, *Remote Sensing of Environment*, 68, 281-316, 1999.

39. Gastellu-Etchegorry, J.P., Gascon F., and P. Estève, An interpolation procedure for generalizing a look-up table inversion method. *Remote Sensing of Environment*, 87, 55-71, 2003.

40. Gastellu-Etchegorry, J.P., Martin E., and F. Gascon, DART: a 3-D model for simulating satellite images and surface radiation budget, *International Journal of Remote Sensing*, 25(1), 75-96, 2004.

41. Gastellu-Etchegorry, J.P., Hapke handbook, CESBIO, Toulouse, France (http://www.cesbio .ups-tlse.fr), 15pp., 2006.

42. Gemmel, F., and J. Varjo, Utility of reflectance model inversion versus two spectral indices for estimating biophysical characteristics in a boreal forest test site, *Remote Sensing of Environment*, 68, 95-111, 1999.

43. Goel, N.S., and R.L. Thompson, A snapshot of canopy reflectance models, and a universal model for the radiation regime, *Remote Sensing Reviews*, 18, 197-225, 2000.

44. Gopal S., and C. E. Woodcock, Accuracy of Thematic Maps using fuzzy sets I: Theory and methods, *Photogrammetric Engineering and Remote Sensing* 58, 35-46, 1994.

45. Govaerts Y., and M.M. Verstraete, Raytran: A Monte Carlo ray tracing model to compute light scattering in three-dimensional heterogeneous media. *IEEE Transactions on Geoscience and Remote Sensing*, 36(2), 493-505, 1998.

46. Green, D.R., and W. Hartley, Integrating photo-interpretation and GIS for vegetation mapping: some issues of error, *Vegetation Mapping from Patch to Planet* (Alexander, R. and A.C. Millington, editors), John Wiley & Sons Ltd., 103-134, 2000.

47. Haack, B. and S. Jampoler, Agricultural classification comparisons using Landsat thematic mapper data, *ITC Journal*, 2(4), 113-118, 1994.

48. Haack, S. Deviant logic, fuzzy logic. University of Chicago Press, Chicago, 146p., 1996.

49. Hagen, A., Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science*, 17 (3), 235-249, 2003.

50. Hagen-Zanker, A., Straatman B., and I. Uljee. Further developments of a fuzzy set map comparison approach, *International Journal of Geographical Information Science* 19 (7), 769-785, 2005.

51. Hapke B., Theory of Reflectance and Emittance Spectroscopy, Cambridge, 455 p., 1993.

52. Helmer, E.H., Brown S., and W.B. Cohen, Mapping montane tropical forest successional stage and land use with multi-date Landsat imagery, *International Journal of Remote Sensing*, 21 (11), 2163-2183, 2000.

53. Hely, C., L. Bremond, S. Alleaume, B. Smith, M.T. Sykes, and J. Guiot. Sensitivity of African biomes to changes in the precipitation regime, *Global Ecology and Biogeography*, 15(3), 258-270, 2006.

54. Houghton, R.A., D.L. Skole, C.A. Nobre, J.L. Hackler, K.T. Lawrence, and W.H. Chomentowski. Annual fluxes or carbon from deforestation and regrowth in the Brazilian Amazon, *Nature*, 403 (6767), 301-304, 2000.

55. Hunter, M.L., Jr., *Fundamentals of conservation biology*, Blackwell Scientific, Cambridge 482 p. , 1996.

56. Idso, B. and C.T. de Wit, Light relations in plant canopies, *Applied Optics*, 9, 177-184, 1970.

57. INEGI, Sistema de Clasificación de Tipos de Agricultura y Tipos de Vegetación de México para la Carta de Uso del Suelo y Vegetación del INEGI, escala 1:125 000. Instituto Nacional de Estadística, Geografía e Informática, Aguascalientes, Ags, México, 1980.

58. Jacquemoud, S., Baret F., and J.F. Hanocq, Modelling spectral and bi-directional soil reflectance, *Remote Sensing of Environment*, 41, 123-132, 1992.

59. Jonckheere, I., Fleck S., Nackaerts K., Muys B., Coppin P., Weiss M., and F. Baret, Review of methods for in situ leaf area index determination, Part I. Theories, sensors and hemispherical photography, *Agricultural and Forest Meteorology*, 121, 19-35, 2004.

60. Kailath, T., The divergence and Bhattacharyya distance measures in signal selection, *IEEE Transactions in Communication Theory*, 15, 152-160, 1967.

61. Kimes, D.S. and J.A. Kirchner, Radiative transfer model for heterogeneous 3-D scenes, *Applied Optics*, 21, 4119-4129, 1982.

62. Laba, M., Gregory SK, Braden J et al. Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map. *Remote Sensing of Environment* 81, pp. 443-455, 2002.

63. Landgrebe, D.A., Information extraction principles and methods for multispectral and hyperspectral image data, in *Information Processing for Remote Sensing, Chap. 1*, edited by C.H. Chen, World Scientific Publishing, River Edge, New Jersey, 2000.

64. Lira, J., and Maletti G., A supervised contextual classifier based on a region-growth algorithm, *Computers & Geosciences*, 28, 951-959, 2002.

65. López E, Bocco G, Mendoza M, Velázquez A and Aguirre-Rivera R, Peasant emigration and land-use change at the watershed level: A GIS-based approach in central Mexico, *Agricultural Systems* 90, pp. 62-78, 2006.

66. Lu, D., Moran E., and M. Batistella, Linear mixture model applied to Amazonian vegetation classification, *Remote Sensing of Environment*, 87, 456-469, 2003.

67. Lu, D., Integration of vegetation inventory data and Landsat image for vegetation classification in the western Brazilian Amazon, *Forest Ecology and Management*, 213, 369-383, 2005.

68. Magnussen, S., Boudewyn P., and M. Wulder, Contextual classification of Landsat TM images to forest inventory cover types, *International Journal of Remote Sensing*, 25 (12), 2421-2440, 2004.

69. Maletti, G., Ersbøll B., and K. Conradsen, A contextual classifier that only requires one prototype pixel for each class, *IEEE Transactions on Nuclear Science*, 49 (3), 700-706, 2002.

70. Martin E., DART: modélisation d'images optiques de télédétection. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, 196p., 2006.

71. Mas J.F. and I. Ramírez, Comparison of land use classifications obtained by visual interpretation and digital processing, *ITC Journal*, 3(4), 278-283, 1996.

72. Mas JF, Velázquez A, Palacio-Prieto JL, Bocco G, Peralta A, Prado J Assessing forest resources in Mexico: Wall-to-wall land use/ cover mapping. *Photogrammetric Engineering and Remote Sensing* 68 (10), 966-969, 2002 (a).

73. Mas J.-F., A. Velázquez, J.R. Díaz, R. Mayorga, C. Alcántara, R. Castro, and T. Fernández, Assessing land use/cover change in Mexico, Proceedings of the 29th International Symposium on Remote Sensing of Environment (CD), Buenos Aires, Argentina, 8-12/04/2002, 2002 (b).

74. Mas, JF., Change estimates by map comparison: a method to reduce erroneous changes due to positional error, *Transactions in GIS*, 9(4), 619-629, 2005.

75. Miller, K.E., E. Chang, and N. Johnson, *Defining common ground for the Mesoamerican Biological Corridor*, World Resources Institute, Washington DC 45 pp., 2001.

76. Moreno-Sánchez, R., A fuzzy sets approach to the definition of boundaries for land use planning in Central Mexico. Metropolitan State College of Denver, Boulder, CO., USA, 2000.

77. Myneni, R. B., Asrar G., and E.T. Kanemasu The theory of photon transport in leaf canopies, *Theory and applications of optical remote sensing*, A.G. Eds Wiley, New-York, 167-265, 1989.

78. Nichol, J., Hang L.K., and W.M. Sing, Empirical correction of low Sun angle images in steeply sloping terrain: a slope-matching technique, *International Journal of Remote Sensing*, 27 (3-4), 629-635, 2006.

79. North, P.R.J., Three-dimensional forest light interaction model using a Monte Carlo method, *IEEE Transactions on Geoscience and Remote Sensing*, 34 (4), 946-956, 1996.

80. North, P.R.J., Estimation of fAPAR, LAI, and vegetation fractional cover from ATSR-2 imagery, *Remote Sensing of Environment*, 80, 114-121, 2002.

81. Palacio-Prieto JL, Bocco G, Velázquez A, Mas JF, Takaki-Takaki F, Victoria A, Luna-González L et al. La condición actual de los recursos forestales en México: resultados del Inventario Forestal Nacional 2000, *Investigaciones Geográficas* 43, 183-202, 2000.

82. Pinty, B. and M.M. Verstraete, Extracting information on surface properties from bi-directional reflectance measurements, *Journal Geophysical Research*, 96(D2), 2865-2874, 1991.

83. Pinty, B., Verstraete M.M., and N. Gobron, The effect of soil anisotropy on the radiance field emerging from vegetation canopies, *Geophysical Research Letters*, 25 (6), 797-800, 1998.

84. Pontius R.G., and M.L. Cheuk, A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions, *International Journal of Geographical Information Science*, 20 (1), 1-30, 2006.

85. Potter E., Wood J., and C. Nicholl, Sunscan Canopy Analysis System, User Manual v1.05, Delta-T Devices Ltd, Cambridge, UK (http://www.delta-t.co.uk/manual.html), 1996.

86. Powell, R.L., Matzke N., de Souza C., Clark M., Numata I., Hess L.L., and D.A. Roberts, Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon, *Remote Sensing of Environment*, 90, 221-234, 2004.

87. Power, C., Simms A., and R. White, Hierarchical fuzzy pattern matching for the regional comparison of land use maps, *International Journal of Geographical Information Science*, 15 (1), 77-100, 2001.

88. Remmel TK, Csillag F, Mitchell S, and M. Wulder, Integration of forest inventory and satellite imagery: a Canadian status assessment and research issues. *Forest Ecology and Management* 207, pp. 405-428, 2005.

89. Remmel T.K., and F. Csillag, Mutual information spectra for comparing categorical maps, *International Journal of Remote Sensing*, 27 (7), 1425-1452, 2006.

90. Riaño, D., Chuvieco E., Salas J., and I. Aguado, Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types, *IEEE Transactions on Geoscience and Remote Sensing*, 41 (5), 1056-1061, 2003.

91. Rosenfield G.H. & K. Fitzpatrick-Lins, A coefficient of agreement as a measure of thematic classification accuracy, *Photogrammetric Engineering and Remote Sensing* 52 (2), 223-227, 1986.

92. Ross, J., The radiation regime and the architecture of plant stands. Boston, W. Junk, 1981.

93. Roy, P.S., Ranganath B.K., Diwakar P.G., Vohra T.P.S., Bhan S.K., Singh I.J. and V.C. Pandian, Tropical forest type mapping and monitoring using remote sensing, *International Journal of Remote Sensing*, 12 (11), 2205-2225, 1991.

94. Rzedowski J., Vegetación de México (México, DF: Limusa), 1978.

95. Salovaara, K.J., Thessler S., Malik R.N., and H. Tuomisto, Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery, *Remote Sensing of Environment*, 97, 39-51, 2005.

96. Särndal, C.E., Swensson V., and J. Wretman, *Model-assisted survey sampling*, New-York: Springer-Verlag, 1992.

97. Silbernagel, J. and M. Moeur, Modeling canopy openness and understory gap patterns based on image analysis and mapped tree data, *Forest Ecology and Management* 149, 217-233, 2001.

98. Snedecor, G.W., and W.F. Cochran, Statistical methods, State University Press, Ames, Iowa, 728 pp., 1967.

99. Soares-Filho, B.S., D.C. Nepstad, L.M. Curran, G.C. Cerqueira, R.A. Garcia, C.A. Ramos, E. Voll, A. McDonald, P. Lefebvre and P. Schlesinger, Modelling conservation in the Amazon basin, *Nature* 440:520-523, 2005.

100. Stehman S.V. and R.L. Czaplewski, Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, 64, pp. 331-344, 1998.

101. Stehman, S.V., Statistical rigor and practical utility in thematic map accuracy assessment, *Photogrammetric Engineering and Remote Sensing*, 67, 727-734, 2001.

102. Stehman S.V., Wickham JD, Smith JH, and Yang L Thematic accuracy of the 1992 National Land-Cover Data for the eastern United-States: Statistical methodology and regional results. *Remote Sensing of Environment* 86, pp. 500-516, 2003.

103. Stehman S.V., Wickham J.D., Yang L., and J.H. Smith, Assessing the Accuracy of Large-Area Land Cover Maps: Experiences from the Multi-Resolution Land-Cover

Characteristics (MRLC) Project, *4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Amsterdam, 601-608, 2000.

104. Strahler, A.H., Vegetation canopy reflectance modelling: Recent developments and remote sensing perspectives, *Remote Sensing Reviews*, 15, 179-194, 1996.

105. Svoray, T., and Y. Carmel, Empirical method for topographic correction in aerial photographs, *IEEE Geoscience and Remote Sensing Letters*, 2 (2), 2005.

106. Tang, X., Fang Y., and W. Kainz, Fuzzy topological relations between fuzzy spatial objects, *Lecture Notes in Computer Sciences*, LNAI 4223, 97, Springer-Verlag, Germany 324-333, 2006.

107. Thenkabail, P.S., Enclona, E.A., Ashton, M.S., Legg, C., and M.J. De Dieu, Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests, *Remote Sensing of Environment*, 90, 23-43, 2004.

108. Trichon, V., Ducrot D., and J.P. Gastellu-Etchegorry, SPOT 4 potential for the monitoring of tropical vegetation. A case study in Sumatra, *International Journal of Remote Sensing*, 20, 2761-2785, 1999.

109. Tuomisto, H., Linna A., and R. Kalliola, Use of digitally processed satellite images in studies of tropical rain forest vegetation, *International Journal of Remote Sensing*, 15 (8), 1595-1610, 1994.

110. Velázquez, A., Mas J.-F., Díaz J.R., Mayorga-Saucedo R., Alcántara P.C., Castro R., Fernández T., Bocco G., Escurra E., and J.L. Palacio, Patrones y tasas de cambio de uso del suelo en México, *Gaceta Ecológica*, INE-SEMARNAT, 62, 21-37, 2002.

111. Verhoef, W., Light scattering by leaf layers with application to canopy reflectance modelling: the SAIL model. *Remote Sensing of Environment*, 16, 125-141, 1984.

112. Visser, H., T. de Nijs, The map comparison kit, *Environmental Modelling & Software*: 21, 346-358, 2006.

113. Vogelmann, J.E., Howard S.M., Yang L., Larson C.R., Wylie, B.K., and N. Van Driel, Completion of the 1990s National Land Cover Data set for the conterminous

United States from Landsat thematic mapper data and ancillary data sources, *Photogrammetric Engineering and Remote Sensing*, 67, 650-662, 2001.

114. Wang, F., Improving remote sensing image analysis through fuzzy information representation, *Photogrammetric Engineering and Remote Sensing*, 56 (9), 1163-1168, 1990.

115. Wang, L., Sousa W.P., Gong P., and G.S. Biging, Comparison of IKONOS and Quickbird images for mapping mangrove species on the Caribbean coast of Panama, *Remote Sensing of Environment*, 91, 432-440, 2004.

116. Wickham, J.D., Stehman S.V., Smith J.H., and L. Yang, Thematic accuracy of the 1992 National Land-Cover Data for the western United-States, *Remote Sensing of Environment*, 91, 452-468, 2004.

117. Wilkinson, G. G., Results and implications of a study of fifteen years of satellite image classification experiments, *IEEE Transactions on Geoscience and remote sensing*, 43 (3), 433-440, 2005.

118. Wulder, M.A., Deshka J.A., Gillis, M.A., Luther, J.E., Hall, R.J., Beaudouin, A., and S.E. Franklin, Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program, *Forest Chronicles* 79, 1075-1083, 2003.

119. Wulder, M.A., Franklin S.F., White J.C., Linke J., and S. Magnussen, An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data, *International Journal of Remote Sensing*, 27(4), 663-683, 2006.

120. Xu, B., Gong P., Seto E., and R. Spear, Comparison of gray-level reduction and different texture spectrum encoding methods for land-use classification using a panchromatic IKONOS image, *Photogrammetric Engineering and Remote Sensing*, 69 (5), 529-536, 2003.

121. Zadeh, L., Fuzzy sets. Information and control, 8, 338-353, 1965.

122. Zhu Z., Yang L., Stehman S.V., and R.L. Czaplewski, Accuracy Assessment for the U.S. Geological Survey Regional Land-Cover Mapping Program: New York and

New Jersey Region, *Photogrammetric Engineering & Remote Sensing,* 66, 1425-1435, 2000.

Annexes

<u>Annex A:</u> Oak forest in the El Cerro ejido near to the Cuitzeo lake, Michoacán.

<u>Annex B:</u> Avocado plantation, pine and fir forests in the community of San Juan nuevo, near to the Tancítaro volcano, Michoacán.

Photograph A.1: Annual agricultural fields viewed from the oak forest site, Cuitzeo.

Photograph A.2: The Cuitzeo lake viewed from the oak forest site.

Photograph A.3: Oak forest plot on steep terrain and Cuitzeo lake.

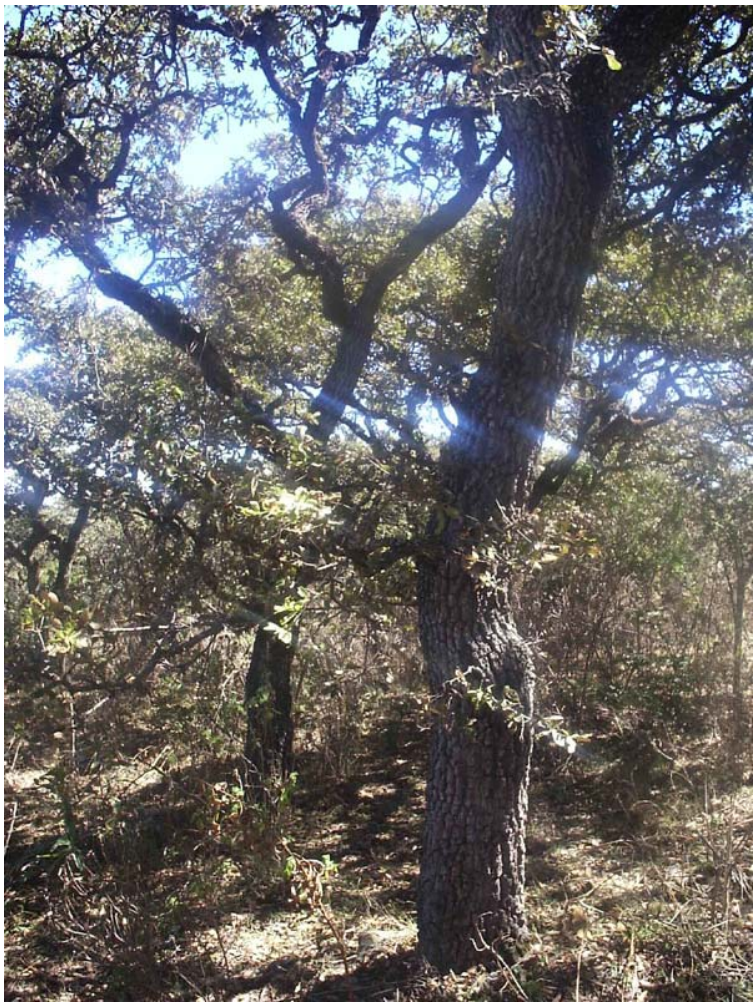
Photograph A.4: Oak (*Quercus Deserticola*) forest plot on steep terrain, Pavka Patiño and Rafael Pompa.

Photograph A.5: Oak (*Quercus Deserticola*) forest plot on steep terrain, canopy structure.

Photograph A.6: Oak (*Quercus Deserticola*) forest plot on flat terrain, canopy structure.

Photograph A.7: Oak (*Quercus Deserticola*) tree on flat terrain, crown structure.

Photograph A.8: Oak (*Quercus Deserticola*) tree on flat terrain and understory.



Photograph A.9: Oak forest plot on steep terrain, soil and understory.

Photograph A.10: Spectroradiometer and ceptometer on soil of oak forest plot, Cuitzeo.

Photograph B.1: Avocado plantation plot, Tancítaro site.

Photograph B.2: Avocado (*Persea Americana*) plantation plot, understory.

Photograph B.3: Pine (*Pinus Montezumae*) forest plot, canopy structure.

Photograph B.4: Pine (*Pinus Montezumae*) forest, understory.

Photograph B.5: Radiometric measurement of pine (*Pinus Montezumae*) forest understory.

Photograph B.6: Radiometric measurement of fir (*Abies Religiosa*) forest understory. Isela Zermeño.

Photograph B.7: Target radiometric measurement over fir (*Abies Religiosa*) forest understory.

Photograph B.8: Reference radiometric measurement over fir (*Abies Religiosa*) forest understory.

Photograph B.9: LAI measurement under fir (*Abies Religiosa*) forest canopy. Moises Méndez.