

Evaluación de confiabilidad del mapa del Inventario Forestal Nacional 2000: diseños de muestreo y caracterización difusa de paisajes

Recibido: 9 de abril de 2007. Aceptado en versión final: 22 de abril de 2008.

Stéphane Couturier*
Álvaro Vega*
Jean-François Mas**
Valdemar Tapia*
Erna López-Granados***

Resumen. Es casi inexistente la evaluación de confiabilidad de los mapas detallados de cobertura vegetal y utilización de suelo (CVUS) en zonas subtropicales. Por la alta biodiversidad y la presencia de clases muy fragmentadas, los diseños de muestreo, aplicados principalmente a zonas templadas, no son transferibles a regiones subtropicales. Este trabajo presenta un método de evaluación de la confiabilidad, a nivel comunitario, para el mapa del Inventario Forestal Nacional (IFN) mexicano del 2000. Para ello, un estudio piloto fue conducido en la cuenca del lago de Cuitzeo de 400 000 ha de superficie, en el cual se probaron varios diseños de muestreo para incluir la totalidad de las clases escasamente representadas (clases escasas) y lograr una buena distribución espacial

del muestreo, manteniendo así bajos costos de operación. Un muestreo doble, en donde la selección de las unidades primarias de muestreo (UPM) se hace separadamente para clases comúnmente distribuidas y clases escasas, dio las mejores características de representación y un total de 2 023 unidades secundarias de muestreo fueron verificadas para la evaluación. Se analizaron las tendencias de confusión entre clases y las perspectivas del método de evaluación.

Palabras claves: Muestreo doble, clase escasa, sistema de clasificación, Landsat, interpretación estereoscópica, fotografía aérea, caracterización difusa.

Accuracy assessment of the National Forest Inventory map of Mexico: sampling designs and the fuzzy characterization of landscapes

Abstract. There is no record so far in the literature of a comprehensive method to assess the accuracy of regional scale Land Cover/ Land Use (LCLU) maps in the sub-tropical belt. The elevated biodiversity 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

*Laboratorio de Análisis Geo-Espacial (LAGE), Instituto de Geografía, Universidad Nacional Autónoma de México, Circuito Exterior, Ciudad Universitaria, Coyoacán, 04510, México, D. F. E-mail: andres@igg.unam.mx

**Centro de Investigaciones en Geografía Ambiental (CIGA), Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro No. 8701, Col. Ex Hacienda de San José de la Huerta, 58190, Morelia Michoacán, México. E-mail: jfmas@ciga.unam.mx

***Instituto de Investigaciones Metalúrgicas, Universidad Michoacana de San Nicolás de Hidalgo. E-mail: erna@ciga.unam.mx

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 (PSU) was done under separate schemes for commonly and scarcely distributed classes, showed best characteristics. A total of 2023 punctual secondary sampling units were verified against their NFI map label. Issues regarding the assessment strategy and trends of class confusions are devised.

Key words: Double sampling, rare class, classification system, Landsat, stereoscopic interpretation, aerial photography, fuzzy classification.

INTRODUCCIÓN

La cartografía del Inventario Forestal Nacional (IFN) 2000 fue generada en el Instituto de Geografía, UNAM, con base en la interpretación visual a escala 1:250 000 de imágenes satelitales Landsat del 2000 (Mas *et al.*, 2002). El proyecto de cartografía se propuso actualizar al 2000 las cartas anteriores de cobertura vegetal y uso de suelo (CVUS) del Instituto Nacional de Estadística, Geografía e Informática (INEGI). Una característica principal del proyecto IFN 2000 fue la propuesta de un esquema jerárquico de clasificación como un estándar conveniente de CVUS para la futura cartografía basada en percepción remota en México. Cuatro niveles de agregación fueron considerados, y el nivel más fino consistió en 75 clases (Palacio *et al.*, 2000), número mucho mayor, por ejemplo, al del nivel 'sub-clase' del National Vegetation Classification System (NVCS, véase FGDC, 1997), de resolución taxonómica similar, empleado en la cartografía de CVUS en los Estados Unidos. Por otro lado, la estrategia empleada para el IFN en México generó productos cartográficos a escalas mucho más pequeñas (1:250 000) que los productos cartográficos en Estados Unidos o Canadá (resolución del orden del píxel de la imagen Landsat). Recientemente se ha expresado la necesidad, por parte de agencias cartográficas de países industrializados, de proporcionar estimaciones estadísticamente robustas de la confiabilidad de sus productos. Por ejemplo,

Laba *et al.* (2002), Wickham *et al.* (2004), para Estados Unidos y Rimmel *et al.* (2005), para Canadá, presentan la evaluación de la confiabilidad de varios mapas de CVUS (de 1992 en Estados Unidos y 2000 en Canadá), basadas en la clasificación de imágenes de Landsat. Los resultados de estos estudios evidencian numerosas confusiones entre clases. Por lo tanto, estos estudios constituyen una valiosa información para el uso de las cartas evaluadas y para inducir estrategias de mejoramiento de la futura cartografía. La cartografía generada en estos últimos países difiere en diversidad taxonómica y resolución espacial con la cartografía del IFN 2000, por lo cual un ejercicio similar para México implica la concepción de una metodología diferente que la empleada en Estados Unidos y Canadá.

En México se realizó una evaluación parcial de confiabilidad inmediatamente después de la elaboración del mapa, con base en la interpretación de fotografías aéreas. Sin embargo, dicha evaluación permitió estimar niveles de confiabilidad sólo para algunas clases, abundantes en la parte norte del país (Mas *et al.*, 2002). Se propone aquí desarrollar una metodología que permita la evaluación generalizada de la confiabilidad del IFN, adaptada a las especificidades de los paisajes de México.

Primero, el alto número de clases a nivel comunidad (75) implica que muchas de estas clases se encuentren escasamente representadas (clases escasas) en el mapa, situación difícil para un muestreo eficiente en términos de costos y estadísticamente válido. Se desarrolló un diseño novedoso para asegurar el muestreo de la totalidad de estas clases escasas, dentro de un protocolo de muestreo probabilístico (*sensu* Stehman y Czaplewski, 1998).

Segundo, la resolución taxonómica fina del mapa enfrenta situaciones típicas de la realidad de los paisajes mexicanos, como son las siguientes:

- ambientes tropicales y templados;
- áreas con actividad agrícola de roza, tumba y quema;
- paisajes muy fragmentados;
- transición temporal entre diferentes tipos de vegetación;
- transición espacial entre diferentes tipos de vegetación (ecotonos);

- tipos de vegetación cuya clasificación es ambigua.

El conjunto de estas situaciones hace muy probable el riesgo de ambigüedades en el proceso de etiquetar polígonos en la imagen Landsat (soporte de la producción del mapa), así como en el proceso de interpretar fotografías aéreas (producción del material de referencia para la evaluación). De forma adicional, la excepcional abundancia de clases de bosque en el IFN agudiza el riesgo de ambigüedades. En consecuencia, como recomendado por Laba *et al.* (2002), se asocia a la evaluación una caracterización difusa ('fuzzy') del paisaje durante la producción del material de referencia. Se propone una versión simplificada del diseño de evaluación lingüística difusa de Gopal y Woodcock (1994).

Se aplicó el método a la evaluación de confiabilidad del mapa de CVUS del IFN 2000 en la cuenca del lago de Cuitzeo (Figura 1), zona que abarca 400 000 ha y 21 clases a nivel comunidad

(Tabla 1). Esa área del país comprende las situaciones de paisajes arriba mencionadas, ilustrando la complejidad del territorio nacional.

En la sección 'Método' se presenta la metodología de la evaluación para el caso de la cuenca del lago de Cuitzeo. Se da especial énfasis en el diseño del muestreo, donde se explican las varias simulaciones que permitieron, después de comparar su factibilidad, destacar una y aplicarla a la zona piloto. Otra sub-sección trata del diseño de verificación, con descripción difusa de polígonos, que permitió distinguir entre verdaderos errores y efectos de fragmentación, definición de clase o escala. Se especifica luego la manera de calcular las estadísticas para formar la matriz de confusión del mapa, tomando en cuenta las características del diseño de muestreo.

Finalmente, en 'Resultados' se discute la eficiencia de cada diseño de muestreo y las razones de la selección de un muestreo óptimo. Se detalla también el patrón principal de confusiones en la cuenca del lago a partir de la información colectada



Figura 1. Ubicación del área piloto del proyecto: la evaluación de confiabilidad se hizo para la zona rellena de gris.

Tabla 1. Distribución de las clases del Inventario Forestal Nacional (nivel comunidad) en la cuenca del lago de Cuitzeo

Clase	Nombre	Formación	Frac. del área	Área (km ²)
100	Agricultura de riego	Agricultura	0.1411	564.97
110	Agricultura de humedad		0.0048	19.04
200	Agricultura de temporal permanente		0.0021	8.27
210	Agricultura de temporal anual		0.2356	943.14
300	Plantación forestal		0.0071	28.24
410	Bosque de oyamel	Bosque templado	0.0037	14.72
420	Bosque de pino		0.0041	16.32
421	Bosque de pino modificado		0.0036	14.31
510	Bosque de pino-encino		0.0958	383.34
511	Bosque de pino-encino modificado		0.0325	130.29
600	Bosque de encino		0.0232	92.88
601	Bosque de encino modificado		0.0553	221.54
700	Bosque mesófilo de montaña	Bosque tropical	0.0029	11.73
920	Matorral sub-tropical	Matorral	0.0194	77.58
921	Matorral sub-tropical modificado		0.0768	307.25
1000	Mezquital		0.0004	1.51
1200	Chaparral ^a			
1330	Pastizal inducido	Pastizal	0.1594	638.04
1410	Popal-tular	Vegetación Higrófila	0.0209	83.50
1510	Vegetación halófila	Otro tipo de vegetación	0.0069	27.78
1600	Sin vegetación aparente ^a	Otro tipo de cobertura		
1700	Asentamiento humano		0.0250	100.02
1800	Cuerpo de agua		0.0796	318.75
Total			1.0000	4003.23

^a Esta clase no está presente en el mapa pero sí en los datos de referencia (aparece en la Tabla 3).

en la matriz de confusión. En las conclusiones se retoman los patrones generales de la evaluación de confiabilidad en Cuitzeo y se discuten las perspectivas del método.

ÁREA EN ESTUDIO Y MATERIALES

El lago de Cuitzeo está ubicado en el Eje Volcánico Transversal al noreste del estado de Michoacán (Figura 1). La cuenca cerrada del lago está cubierta por vegetación templada subhúmeda y tropical

seca. La vegetación original fue bosque de encino, pino-encino y pino hacia el sur-suroeste, y matorral sub-tropical hacia el nor-noreste. En el 2000, de acuerdo con el IFN, algunos usos comunes del suelo fueron la agricultura de temporal anual, la agricultura de riego y el pastizal inducido (Tabla 1).

Además de la diversidad de CVUS, una motivación para seleccionar el área en estudio fue la disponibilidad de una cubierta completa de 244 fotografías aéreas en papel a escala 1:37 000 adquiridas en 1999, siendo casi sincronas con las imágenes Landsat del mapa IFN.

Otros datos de referencia que se reunieron y/o procesaron fueron los siguientes: *a*) 12 cartas topográficas INEGI escaneadas, escala 1:50 000, y *b*) un compuesto a color 541 del mosaico de imágenes Landsat del año 2000 de la cuenca, geo-rectificado a las cartas topográficas 1:50 000 del INEGI.

MÉTODO

El proceso de evaluación de la confiabilidad, representado en un diagrama de flujo (Figura 2), se divide en tres partes:

- *Diseño de muestreo*: selección de los sitios de verificación.
- *Diseño de verificación*: determinación de la categoría correspondiente a los sitios con base en observación de campo o interpretación de fotografías aéreas.
- *Síntesis de evaluación*: comparación entre la información de verificación y la información del mapa a través de la elaboración de las matrices de confusión y del cálculo de índices de confiabilidad.

Diseño del muestreo

El diseño del muestreo tuvo como objetivo la selección de sitios estadísticamente representativos de cada clase del mapa a nivel comunidad. Cinco estrategias fueron contempladas y probadas: una de muestreo aleatorio simple estratificado por clase (Muestreo Simple, o MS) y cuatro muestreos dobles.

Los muestreos dobles fueron operados en dos etapas: primero con la selección de unidades primarias de muestreo (UPM) que corresponden a un área útil fotointerpretable de fotografía aérea y, segundo, con la selección de unidades secundarias de muestreo (USM), que son puntos seleccionados dentro de las UPM seleccionadas en la primera etapa. La Figura 3 representa un mosaico de imágenes Landsat que sirvieron a la elaboración del mapa IFN, geo-rectificadas a las cartas topográficas 1:50 000 del INEGI. El centro de cada fotografía aérea fue primero ubicado en los mapas topográficos

a escala 1:50 000. Para obtener la cuadrícula base visible en la Figura 3, se construyó una malla a partir de los puntos equidistantes a dos centros contiguos de fotografía. La cuadrícula se dibujó juntando con líneas rectas los nodos de esta malla; luego, si su centro caía fuera de la cuenca, se eliminaron los cuadros de la cuadrícula base. Así, el carácter aleatorio de las selecciones fue preservado a la orilla de la cuenca (Zhu *et al.*, 2000). Al final, la cuadrícula base comprendió 218 fotografías (Figura 3).

Los muestreos dobles (MD) se diferenciaron por el modo de selección de una UPM, definiendo la probabilidad de inclusión en el muestreo del área que se representó en el mapa. Por ejemplo, una selección aleatoria simple fue equivalente a una probabilidad de inclusión uniforme (constante) entre las UPM de la cuadrícula base. Las USM son puntos seleccionados de una malla regular cuadrada, generada vía sistema de información geográfica (SIG) dentro de cada UPM. El cálculo de la confiabilidad fue regido por la probabilidad de inclusión de cada punto (USM) verificado (véase subsección Síntesis de Evaluación).

El primer muestreo (MD1) se definió por la selección aleatoria simple de las UPM, y aleatoria, estratificada por clase, de las USM, como en Laba *et al.* (2002). MD2 se caracterizó por una selección estratificada por clase de las UPM, y aleatoria simple de las USM para cada clase, como en Stehman *et al.* (2003), MD3 por una selección proporcional estratificada por clase; la probabilidad de inclusión de una UPM es proporcional a la abundancia de la clase en la UPM (tal abundancia es obtenida a partir de cálculos de atributos en un SIG. Luego, para simplificar el cálculo de confiabilidad, se define la probabilidad de inclusión de las USM como inversamente proporcional a la abundancia de la clase en la UPM. Algunas características de MD3 están descritas en Stehman *et al.* (2000), pero la estrategia de muestreo nunca fue aplicada, probablemente debido a que no era necesaria para sistemas clasificatorios en países templados.

Finalmente, un diseño híbrido (MD4) novedoso incluye una selección aleatoria simple (como en MD1) para clases comunes (fracción de área superior a 5 %, 7 clases), y una selección proporcional

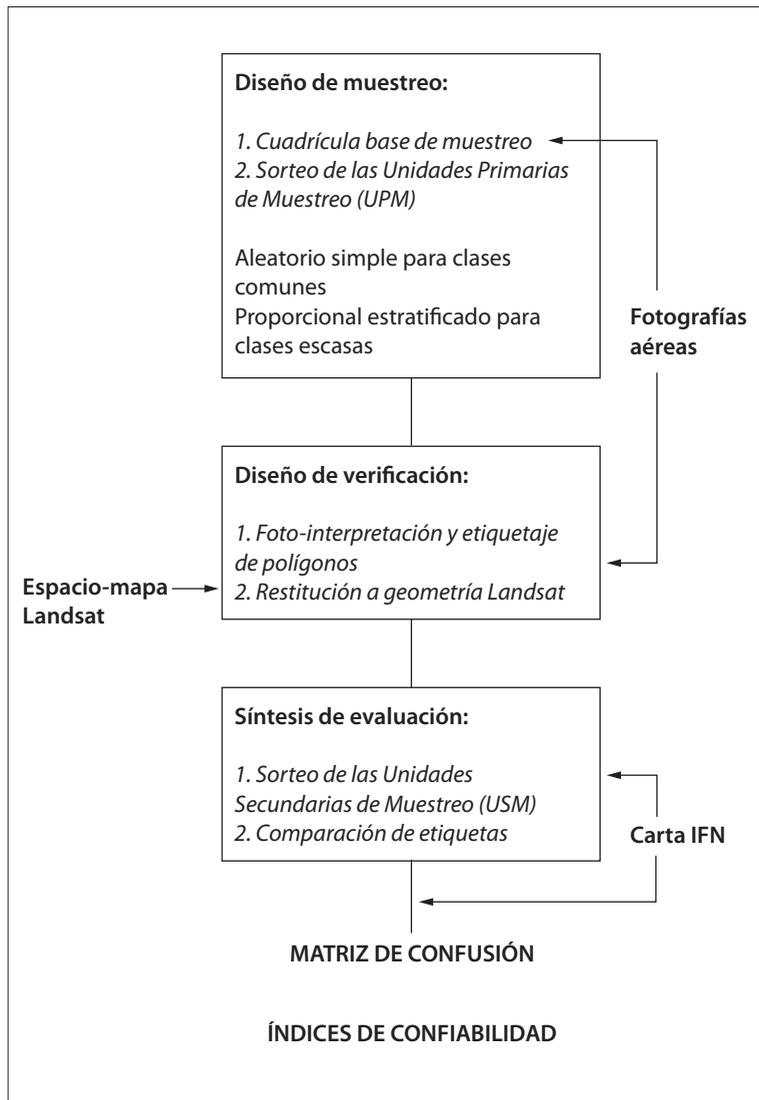


Figura 2. Algoritmo de evaluación del Inventario Forestal Nacional (IFN).

estratificada (como en MD3) para clases escasas (fracción de área inferior a 5 %, 14 clases).

Diseño de verificación

El diseño de verificación se basó en la fotointerpretación difusa de estereo-pares analógicos. Un primer intérprete hizo la delineación de polígonos homogéneos a escala 1:125 000 (escala más grande que la escala del mapa del inventario, y adecuada para foto-interpretación). Posteriormente, la restitución geométrica en pantalla, relativa al mosaico Landsat, permitió minimizar la interferencia de

errores ficticios de posición entre el mapa y el material de referencia. Las Figuras 4 y 5 ilustran, sobre la base geométrica del compuesto a color Landsat, el proceso de delineación de polígonos en una fotografía aérea y el proceso de restitución en pantalla de los polígonos fotointerpretados. Cada punto de muestreo (USM) seleccionado se ubicó primero en el compuesto a color del mosaico Landsat en pantalla, y luego visualmente transferidos en la mica de fotointerpretación.

A cada polígono se le atribuyó una calificación lingüística difusa (fuzzy) de las cuatro clases del

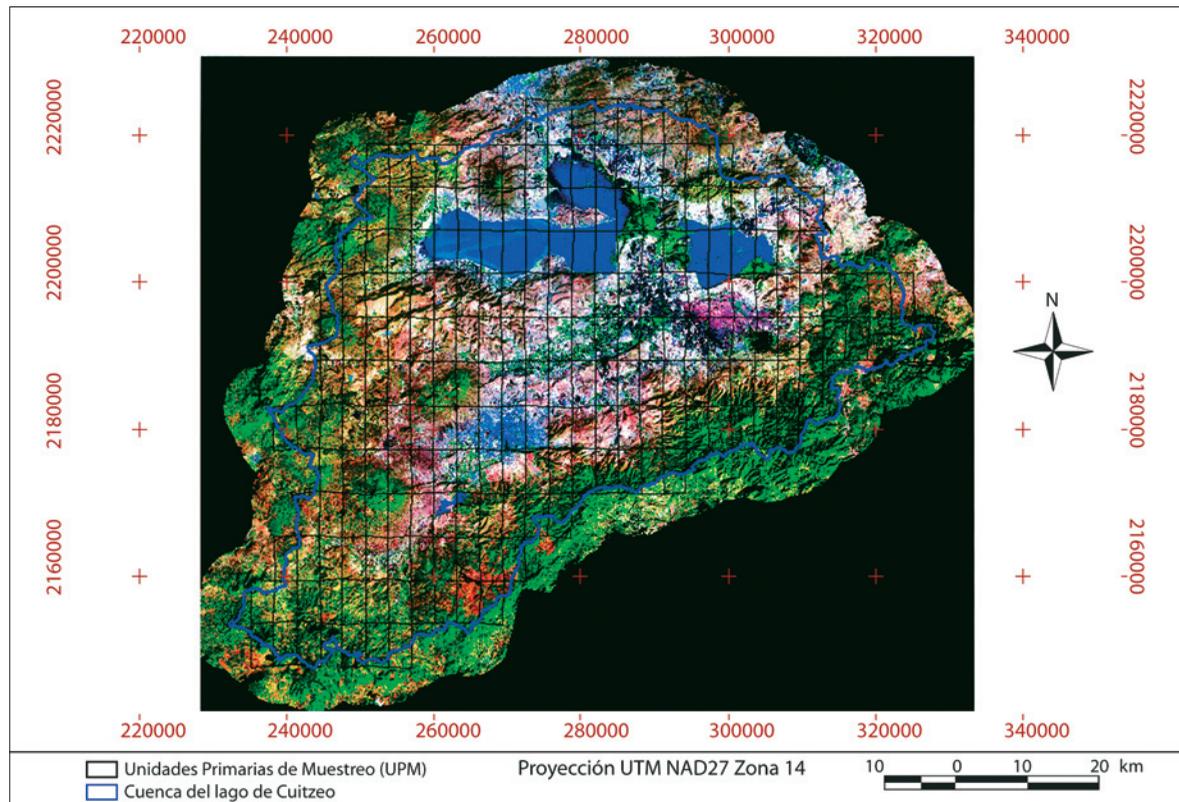


Figura 3. Espacio-mapa de la cuenca del lago de Cuitzeo y cuadrícula base de muestreo doble. El compuesto a color 541 es de un mosaico de imágenes Landsat geo-rectificado a las cartas del INEGI. La cuadrícula irregular de cuadros de aproximadamente 3 por 6 km (un área útil de fotografía aérea), está procesada con base en la posición relativa de los centros de las fotografías aéreas (elaboración propia con imágenes individuales Landsat adquiridas desde el proyecto de elaboración del IFN).

inventario más probablemente representadas en el polígono (categorías de Gopal y Woodcock, 1994). Las calificaciones difusas son las siguientes:

5. Esta clase corresponde perfectamente al polígono.
4. Esta clase corresponde bien al polígono.
3. Esta clase no es la mejor opción pero se encuentra parcial o probablemente en el polígono.
2. Esta clase sería una mala clasificación del polígono.
1. Esta clase no tiene que ver con el polígono.

Esta versión de la calificación lingüística fue fácil de usar para el fotointérprete y compatible con las operaciones automáticas subsiguientes del SIG en el

algoritmo de evaluación. Finalmente, se aplicó un grado de certidumbre a la fotointerpretación:

4. Muy buena certidumbre.
3. Buena certidumbre.
2. Mediana certidumbre.
1. Incertidumbre fuerte.

Por otro lado, una fotointerpretación muy detallada del conjunto de las 244 fotografías aéreas a escala 1:50 000, con la ayuda de información intensiva de terreno, fue conducida de manera independiente por otro fotointérprete (López *et al.*, 2006). De esta manera, cuando el grado de certidumbre del primer fotointérprete era 1 o 2, su interpretación se confirmaba o corregía con la segunda fotointerpretación.

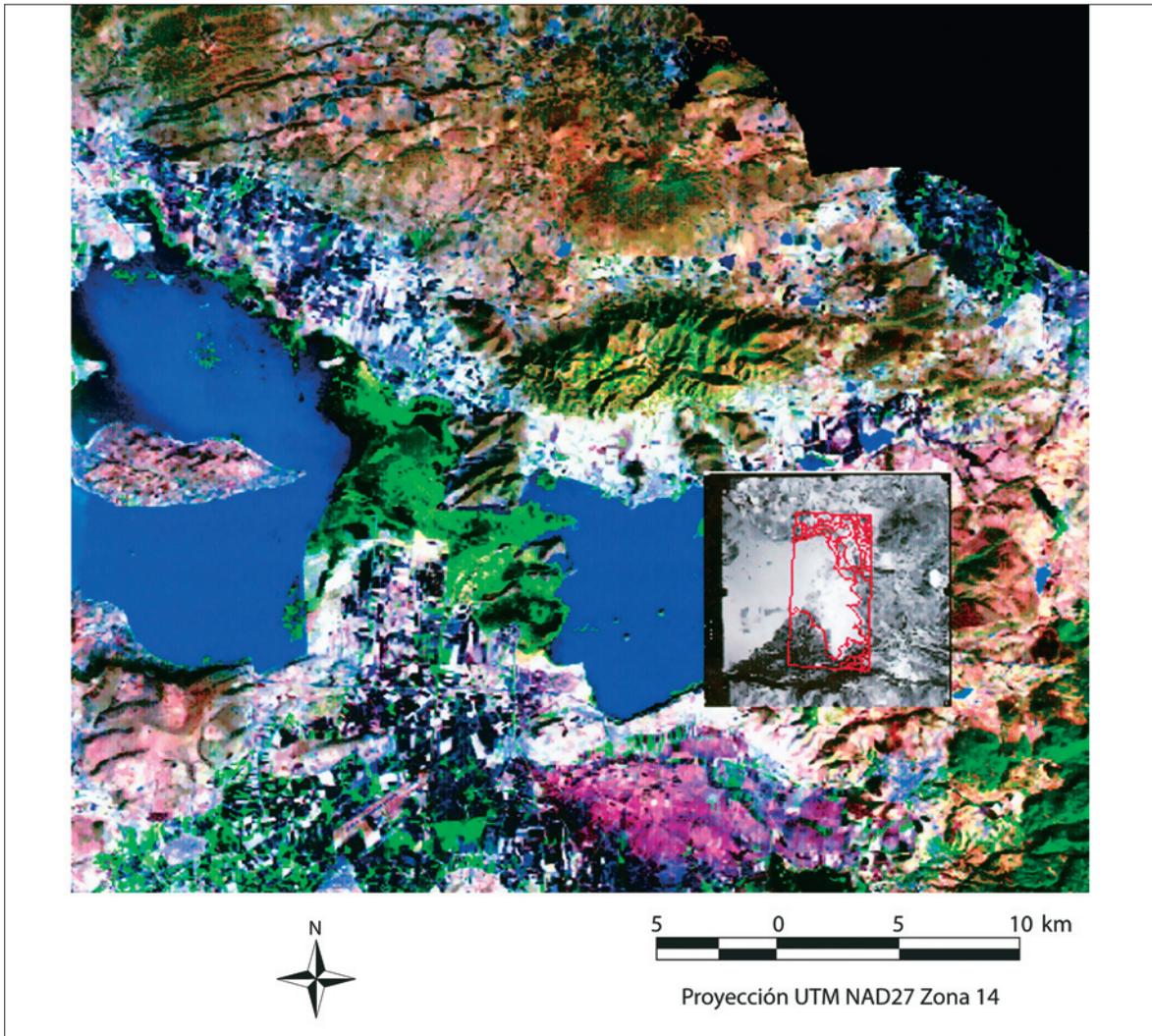


Figura 4. Fotografía aérea de la laguna este de Cuitzeo, segmentada en polígonos de verificación y sobrepuesta en el mosaico de imágenes Landsat. La restitución de la fotointerpretación se hizo con el soporte geométrico de las cartas topográficas del INEGI 1:50 000; el mosaico de imágenes de Landsat está también geo-rectificado a las cartas del INEGI.

Síntesis de evaluación

En todos los puntos (USM) muestreados se hizo una verificación entre las etiquetas difusas de referencia y las etiquetas del mapa IFN vía SIG; para tomar en cuenta la escala del mapa, se aplicó una tolerancia de 500 m entre la posición del USM y la posición de los polígonos cartografiados.

Con el diseño MD4, seleccionado para los resultados de este análisis, se propuso calcular índices de confiabilidad para la evaluación difusa. En el caso de las clases escasas, un diseño doble estratificado fue

aplicado, lo que significa que las probabilidades de inclusión de una USM de la clase k fue la de un muestreo estratificado simple (Stehman *et al.*, 2000): $p_k = f_k$, f_k siendo la frecuencia de presencia de la clase k en el mapa. En consecuencia, la fórmula para muestreo estratificado simple fue utilizada para los estimadores globales. En el caso de una clase común, el cálculo está regido también por la teoría de las probabilidades de inclusión (Snedecor y Cochran, 1967).

La probabilidad de inclusión de un UPM en la primera etapa de muestreo es $p_{1k} = K/N$, donde K

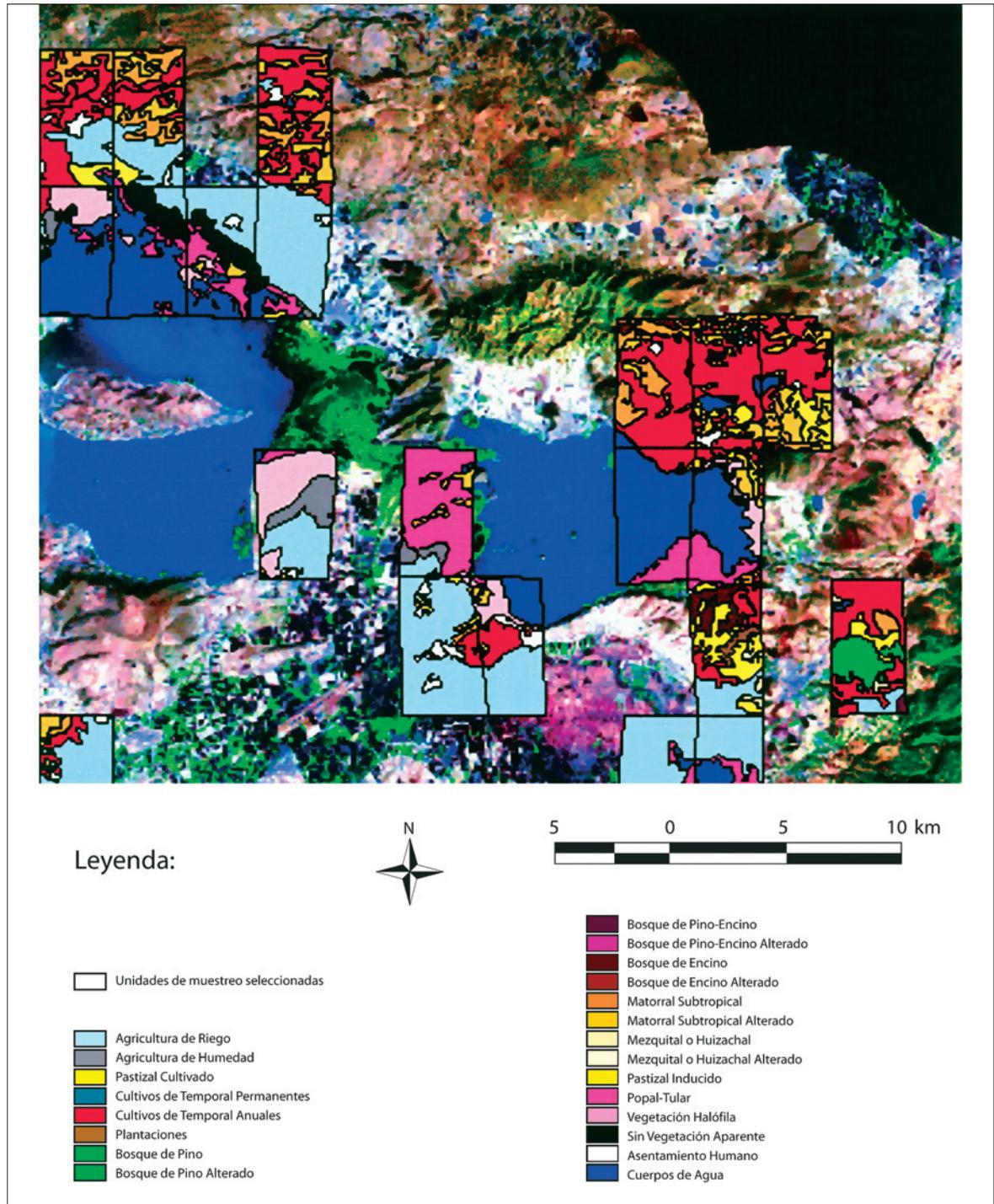


Figura 5. Verificación de Unidades Primarias de Muestreo a partir de la interpretación de fotografías aéreas en la zona este del lago de Cuitzeo. La restitución de la fotointerpretación se hizo con el soporte geométrico de las cartas topográficas del INEGI 1:50 000 y el mosaico de imágenes de Landsat también geo-rectificado a las cartas del INEGI.

es el número de UPM seleccionado para todas las clases comunes, y N es el número total de UPM en la cuadrícula base. En la segunda etapa, la probabilidad de inclusión de cada USM es:

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

donde $p_{2|1}$ es la probabilidad condicional. Así, $p_{2|1}$ es igual a n_k/N'_k donde n_k es el tamaño de muestreo y N'_k es el número de USM de clase k en las K UPM seleccionadas. En consecuencia, pesos diferenciales de $n_k/N'_k * K/N$ fueron atribuidos a clases comunes para calcular los estimadores globales.

Se calculó también la precisión de la estimación. De acuerdo con la teoría de la distribución binomial, el intervalo de confianza del estimador de la confiabilidad para muestreo estratificado simple depende del tamaño de la muestra y de la confiabilidad de la clase en la siguiente forma (Snedecor y Cochran, 1967, citado por Fitzpatrick-Lins, 1981):

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

donde d es el error sobre la confiabilidad (medio intervalo de confianza), n es el número de puntos de muestreo, p es la confiabilidad de la clase, y $t = 1.96$ para una probabilidad de ocurrencia en la extremidad de la curva Gaussiana inferior a 0.05.

Para obtener una precisión aceptable ($d < 0.10$) de acuerdo con la ecuación (2), se fijó el tamaño de muestra (el número de Unidades Secundarias de Muestreo) a $n = 100$ por clase cartografiada (Mas *et al.*, 2003). Este valor de $n = 100$ es un valor típico para estudios de confiabilidad a escala regional con diseños de muestreo doble (Stehman y Czaplewski, 1998).

Sin embargo, la ecuación (2) no toma en cuenta el fenómeno de auto-correlación espacial de los errores en el mapa, si bien la información colectada en dos sitios de verificación lejanos tenderá a ser independiente una de otra, al contrario, la respuesta de dos sitios cercanos será similar (no independiente). Para tomar en cuenta este fenómeno que afecta la muestra, especialmente en el caso de muestreos dobles, se utilizó la fórmula de

la función SURVEY-MEANS del Statistical Analysis Software (SAS, Versión 8.3, 2001), aplicable en el caso de muestreos dobles (Särndal *et al.*, 1992, sección 9.4):

$$d = t * \sqrt{\sum_{k=1}^q \frac{n_k (1 - n_k / N_k)}{(n_k - 1)} \sum_{i=1}^{n_k} (g_{ik} - \bar{g}_{k..})^2} \quad (3)$$

donde:

$$g_{ik} = \frac{\sum_{j=1}^{m_{ik}} w_{ijk} (1 - x_{ijk} P_C)}{\sum_{k=1}^q \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ik}} w_{ijk} x_{ijk}} \quad \text{y} \quad g_{k..} = \left(\sum_{i=1}^{n_k} g_{ik} \right) / n_k \quad (4)$$

Siendo n_k y N_k respectivamente el tamaño de muestra y la población total de la clase k , w_{ijk} la probabilidad de inclusión de la unidad de muestreo i en el conglomerado j para la clase k , y x_{ijk} toma el valor 0 o 1 si el mapa es erróneo o correcto en esta unidad de muestreo. El error estándar así calculado toma en cuenta la varianza de los resultados entre conglomerados, asumiendo una selección estratificada por clase en la segunda etapa del muestreo doble (*Ibid.*). De acuerdo con estas condiciones, fue posible calificar la eficiencia de los diseños de muestreo.

RESULTADOS

Diseños de muestreo

Con el objetivo de establecer 100 USM para las 21 clases, se hicieron diez simulaciones para cada uno de los cinco diseños de muestreo. Algunas características de los resultados para evaluar la eficiencia de cada diseño, basadas en la ecuación (1), están descritas en la Tabla 2.

El método de selección estratificada simple (MS) mostró una dispersión sistemática de los puntos de muestreo sobre más de 85% de las UPM (Tabla 2). La restricción presupuestal del proyecto piloto en Cuitzeo impidió interpretar más de un cuarto del total de fotografías (un cuarto de las 218, o sea 54 fotografías), un monto que sirvió también

de objetivo para probar la capacidad del diseño para mantener costos bajos. Por otra parte, con la ecuación (1) no se permitió un intervalo de confianza de más de 15 % sobre el estimador de confiabilidad, porque en este caso, el tamaño de muestra fue considerado insuficiente para medir la confiabilidad de una clase.

Tomando en cuenta estas restricciones, las simulaciones de MD1 dejaron sistemáticamente tres clases escasas o más (Tabla 2), con un tamaño de muestra no representativo estadísticamente. Este diseño es conocido por la característica ventajosa de generar una gran dispersión del muestreo a pesar de su carácter de muestreo por conglomerados (muestreo doble). Por esta razón, Stehman *et al.* (2003) lo utilizaron para la evaluación de confiabilidad del National Land-Cover Data (NLCD) 1992 de Estados Unidos; sin embargo, fue descartado para la zona piloto del IFN en México porque no pudo atender satisfactoriamente el gran número de clases del sistema clasificatorio.

En el diseño MD2, la selección de las UPM se operó separadamente para cada clase, permitiendo que todas las UPM, donde la clase estuvo presente, tuvieran igual probabilidad de inclusión. En este diseño, si una UPM contenía solamente un punto (USM) de la clase k , la probabilidad de estar seleccionada era igual que otra con 50

puntos de la clase k . En el caso de una clase escasa muy fragmentada, se necesitó entonces un alto número de UPM para alcanzar el tamaño de 100 USM. Por esta razón, las simulaciones mostraron que más de 80 UPM en total fueron necesarias para cubrir las 21 clases, lo que rebasaba la cantidad admitida de fotografías (54). Aunque más viable que MD1, el diseño MD2 resultó inadecuado para mantener costos bajos. Por otra parte, la computación de las probabilidades de inclusión en la segunda etapa del diseño resultó complicada (dependía de la frecuencia de la clase en cada una de las UPM) y no se pudo hacer de manera automática vía SIG.

Por contraste, MD3 es un diseño que mantiene probabilidades de inclusión iguales entre UPM para conservar un nivel de complejidad bajo en el cálculo de índices de confiabilidad (*i.e.* fórmulas estándares de muestreo aleatorio estratificado; Stehman *et al.*, 2000:604); la probabilidad de inclusión de una UPM es proporcional a la abundancia de la clase k en la misma, y la selección de las USM se hace con una probabilidad inversamente proporcional a esta abundancia. Sin embargo, la selección de las UPM, como en MD2, se opera separadamente para cada una de las 21 clases.

Con una restricción de 54 fotografías, fue limitado a tres el máximo de UPM por clase, aplicando

Tabla 2. Evaluación de los resultados de simulaciones para varios diseños de muestreo. MS es el muestreo simple estratificado por clase. MD designa un muestreo doble. MD1 procede a una selección aleatoria simple de las unidades primarias de muestreo. MD2 procede a una selección estratificada por clase, MD3 a una selección proporcional, y MD4 separa clases comunes (a la cuales aplica MD1) y clases escasas (a las cuales aplica MD3)

	Número total de UPM ^a seleccionadas	Número de clases mal representadas (intervalo de confianza superior a 15%)	Número promedio de USM ^b por UPM ^a para clases escasas	Número promedio de USM ^b por UPM ^a para clases comunes
MS ^c	167-193	0	14-20	9-13
MD1	54	3-8	21-25	14-21
MD2	54	2-4	12-23	19-25
MD3	54	0-4	20-33	26-36
MD4	54	0	17-23	16-22

^a unidad primaria de muestreo.

^b unidad secundaria de muestreo.

^c muestreo simple.

así las simulaciones. Con una mayoría de clases escasas, tres UPM fueron suficientes para muestrear y dar un patrón de distribución espacial cercano a su distribución real. Sin embargo, para clases comunes, la limitación a tres UPM generó bastante agregación espacial de puntos con respecto a la distribución general de la clase (el número promedio de USM en una UPM fue de 26-36, muy elevado, véase Tabla 2). Sin embargo, el objetivo de las 100 USM fue, en general, logrado.

Con el afán de resolver la desventaja del diseño MD3, en previsión de una alta auto-correlación espacial de los errores, MD4 fue diseñado como una estrategia novedosa: juntaría las ventajas conocidas de MD1 (estrategia con buena dispersión espacial, empleada en la evaluación del CVUS en los Estados Unidos) para las clases comunes, y las ventajas de MD3 (inclusión de todas las clases y fácil cálculo estadístico) para las clases escasas. De esta manera, la selección de un grupo de UPM exclusivamente para clases comunes permitió una mejor dispersión de estas clases, cercana a la dispersión de MD1 (Tabla 2). Al mismo tiempo, se extendió a más de tres las UPM para aquellas clases escasas que no obtuvieron un tamaño de muestreo suficiente con tres UPM. Con lo anterior, mediante inspección visual, la mejor dispersión de USM se obtuvo con el diseño MD4.

Una secuencia de rutinas en ArcView y Excel-Visual Basic, operando sobre los atributos de los USM, se desarrolló especialmente para implementar automáticamente la selección aleatoria proporcional para ambas etapas. La Figura 6 muestra el conjunto de UPM seleccionadas por MD4 como resultado de una simulación. La Figura 7 muestra la dispersión de las USM en la zona este del lago de Cuitzeo. Estos datos sirvieron al cálculo de la evaluación de confiabilidad.

Índice Global de Confiabilidad

La confiabilidad global y los patrones de confusión entre clases del IFN fueron registrados en forma de una matriz de confusión. Esta matriz aparece en la Tabla 2, con las mismas convenciones sobre índices de confiabilidad que las de Stehman y colaboradores (2003). De hecho, la normalización de los elementos de la matriz a fracciones de área

total permite el análisis de patrones de confusión entre áreas comparables de clases.

El índice global de confiabilidad (IGC) de 81.3 %, refleja principalmente el nivel de confiabilidad de la clase dominante 'cultivo anual' (89 %). Los cuerpos de agua y asentamientos humanos son clases que registraron muy alta confiabilidad. En compuestos a color de imágenes Landsat, la separabilidad espectral entre estas dos clases y otras del sistema clasificatorio es muy alta, lo que contribuye a explicar su alta confiabilidad. Por contraste, hubo alta confusión entre cubiertas no arboladas de vegetación acuática (popal-tular y pasto halófito). La ambigüedad y variabilidad espectral (entre fases de inundación) de estos tipos de vegetación acuática son probablemente factores claves para explicar la alta confusión observada. Examinando las cartas de la serie II del INEGI en el sitio de estos errores, se observó que el IFN 2000 en general reportaba falsos cambios con respecto a la información anterior.

Aportación de la caracterización difusa del paisaje

La confiabilidad de la clase común pino-encino (98 %), y de clases escasas como matorral subtropical (86 %), bosque de encino (95 %), bosque de pino (81 %) y bosque de oyamel (84 %) fueron altas. Así, las clases de vegetación semi-árida y templada poco modificada están caracterizadas por un nivel bajo de confusión. Se tiene como excepciones los dos únicos fragmentos de mezquital y de bosque mesófilo cartografiados en el área, pues ambos correspondieron a otras clases en el terreno (0 % de confiabilidad, doblemente verificados con visitas al campo). Estos errores del IFN aparecieron claramente como omisiones al reportar cambio y rectificar un error, respectivamente, de cartas previas del INEGI.

En contraste con las altas confiabilidades obtenidas para las clases de bosques templados poco modificados, numerosas confusiones fueron registradas entre las clases denominadas bosques templados con vegetación secundaria (más específicamente con vegetación secundaria, herbácea y/o arbustiva, de acuerdo con el esquema de clasificación del IFN, lo cual referiremos como 'bosque modificado').

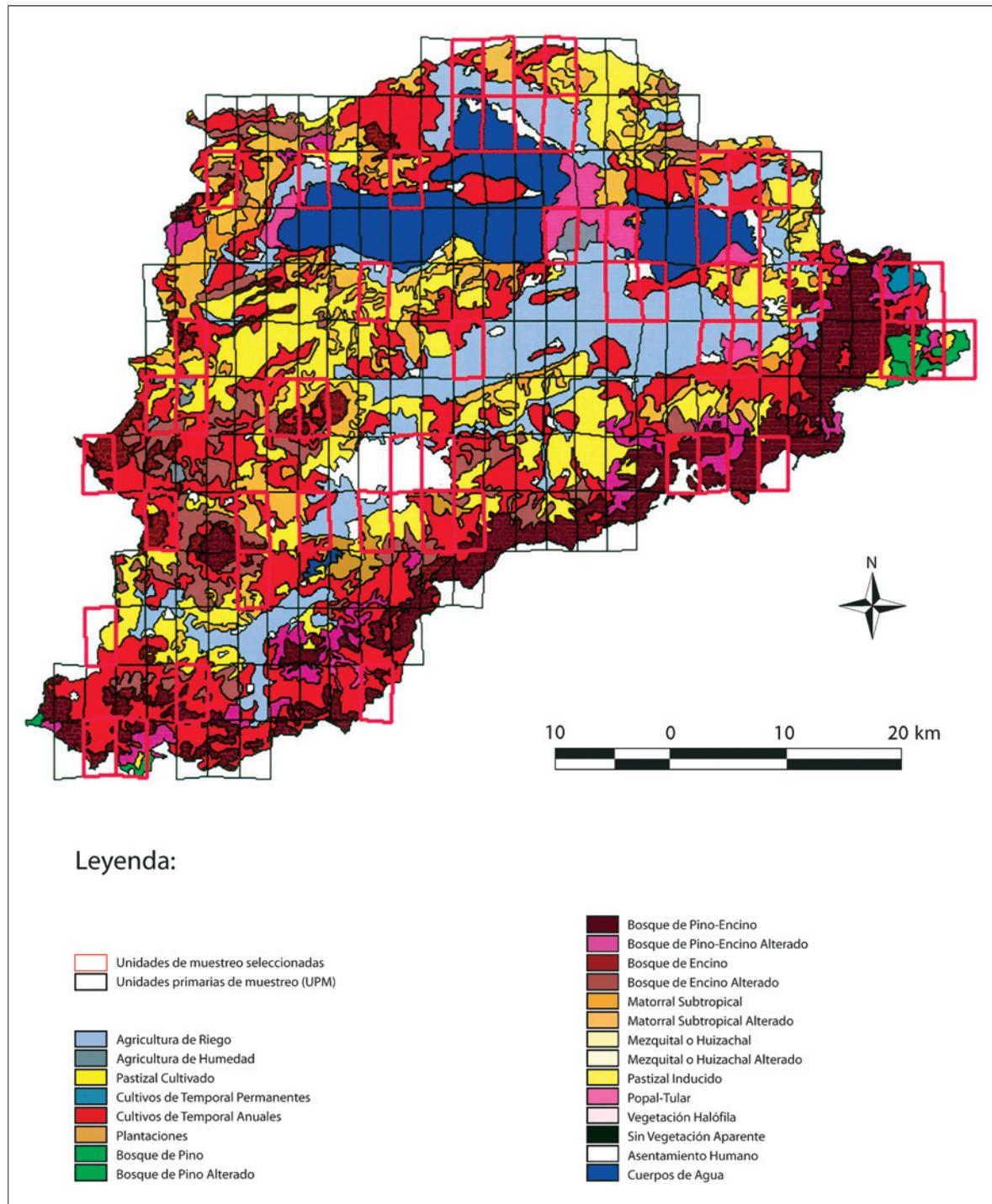


Figura 6. Inventario Forestal Nacional (IFN) de la cuenca del lago de Cuitzeo y Unidades Primarias de Muestreo (UPM) seleccionadas de la cuadrícula base de fotografías, a través del diseño de muestreo MD4. Este diseño está basado en un proceso híbrido para las UPM, aleatorio simple para clases comunes y estratificado proporcional para clases escasas (elaboración propia a partir de la carta nacional del IFN).

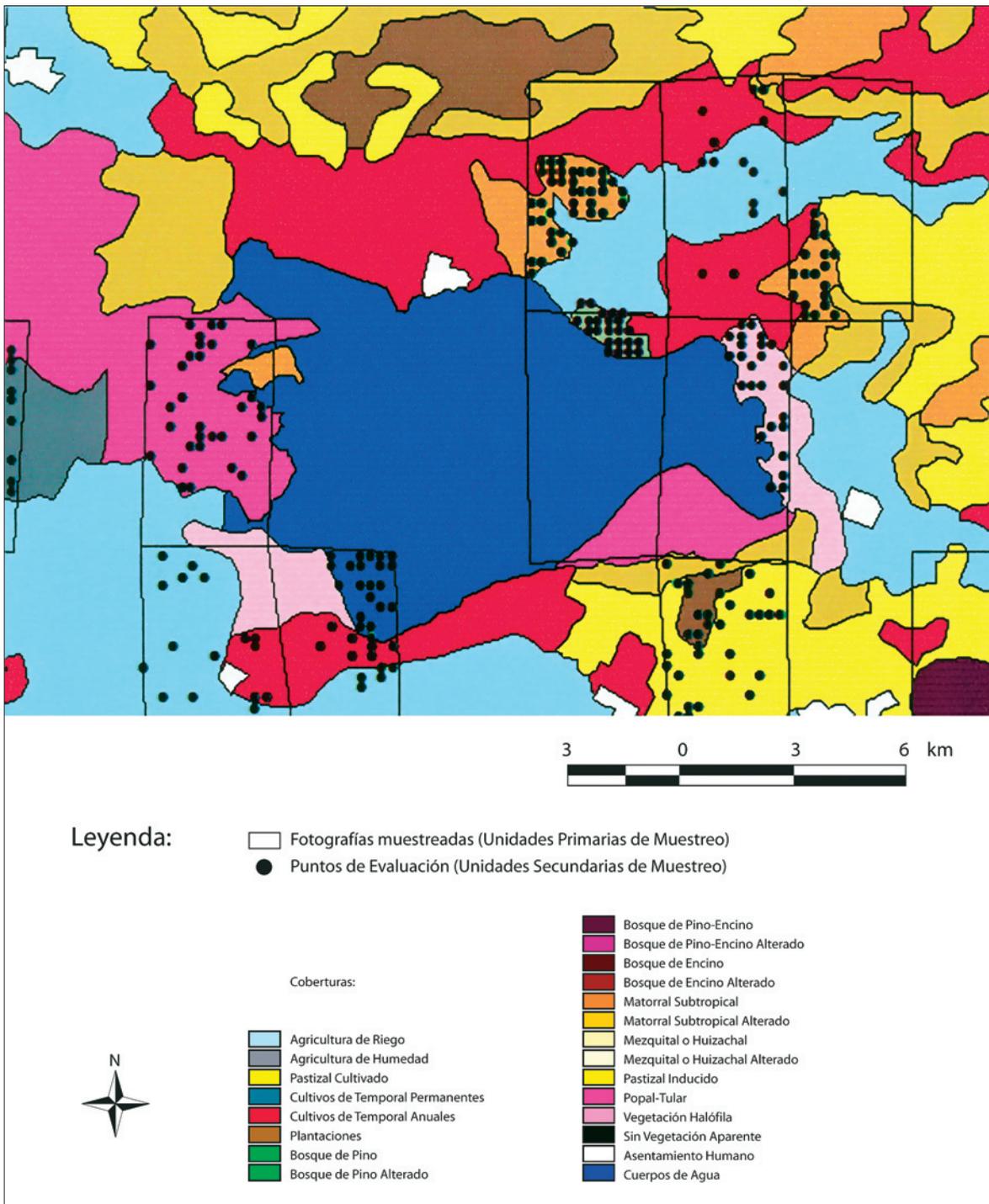


Figura 7. Puntos de evaluación seleccionados como Unidades Secundarias de Muestreo (USM) en las UPMs del diseño de muestreo MD4 en la zona este del lago de Cuitzeo. Este diseño está basado en un proceso híbrido para las USM, estratificado simple para clases comunes y estratificado proporcional para clases escasas (elaboración propia a partir de la carta nacional del IFN).

Tabla 3. Matriz de confusión del mapa del inventario nacional forestal a nivel comunidad en la cuenca del lago de Cuitzeo

	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	Área frac ^b	User ^c	StdE ^d	Binom ^e
100	12.4			0.56															0.28			0.28	0.28	100	14.1	0.88	0.070	0.032
110		0.39		0.04															0.28	0.01		0.01	0.01	100	0.48	0.82	0.065	0.038
200			0.20																	0.01		0.01		100	0.21	0.99	0.005	0.010
210				21.0																0.71	0.24	0.47		100	23.6	0.89	0.035	0.031
300				0.01	0.59				0.04											0.01				100	0.70	0.84	0.091	0.037
410						0.31			0.05											0.01				100	0.37	0.84	0.110	0.037
420				0.01					0.03	0.03														100	0.41	0.81	0.017	0.039
421				0.03					0.01	0.15												0.01	0.01	100	0.36	0.10	0.032	0.030
510									0.10	0.10														100	9.58	0.98	0.016	0.014
511				0.19						1.50	1.20													100	3.25	0.46	0.115	0.050
600				0.05					0.02	2.20														100	2.32	0.95	0.025	0.022
601				0.50	0.11				0.66	0.33		2.99										0.11		100	5.53	0.54	0.124	0.050
700				0.02	0.01				0.22	0.01			0.00											100	0.29	0.00	0.000	0.000
920				0.23									1.67											100	1.94	0.86	0.051	0.035

De acuerdo con los datos de fotointerpretación, extensas superficies de bosque de pino modificado fueron erróneamente cartografiadas como encino modificado, al igual que el bosque modificado mixto (pino-encino) que fue erróneamente cartografiado como pino modificado, lo mismo sucedió con el bosque de encino al ser erróneamente cartografiado como bosque mixto modificado. Estas confusiones explican buena parte de las confiabilidades bajas y muy bajas registradas para bosques de encino modificado (IGC = 54%), pino modificado (IGC = 10%) y mixto modificado (IGC = 46%). Dichas observaciones pueden dar lugar a varias interpretaciones: si se asume que los datos de referencia reflejan la realidad, la presencia más pronunciada de la actividad humana en zonas modificadas hace que una gran proporción de los bosques modificados ocupen zonas de ecotonos más difíciles de definir que las áreas de bosque poco modificado (incertidumbre temática). Otra interpretación surge de la consideración de la digitalización de los bordes entre polígonos contiguos marcando clases de bosques modificados o poco modificados (simplificación de bordes intrincados, incertidumbre posicional). El estudio más preciso de estas interpretaciones en el mapa, permitido por el enfoque difuso, es el objeto de mayor investigación por otra parte (Couturier, 2007).

La estrategia de caracterización difusa de los paisajes probablemente disminuyó el nivel de confusión entre bosques modificados, sin embargo, no al punto de reducir estas confusiones a nivel bajo o de eliminarlas. Esta observación reveló la dificultad de clasificar tipos de vegetación secundaria con las herramientas de percepción remota disponibles para la producción del mapa. Tomados en conjunto, cierto equilibrio resultó de la mayor parte de la confusión generada por estos errores reales o ficticios (resultados de una ambigüedad de definición); las confusiones entre modificado y poco modificado y entre tipos de bosques comprendieron varios círculos de retroalimentación. Los totales en la confiabilidad del productor (*producer's accuracy*) sugieren tendencias generales como la subestimación del bosque de encino poco modificado, y la sobreestimación del bosque de encino

modificado, ambos por alrededor de 3 % del área total del Cuitzeo.

CONCLUSIÓN

Faltan métodos para evaluar la confiabilidad de mapas de Cobertura Vegetal y Utilización de Suelo (CVUS) a escala regional, adaptados a zonas subtropicales. El objetivo del artículo fue presentar un método para la evaluación del mapa del Inventario Forestal Nacional de México (IFN), y proporcionar resultados de su implementación sobre la cuenca del lago de Cuitzeo. Se probaron cinco diseños de muestreo para evaluar el mapa a nivel comunidad, de los cuales, los diseños de muestreo doble lograron mayor control espacial sobre la muestra que el muestreo simple, permitiendo controlar los costos de adquisición e interpretación de fotografías aéreas.

El muestreo doble con selección aleatoria simple para UPM (MD1) tiene la ventajosa característica de generar una gran dispersión del muestreo a pesar de su carácter de muestreo por conglomerados (muestreo doble). Por esta razón, MD1 está empleado por Stehman *et al.* (2003) para la evaluación de confiabilidad del National Land-Cover Data (NLCD) 1992 en EUA; sin embargo, MD1 fue descartado para la zona piloto del IFN en México porque no permitió atender satisfactoriamente el gran número de clases escasamente distribuidas (clases escasas) del sistema clasificatorio. Para la evaluación del mapa a escala nacional, fue también rechazado dada la alta biodiversidad y la presencia de clases muy fragmentadas; el número de clases escasas en el IFN es mucho mayor que en el mapa de EUA, tomando en cuenta la misma resolución taxonómica.

Se propone, por lo tanto, un diseño híbrido que reúne las ventajas conocidas de MD1 (estrategia con buena dispersión espacial) para las clases comunes, y las ventajas de un muestreo proporcional (inclusión de todas las clases y fácil cálculo estadístico) para las clases escasas.

La matriz de confusión resultante de una evaluación con enfoque difuso y del diseño de muestreo mencionado fue proporcionada y su lec-

tura ilustrada con un análisis de paisajes difusos. El índice de confiabilidad global fue alto (81 %). En general, las clases de bosque poco modificado y 'otros tipos de cobertura' fueron representadas con muy alta confiabilidad en el área de Cuitzeo. En contraste, se registró una alta confusión para las clases (en general escasamente distribuidas) de bosque modificado. Este patrón de confusión se debió probablemente a una combinación de ambigüedad generada por la forma de inclusión de la vegetación secundaria en el sistema clasificatorio, y de prácticas de digitalización en el proceso de interpretación visual. Indudablemente, una virtud del esquema clasificatorio propuesto en el proyecto IFN 2000 es su compatibilidad con los esquemas de los mapas anteriores del INEGI. La evaluación fue útil en este caso para estimar el patrón de confusión entre estas clases y las tendencias de los errores.

El método presentado en esta investigación permitió la evaluación de la confiabilidad del mapa en el área piloto de acuerdo con un esquema de muestreo probabilístico. Un reto para la aplicación del diseño al área de Cuitzeo fue la distribución geográfica limitada de los puntos de verificación, que generó una alta autocorrelación espacial entre las muestras, pues la cartografía tiene una escala de 1:250 000 y la zona, en sí, contiene un alto número de clases. Sin embargo, este diseño fue concebido para la evaluación de cualquier área, y especialmente áreas más amplias en el territorio nacional en donde sería menor el efecto de la autocorrelación. En particular, el diseño está listo para aplicarse a escala nacional, con los datos y capacidades típicamente disponibles en México; estratificaciones adicionales serían necesarias a esa escala, por estado administrativo y por la cobertura de un sobrevuelo sistemático (por líneas) del país.

AGRADECIMIENTOS

Esta investigación se realizó en el ámbito de la tesis de doctorado en Geografía del primer autor, quien recibió una beca CONACYT con matrícula 205000

y participó en los proyectos CONACYT núm. 38965T y CONAFORT núm. 47741 coordinados por el tercer autor. Se agradece a varios ejidatarios ancianos del pueblo El Cerro por proveer información sobre dinámicas claves del paisaje en la zona tropical seca de la cuenca. También a Manuel Mendoza y su equipo (proyecto de hidrología) por su ayuda y comentarios en las visitas de campo, y a muchos más de la ciudad de Morelia.

REFERENCIAS

- Couturier, S. (2007), *Evaluación de errores de cartas de cobertura vegetal y usos del suelo con enfoque difuso y con la simulación de imágenes de satélite*, tesis de Doctorado, UNAM/Université Paul Sabatier de Toulouse, Francia (UPS), México.
- FGDC (1997), *FGDC vegetation classification and information standard*, June 3, 1996 draft, Federal Geographic Data Committee, Vegetation Subcommittee (FGDC-VS), FGDC Secretariat, Reston, VA, USA.
- Fitzpatrick-Lins, K. (1981), "Comparison of sampling procedures and data analysis for a land-use and land-cover map", *Photogrammetric Engineering and Remote Sensing*, no. 47, pp. 343-351.
- Gopal, S. and C. E. Woodcock (1994), "Accuracy of thematic maps using fuzzy sets I: theory and methods", *Photogrammetric Engineering and Remote Sensing*, no. 58, pp. 35-46.
- 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*, no. 81, pp. 443-455.
- López-Granados, E., G. Bocco, M. Mendoza, A. Velázquez and R. Aguirre-Rivera (2006), "Peasant emigration and land-use change at the watershed level: a GIS-based approach in central Mexico", *Agricultural Systems*, no. 90, pp. 62-78.
- Mas, J.-F., 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*, no. 68(10), pp. 966-969.
- Mas, J.-F., J. Reyes Díaz-Gallegos y A. Pérez Vega (2003), "Evaluación de la confiabilidad temática de mapas o de imágenes clasificadas: una revisión", *Investigaciones Geográficas, Boletín*, núm. 51, Instituto de Geografía, UNAM, México, pp. 53-72.

- Palacio-Prieto, J. L., G. Bocco, A. Velázquez, J.-F. Mas, F. Takaki-Takaki, A. Victoria, L. Luna-González *et al.* (2000), "La condición actual de los recursos forestales en México: resultados del Inventario Forestal Nacional 2000", *Investigaciones Geográficas, Boletín*, núm. 43, Instituto de Geografía, UNAM, México, pp. 183-202.
- Rommel, T. K., F. Csillag, S. Mitchell and M. Wulder (2005), "Integration of forest inventory and satellite imagery: a Canadian status assessment and research issues". *Forest Ecology and Management*, no. 207, pp. 405-428.
- Särndal, C. E., V. Swensson and J. Wretman (1992), *Model-assisted survey sampling*, Springer-Verlag, New-York.
- Snedecor, G. W. and W. F. Cochran (1967), *Statistical methods*, State University Press, Ames, Iowa.
- Stehman, S. V. and R. L. Czaplewski (1998), "Design and analysis for thematic map accuracy assessment: fundamental principles", *Remote Sensing of Environment*, no. 64, pp. 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*, no. 86, pp. 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", *4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Amsterdam, pp. 601-608.
- Wickham, J. D., S. V. Stehman, J. H. Smith and L. Yang (2004), "Thematic accuracy of the 1992 National Land-Cover Data for the western United-States", *Remote Sensing of Environment*, no. 91, pp. 452-468.
- 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*, no. 66, pp. 1425-1435.