

Evaluación de técnicas geostatísticas como apoyo para la validación de productos en teledetección

Martínez Díaz, Beatriz^{1a}; Cassiraga, Eduardo^b; Camacho-de Coca, Fernando^c; García-Haro, Javier^a

^(a)Departament de Física de la Terra i Termodinàmica, Facultat de Física, Universitat de València.
Dr Moliner, 50, 46100, Valencia, España.

^(b) Departamento de Ingeniería Hidráulica y Medio Ambiente, Universidad Politécnica de Valencia,
Camino de Vera, s/n, 46022, Valencia, ESPAÑA

^(c) EOLAB. ICMUV P.O.

Box 22085. E-46071, Valencia, ESPAÑA

RESUMEN

En este trabajo se evalúa el potencial de las técnicas geostatísticas en la estimación de mapas verdad-terreno de parámetros biofísicos, como el índice de superficie foliar (LAI), a partir de medidas in-situ para su uso en la validación directa de productos a alta y media resolución espacial. Para ello, se ha ensayado con tres modelos geostatísticos: el krigeado ordinario, el cokrigeado colocalizado y el krigeado con deriva externa. La metodología se ha aplicado a un área de cultivos de 5x5 km² en la zona de Barrax (Albacete, España) con medidas estimadas del LICOR LAI2000.

Palabras claves: Validación de productos, geoestadística, índice de superficie foliar.

ABSTRACT

This paper evaluates the performance of geostatistical methods to estimate leaf area index (LAI) fields from ground based measurements. These fields are necessary for the direct validation of high and medium spatial satellite derived products. Three models variants of the kriging technique the ordinary kriging (OK), the collocated cokriging (CKC) and kriging with an external drift (KED) are used. The methodology has been applied to the 5.5 km² agricultural area of Barrax using LAI estimates from LICOR LAI2000 measurements.

Keywords: satellite derived product, validation, geostatistic methods, leaf area index.

Introducción

El concepto de validación se entiende como el proceso de evaluar mediante técnicas independientes la calidad de los productos derivados a partir de satélite (Justice et al., 2000). El objetivo final de la validación es, además de asociar una incertidumbre, asignar al producto un alto nivel de consistencia espacial y temporal para que pueda ser utilizado en cualquier contexto, momento y con diferentes productos derivados de otros sensores independientemente de sus características espectrales y espaciales.

La validación se ha entendido desde un principio como la comparación con datos de referencia, también conocida como validación directa. Sin embargo para productos de media y baja resolución, este procedimiento es difícil de llevar a cabo debido a la heterogeneidad de la variable a caracterizar, adquiriendo más atención el desarrollo de estrategias que resuelvan el problema del escalado de la información *in-situ* hasta la resolución del producto a validar. El objetivo de este trabajo consiste en evaluar el potencial y las limitaciones de las técnicas.

¹ e-mail:beatriz.martinez@uv.es

Copyright: Universidad Nacional del Centro de la provincia de Buenos Aires, Argentina.

cas geoestadísticas para derivar mapas verdadero a partir de información *in-situ* que sirvan para validar productos biofísicos de la cubierta vegetal, como el índice de superficie foliar (LAI). Para ello, se ha ensayado con tres modelos geoestadísticos, el krigeado ordinario (KO), el cokrigeado colocalizado (CKC) y el krigeado con deriva externa (KDE). Por otra parte, se han escogido diferentes índices de vegetación como variable auxiliar en la predicción de los mapas y se ha discutido otro aspecto importante de la validación directa, como es el diseño de un muestro espacial óptimo que caracterice la variabilidad de la zona de estudio.

Metodología

La idea básica del krigeado es estimar el valor de un atributo en una localización donde no se lo conoce, como una combinación lineal de los valores registrados de dicha variable (ver por ejemplo Isaaks y Srivastava, 1989). Para representar el parámetro de interés adoptamos un modelo probabilístico que consiste en admitir que los valores que puede tomar la variable dentro del área de estudio vienen dados por una función aleatoria. Los tres estimadores evaluados en este trabajo dependen del modelo considerado para la función aleatoria y de si consideran o no información auxiliar.

En primer lugar se ha considerado el método de KO por ser el más sencillo de los tres en cuanto a que no necesita información secundaria y supone un modelo de tendencia constante y no conocido en un cierto radio de vecindad. Este modelo permite evaluar el muestreo espacial de la variable medida (i.e variable primaria) en las predicciones.

Por otra parte los modelos de CKC y KDE permiten evaluar la influencia de la variable secundaria. El KDE considera que la tendencia que sigue la variable principal es una función lineal definida en cada localización a partir de una variable secundaria. Los coeficientes del modelo de tendencia se calculan en cada estimación de la variable primaria, proporcionando un mayor detalle de dicha variable a escala local. El estimador por CKC incluye en su expresión una combinación lineal de los datos más cercanos de la variable principal y el dato secundario localizado en la posición donde se va a estimar la variable primaria (i.e. dato colocalizado). El CKC requiere conocer el coeficiente de correlación lineal entre la variable primaria y secundaria.

El procedimiento para aplicar estos modelos implica como primera etapa el modelado del patrón espacial en base a las medidas. Este

análisis consiste en describir y cuantificar la relación entre observaciones de la misma o diferentes variables en diferentes posiciones a partir de herramientas como el variograma. Posteriormente, hay que ajustar un modelo al variograma, que constituye el modelo de variabilidad espacial utilizado como entrada en los algoritmos de estimación empleados para obtener los campos de LAI.

Experimento

La metodología se ha aplicado a un área de cultivos de 5x5 km² en la zona de estudio de Barrax (Albacete, España), incluida dentro del proyecto VALERI como zona para la validación de productos biofísicos (<http://www.avignon.inra.fr/valeri/>).

Variable primaria

Como variable primaria se disponen de medidas de LAI no destructivas estimadas con el LICOR LAI2000. La adquisición de las medidas utilizadas para la validación directa se realizó durante la campaña de SPARC (SPectra BARax Campaign), los días, del 12 al 14 de Julio del 2003 (<http://gpds.uv.es/sparc/>). El muestreo realizado por ambos instrumentos consistió en un muestreo estratificado por clases donde, el objetivo principal fue el de caracterizar lo mejor posible la variabilidad presente en la zona de estudio, es decir, tanto la descrita por todo el conjunto de cultivos de la zona como la mostrada dentro de cada campo. La localización de los datos registrados se puede observar en la imagen superior de la figura 1.

Variable secundaria

Se ha utilizado como información auxiliar una imagen SPOT HRVIR2 de 20 m de resolución espacial adquirida el 3 de Julio de 2003. La imagen se corrigió geoméricamente a la proyección UTM 30N-ED50 y se convirtió a reflectividades en el techo de la atmósfera. Dada la alta correlación que existe entre los índices de vegetación y el LAI (Carlson y Ripley, 1998), se ha escogido el índice NDVI como variable secundaria. Por otra parte y dado que la zona de estudio presenta una alta proporción de suelo desnudo, se ha considerado el índice SAVI (Huete, 1988) diseñado para corregir el efecto del suelo como variable secundaria y así evaluar su influencia en las estimaciones. En la imagen inferior de la figura 1 se puede ver el parámetro NDVI derivado de la imagen SPOT para el área de estudio.

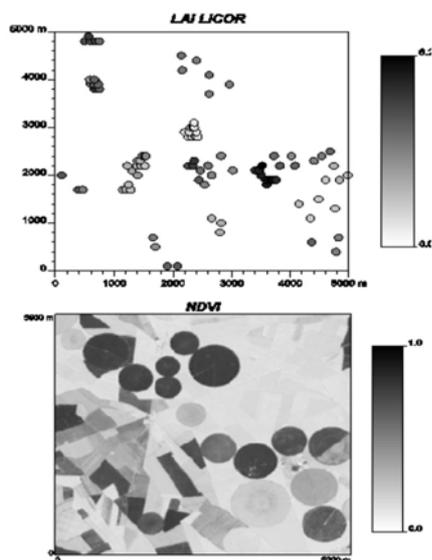


Fig. 1. Localización de las medidas *in-situ* de LAI y NDVI derivado de la imagen SPOT.

Criterios de evaluación

Además del conjunto de medidas de LAI2000, se dispone de medidas de LAI derivadas a partir de fotografías digitales hemisféricas (Digital Hemispherical Photographs, DHP). Ambos métodos han demostrado ser compatibles en el área de estudio, con discrepancias equivalentes a las que se obtienen con diferentes instrumentos LICORs (Martínez, 2006), por lo que este conjunto de datos se utilizará para evaluar los mapas estimados. También se utilizará el producto de LAI derivado de la misma imagen SPOT a partir de la metodología propuesta en VALERI para obtener mapas verdad-terreno a partir de medidas *in-situ*. Este método consiste en aplicar a la imagen de radiancias una función de transferencia (FT) obtenida al relacionar las medidas *in-situ* de LAI con los valores radiométricos de la imagen de alta resolución. En este caso, la FT se obtiene a partir de un método iterativo de regresión multivariada por mínimos cuadrados (Martínez et al., 2007). El mapa derivado a partir de esta FT (LAI_VALERI) se comparará con las estimaciones derivadas de los métodos geoestadísticos.

Muestreo espacial

Dado que el muestreo espacial es uno de los aspectos más críticos en la caracterización de

la variabilidad sobre todo en zonas heterogéneas, proponemos un estudio sobre la influencia del muestreo espacial en la estimación de la variable primaria. Este estudio consistirá en extraer de la imagen LAI_VALERI distintos conjuntos de datos caracterizados por muestreos diferentes e intentar reproducir el producto LAI a partir de los tres métodos geoestadísticos. Por último y dado que el objetivo de este trabajo es el de derivar mapas verdad-terreno para validar productos de diferentes resolución espacial, también se evaluará el muestreo óptimo según la resolución espacial del producto a validar.

Resultados

Continuidad espacial

La figura 2 muestra los variogramas correspondientes a las direcciones 30° NE y 60° NO, respectivamente, a los que se les ha ajustado un modelo gaussiano. Este modelo proporciona las direcciones máxima y mínima de continuidad espacial, en nuestro caso 1600 m y 1000 m, así como el parámetro de efecto pepita (0.05).

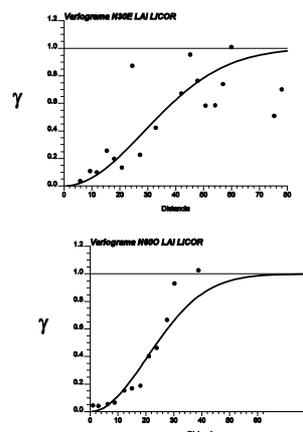


Fig. 2. Variogramas direccionales de LAI.

Variable secundaria

La figura 3 muestra la relación entre la variable primaria y la secundaria a través de un diagrama de dispersión. Se aprecia una importante correlación lineal entre éstas, superior a 0.8. Sólo se ha mostrado el NDVI debido a que no se han observado diferencias importantes con el SAVI.

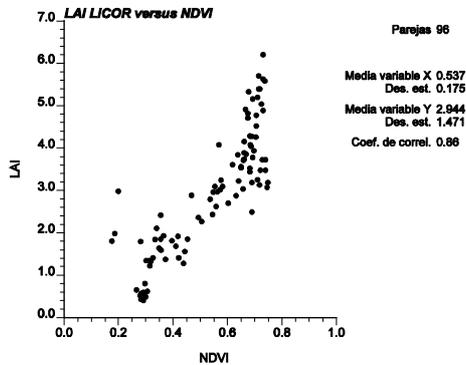


Fig. 3. Diagrama de dispersión entre la variable primaria (LAI) y secundaria (NDVI).

Mapas estimados

En la figura 4 se pueden ver los mapas estimados por KO, CKC y KDE, respectivamente. Sólo se incluyen los mapas estimados considerando como variable secundaria el NDVI.

Tal como se esperaba, el resultado obtenido por KO está totalmente condicionado por el número de medidas y por la localización de la variable primaria. Este modelo identifica sólo los grupos más grandes de cultivos y no proporciona información sobre la variabilidad a escalas finas. En el caso de CKC y KDE, se observa el impacto de la variable secundaria en la obtención del patrón de continuidad de la zona de estudio. Sin embargo, el modelo CKC no resuelve correctamente el problema asociado a una variable no estacionaria y no consigue reproducir el cambio tan fuerte en el LAI que se observa entre parcelas, como es el caso del modelo KDE.

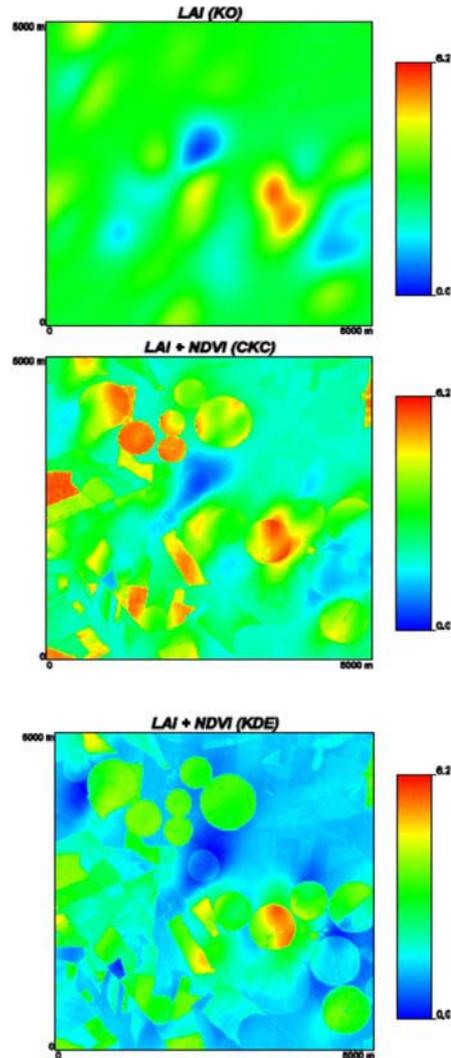


Fig. 4. Mapas de LAI estimados por los métodos de KO, CKC y KDE.

La tabla 1 muestra los valores medios, desviación estándar y primer y tercer cuartil de cada estimación. El modelo KO es el que presenta la media más alta, menor variabilidad en las estimaciones y peor caracterización de los valores bajos de LAI dado el alto valor del primer cuartil. El modelo KDE es el que menor media presenta y valor más bajo del primer cuartil, tal como cabía esperar dada la alta contribución de suelo desnudo en la zona de estudio.

Tabla 1. Media, desviación estándar, primer y tercer cuartil de las estimaciones por KO, CKC y KDE.

	Media	σ	1 ^{er} C	3 ^{er} C
KO	2.89	0.65	2.66	3.17
CKC	2.75	1.12	1.93	3.40
KDE	1.94	1.01	1.22	2.63

Criterios de evaluación

La tabla 2 muestra la pendiente, ordenada al origen, coeficiente de correlación y RMSE derivados al comparar las estimaciones por KO, CKC y KDE con los datos derivados de la DHP. Los resultados muestran como los modelos KO y CKC proporcionan el error más alto y coeficientes de correlación menores. El modelo KDE es el que proporciona una pendiente de casi uno, menor ordenada y menor RMSE.

Por otra parte, la tabla 3 muestra los mismos estadísticos, pero entre las estimaciones por KO, CKC y KDE frente al mapa de LAI_VALERI. Se observa como el KO es el que peor correspondencia presenta con el mapa de LAI_VALERI, dada por una correlación muy baja y un RMSE muy alto. El mapa obtenido por KDE muestra ser el más comparable con el mapa LAI_VALERI, con una pendiente alrededor de uno y un coeficiente de correlación alto (0.89).

Tabla 2. Comparación entre las estimaciones por KO, CKC y KDE con las medidas de LAI derivadas de DHP.

	Pendiente	Ordenada	r	RMSE
KO	0.92	0.69	0.76	0.91
CKC	0.83	0.85	0.82	0.92
KDE	0.97	0.71	0.81	0.81

Tabla 3. Comparación entre las estimaciones por KO, CKC y KDE con el mapa de LAI_VALERI.

	Pendiente	Ordenada	r	RMSE
KO	2.75	0.10	0.24	2.2
CKC	1.99	0.58	0.82	1.7
KDE	1.19	0.57	0.89	1.0

Muestreo espacial

En la evaluación del muestreo espacial óptimo, se han considerado cuatro muestreos: dos sistemáticos y dos estratificados. Los patrones sistemáticos definen una malla regular la cual permite muestrear uniformemente toda la población de la zona de estudio. Se han elegido

dos muestreos, SP1 y SP2, con distancias de separación 200 y 500 m, respectivamente. Estas distancias son menores que el alcance definido para el variograma del NDVI.

Los muestreos estratificados se han diseñado considerando el tipo de cubierta y su variabilidad espacial derivada de la imagen de NDVI e información recogida durante la campaña de campo. Además, se han incluido tres consideraciones derivadas de estudios anteriores (Martínez, 2006), como que (i) alrededor de 50 medidas son suficientes para caracterizar la zona de estudio, (ii) las medidas deben extenderse en toda el área y (iii) la necesidad de caracterizar mejor el suelo. La diferencia entre los muestreos estratificados reside en el número de medidas utilizadas para caracterizar la variabilidad de un mismo cultivo. El muestreo estratificado SP3 presta más atención a caracterizar la variabilidad de aquellos campos de un mismo cultivo que presentan estado fenológico diferente (i.e. distinto NDVI) con tres medidas distribuidas en su interior, mientras que el muestreo SP4 caracteriza la mayoría de los campos pero sólo con una medida en su interior. La figura 5 muestra los cuatro muestreos espaciales diseñados.

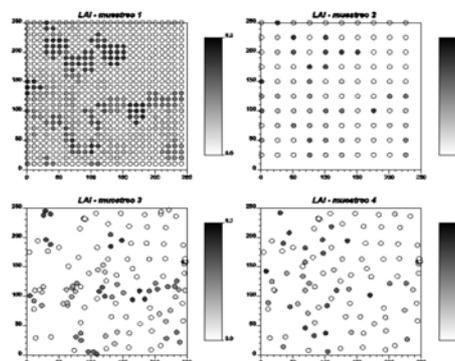


Fig 5. Muestreos espaciales diseñados a partir del mapa LAI_VALERI.

Al igual que se hizo con las estimaciones derivadas de los datos LICOR LAI2000, se ha evaluado el error de estimación asociado a estos muestreos espaciales con los datos *in-situ* derivados de la DHP. La tabla 4 muestra la comparación de las estimaciones por KO, CKC y KDE para los cuatro muestreos.

Tabla 4. Evaluación de las estimaciones derivadas por KO, CKC y KDE respecto a las medidas *in-situ* de DHP para los cuatro muestreos.

	Pendiente	Ordenada	r	RMSE
KO				
SP1	0.6	0.4	0.8	1.1
SP2	0.3	0.8	0.6	1.4
SP3	0.6	0.5	0.7	1.4
SP4	0.5	0.4	0.7	1.3
CKC				
SP1	0.8	0.2	0.8	0.9
SP2	0.7	0.4	0.8	0.9
SP3	0.8	0.4	0.8	1.2
SP4	0.9	0.2	0.8	0.9
KDE				
SP1	0.9	0.3	0.9	0.8
SP2	1.0	0.16	0.9	0.8
SP3	0.9	0.2	0.9	0.8
SP4	0.9	0.2	0.9	0.7

	Pendiente	Ordenada	r	RMSE
KO				
SP1	0.8	0.3	0.9	0.8
SP2	0.4	0.8	0.6	1.3
SP3	0.4	0.6	0.5	1.5
SP4	0.5	0.7	0.7	1.1
CKC				
SP1	0.9	0.04	0.9	0.7
SP2	0.8	0.3	0.9	0.8
SP3	0.8	0.13	0.8	0.9
SP4	0.9	0.14	0.9	0.7
KDE				
SP1	0.9	0.05	1.0	0.18
SP2	0.9	0.04	0.9	0.3
SP3	0.9	-0.02	0.9	0.4
SP4	0.9	0.06	0.9	0.2

El modelo KDE junto con el muestreo SP4 es el que mejor se ajusta a las estimaciones *in-situ*. Por otra parte, el modelo CKC junto con el muestreo SP4 también muestra un error aceptable (0.9) y pendiente casi uno. El modelo KO es el que mayor diferencia presenta con las estimaciones *in-situ*, siendo el muestreo SP1 el que menor RMSE (1.1) y mayor correlación (0.8) proporciona del resto.

Por otra parte, se han comparado las estimaciones para los cuatro muestreos con el mapa LAI_VALERI. La tabla 5 muestra los estadísticos obtenidos al comparar cada mapa estimado por KO, CKC y KDE con el mapa LAI_VALERI.

Las diferencias más importantes se observan para los muestreos SP2 y SP3 y los algoritmos de KO y CKC. El muestreo SP1 se presenta como el más adecuado para reproducir el mapa LAI_VALERI con los tres algoritmos, sin embargo no es el diseño más óptimo para trasladar al campo debido a su muestreo tan intensivo. Por otra parte, el muestreo SP4 presenta resultados comparables al SP1 con CKC y KDE, siendo éste muestreo mucho más fácil de trasladar al campo que el SP1. Cabe destacar la alta correspondencia entre las estimaciones por KDE con el mapa LAI_VALERI para los cuatro muestreos.

Tabla 5. Evaluación de las estimaciones derivadas por KO, CKC y KDE respecto al mapa de LAI_VALERI para los cuatro muestreos.

Por último, se ha evaluado la influencia de la resolución espacial en el diseño del muestreo óptimo. Para ello, se han considerado tres resoluciones espaciales, alta (20 m), media (300 m) y baja (1 km). Los mapas a 300 m y 1 km se han obtenido al degradar mediante la media aritmética los mapas derivados a 20 m. La tabla 6 muestra el coeficiente de correlación y el RMSE entre los mapas estimados para los cuatro muestreos y el mapa LAI_VALERI.

En general, se observa cómo el algoritmo de KDE es el que mayor coeficiente de correlación y RMSE presenta para todos los muestreos y resoluciones espaciales (RMSE<0.39 y r>0.96), obteniéndose resultados similares tanto para el muestreo más intensivo (SP1) como para el menos intensivo (SP4) cuando se aplica KDE. De este modo, si se dispone de información secundaria el KDE resulta adecuado para estimar mapas a diferentes resoluciones espaciales minimizando los esfuerzos en el campo. A pesar de que el KDE es el que mejor se correlaciona con el mapa LAI_VALERI, cabe destacar otras configuraciones (algoritmo y muestreo) que también pueden ser adecuadas a otras resoluciones espaciales. Tal es el caso del CKC con los muestreos SP1 y SP4, el cual proporciona un coeficiente de correlación alto (>0.9) y RMSE razonable (<0.7) para todas las resoluciones espaciales. El muestreo SP3 es el que peor correlación muestra con el mapa LAI_VALERI y mayor RMSE proporciona para todos los modelos y todas las resoluciones.

Aunque, el KO es el que peor correspondencia muestra a alta resolución, la configuración formada por KO y muestreo SP4 proporciona resultados comparables a los proporcionados por el muestreo intensivo SP1 (r>0.85 y RMSE<0.65) para resoluciones media y baja.

Este resultado es de gran relevancia dado que permitiría con un muestreo inferior a 50 ESUs y sin necesidad de información secundaria obtener un mapa de LAI sobre el área 5x5 km² para validar productos a media y baja resoluciones.

Tabla 6. Coeficiente de correlación y RMSE entre las estimaciones por KO, CKC y KDE y el mapa LAI_VALERI a alta, media y baja resolución espacial.

	SP1	SP2	SP3	SP4
	r/RMSE			
KO				
20m	0.87/ 0.80	0.63/ 1.27	0.51/ 1.50	0.72/ 1.12
300m	0.96/ 0.32	0.74/ 0.83	0.64/ 1.01	0.85/ 0.65
1km	0.98/ 0.13	0.87/ 0.33	0.51/ 0.68	0.88/ 0.30
CKC				
20m	0.92/ 0.66	0.86/ 0.83	0.84/ 0.91	0.91/ 0.68
300m	0.97/ 0.33	0.89/ 0.61	0.87/ 0.65	0.94/ 0.48
1km	0.98/ 0.19	0.93/ 0.34	0.76/ 0.48	0.95/ 0.30
KDE				
20m	0.99/ 0.18	0.98/ 0.24	0.97/ 0.39	0.98/ 0.24
300m	0.99/ 0.13	0.99/ 0.16	0.98/ 0.24	0.99/ 0.19
1km	0.98/ 0.12	0.99/ 0.11	0.96/ 0.16	0.99/ 0.13

Conclusiones

En este trabajo se han evaluado las técnicas geostatísticas de KO, CKC y KDE para derivar mapas de LAI verdad-terreno a partir de medidas *in-situ*. La técnica de KDE ha mostrado una alta consistencia para la estimación del LAI, dada la alta correlación (0.9) y RMSE razonable (0.8) respecto a las medidas *in-situ* derivadas con la DHP.

También se ha evaluado la influencia del muestreo óptimo, donde el diseño estratificado SP4 ha presentado resultados comparables al muestreo intensivo SP1 con los modelos CKC y KDE a alta resolución. El modelo KDE es el que ha proporcionado mejor consistencia con las medidas *in-situ* y el mapa de LAI_VALERI para todos los muestreos.

En cuanto a la influencia de la resolución espacial, el modelo KDE junto con el muestreo SP4 ha mostrado ser adecuado para estimar mapas a diferentes resoluciones espaciales cuando se dispone de información secundaria. Por otra parte, la técnica del KO junto con el muestreo SP4 ha demostrado ser especialmente útil para estimar a resoluciones espaciales bajas, donde

prima el valor medio y no tanto los detalles y la variabilidad espacial a nivel subpixel. Este resultado es de gran relevancia dado que permitiría con un muestreo inferior a 50 ESUs y sin necesidad de información secundaria obtener un mapa de LAI para validar productos a la resolución espacial baja.

Agradecimientos

Este trabajo ha sido realizado en el marco de los proyectos DESURVEY (EC-003950), DULCINEA (CGL2005-04202) y LSA SAF (Eumetsat).

Referencias

- Carlson, T. N. and Ripley, D. A. 1998. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, 62: 241-252.
- Huete, A. R. 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of the Environment*, 25: 295-309.
- Isaaks, E. and Srivastava, R. (1989). An Introduction to Applied Geostatistics, Oxford University Press, New York.
- Justice, C.O., Belward, A., Morisette, J., Lewis, P., Privette, J. and Baret, F. 2000. Developments in the validation of satellite sensor products for the study of the land surface. *International Journal of Remote Sensing*, 21 (17): 3383- 3390.
- Martínez, B. 2006. Caracterización espacial de parámetros biofísicos de la cubierta vegetal para la validación de productos derivados mediante teledetección. Aplicación de técnicas geoestadísticas. Tesis Doctoral, Universidad de Valencia, pp. 242.
- Martínez, B., García-Haro, F.J. and Camachode Coca, F. 2007. Derivation of high-resolution leaf area index maps in support of validation activities. Application to the cropland Barrax site. Agricultural Forest and Meteorology, In review.