COMPARACIÓN DE MATRICES DE CONFUSIÓN CELDA A CELDA MEDIANTE BOOTSTRAPING

Ariza, F.J.; Pinilla, C.; García, J.L.
Grupo de Investigación en Ingeniería Cartográfica.
Universidad de Jaén.
C/ Virgen de la Cabeza nº 2, 23071-Jaén (Spain)
fjariza@ujaen.es

RESUMEN. Este trabajo presenta cómo aplicar la simulación para poder realizar contrastes de hipótesis, celda a celda, entre dos matrices de confusión. Junto con las bases estadísticas del bootstrapping se presenta un ejemplo numérico que permite entender el alcance de los resultados que se pueden obtener con esta metodología.

ABSTRACT. This paper shows how the simulation can be applied in order to carry out hypothesis contrasts, in a cell by cell basis, between two confusion matrices. A numeric example is shown in order to understand the extent of the results that can be obtained with this methodology.

INTRODUCCIÓN

En Teledetección, el análisis de la calidad de los trabajos de clasificación se suele acometer mediante el uso de la matriz de confusión, las cuales presentan una visión general de las asignaciones correctas (en la diagonal) e incorrectas (errores de omisión y comisión, en los valores fuera de la diagonal) (Pinilla, 1995). No obstante, el uso directo de la matriz de confusión tiene algunos inconvenientes, de los cuales, los más destacados pueden ser:

- La matriz es un conjunto de valores, por lo que resulta más cómodo el uso de índices, como el porcentaje de acuerdo *"Pa"* o el índice de acuerdo "*Kappa*", que dan una visión global de todo el proceso.

- En caso de querer comparar dos trabajos, el uso de sus respectivas matrices de confusión como base para la comparación no es tarea cómoda.

Respecto al primer inconveniente, poco se puede hacer. La matriz da una visión general que no puede ser sintetizada, sin pérdida de información, por un índice global. La solución común en los trabajos es, pues, acompañar al índice con la matriz y a la matriz con el índice.

El segundo de los inconvenientes señalados muestra, sin embargo, la necesidad de métodos capaces de permitir una comparación más cómoda de matrices de confusión.

En esta última línea, destacan los trabajos de Feinberg y Holland (1970), que proponen el uso normalizado de dichas matrices, lo cual significa que todos los valores presentes en la matriz se muestran en tantos por uno. Esto presenta ventajas (Hardin, 1997):

- Para todas las clases, los valores presentes en la diagonal dan idea directa de los errores de comisión y omisión dado que tras la normalización las marginales[1], valen la unidad.

- Facilita la labor de comparación de matrices obtenidas por diversos métodos o fuentes (Zhuang, 1992, Congalton, 1998).

- Facilita la comparación, celda a celda, entre distintas matrices.

La normalización consiste en un proceso iterativo de compensación en el que se va consiguiendo el valor unidad en las marginales, tanto en filas como en las columnas. Por este motivo algunos autores denominan el proceso *margfit* (Congalton, 1998) de *marginal fiting*.

Lo ventajoso de la aplicación del método es que las celdas de la matriz así obtenidas son directamente comparables, unas con otras y con otras matrices, y los valores obtenidos fuera de la diagonal dan una mejor idea de los errores de omisión y comisión presentes. En este proceso se modifican los valores de los parámetros de tipo global que se pueden derivar de la matriz (por ejemplo, el porcentaje de acuerdo), pero también de parámetros tradicionales como la exactitud del usuario o del productor.

Si bien los valores normalizados son más fáciles de comparar que los valores originales, desgraciadamente no se conoce la función de distribución de los mismos, por lo que no existe método estadístico paramétrico para decidir si existen o no diferencias celda a celda. No obstante, dado que las matrices de error no son en sí mas que un suceso aleatorio, convendría realmente disponer de un número adecuado de matrices de error para poder así derivar conclusiones con significación estadística.

El *bootstrapping* es la técnica estadística adecuada para alcanzar la meta planteada en el párrafo anterior: con un único muestreo - la matriz de confusión con la que se trabaja- se pueden derivar - simular- distintas matrices de confusión, lo que permite estimar las funciones de distribución de los valores presentes en cada una de las celdillas y proceder, de esta forma al análisis de la existencia de diferencias significativas mediante test estadísticos (Hardin, 1997).

En este trabajo se muestran las bases teóricas del bootstrapping, la manera de implementarlo en algoritmos sobre matrices de confusión, y un ejemplo de uso que permite aclarar las ventajas de su uso: con todo ello se pretende abrir las puertas a métodos de control de calidad de mayor rigor estadístico y de mayor capacidad de análisis y decisión.

[1] Las marginales son la suma de todos los elementos de una fila o columna de la matriz de error.

BASES DEL BOOTSTRAPPING

El *Bootstrapping* es un método de Monte Carlo que permite estimar la distribución muestral de un estadístico cuando no existe un estimador paramétrico. Consiste en el remuestreo aleatorio de la muestra original un elevado número de veces, de forma que en cada iteración el estadístico de interés es calculado y almacenado. Así se consigue obtener una aproximación de la distribución, que vendrá dada por la distribución frecuencial del estadístico, y se podrá estimar, por ejemplo, la desviación típica.

El proceso de *Bootstrapping* puede resumirse en los siguientes pasos (Hardin y Shumway, 1997):

- Extraer aleatoriamente una muestra de tamaño n a partir de la población utilizando una estrategia de muestreo apropiada.
- Extraer m muestras aleatorias a partir de la muestra original, siendo m elevado ($m > 200$). Cada muestra también debe ser de tamaño n.
- Determinar el estadístico de interés en cada muestra. Ordenar el conjunto de m valores del estadístico para crear una distribución, que será una aproximación a la distribución muestral del estadístico.

Como se deduce de lo anterior, se trata de un método de simulación que permite el remuestreo. Este proceso requiere poco análisis pero, sin embargo, necesita un importante esfuerzo computacional, de manera que interviene la simulación para facilitar el cálculo; además, este método es interesante en el análisis de resultados de la simulación (Ríos y Col, 1997). Es una técnica adecuada para el cálculo de errores medios cuadráticos (ECM) como se puede observar en diversas referencias (Ross, 1999; Ríos y Col, 1997).

Veamos su base analítica utilizando el caso del EMC. Sea una muestra $A_1,...,A_n$ independiente procedente de una variable aleatoria A con función de distribución de probabilidad F, que depende de un parámetro desconocido q, que se pretende estimar mediante una función dependiente de la muestra $h(a_1,...,a_n)$. Para estudiar su bondad como estimador, se calcula su error cuadrático medio según:

$$ECM(F) = E_F[(h(a_1,...,a_n) - q_F)^2] \qquad (1)$$

Si la distribución F fuera conocida, se podría calcular, sin problemas, el error cuadrático medio, pero dicha distribución es desconocida. Pero cuando el tamaño de la muestra n es grande, la función de distribución F se puede aproximar mediante su función de distribución empírica, definida como:

$$F_n(a) = \frac{\text{número de valores muestrales} \leq a}{n} \qquad (2)$$

Entonces, el parámetro $q_{F_n} = q_n$ está próximo a q_F y el ECM(F) se aproxima mediante la técnica de *bootstrap* del error cuadrático medio, dado por:

$$\mathrm{ECM}(F_n) = E_{F_n}\left[(h(a_1^*,...,a_n^*) - q_n)^2\right] \qquad (3)$$

donde $(a_1^*,...,a_n^*)$ es una muestra aleatoria simple de F_n.

Debido a la construcción de F_n, el vector $(a_1^*,...,a_n^*)$ tiene la misma probabilidad de tomar cualquiera de los n^n valores posibles $(a_{i_1},...,a_{i_n})$ para $i_j \in \{1,...,n\}$, $j=1, ..., n$, por lo que:

$$\mathrm{ECM}(F_n) = \sum_{i_n} ... \sum_{i_1} \frac{\left(h(a_{i_1},...,a_{i_n}) - q_n\right)^2}{n^n} \qquad (4)$$

lo cual necesita la suma de n^n términos, que es imposible para n grande.

Pero se puede calcular el promedio de un gran número de términos mediante simulación, generando n variables aleatorias independientes $(a_{1_i}^*,..., a_{n_i}^*)$ cada una con distribución F_n, y después calcular:

$$b_1 = \left[h(a_{1_i}^*,...,a_{n_i}^*) - q_n\right]^2 \qquad (5)$$

De esta forma, se repite el proceso, hasta obtener las variables $b_1,...,b_r$ (r =100 se suele considerar suficiente para obtener una buena aproximación). Entonces, se puede estimar el error cuadrático medio mediante el promedio de las variables b:

$$\mathrm{E\hat{C}M}(F_n) = \frac{\sum_{i=1}^{r} b_i}{r} \qquad (6)$$

BOOTSTRAPPING APLICADO A LAS MATRICES DE CONFUSIÓN

Como se ha explicado anteriormente, el disponer de los valores normalizados de la matriz de confusión facilita la comparación celda a celda. Sin embargo, dado que no se conoce la función de distribución de los mismos, no existe método estadístico paramétrico para decidir si las diferencias existentes son estadísticamente significativas (Ariza, 2002; Hardin y Shumway, 1997), por lo que el usuario queda limitado a la inspección visual.

La utilidad de la aplicación del método de bootstrapping a las matrices de confusión reside en la posibilidad de estimar la distribución de cada una de las celdas, con lo que ya se pueden contrastar posteriormente hipótesis estadísticas que permitan saber si es significativa la diferencia entre una misma celda en dos matrices diferentes.

El proceso a aplicar para una determinada celda de la matriz sería el siguiente:

A) Obtener la matriz de confusión procedente de la elaboración de cartografía temática, utilizando una técnica de muestreo apropiada (ver ejemplo en Tabla 1).

Tabla 1.- Ejemplo de matriz de confusión

	1	2	3
1	4	1	2
2	1	5	0
3	2	1	3

B) Convertir la matriz de confusión en una lista donde el número de elementos es igual a *n*. Cada celda de la matriz introduce en la lista tantos elementos como el valor de la celda. Cada elemento debe tener asociado un indicador de la fila y columna a la que pertenece en la matriz de confusión (ver ejemplo en Tabla 2 donde cada elemento de la lista es de la forma fila-columna).

C) Extraer una muestra aleatoria con reemplazo de tamaño *n* de la lista y construir una nueva matriz de confusión sirviéndose de los indicadores de fila y columna.

D) Aplicar a la matriz el método de pseudoceros (Feinberg y Holland, 1970).

E) Normalizar la matriz usando el método de Feinberg y Holland (1970).

F) Guardar el valor de la celda.

G) Repetir los pasos C a F *m* veces, siendo *m* un número lo suficientemente elevado.

H) Ordenar los valores almacenados de la celda para construir el histograma de frecuencias, que será la estimación de la distribución.

I) Contrastar la normalidad de la distribución estimada.

J) Estimar la desviación típica de la celda.

Una vez se conocen las desviaciones típicas en dos celdas de dos matrices que se pretende comparar, se puede aplicar un contraste de hipótesis de igualdad de valores:

$$Z = \frac{x_{i,j_1} - x_{i,j_2}}{\sqrt{\hat{s}^{\,2}_{i,j_1} + \hat{s}^{\,2}_{i,j_2}}} \qquad \textit{Ec.7}$$

siendo:

Z el estadístico para contrastar la hipótesis.

x_{i,j_1} el valor original de la celda (i,j) en la matriz 1.

x_{i,j_2} el valor original de la celda (i,j) en la matriz 2.

$\hat{s}^{\,2}_{i,j_1}$ la desviación típica estimada para la celda (i,j) en la matriz 1.

$\hat{s}^{\,2}_{i,j_2}$ la desviación típica estimada para la celda (i,j) en la matriz 2.

EJEMPLO DE APLICACIÓN

Se va a considerar la existencia de tres matrices de confusión (Tabla 3) relativas a una misma situación pero que proceden de tres trabajos de control distintos.

Tabla 2.- Matriz de confusión de la Tabla 1 en forma de lista

1-1	1-1	1-1	1-1	1-2	1-3	1-3	2-1	2-2	2-2	2-2	2-2	2-2	3-1	3-1	3-2	3-3	3-3	3-3

Tabla 3.- Matrices de confusión relativas a XXX

Caso A: Matriz Original								Caso B: Matriz con poca variación								Caso C: Matriz con mucha variación						
326	10	0	0	3	80	0		326	10	0	0	2	80	0		296	9	0	1	2	80	0
18	805	0	0	70	150	5		18	813	0	0	71	148	5		16	813	2	0	74	144	5
0	0	51	1	0	0	0		0	0	49	1	0	0	0		0	0	45	1	1	0	0
0	4	0	3	0	1	5		0	4	0	3	0	1	5		1	4	1	3	0	7	5
3	99	1	0	210	0	0		3	100	2	0	209	0	0		0	110	4	0	215	0	0
70	127	0	3	30	555	0		70	118	0	4	30	555	0		103	106	0	4	20	554	0
1	2	0	5	1	0	3		1	2	0	5	1	0	3		2	4	0	4	1	0	2

La calidad global de cada trabajo puede estimarse mediante el Porcentaje de Acuerdo (ver última fila Tabla 3). De esta forma, el estudio de la similitud estadística de los niveles de calidad alcanzados en los trabajos puede realizarse utilizando este índice y planteando un contraste de hipótesis para verificar si el comportamiento global de las tres matrices es similar, dos a dos. La Tabla 4 presenta el resultado de este análisis.

Lo apuntado hasta ahora supone una aproximación global al problema cuando el Bootstrapping permite, justamente, descender a nivel de lo que ocurre en cada celda. La simple observación de las matrices (Tabla 3) nos permite afirmar que los Casos A y B son bantante similares en cuanto a los valores numéricos que se presentan en las celdas; sin embargo, la matriz correspondiente al Caso C presenta en sus celdillas, respecto a los Casos A y B, una variación de valores sensiblemente mayor que la observada en la comparación anterior. Conviene advertir que, en este caso, la apreciación anterior se puede realizar cómodamente dado que las tres matrices que se presentan (Tabla 3) tienen en sus celdillas valores en un mismo orden de magnitud, en caso contrario, la normalización de las matrices podría ser una técnica útil de comparación celda a celda.

Tabla 4.- Igualdad estadística de los comportamientos globales de tres matrices de confusión

	Caso A	Caso B	Caso C
Caso A	---	No	No
Caso B	SI	---	No
Caso C	No	No	---

Contraste de hipótesis:
Ho: Pa (Caso X) = Pa (Caso Y)
H1: Pa (Caso X) ¹ Pa (Caso Y)
Nivel de significación α = 5% (Z_{tabla} = 1,96)

Tabla 5.- Matriz normalizada (Caso A: Original)

0,8265	0,0146	0,0006	0,0001	0,0094	0,1486	0,0002
0,0277	0,6386	0,0008	0,0002	0,1195	0,1626	0,0505
0,0001	0,0002	0,9778	0,0221	0,0001	0,0002	0,0000
0,0002	0,0376	0,0001	0,3548	0,0001	0,0129	0,5942
0,0106	0,1755	0,0199	0,0001	0,7927	0,0010	0,0002
0,1175	0,1147	0,0007	0,0341	0,0582	0,6745	0,0002
0,0175	0,0188	0,0001	0,5885	0,0200	0,0002	0,3548

Pa = 65.99; Sigma Pa = 0.03206

Tabla 6.- Matriz normalizada (Caso B: poca variación)

0,8273	0,0148	0,0004	0,0001	0,0067	0,1504	0,0001
0,0271	0,6430	0,0006	0,0002	0,1234	0,1593	0,0464
0,0002	0,0002	0,9707	0,0286	0,0001	0,0002	0,0000
0,0002	0,0409	0,0001	0,3488	0,0002	0,0140	0,5959
0,0102	0,1733	0,0277	0,0001	0,7876	0,0010	0,0001
0,1162	0,1072	0,0005	0,0413	0,0597	0,6749	0,0002
0,0188	0,0205	0,0001	0,5808	0,0223	0,0002	0,3572

Pa = 65.85, Sigma Pa = 0.03212

Tabla 7.- Matriz normalizada (Caso C: mucha variación)

0,7841	0,0165	0,0003	0,0253	0,0084	0,1652	0,0002
0,0203	0,6348	0,0080	0,0003	0,1302	0,1365	0,0699
0,0002	0,0004	0,9272	0,0621	0,0097	0,0003	0,0001
0,0103	0,0265	0,0321	0,2896	0,0002	0,0555	0,5858
0,0006	0,1802	0,0321	0,0002	0,7859	0,0009	0,0002
0,1544	0,1030	0,0003	0,0567	0,0439	0,6414	0,0003
0,0300	0,0387	0,0001	0,5659	0,0217	0,0002	0,3435

Pa = 62.95, Sigma Pa = 0.03332

Tabla 8.a.- Matriz de Valores medios del Caso A

0,8258	0,0139	0,0000	0,0000	0,0096	0,1507	0,0000
0,0269	0,6402	0,0000	0,0000	0,1199	0,1620	0,0509
0,0000	0,0000	0,9709	0,0292	0,0000	0,0000	0,0000
0,0000	0,0371	0,0000	0,3544	0,0000	0,0143	0,5942
0,0103	0,1737	0,0291	0,0000	0,7869	0,0000	0,0000
0,1137	0,1135	0,0000	0,0411	0,0575	0,6730	0,0012
0,0233	0,0217	0,0000	0,5752	0,0261	0,0000	0,3537

Tabla 8.b.- Matriz de desviaciones tipicas del Caso A						
0,0017	0,0004	0,0000	0,0000	0,0007	0,0015	0,0000
0,0005	0,0026	0,0000	0,0000	0,0015	0,0016	0,0030
0,0000	0,0000	0,0028	0,0027	0,0000	0,0000	0,0000
0,0000	0,0025	0,0000	0,0151	0,0000	0,0025	0,0141
0,0006	0,0020	0,0028	0,0000	0,0026	0,0000	0,0000
0,0012	0,0011	0,0000	0,0022	0,0006	0,0024	0,0010
0,0024	0,0021	0,0000	0,0148	0,0036	0,0000	0,0155

Tabla 9.a.- Matriz de Valores medios del Caso B						
0,8256	0,0144	0,0001	0,0000	0,0061	0,1536	0,0002
0,0258	0,6439	0,0006	0,0006	0,1196	0,1559	0,0536
0,0000	0,0001	0,9487	0,0473	0,0029	0,0009	0,0001
0,0005	0,0370	0,0268	0,3325	0,0006	0,0159	0,5866
0,0093	0,1770	0,0235	0,0000	0,7889	0,0000	0,0013
0,1128	0,1086	0,0003	0,0445	0,0587	0,6736	0,0015
0,0261	0,0190	0,0000	0,5751	0,0232	0,0000	0,3567

Tabla 9.b.- Matriz de desviaciones tipicas del Caso B						
0,0019	0,0005	0,0003	0,0000	0,0006	0,0019	0,0007
0,0006	0,0027	0,0006	0,0008	0,0019	0,0016	0,0038
0,0001	0,0002	0,0052	0,0052	0,0010	0,0004	0,0007
0,0009	0,0029	0,0060	0,0195	0,0011	0,0020	0,0195
0,0007	0,0026	0,0028	0,0000	0,0024	0,0001	0,0018
0,0018	0,0012	0,0004	0,0022	0,0010	0,0022	0,0011
0,0033	0,0021	0,0000	0,0190	0,0034	0,0000	0,0200

Tabla 10.a.- Matriz de valores medios del Caso C						
0,7913	0,0174	0,0001	0,0157	0,0081	0,1665	0,0009
0,0199	0,6369	0,0095	0,0003	0,1334	0,1349	0,0651
0,0000	0,0003	0,9225	0,0680	0,0081	0,0000	0,0011
0,0100	0,0301	0,0301	0,3003	0,0000	0,0601	0,5693
0,0001	0,1743	0,0364	0,0000	0,7873	0,0000	0,0018
0,1519	0,1025	0,0002	0,0541	0,0466	0,6385	0,0062
0,0267	0,0383	0,0011	0,5616	0,0166	0,0000	0,3556

Tabla 10.b.- Matriz de desviaciones tipicas del Caso C						
0,0024	0,0009	0,0003	0,0037	0,0010	0,0019	0,0016
0,0006	0,0022	0,0024	0,0005	0,0023	0,0029	0,0037
0,0000	0,0003	0,0096	0,0092	0,0015	0,0000	0,0024
0,0018	0,0027	0,0075	0,0191	0,0000	0,0042	0,0175
0,0002	0,0023	0,0026	0,0000	0,0024	0,0000	0,0020
0,0015	0,0019	0,0004	0,0035	0,0017	0,0025	0,0020
0,0022	0,0043	0,0025	0,0170	0,0045	0,0000	0,0182

Las Tablas 5, 6 y 7 presentan los resultados de la normalización de las matrices correspondientes a los tres casos bajo consideración. El proceso se ha realizado con el programa CTem (Ariza y Col, 2002) mediante un proceso de 100 iteraciones. Mediante esta técnica todas las celdas presentan ya valores en el intervalo *[0, 1]* y de esta forma resulta bien fácil la comparación, vis a vis, de una celda *i,j* de un trabajo con la misma celda *i,j* de otro trabajo. Si los valores de las celdas de las matrices que se presentan en las Tablas 5, 6 y 7 se multiplican por 100, las diferencias entre ellos son tantos por ciento de discrepancia entre los trabajos para esa posición *i,j* de la matriz. Por ejemplo, para la celda *i=1; j=1*, se tiene:

Diferencia Caso A/Caso B: *82,65 - 82,73 = -0.008 %*

Diferencia Caso A/Caso C: *82,65 - 78,41 = 4.24 %*

Diferencia Caso B/Caso C: *82,73 - 78.41 = 4.32 %*

Las Tablas 8a y 8b; 9a y 9b; 10a y 10b presentan los resultados del bootstrapping cuando se aplica a las matrices (Tabla 3) con un número de iteraciones en el bootstrap de 50. Al igual que en el caso anterior, el proceso se ha realizado con el programa CTem. Como se observa, aquí se presentan parejas de tablas pues se muestran los valores medios y las desviaciones de cada celda.

Dado que las Tablas 8, 9 y 10 caracterizan estadísticamente el comportamiento de cada celda, se puede plantear una comparación celda a celda. Esto es lo que se ha realizado en las Tablas 11, 12 y 13. Cada una de ellas representa el valor del estadístico calculado según la ecuación 7 anterior.

Se puede plantear el siguiente contraste para cada pareja de celdas de igual índice entre cada una de las Tablas 8, 9 y 10:

H_0: $a_{ij} – a'_{ij} = 0$; $(a_{ij} = a'_{ij})$

H_1: a_{ij} - a'_{ij} ¹ 0;

Tomando un nivel de significación α = 5%, en este caso el test es a dos colas, por lo que la regla de decisión es: H_0 se rechaza si Z ³ $|Z_{a/2}|$ = 1,96. Las Tablas 11, 12 y 13 presentan en colores el resultado de este contraste:

¬ Gris claro: supone un comportamiento similar.

¬ Gris oscuro: supone que no existe un comportamiento similar desde un punto de vista estadístico.

¬ Sin color: no se puede decir nada.

El simple análisis visual de las Tablas 11, 12 y 13 es claro en cuanto a la similitud estadística, celda a celda, de las matrices de confusión que se presentan en la Tabla 3, en las Tablas 4, 5 y 6 de manera normalizada, cuando se comparan dos a dos.

Aquí no cabe ya la intuición del intérprete, ahora si el valor que aparece en cada celda, $Z_{observado}$ de las matrices presentadas en las Tablas 11, 12 y 13 supera el umbral marcado por el estadístico correspondiente al nivel de siginificación, $Z_{observado}$ ³ 1,96 = $|Z_{0,025}|$, se puede considerar que a_{ij} ¹ a'_{ij}. En este caso el valor representado en las celdas a_{ij} del las matrices de las Tablas 11, 12 y 13 son valores absolutos, pero su inclusión con signo podrían servir para aportar mayor información sobre los problemas de desajuste que existen entre las matrices que se comparan.

Tabla 11.- Valor del estadístico para las diferencias entre los Casos A y B

0,0784	0,7809	0,3333	--	3,7963	1,1980	0,2857
1,4084	0,9871	1,0000	0,7500	0,1239	2,6958	0,5577
0,0000	0,5000	3,7589	3,0892	2,9000	2,2500	0,1429
0,5556	0,0261	4,4667	0,8880	0,5455	0,4998	0,3158
1,0847	1,0060	1,4142	--	0,5652	0,0000	0,7222
0,4160	3,0100	0,7500	1,0928	1,0290	0,1843	0,2018
0,6862	0,9091	--	0,0042	0,5856	--	0,1186

Tabla 12.- Valor del estadístico para las diferencias entre los Casos A y C

11,7304	3,5537	0,3333	4,2432	1,2288	6,5269	0,5625
8,9626	0,9689	3,9583	0.6000	4,9164	8,1821	2,9811
--	1,0000	4,8400	4,0467	5,4000	--	0,4583
5,5556	1,9023	4,0133	2,2220	--	9,3704	1,1080
16,1276	0,1969	1,9105	--	0,1130	--	0,9000
19,8861	5,0104	0,5000	3,1447	6,0462	9,9551	2,2361
1,0443	3,4689	0,4400	0,6034	1,6485	--	0,0795

Tabla 13.- Valor del estadístico para las diferencias entre los Casos B y C

11,2053	2,9139	0,0000	4,2432	1,7150	4,8009	0,4008
6,9532	2,0099	3,5976	0,3180	4,6258	6,3404	2,1683
0,0000	0,5547	2,3997	1,9588	2,8844	2,2500	0,4000
4,7206	1,7414	0,3436	1,1797	0,5455	9,5015	0,6603
12,6372	0,7778	3,3761	--	0,4714	0,0000	0,1858
16,6875	2,7145	0,1768	2,3222	6,1349	10,5400	2,0591
0,1513	4,0331	0,4400	0,5295	1,1702	--	0,0407

CONCLUSIONES

En esta trabajo se ha presentado una metodología de aplicación de las técnicas de bootstrapping para la mejora en los análisis basados en celdas de la matriz de confusión. Junto a la metodología anterior se ha desarrollado un caso numérico para ejemplizar, de una forma muy simple, los resultados y aportación que se obtienen con el método propuesto.

Los contrastes de hipótesis planteados en el ejemplo numérico lo son para celdas de iguales índices i,j de dos matrices de confusión distintas, pero el bootstrapping puede también servir para plantear hipótesis sobre cotas a valores suma en los errores de omisión/comisión de una matriz dada y cualquiera otra hipótesis estadística que se quiera probar tomando como base los valores medios y desviaciones calculados. Esto da una gran flexibilidad en el uso práctico del método.

Las matrices de valores medios, desviaciones derivadas del boostrap y los resultados de los contrastes de hipótesis no han de considerarse como los elementos únicos en un análisis en el que se comparen dos matrices de confusión, más bien deben entenderse como elementos cuantitativos que permiten entender y apoyar otras hipótesis más generales sobre la realización del trabajo y del comportamientos de clases particulares.

Desde un punto de vista operativo, la simulación que supone el bootstrap no suele estar incorporada en ningún programa de teledetección ni de análisis estadístico (p.e. SPSS, StatGraphics, etc.) pero su programación es bastante inmediata.

BIBLIOGRAFÍA

Ariza, F.J. 2002. *Calidad en Producción Cartográfica*. Ed Ra-Ma, Madrid.

Ariza, F.J.; García, J.L.; Atkinson, A. 2002. *Ctem: Control de Calidad Temática*.

Congalton, R.; Green, K. 1998. *Assesing the accuracy of remotely sensed data: Principles and Practices*. Lewis Publishers.

Feinberg, S. E; Holland 1970. An iterative procedure for estimation in contingency tables. *Annals of Mathematical Statistics*, 42(3).

Hardin, P. 1997. Statistical significance and normalized confusion matrices. *Photogrammetric Engineering & Remote Sensing*, 63(6).

Pinilla, C. 1995. *Elementos de Teledetección*. Ed. Ra-Ma, Madrid.

Ríos, D.; Ríos, S., Martín, J. 1997. *Simulación, métodos y aplicaciones*. Ed Ra-Ma, Madrid.

Ross, S. 1999. *Simulación* (2ª Edición). Pearson-Prentice Hall. México.

Zhenkul, M. 1995. Tau Coefficients for Accuracy Assessment of Classification of Remote Sensing Data. *Photogrammetric Engineering and Remote Sensing*, 61(4): 435-439.

Zhuang, M. 1992. Accuracy of Spatial Data Used in Geographic Information Systems. *Photogrammetrig Engineering and Remote Sensing*, 58(6): 835-841.