

## Unas herramientas de uso libre para evaluar la fiabilidad temática de datos espaciales

Jean-François Mas<sup>1</sup>  
Azucena Pérez-Vega<sup>2</sup>  
Adrián Ghilardi<sup>1</sup>  
Silvia Martínez<sup>1</sup>  
Jaime Octavio Loya-Carrillo<sup>1</sup>  
Ernesto Vega<sup>3</sup>

<sup>1</sup> Centro de Investigaciones en Geografía Ambiental,  
Universidad Nacional Autónoma de México,  
Antigua Carretera a Pátzcuaro 8701, Col Ex-Hda de San José de La Huerta,  
58190 Morelia, Mich, México  
jfmas, aghilardi@ciga.unam.mx, silviamrz@gmail.com

<sup>2</sup> Dirección de la División de Ingenierías, Universidad de Guanajuato,  
Avenida Juárez 77, Zona Centro, 36000 Guanajuato, Gto, México  
azu\_pvega@hotmail.com

<sup>3</sup> Centro de Investigaciones en Ecosistemas,  
Universidad Nacional Autónoma de México,  
Antigua Carretera a Pátzcuaro 8701, Col Ex-Hda de San José de La Huerta,  
58190 Morelia, Mich, México  
evega@cieco.unam.mx

**Abstract.** Thematic maps are widely used in land management, however few studies assess its accuracy using robust methods. Accuracy assessment is usually performed by comparing the map category with the category assigned to reference sites through the analysis of high-resolution images or fieldwork. These reference sites are selected based on sampling and the value and variance of the estimators of the accuracy indices depend on the sampling design. However, the tools available in the main GIS and image processing packages computed these indices without taking into account the type of sampling and give inconsistent estimates. We present a suite of free tools that enables users to calculate accuracy and evaluate the areas covered by certain categories with their respective confidence intervals. These tools were implemented in the package Dinamica-EGO, a platform for spatial modeling as well as a plugin of the open source geographic information system Q-GIS and the open source statistical analysis package R. We present an application of these tools to the evaluation of a land use / land cover map.

**Keywords:** thematic maps, accuracy assessment, confusion matrix mapas temáticos, evaluación de la fiabilidad, matriz de confusión.

### 1. Introducción

Aunque los mapas temáticos como, por ejemplo, los mapas de uso / cubierta del suelo son ampliamente utilizados en el diseño de políticas medioambientales, en la gestión del territorio o en trabajos académicos, pocos estudios incluyen la evaluación de la fiabilidad de los mapas utilizados mediante métodos robustos y completos (Olofsson et al., 2013). La evaluación de la fiabilidad de los mapas temáticos se realiza normalmente mediante la comparación de la categoría del mapa con la categoría asignada a sitios de referencia (o de verificación), la cual es obtenida a través de trabajo de campo o bien del análisis de imágenes de muy alta resolución espacial. Para seleccionar estos sitios de referencia, se utiliza una amplia gama de diseños

de muestreo, incluyendo muestreos aleatorios simples, sistemáticos simples o bien muestreos aleatorios o sistemáticos estratificados, entre otros.

Generalmente, se elabora una matriz de confusión o matriz de error que es una tabla de doble entrada en la cual se reporta el número de sitios de verificación para cada combinación de categoría en el mapa y en los datos de referencia. Es común calcular directamente los índices de fiabilidad con base en esta matriz, sin embargo eso es válido únicamente si el número de sitios por categoría del mapa es proporcional a la superficie de esta categoría en el mapa como en el caso de un muestreo aleatorio simple. En caso de muestreos que no conservan necesariamente esta proporción como los muestreos estratificados, es necesario llevar a cabo un ajuste de los valores de la matriz antes de calcular los índices (Card, 1982). Es importante también evaluar la certidumbre del valor estimado de estos índices a través la construcción de un intervalo de confianza. Por ejemplo, una estimación de la fiabilidad global de un mapa de 70% con un intervalo de confianza de 30% es de poca utilidad debido a que la fiabilidad puede fluctuar entre 40 y 100%. En este caso sería por lo tanto necesario aumentar el número de sitios de verificación para reducir el intervalo de confianza. Finalmente, los errores de cartografía afectan la estimación de las áreas cubiertas por las diferentes categorías. Por ejemplo, el área de una categoría que presenta muchos errores de omisión es subestimada en las estadísticas directamente obtenidas del mapa. Con base en la matriz de confusión, es posible estimar las áreas de las diferentes categorías compensando los sesgos relacionados con los errores de clasificación.

Las herramientas para evaluar los índices de fiabilidad disponibles en los principales paquetes de SIG y de procesamiento de imágenes de satélite calculan los índices de fiabilidad sin tomar en cuenta el tipo de muestreo y dan estimaciones erróneas. Además, no permiten calcular los intervalos de confianza de dichos índices ni tampoco de llevar a cabo la corrección de las estadísticas de área. Como alternativa, presentamos en este artículo unas herramientas de uso libre que permiten a los usuarios calcular los índices de fiabilidad y evaluar las áreas cubiertas por determinadas categorías con sus respectivos intervalos de confianza. Presentamos un ejemplo de aplicación práctica de la herramienta de Q-GIS en la evaluación de un mapa de uso/cubierta del suelo.

## 2. Metodología

Estas herramientas se implementaron en tres ambientes i) el programa Dinamica EGO, una plataforma gratuita para modelación espacial ([www.csr.ufmg.br/dinamica](http://www.csr.ufmg.br/dinamica)), ii) el sistema de información geográfica de código abierto Q-GIS ([www.qgis.org](http://www.qgis.org)) y iii) como paquete del software de análisis estadístico R (también de código abierto) ([www.r-project.org](http://www.r-project.org)).

### 2.1. Evaluación de la fiabilidad

Si consideramos la matriz de confusión con  $q$  columnas y  $q$  líneas (Tabla 1), en la cual cada celda representa el número  $n_{ij}$  de sitios de verificación identificados como de la categoría  $i$  y  $j$  respectivamente en el mapa y en los datos de referencia obtenidos en campo o usando datos muy fiables para determinar la categoría. Dependiendo del tipo de muestreo utilizado, el número de sitios para cada categoría del mapa no es necesariamente proporcional a la superficie de estas categorías en el mapa. En otras palabras, ciertas categorías son sobre-representadas en el muestreo respecto a su representación en el mapa, otras, al contrario, son sub-representadas en el muestreo. Card (1982) propone un método para ajustar la matriz bruta (la matriz basada en el número de sitios) reemplazando el valor  $n_{ij}$  de cada celda por  $\hat{p}_{ij}$  que es un estimador no sesgado de la proporción del área (ecuación 1).

Table 1: Matriz de confusión en número de sitios de verificación.

Mapa	Referencia						Suma
	1	2	...	j	...	q	
1	$n_{11}$	$n_{12}$	...	$n_{1j}$	...	$n_{1q}$	$n_{1+}$
2	$n_{21}$	$n_{22}$	...	$n_{2j}$	...	$n_{2q}$	$n_{2+}$
...	...	...	...	...	...	...	...
i	$n_{i1}$	$n_{i2}$	...	$n_{ij}$	...	$n_{iq}$	$n_{i+}$
...	...	...	...	...	...	...	...
q	$n_{q1}$	$n_{q2}$	...	$n_{qj}$	...	$n_{qq}$	$n_{q+}$
Suma	$n_{+1}$	$n_{+2}$	...	$n_{+j}$	...	$n_{+q}$	

$$\hat{p}_{ij} = \frac{\pi_i n_{ij}}{n_{i+}} \tag{1}$$

donde  $\pi_i$  es la proporción del área total cubierta por la categoría i en el mapa,  $n_{ij}$  el número de sitios de verificación cartografiados como i pero identificados como j en los datos de referencia, y  $n_{i+}$  el número de sitios mapeados como i en el mapa. En la matriz ajustada, cada elemento  $\hat{p}_{ij}$  representa la probabilidad que una área seleccionada de forma aleatoria sea clasificada como categoría i siendo realmente j en los datos de referencia. La suma de las celdas de cada fila  $\hat{p}_{i+}$  es por lo tanto igual a  $\pi_i$ , proporción de la categoría i en el mapa. Basado en esta matriz, el cálculo de la fiabilidad global, del usuario y del productor se llevan a cabo con base en las ecuaciones 2, 3 y 4. Este cálculo compensa eventuales sesgos relacionados con el muestreo.

La fiabilidad global  $\hat{O}$  es la proporción de área correctamente clasificada y se calcula sumando los elementos de la diagonal de la matriz ajustada (ecuación 2).

$$\hat{O} = \sum_{K=1}^q \hat{p}_{kk} \tag{2}$$

donde q es el número de categorías.

Las fiabilidades del usuario  $\hat{U}_i$  y del productor  $\hat{P}_j$  son respectivamente calculadas utilizando las ecuaciones 3 y 4. La fiabilidad del productor, relacionada a los errores de omisión, expresa la proporción de sitios de referencia de una categoría que es correctamente clasificada en el mapa. La fiabilidad del usuario, relacionada a los errores de comisión, es la proporción de sitios mapeados como cierta categoría que fueron correctamente clasificados.

$$\hat{U}_i = \frac{\hat{p}_{ii}}{\hat{p}_{i+}} \tag{3}$$

$$\hat{P}_j = \frac{\hat{p}_{jj}}{\hat{p}_{+j}} \tag{4}$$

Para muestreos estratificados, los intervalos de confianza de las estimaciones de la fiabilidad global, del productor y del usuario son calculados con base en las ecuaciones 5 a 7 (Card, 1982):

$$HCI_{\hat{O}} = z \sqrt{\sum_{i=1}^q \frac{\hat{p}_{ii}(\pi_i - \hat{p}_{ii})}{n_{i+}}} \quad (5)$$

donde  $HCI_{\hat{O}}$  es el medio intervalo de confianza de la fiabilidad global y  $z$  corresponde al número de desviación estándar de una distribución normal (para un nivel de confianza de 95%,  $z = 1.96$ ).

$$HCI_{\hat{U}_i} = z \sqrt{\frac{\hat{p}_{ii}(\pi_i - \hat{p}_{ii})}{\pi_i^2 n_{i+}}} \quad (6)$$

donde  $HCI_{\hat{U}_i}$  es el medio intervalo de confianza de la fiabilidad del usuario para la categoría  $i$ .

$$HCI_{\hat{P}_j} = z \sqrt{\hat{p}_{jj}\hat{p}_{+j}^{-4} \left[ \hat{p}_{jj} \left( \sum_{i \neq j}^q \hat{p}_{ij}(\pi_i - \hat{p}_{ij})/n_{i+} \right) + (\pi_j - \hat{p}_{jj})(\hat{p}_{+j} - \hat{p}_{jj})^2/n_{j+} \right]} \quad (7)$$

donde  $HCI_{\hat{P}_j}$  es el medio intervalo de confianza de la fiabilidad del productor para la categoría  $j$ .

## 2.2. Evaluación de área

Debido a los errores asimétricos, el área de una categoría en el mapa puede presentar sesgos (sobre o subestimación) (Stehman, 1995; Stehman y Czaplewski, 1998). Por ejemplo, el área de una categoría sistemáticamente afectada por errores de comisión será sobreestimada. Por lo tanto, las estimaciones de área derivadas del mapa deben ajustarse para tomar en cuenta estos sesgos. Estas estimaciones de área deben además ser acompañadas por una estimación de su certidumbre (intervalos de confianza). En la matriz de confusión ajustada, la suma  $p_{+j}$  de los elementos de la columna  $j$  es un estimador sin sesgo de la proporción de la categoría  $j$ . Por lo tanto,  $A_j$  el área de la categoría  $j$  se obtiene con base en la ecuación 8 (Card, 1982; Olofsson et al., 2013; Stehman, 2013).

$$\hat{A}_j = \hat{A}_{tot} \hat{p}_{+j} \quad (8)$$

donde  $\hat{A}_{tot}$  es el área total mapeada. La ecuación 9 permite calcular el medio intervalo de confianza para la proporción  $\hat{p}_{+j}$

$$HCI_{\hat{p}_{+j}} = z \sqrt{\sum_{i=1}^q \left( \frac{\hat{p}_{ij} n_{i+}}{n_{ij}} \right)^2 \frac{n_{ij}/n_{i+}(1 - n_{ij}/n_{i+})}{n_{i+} - 1}} \quad (9)$$

## 3. Resultados

Para ilustrar el procedimiento de evaluación de la fiabilidad con el complemento de QGIS tomamos el ejemplo de un mapa de cubierta / uso del suelo con cuatro categorías (Figura 1). Se pueden observar que algunas categorías cubren más superficie que otras en el mapa. En nuestro ejemplo, las cuatro categorías presentan respectivamente las proporciones 0.3294, 0.4340, 0.1505 y 0.0861. Se seleccionaron 80 puntos de verificación con base en un muestreo aleatorio estratificado con base en las categorías del mapa (20 sitios por categorías). En la figura 1, se representan los sitios de verificación por pequeños cuadrillos, el color dentro del

cuadrado indican la categoría determinada en campo (información de referencia).

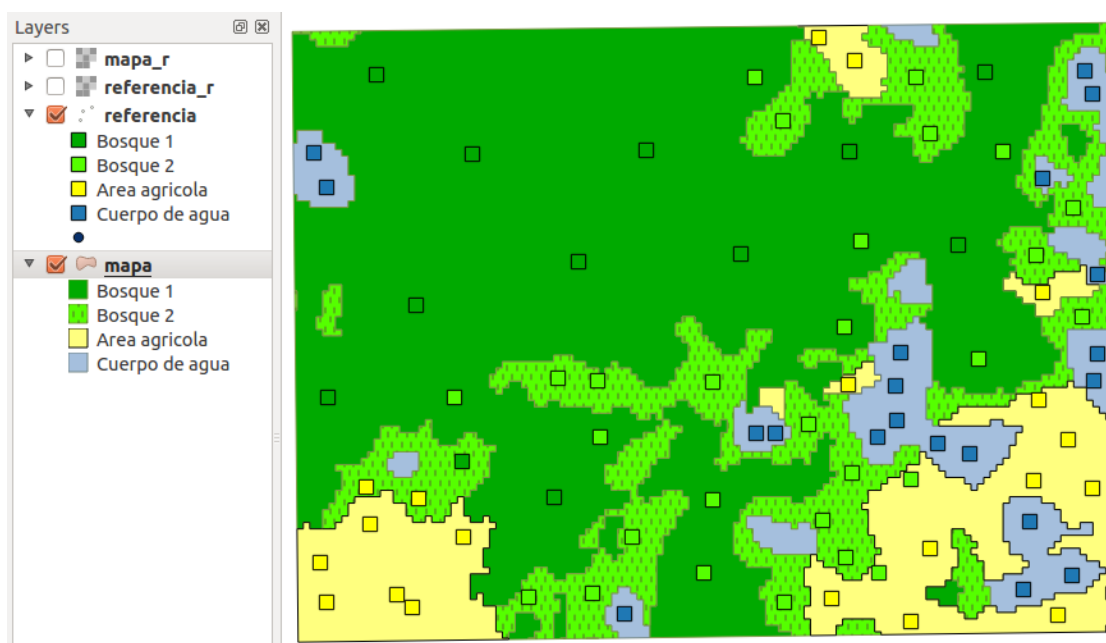


Figure 1: Mapa evaluado y sitios de verificación.

Las tablas 3 y 3 corresponden respectivamente a las matrices de confusión bruta y ajustada. Las tablas 4 y 5 muestran respectivamente los índices de fiabilidad y las áreas corregidas tomando en cuenta los errores reportados en la matriz de confusión.

Table 2: Matriz de confusión bruta (números de sitios de verificación).

		Referencia			
Mapa		1	2	3	4
1		11	9	0	0
2		1	17	2	0
3		0	2	18	0
4		0	0	0	20

Table 3: Matriz de confusión ajustada (proporción del área).

		Referencia				
Mapa		1	2	3	4	Suma
1		0.3201	0.2619	0.0000	0.0000	0.5820
2		0.0093	0.1574	0.0185	0.0000	0.1852
3		0.0000	0.0147	0.1320	0.0000	0.1467
4		0.0000	0.0000	0.0000	0.0861	0.0861
Suma		0.3294	0.4340	0.1505	0.0861	1.0000

Table 4: Índices de fiabilidad por categoría.

Categoría	1	2	3	4
Fiabilidad del usuario	0.5500	0.8500	0.9000	1.0000
Medio IC	0.2180	0.1565	0.1315	0.0000
Fiabilidad del productor	0.9719	0.3627	0.8770	1.0000
Medio IC	0.0533	0.1154	0.1427	0.0000

Table 5: Estimaciones de las áreas de cada categoría.

Categoría	1	2	3	4
Proporción en el mapa	58.20	18.52	14.67	8.61
Proporción ajustada	32.94	43.40	15.05	8.61
Medio IC	13.15	13.50	3.19	0.00

#### 4. Discusión

La herramienta propuesta permite calcular los índices de fiabilidad con sus respectivos intervalos de confianza tomando en cuenta el tipo de muestreo. Es importante notar que en muchos paquetes no se lleva a cabo ningún ajuste, ya que se realiza el cálculo de los índices directamente sobre la matriz bruta, y los resultados obtenidos son erróneos. Por ejemplo, la fiabilidad global del mapa es 69.6%, sin embargo con base en la matriz bruta (antes de la corrección de Card) se obtiene 82.5% (66 sitios de 80). La diferencia se debe a que categorías como cuerpos de agua con una muy alta fiabilidad son sobre-representadas en el muestreo y tienden a sesgar la evaluación de la fiabilidad de forma optimista. La evaluación basada en la matriz ajustada compensa estas diferencias de representación en el muestreo y da una evaluación sin sesgo de la proporción del mapa correctamente clasificada.

El método propuesto por Card (1982) permite también llevar a cabo correcciones de las áreas obtenidas a partir del mapa. En la figura 1, se puede observar que las categorías 1 y 2 son respectivamente sobre y subestimadas en el mapa debido a una confusión entre ambas categorías. Muchas áreas mapeadas como bosque 1 pertenecen en realidad a la categoría Bosque 2 de acuerdo a la información de los sitios de verificación. En la tabla 5, se presentan estimaciones de áreas que corrigen el efecto de estos errores en el mapa. Debido a que todas estas estimaciones están basadas en un muestreo, presentan una cierta incertidumbre, expresada a través el intervalo de confianza. En algunos casos, este intervalo de confianza es grande (por ejemplo la fiabilidad del usuario de la categoría 1 es superior a 20%) y sería necesario aumentar el número de sitios de verificación para reducir esta incertidumbre.

Para tipos de muestreos más complejos como doble muestreo (en dos pasos) o por conglomerado, el efecto de la autocorrelación espacial en la independencia de los sitios muestreados es más importante y la forma de calcular los intervalos de confianza tendría que tomar en cuenta estos efectos (Mas y Couturier, 2011).

Finalmente, estos métodos pueden aplicarse a la evaluación de mapas de cambio (por ejemplo mapas de deforestación). En este caso particular el muestreo se hace con base en el mapa de cambio y tendrá como objetivo enfocar el esfuerzo de muestreo en las categorías de cambio que generalmente cubren una área mucho menor que las de permanencia. La verificación de estos sitios se realiza usando insumos de dos fechas para poder evaluar los cambios. Recomendaciones muy útiles para este tipo de ejercicio se encuentran en Olofsson et al. (2013 y 2014).

## 5. Conclusiones

Esperamos que estas herramientas, en particular complemento de Q-GIS, permitan a usuarios de sistemas de información geográfica llevar a cabo evaluaciones de la calidad de los mapas generados de forma más robusta. Una descripción más detallada de las herramientas se encuentra en Mas et al. (2014).

## Agradecimientos

Este trabajo se realizó con el apoyo del proyecto SEP-CONACyT 178816.

## Referencias

- Card, D.H. Using known map category marginal frequencies to improve estimates of thematic map accuracy. **Photogrammetric Engineering and Remote Sensing**, v. 48, p. 431-439, 1982.
- Mas, J.F.; Couturier, S. Evaluación de bases de datos cartográficas. In: Bautista, F.; Palacio, J.L.; Delfin, H. (editores). **Técnicas de muestreo para manejadores de recursos naturales**. INE, CIGA, IG, UNAM, p. 675-703, 2011. Disponible em: <<http://www.ciga.unam.mx/publicaciones/>> Acceso em: 24.oct. 2014.
- Mas, J.F.; Pérez-Vega, A.; Ghilardi, A.; Martínez, S.; Octavio Loya-Carrillo, J.; Vega, E. A Suite of Tools for Assessing Thematic Map Accuracy. **Geography Journal**, v. 2014. Disponible em: <<http://downloads.hindawi.com/journals/geography/2014/372349.pdf>> Acceso em: 24 oct. 2014.
- Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. **Remote Sensing of Environment**, v. 148, p. 42-57, 2014.
- Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. **Remote Sensing of Environment**, v. 129, p. 122-131, 2013.
- Stehman, S.V. Estimating area from an accuracy assessment error matrix. **Remote Sensing of Environment**, v. 132, p. 202-211, 2013.
- Stehman, S.V. Thematic map accuracy assessment from the perspective of finite population-sampling. **International Journal of Remote Sensing**, v. 16, p. 589-593, 1995.
- Stehman, S.V.; Czaplewski, R.L. Design and analysis for thematic map accuracy assessment: Fundamental principles. **Remote Sensing of Environment**, v. 64, p. 331-344, 1998.