

Estimación y documentación de la calidad de la información geográfica en MiraMon: Estimadores locales y especializados para aspectos geométricos y temáticos.

Xavier Pons



Lluís Pesquer



Guión



- Conceptos
- Espacialización de la calidad
 - posicional
 - temática
 - cuantitativa
 - categórica
- Registro y trazabilidad de la calidad en los metadatos

Ejemplificados mediante procesos asociados y herramientas implementadas en el SIG MiraMon

- **Calidad:** conjunto de características de un producto que le confieren su aptitud para satisfacer necesidades explícitas e implícitas (ISO 9000).

Indicadores de calidad pueden ser (entre otros):

- Completitud
- Consistencia
- Exactitud posicional (planimétrica, altimétrica...)
- Exactitud temporal
- Exactitud temática (cuantitativa, categórica...)
- Linaje y trazabilidad

Quality indicators

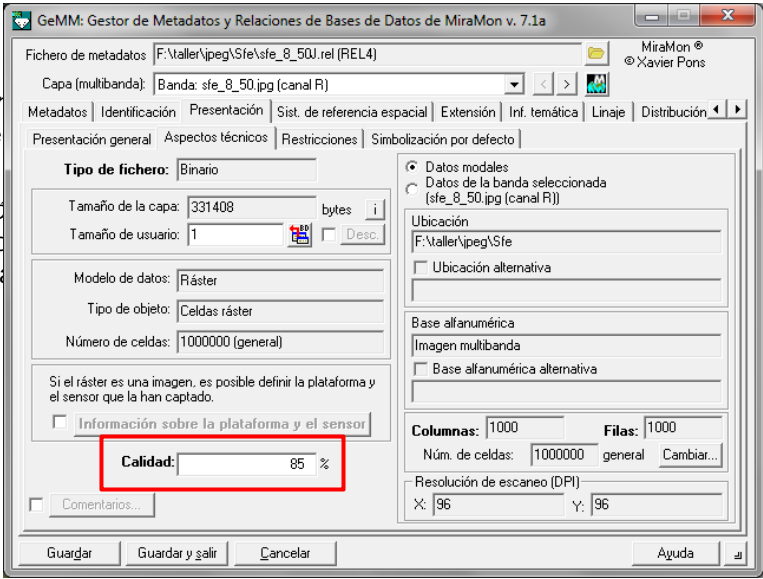
QUALity aware Visualisation for the Global Earth Observation system of systems

www.geoviqua.org

Quality parameter	Quality indicator	Definition
Completeness	Commission	Excess data present in a dataset
	Omission	Data absent from a dataset
Logical consistency	Conceptual consistency	Adherence to rules of the conceptual schema
	Domain consistency	Adherence of values to the value domains
	Format consistency	Degree to which data is stored in accordance with the physical structure of the dataset
	Topological consistency	Correctness of the explicitly encoded topological characteristics of a dataset
Positional accuracy	Absolute or external accuracy	Closeness of reported coordinate values to values accepted as or being true
	Relative or internal accuracy	Closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true
	Gridded data position accuracy	Closeness of gridded data position values to values accepted as or being true
Temporal accuracy	Accuracy of a time measurement	Correctness of the temporal references of an item (reporting of error in time measurement)
	Temporal consistency	Correctness of ordered events or sequences
	Temporal validity	Validity of data with respect to time
Thematic accuracy	Classification correctness	Comparison of the classes assigned to features or their attributes to a universe of discourse (e.g. ground truth or reference dataset)
	Non-quantitative attribute correctness	Correctness of non-quantitative attributes
	Quantitative attribute accuracy	Accuracy of quantitative attributes

Un primer ejemplo...

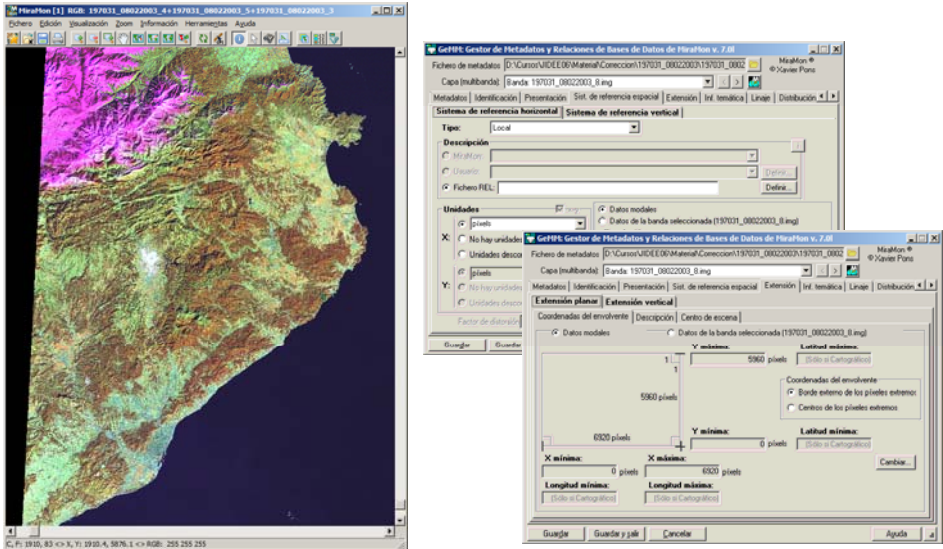
- Calidad
- Problema cabe
- Solución cual
- venta



Estimación y documentación de la calidad ... CREAF

Calidad posicional planimétrica

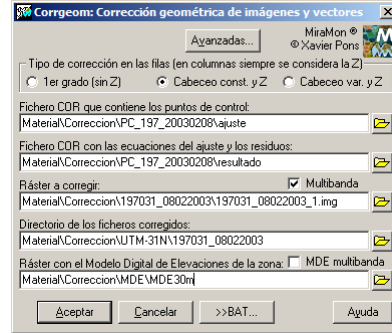
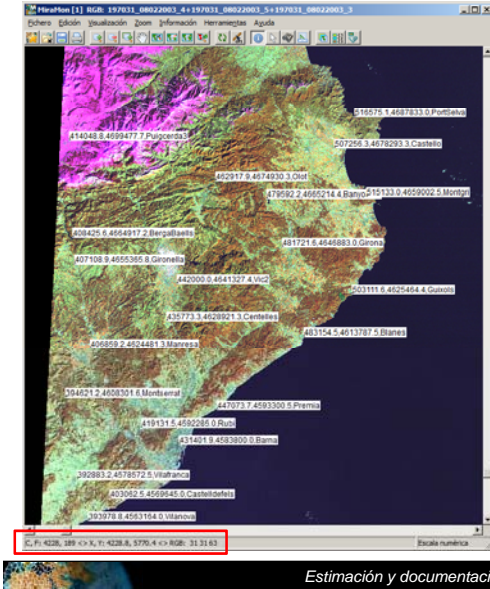
Georeferenciación



Estimación y documentación de la calidad ... CREAF

Calidad posicional planimétrica

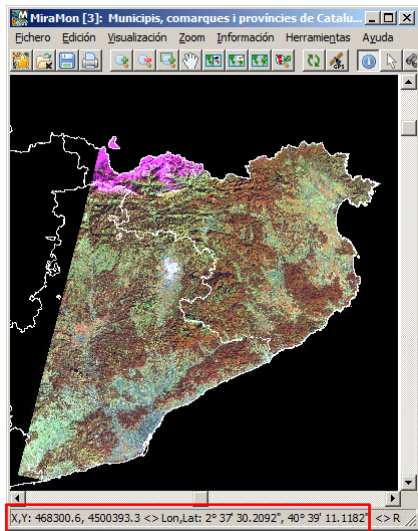
Georeferenciación:



Herramienta de corrección geométrica a partir de puntos de control sobre el terreno que gobiernan la georeferenciación.

Calidad posicional planimétrica

Georeferenciación:



Parámetros globales de calidad planimétrica

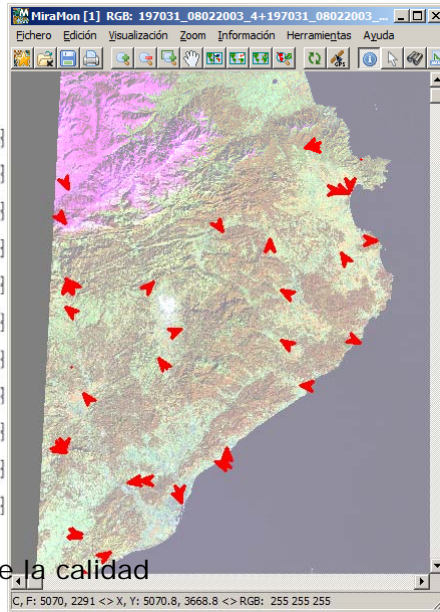
Calidad del sistema de referencia horizontal	
Parámetro (1/1):	Exactitud posicional
Indicador (1/1):	Exactitud horizontal relativa
Medida (1/11):	RMS
Valor de la medida:	7.837271 (m)
Tipo de valor:	Real de doble precisión
Medida (2/11):	RMS en X
Valor de la medida:	0.266566 (pixels)
Tipo de valor:	Real de doble precisión
Medida (3/11):	RMS en Y
Valor de la medida:	0.204649 (pixels)
Tipo de valor:	Real de doble precisión
Medida (4/11):	RMS (ajuste puntos de control)
Valor de la medida:	0.330680 (pixels)
Tipo de valor:	Real de doble precisión

$$RMS = \sqrt{\frac{\sum (z_{cota} - z_{MDE})^2}{n}}$$

Georeferenciación

Diferencias entre los valores OBSERVADOS y ESTIMADOS para cada punto del TEST:

punto 1 del test:	err(x)test = +0.004398 [3930.800000- 3930.795602]
	err(y)test = +0.322956 [4646.000000- 4645.677044]
	err(xy)test = +0.322956
punto 2 del test:	err(x)test = +0.370595 [3224.062500- 3223.691905]
	err(y)test = -0.267420 [2779.187500- 2779.454920]
	err(xy)test = +0.457006
punto 3 del test:	err(x)test = -0.149378 [791.687500- 791.836878]
	err(y)test = +0.004684 [583.312500- 583.307816]
	err(xy)test = +0.149451
punto 4 del test:	err(x)test = -0.298803 [547.600000- 547.898803]
	err(y)test = +0.423041 [4300.400000- 4299.976959]
	err(xy)test = +0.517926
punto 5 del test:	err(x)test = -0.008389 [617.100000- 617.108389]
	err(y)test = +0.054299 [1548.933333- 1548.879034]
	err(xy)test = +0.054943
punto 6 del test:	err(x)test = +0.122612 [3436.933333- 3436.810721]
	err(y)test = +0.035119 [5083.100000- 5083.064881]
	err(xy)test = +0.127543
punto 7 del test:	err(x)test = -0.016542 [692.500000- 692.516542]
	err(y)test = -0.015113 [2530.166667- 2530.181780]
	err(xy)test = +0.022406
punto 8 del test:	err(x)test = -0.001643 [646.375000- 646.376643]
	err(y)test = -0.254663 [3490.750000- 3490.534437]
	err(xy)test = +0.259468
punto 9 del test:	err(x)test = -0.115360 [1627.666667- 1627.782027]
	err(y)test = -0.081787 [3495.000000- 3495.081787]
	err(xy)test = +0.141411
punto 10 del test:	err(x)test = +0.107960 [3851.325000- 3851.217040]
	err(y)test = -0.148932 [3821.675000- 3821.823932]
	err(xy)test = +0.187946
punto 11 del test:	err(x)test = +0.017266 [2503.166667- 2503.153391]
	err(y)test = -0.072884 [1969.853333- 1969.906217]
	err(xy)test = +0.073993



Errores locales ->Espacialización de la calidad

Datos temáticos cuantitativos

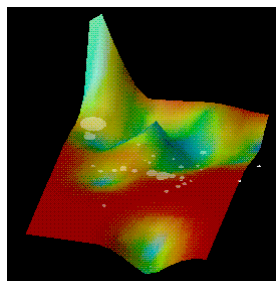
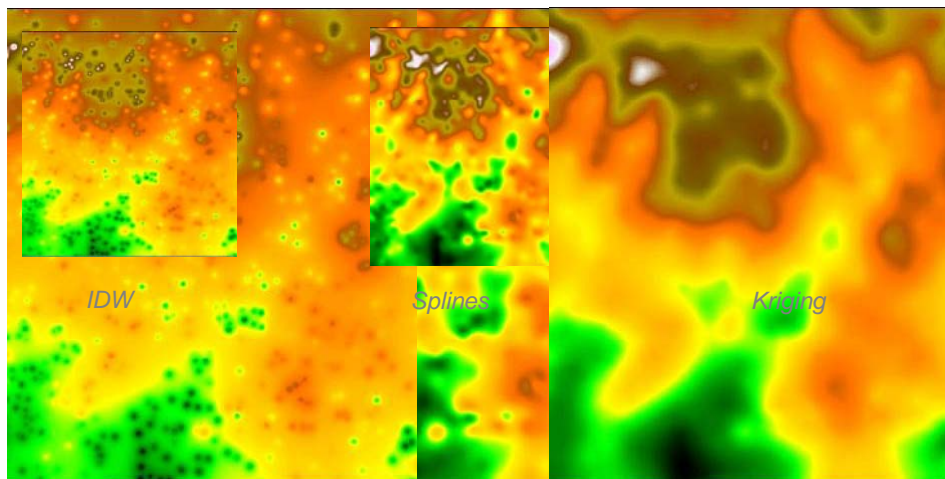


Table 3.1. Comparison of non-geostatistical spatial interpolation methods and kriging as a generic model for geostatistical methods (mainly modified from Burrough and McDonnell (1998)).

Method	Assumption	Univariable/ Multivariable	Deterministic/ stochastic	Local/ global	Exact/ inexact	Abrupt/ gradual	Limitation of the procedure	Computing load	Output data structure	Suitability
Inverse distance weighting (IDW)	underlying surface is smooth	Univariable	Deterministic	Local	Inexact (but can be forced to be exact)	Gradual	No error assessment. Results depend on size of search window and choice of weighting parameter. Poor choice of window can give artefacts when used with high data densities such as digitised contours.	Small	Gridded surface, contours	Quick interpolation from sparse data on regular grid or irregularly spaced samples.
Thin plat splines (TPS)	Underlying surface is smooth everywhere	Univariable/ Multivariable	Deterministic	Local	Exact	Gradual	Goodness of fit possible, but within the assumptions that the fitted surface is perfectly smooth.	Small	Gridded surface, contours	Quick interpolation (univariate or multivariate) of digital elevation data and related attributes to create digital elevation models (DEM) from moderately detailed data
Kriging*	Interpolated surface is smooth. Statistical stationarity and the intrinsic hypothesis.	Univariable/ Multivariable	Stochastic	Local	Exact	Gradual	Error assessment depends on variogram and distribution of data points and size of interpolated blocks. Requires care when modelling spatial correlation structures.	Moderate	Gridded surface	When data are sufficient to compute variograms, kriging provides a good interpolator for sparse data. Binary and nominal data can be interpolated with Indicator kriging. Soft information can also be incorporated as trends or stratification. Multivariate data can be interpolated with co-kriging.

Tabla parcial extraída de: Jin Li and Andrew D. Heap, A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia Record 2008/23

Comparación visual de algunos de los métodos representativos



Kriging:

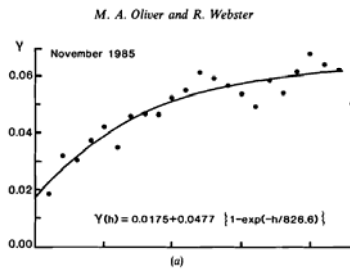
Método geoestadístico de interpolación espacial basado en el análisis del patrón espacial de la varianza: variograma

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [z(\vec{x}_i) - z(\vec{x}_i + h)]^2$$

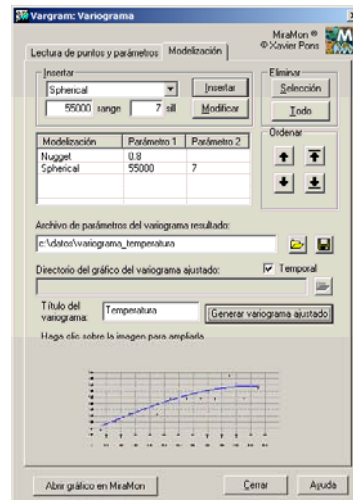
La construcción y ajuste del variograma determina el modelo predictivo y también el modelo espacializado de su error asociado.

$$z(\vec{x}_0) = \sum_{i=1}^n \lambda_i \cdot z(\vec{x}_i)$$

$$\sigma_e^2 = \sum_{i=1}^n \lambda_i \gamma(\vec{x}_i, \vec{x}_0) + \phi$$

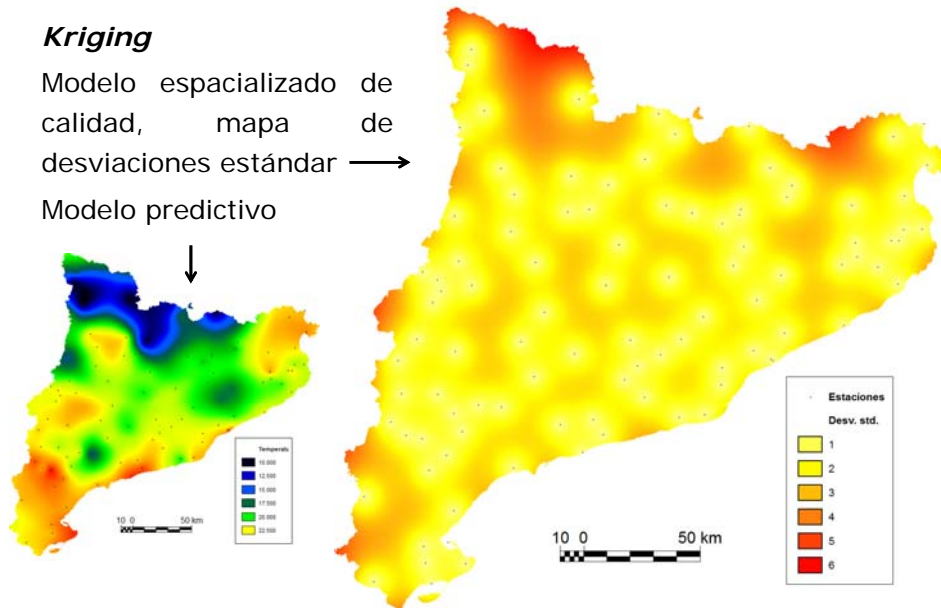


Kriging: Análisis del variograma



Kriging

Modelo espacializado de calidad, mapa de desviaciones estándar → Modelo predictivo

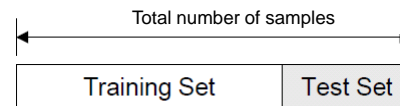


Estimación y documentación de la calidad ...



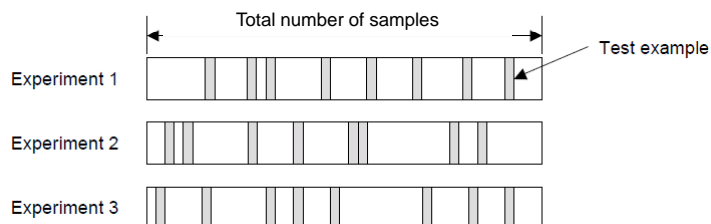
- Test independiente

Se divide la muestra en dos partes. Se aplica el modelo a partir de la submuestra de entrenamiento y para el test se compara su valor con el valor calculado.



- Submuestreo aleatorio:

Se repite el caso anterior a partir de distintas particiones obtenidas aleatoriamente. La validación global se obtiene por promedios de las individuales.

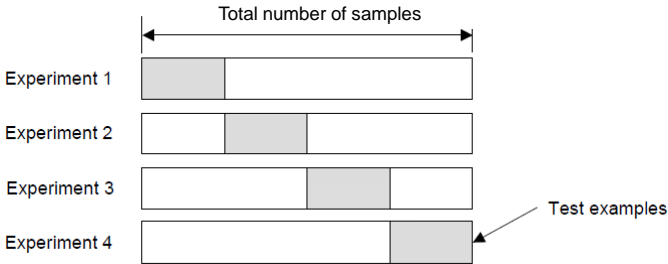


Estimación y documentación de la calidad ...



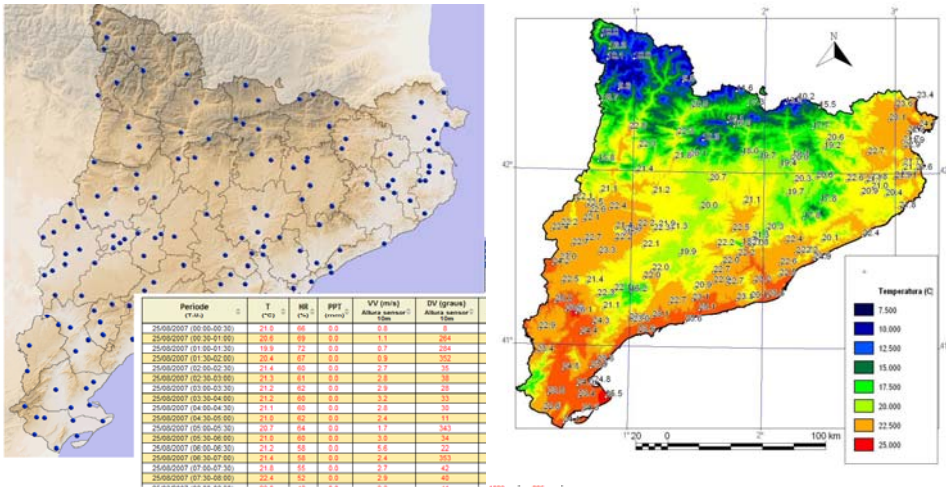
Conceptos

- Validación cruzada
 - Todas las muestras son usadas per generar el modelo y para el test.
 - Un caso muy habitual, consiste en un test de un único punto, pero también puede ser una submuestra de tamaño prefijado (*K-Fold Cross validation*)



Regresión multivariante

Regresión multivariante: análisis de la calidad temática cuantitativa



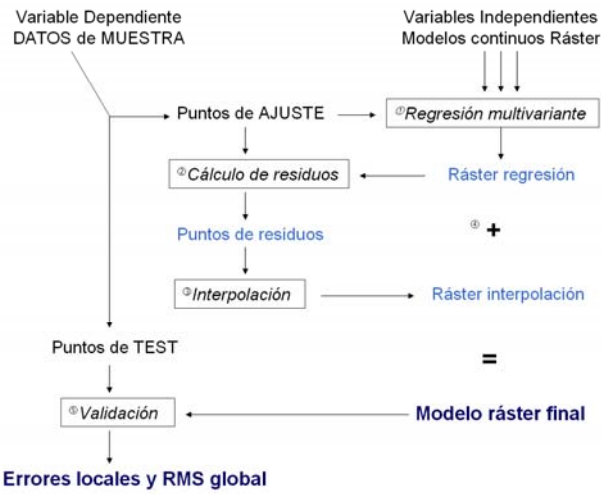
GIS & RS
MIRAMON

Regresión multivariante

$$Y = \sum_{j=1}^N a_j \cdot x_j$$

GIS & RS
MIRAMON

Regresión multivariante



Regresión multivariante
Ajuste por mínimos cuadrados con interpolación de residuos por: Inverso de la distancia (IDW)

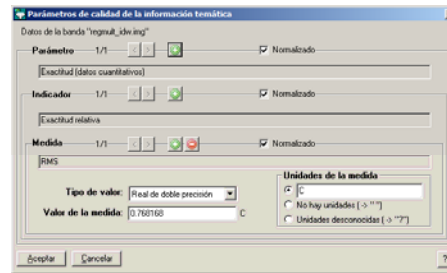
Número total de puntos: 144
Número de puntos de ajuste: 100
Número de puntos de test: 44

Coefficientes del ajuste:
Término independiente: -63.386216
Coeficiente de Altitud: -0.005905, normalizado: -0.980596
Coeficiente de Continentalidad: 0.013614, normalizado: 0.164388
Coeficiente de curvatura: -0.411401, normalizado: -0.080462
Coeficiente de Coseno de la latitud: 116.101377, normalizado: 0.499191
Coeficiente de coseno orientación: -0.134618, normalizado: -0.029949

Parámetros de la regresión seleccionada
R²: 0.931172
R² ajustado: 0.927511
F Fisher: 254.344424
p significación: 0.000000
Cp Mallows: 6.094177
Cp Mallows-(Nvariables+1): 0.094177

variables eliminadas en la regresión seleccionada:
Radiación solar acumulada mensual

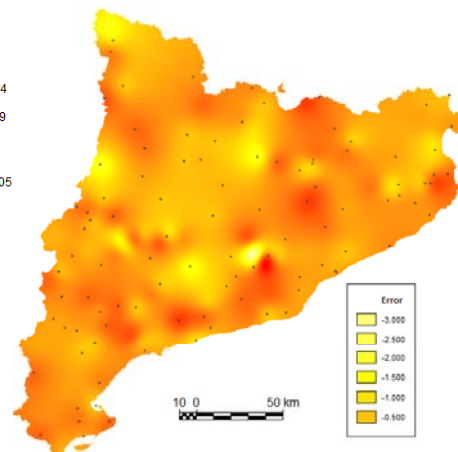
Exactitud (datos cuantitativos):
Exactitud relativa: RMS: 0.768168 (C)



Informe con parámetros de calidad globales.

Gestor de Metadatos

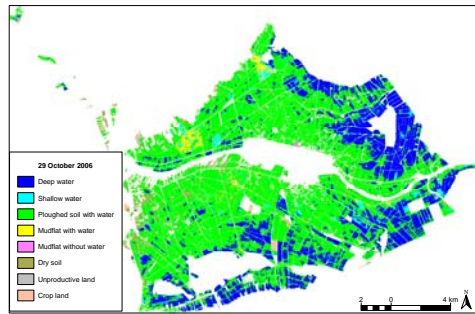
Análisis local del error del modelo



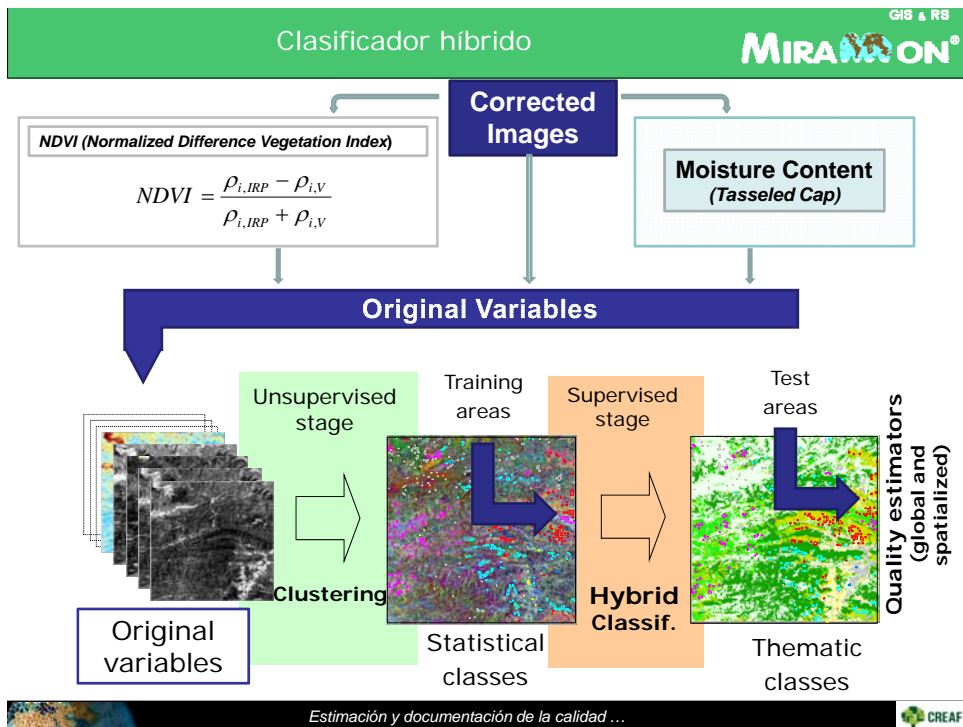
Valores puntuales del error

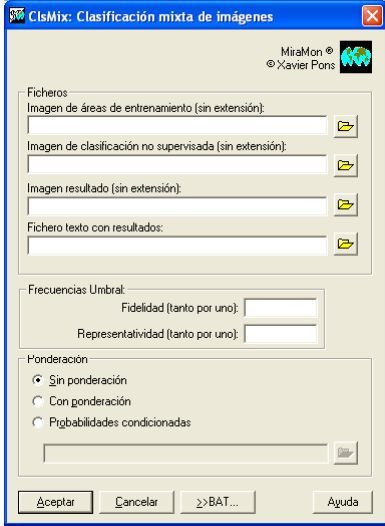
Modelo continuo interpolado

Datos temáticos categoricos



Estimación y documentación de la calidad ...





Evaluación de la clasificación:

matriz de confusión

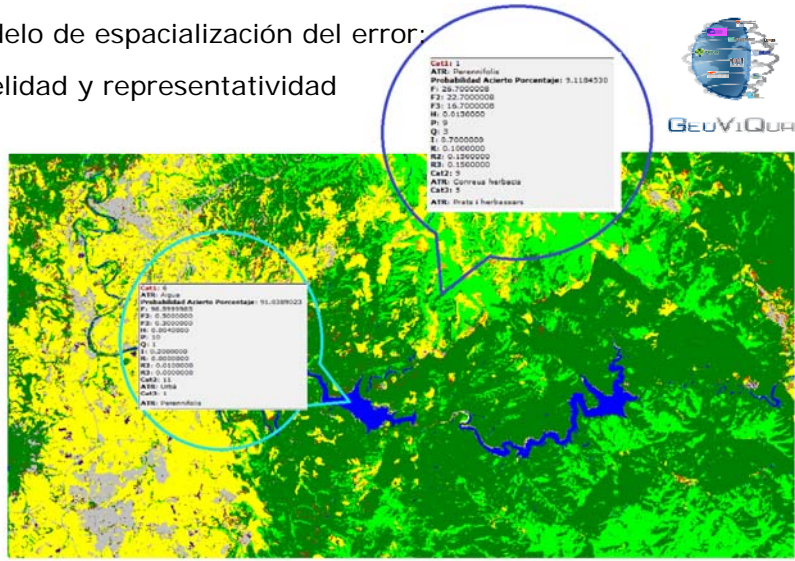
- exactitud del usuario: ratio entre el número de predicciones acertadas y el número total de predicciones para cada clase
- exactitud del productor: ratio entre el número de predicciones correctas y el número total de observaciones para cada clase
- coeficiente Kappa (Congalton 1991): medida de asociación que describe la correspondencia entre dos clasificaciones

Matriz de confusión

Verdad-Terreno	Perennifolios	Caducifolios	Matorrales	Vegetación húmedales	Prados y herbáceos	Frutales	Arrozales	Cultivos herbáceos	Viveros	Urbanos	Vías de comunicación	Agua	Montañas	Errores	Exactitud usuario		
Capa a evaluar	NODATA	11004	11168	2568										2698	1009		
	>NODATA	35	4	31	0	134516	4851	3289	0	359122	6286	5	3205	204	6906		
Generar	Perennifolios	56237	36596	5644	6312	49589	5250	6	451	159154	7939	4	11216	926	18.6	81.4	
Informar en formato texto	Os	19	9	21											9701		
	Caducifolios	13681	73475	2895											24871	24.3	75.7
Informar en formato hoja	s	0	6	7	7	34650	226	1953	0	28296	12	4409	214	80	9948		
	Matorrales	8	37920	69	1194	31592	233	0	54	40649	3921	4	5967	09	43.3	56.7	
Agrupar	Vegetación húmedales	166	39	470	1160	5	594	27	5	63	3	14	13	2549	54.5	45.5	
	Prados y herbáceos	25945	8101	3	0	234946	42	114	0	32540	1	3369	88	79	32.3	67.7	
	Agua	598	137	188	126	62	1	39	0	135	0	264	74	4	7	93	
	Frutales	3543	968	1273	200	1172	18	2	40	12047	83	2265	87	8	55.1	44.9	
	Arrozales	18	4	10	43	4	133	22	9667	256	1	23	2	3	5.1	94.9	
	Cultivos herbáceos	16316	54655	9537	3136	95856	948	8	1547	1879028	4	22	14343	730	26.4	73.6	
	Viveros	5642	120	4581	10	2718	62	1	3	24736	16	3	2495	87	27.4	72.6	
	Urbanos	90661	3264	3	1186	28288	1612	9	116	110078	2	31	87788	848	35.8	64.2	
	Vías de comunicación	2317	33	2203	260	501	2400	314	3	260	170	6800	9860	1	60.7	39.3	
Total		73717	13176	1626	13624	612788	0	64	6	2646333	47	663	135352	2167			
Error		31	48	101	13624	612788	0	64	6	2646333	47	663	135352	2167			
Error comisión		23.7	44.2	65.3	91.5	61.7	43.3	90	18.8	29	38	29.2	92.7				
Exactitud		76.3	55.8	34.7	8.5	38.3	56.7	10	81.2	71	62	70.8	7.8				

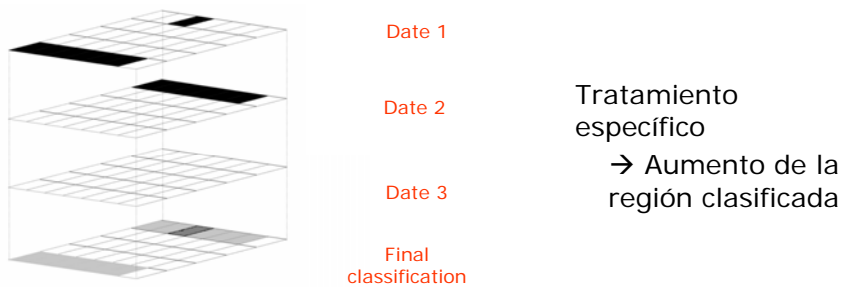
0.5 Kappa
2.36E-08 Varianza de Kappa

Modelo de espacialización del error:
Fidelidad y representatividad

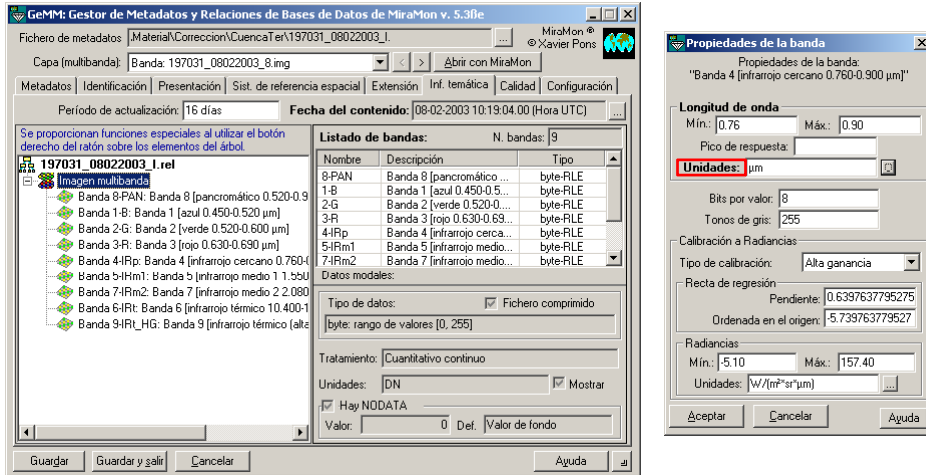


Especial tratamiento y rigor para la ausencia de datos (NODATA).

La estrategia convencional es declarar no clasificado al píxel con algún NODATA en cualquiera de las variables de la serie



Herramientas de importación de los principales formatos de distribución de datos satelitales (CEOS, TIFF-glovis, HDF...) con especial preocupación por incorporar sus metadatos.



Estimación y documentación de la calidad ...



```
[QUALITY:LINEAGE:PROCESS1]
norganismes=1
history=Corrección aplicada: 4
date=20030211

[QUALITY:LINEAGE:PROCESS1:ORGANISME_1]
organisationName=ESA-ESRIN - (Italia)

[QUALITY:LINEAGE:PROCESS2]
norganismes=1
history=Bandas 1,2,3,6,8: Datos de pre-vuelo, Baja ganancia. Bandas 4,5,7,9: Datos de pre-vuelo,
Alta ganancia.
purpose=Calibración radiométrica
date=20030211

[QUALITY:LINEAGE:PROCESS2:ORGANISME_1]
organisationName=ESA-ESRIN - (Italia)

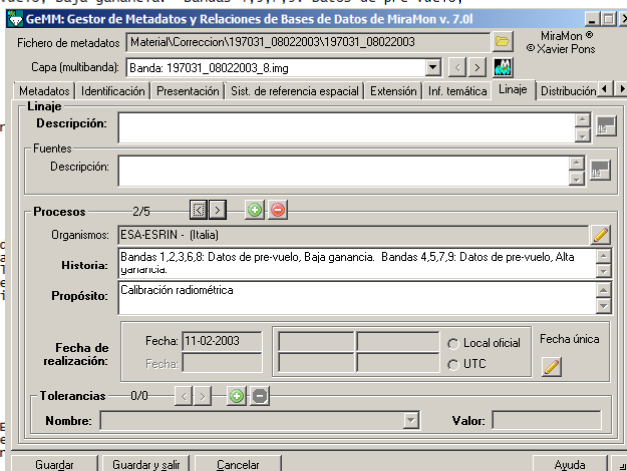
[QUALITY:LINEAGE:PROCESS3]
norganismes=1
history=Todas las bandas: Representación
purpose=Corrección radiométrica
date=20030211

[QUALITY:LINEAGE:PROCESS3:ORGANISME_1]
organisationName=ESA-ESRIN - (Italia)

[QUALITY:LINEAGE:PROCESS4]
norganismes=1
history=Todas las bandas: Alineación bic
escaneo del espejo, Longitud de la línea
sensor (ADS), Sistema de corrección de 1
escaneos, Rotación terrestre, Altitud de
de carga útil (PCD), Efemérides definiti
próximo).
purpose=Corrección geométrica
date=20030211

[QUALITY:LINEAGE:PROCESS4:ORGANISME_1]
organisationName=ESA-ESRIN - (Italia)

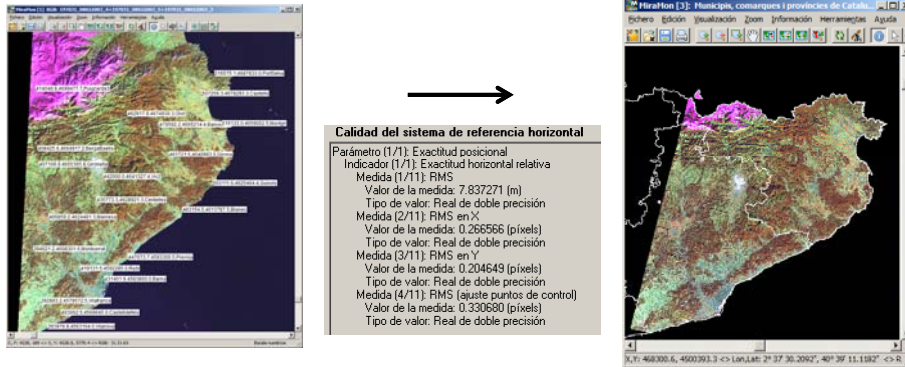
[QUALITY:LINEAGE:PROCESS5]
norganismes=1
history=ceosimg.exe c:\Material\CEOS\SC
/ROW_ORBIT=31 "/GeographicIdentifierCode
purpose=Transformación de rásters en for
date=20061013 11043520+0200
```



Estimación y documentación de la calidad ...



Metadatos: trazabilidad



```
[QUALITY:LINEAGE:PROCESS6]
rorganismes=1
history=corr\geom.exe 3 Material\Correccion\PC_197_20030208\ajust Material\Correccion
\georef\resultado Material\Correccion\197031_08022003\197031_08022003_1.img Material
\Correccion\georef 0 Material\Correccion\MDE\ND_girona15 /MULTIBANDA /MDE_MULTIBANDA
purpose=Correccion geométrica de imágenes y vectores. Corr. geom. con polinomio con Z
(columnas) y polinomio de 1r grado (filas). Remuestreo por vecino más próximo
date=20061013 16423815+0200

[QUALITY:LINEAGE:PROCESS6:ORGANISME_1]
organisationName=C R E A F
```

Estimación y documentación de la calidad ...

Metadatos: trazabilidad

Permite la toma de decisiones automáticas según comparando estimadores globales entre diversos modelos

Data on the "Precip_1011_1.75.img" band

Parameter 1/1: [] [] [] []

Accuracy (quantitative data)

Indicator 1/1: [] [] [] []

Relative accuracy

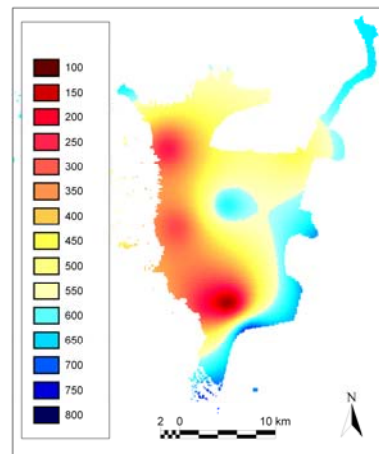
Measure 1/1: [] [] [] []

RMS

Type of value: Double precision real

Measurement value: 208.550811

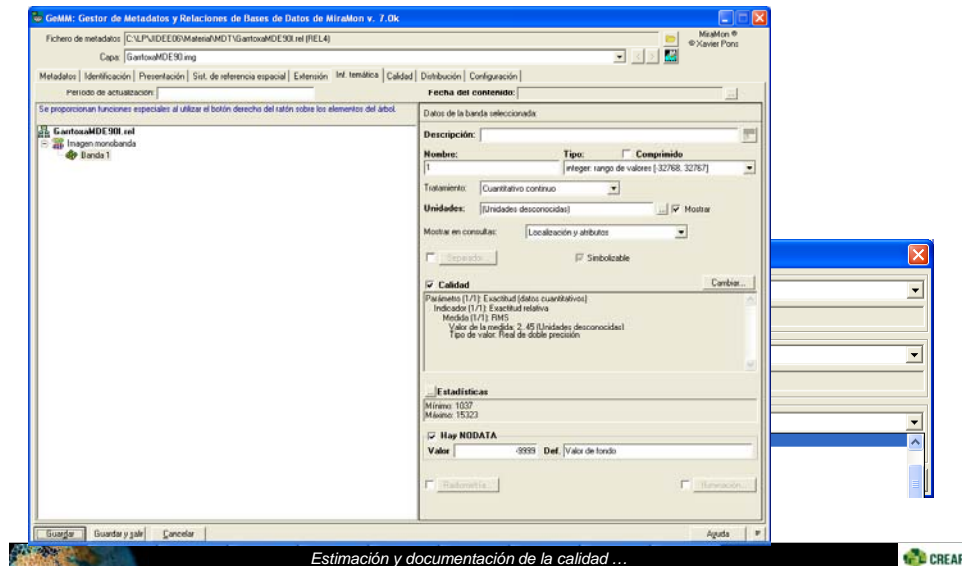
Exp.	IDW		Splines	
	RMS	Ten.	RMS	
1	202.71	25	332.25	
1.25	203.99	50	201.08	
1.50	206.04	75	190.53	
1.75	208.59	100	189.12	
2.00	211.39	125	189.09	
2.25	214.23	150	189.19	
2.50	217.02	175	189.28	
2.75	219.71	200	189.37	



Pesquer L., Prat E., Díaz-Delgado R., Bustamante J., Pons X., Masó J. "Automatic modelling and continuous map generation from georeferenced species census data in an interoperable GIS environment". International Congress on Environmental Modelling and Software Leipzig, Germany.

Estimación y documentación de la calidad ...

A nivel de campo de una base de datos vectorial o de variable, o banda, en un ráster monobanda o multibanda



- Zabala A., Vitulli R., Pons X. (2012) "Impact of CCSDS-IDC and JPEG 2000 compression on **image quality** and classification". *Journal of Electrical and Computer Engineering*, article number: 761067.
- Sevillano E., Díaz P., Ninyerola M., Masó J., Zabala A., Pons X. (2012) "Analysis of the **Quality Metadata** in GEOSS clearinghouse". *EuroGEOSS – Madrid (España)*
- Moré G., Pons X., (2011) "Preliminary considerations about the assessment and visualisation of the **quality on geometric corrections** of satellite imagery depending on the number of ground control points" *7th International Symposium on Spatial Data Quality - Coimbra*
- Pesquer L., Zabala A., Pons X., Serra J. (2010) "**Quality analysis** in N-dimensional lossy compression of multispectral remote sensing time series images". *Proceedings of SPIE Vol. 7810, 78100J* - San Diego (California-USA).
- Díaz P., Masó J., Guimet J. (2010) "Comparative **quality assessment** of metadata. Two regional SDI case studies." *INSPIRE conference*. Krakovia (Polonia).

The screenshot shows the MiraMon website interface. At the top, there's a header with 'GIS & RS' and the 'MIRAMON®' logo. Below it, a navigation bar includes 'Sistema de Información Geográfica y Teledetección' and language options for 'Català' and 'English'. The main content area is divided into several sections:

- Temas:** A list of topics including '¿Qué es MiraMon?', 'Descargas', 'Precios y compras', 'Productos para Internet', 'Notas técnicas', 'Publicaciones, congresos...', 'Investigación y proyectos', 'Cursos', and 'Materiales'.
- Documentos:** A list of documents such as 'Tríptico MiraMon PDF v7' with details for Catalan, Castilian, and English versions.
- Actualidad:** A news section with dates from September 2012 to April 2012, mentioning participation in the XV National Congress of Geographical Information Technologies and the 14th Master in Teledetection and GIS.
- Equipo docente:** Information about the teaching team, including objectives, bibliography, and evaluation methods.
- 5. OTROS MATERIALES:** A section describing the use of the MiraMon commercial SIG version 5, its availability at the CREAF center, and details about the price and updates.

Gracias por su atención



www.creaf.uab.es/miramom
I.pesquer@creaf.uab.cat

XV Congreso Nacional Tecnologías de Información Geográfica
Tecnologías de Información Geográfica en el contexto de Cambio Global