

“CONTRIBUTION OF CLASSICAL AND ROBUST PRINCIPAL COMPONENT ANALYSIS TO DEPTH ESTIMATION”

E. Castillo⁽¹⁾, V. Bayarri, R. Ferrer

⁽¹⁾Universidad de Cantabria. E.T.S.I. Caminos, Canales y Puertos. Dpto. Ingeniería Geográfica y Técnicas de Expresión Gráfica. Área de Ingeniería Cartográfica, Geodesia y Fotogrametría. Avda. Los Castros s/n. Santander, 39005.

ABSTRACT:

The Bay of Santander is a very important environmental element, located in the north of Spain (Cantabria). Human spills to convert sea into land started about 1850 and gradually change the tide prism. This produced the Bay to start filling up and consequently, problems for the safety navigation. Whence the navigational channel must be periodically measured and dredged.

Hydrographic surveying by conventional shipborne sounding techniques is slow, hazardous and expensive. This has produced an increasing interest in the application of remote sensing techniques at last in the critical shallow areas, which are frequently use by ships approaching or leaving ports or harbours.

Traditional satellites and multispectral airborne sensors have been used for this purpose, but the emergence of passive hyperspectral imagery brings the possibility of deriving more accurate bathymetric maps, since past considerations suggested that multispectral techniques could be extended to accommodate this new data.

Keywords: remote sensing, principal component analysis, depth estimation, hyperspectral imagery.

1.- INTRODUCCION

A lo largo del tiempo los métodos utilizados en la captura de datos batimétricos se fundamentan en el método utilizado para posicionar el punto planimétricamente y en la técnica utilizada en la medición del espesor de la lámina de agua. No obstante, una de las mayores dificultades que tienen dicho tipo de trabajos es que en ciertos trabajos la zona de la que se quiere tomar información no es accesible de una forma directa.

Hoy en día la gran mayoría de las batimetrías se realizan a partir de GPS y ecosonda, obteniéndose precisiones mucho mejores que las que venían consiguiendo con anterioridad pero los trabajos batimétricos que se realizan en aguas poco profundas mediante técnicas tradicionales de sondeo, son lentos, caros e incluso en ciertas ocasiones resultan peligrosos, de ahí el interés que está suscitando la captación de profundidades a partir de imágenes de sensores satelitales o aerotransportados. Las batimetrías que se obtienen a partir de dichas imágenes tienen la ventaja de que es fácil disponer de este material para la gran mayoría de las zonas y los precios son bastante razonables pero desafortunadamente el inconveniente que presentan es la precisión que se obtiene en dicha toma de profundidades ya que limitan en gran medida el valor de la batimetría.

El avance que se ha producido en cuanto a tecnología en los últimos años ha generado nuevas metodologías. La llegada de sensores de alta resolución espacial, espectral y radiométrica han generado nuevos procedimientos en la toma de profundidades con resultados muy buenos desde el punto de vista de precisiones obtenidas.

Esta información se requiere en multitud de aspectos de la ingeniería civil (construcción de muelles, diques, emisarios submarinos, control de cauces,...) y dependiendo del objetivo final que se pretenda se buscará una precisión final.

2.- ZONA DE ESTUDIO

La Bahía de Santander está localizada en un lugar privilegiado de la Comunidad Autónoma de Cantabria (España) y se trata de una depresión y un ecosistema de gran riqueza tanto desde el punto de vista biológico como socioeconómico. En su entorno se concentran más de 250.000 habitantes, más del 50% de la población de la comunidad. La Bahía vienen

formándose desde el Mesozoico y es importante para una población que desarrolla diversas actividades, todas ellas generadoras de vertidos de aguas residuales.

Este ecosistema, muy sensible y de alta fragilidad ecológica, presenta en su mayoría sedimentos tales como limos, arcillas y otras partículas procedentes del río Miera. Son numerosos los puntos que recogen los vertidos de los términos municipales que rodean la Bahía (Santander, Camargo, Astillero, Marina de Cudeyo, Ribamontán al Mar y Villaescusa) y que desembocan en ella, ya sea directamente o a través de rías cercanas. Se recogen aguas fluviales, fecales e incluso residuales de muchas industrias.

Como consecuencia de la dinámica costera y de la desembocadura del río Miera, también conocido como estuario de Cubas o ría de cubas, se ha generado una estructura singular denominada “El puntal de Somo”. Se trata de una lengua de arena de unos 2,5 km de longitud con un ancho medio de zona emergida de 250 m aproximadamente.

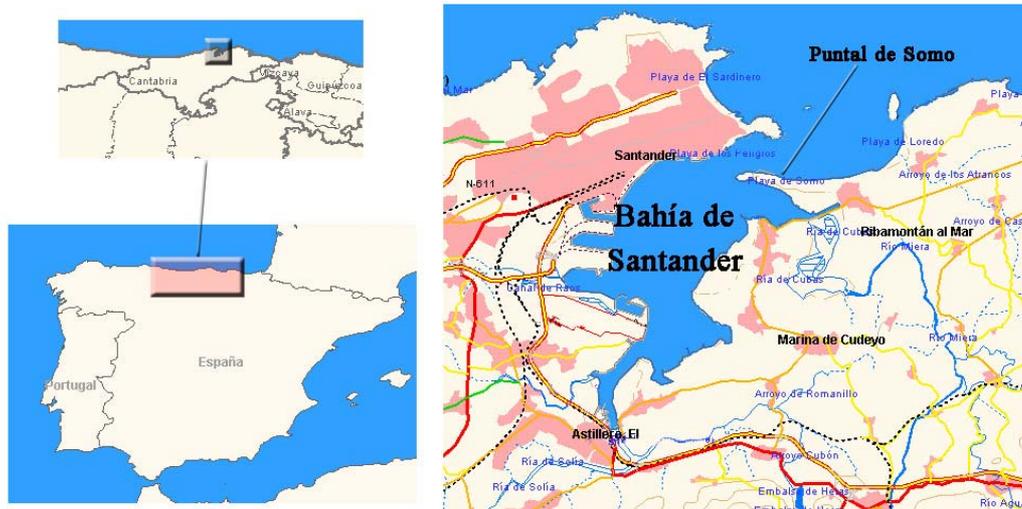


Figura 1.- Localización de la Bahía de Santander en el contexto de España.

Dicha lengua se encuentra ubicada al final del canal de navegación y es observada periódicamente mediante geodesia espacial, utilizando receptores GPS de doble frecuencia en modo RTK, debido al ritmo de crecimiento en longitud que viene experimentando, 5 m por año aproximadamente. En la actualidad se están tomando fotos en tiempo real utilizando sensores semimétricos, pero la determinación métrica de dicho avance es más precisa si se emplean batimetrías obtenidas hasta el momento con ecosonda de doble frecuencia y geodesia espacial. El análisis de componentes principales clásico y robusto se presenta como una alternativa en la toma de profundidades. En los puntos siguientes se analizarán las precisiones obtenidas en cada caso con el fin de validar dichos procedimientos indicando si se presentan como alternativas viables en el seguimiento y vigilancia de la navegación.

3.- MATERIAL

Las imágenes que se han utilizado en el estudio han sido capturadas por el sensor CASI, propiedad del Instituto Cartográfico de Cataluña. El CASI es un sensor óptico de barrido multiespectral basado en un dispositivo CCD bidimensional, con una dimensión espacial (columnas del CCD) y otra espectral (filas del CCD). Permite la selección de bandas para obtener información espacial y espectral. El rango de longitudes de onda sobre las que opera va desde los 405 nm hasta 950 nm. En la dimensión espacial ofrece 288 filas, cada una correspondiente a una franja de 1,8 nm. En cuanto a la dimensión espacial disponemos de 612 columnas, de las que 512 pasan a generar la imagen que trataremos y el resto se utiliza para la calibración de los datos.

La resolución espacial de las imágenes que se capturan depende, en su eje transversal a la línea de navegación, de la altura de vuelo y la focal del sensor (12,5 mm. que hace que el FOV valga 34.5°), obteniéndose resoluciones entre 2 y 10 m. En cuanto al tamaño del píxel en el sentido de la dirección de vuelo, está en función de la velocidad del avión sobre el terreno y del tiempo de integración. A menudo se buscan unos parámetros de altura de vuelo, velocidad y tiempo de integración que proporcionen un tamaño parecido en ambos ejes: transversal y a lo largo de la línea de vuelo. En la Tabla 1 se indican las especificaciones técnicas del sensor y en la Tabla 2 la configuración espectral proporcionada por el mismo.

Especificaciones técnicas del CASI	
Campo de visión	34.5° a lo ancho/0.077° a lo largo
Rango espectral	545nm entre 400 and 945nm
Filas espectrales	288 en intervalos de 1.9nm
Resolución espectral	2.2 nm en los 650nm
Apertura	F/2.8 a F/11.0
Rango dinámico	4096:1 (12 bits)
Ruido de base	1.0 ND
Relación Señal ruido	420:1 pico
Precisión de la calibración	470 - 800 nm +/-2% absoluta 430 - 870 nm +/-5% absoluta
Salida de datos	840 Kbytes a 1Mbyte/segundo

Tabla 1. Especificaciones técnicas del sensor CASI.

En el caso que nos ocupa se han obtenido imágenes cuyo tamaño de píxel es de 4 metros. Dichas imágenes han sido corregidas atmosférica (modelo 6S), radiométrica y geoméricamente con el fin de eliminar todos los posibles errores producidos durante el proceso de captación de los datos.

Los datos empleados para calibrar y validar las profundidades estimadas por teledetección fueron obtenidos mediante un sonar instalado en un barco. El instrumento principal es una sonda Atlas Elektronik Deso 15 que necesitó de diversos accesorios, como un pie de sonda con sujeción para la antena GPS, baterías y un generador de corriente alterna para alimentar los instrumentos además de un PC para almacenar los datos medidos (Figura1). Dicha sonda es capaz de medir un rango de profundidades comprendido entre los 0,2 y los 650 metros y dispone de dos canales de frecuencia (33 Khz y 210 Khz respectivamente), con una velocidad del sonido en el agua que se puede ajustar entre los 1400 y los 1600 m/s. Su peso aproximado es de 18 Kg.

Número Canal	Inicio (nm)	Final (nm)	Número Canal	Inicio (nm)	Final (nm)
1	408.6735	423.3988	19	678.8560	694.1093
2	423.3988	438.1647	20	694.1093	709.3793
3	438.1647	452.9698	21	709.3793	724.6648
4	452.9698	467.8127	22	724.6648	739.9645
5	467.8127	482.6924	23	739.9645	755.2770
6	482.6924	497.6072	24	755.2770	770.6011
7	497.6072	512.5560	25	770.6011	785.9353
8	512.5560	527.5374	26	785.9353	801.2784
9	527.5374	542.5502	27	801.2784	816.6290
10	542.5502	557.5929	28	816.6290	831.9860
11	557.5929	572.6643	29	831.9860	847.3478
12	572.6643	587.7631	30	847.3478	862.7132
13	587.7631	602.8878	31	862.7132	878.0809
14	602.8878	618.0373	32	878.0809	893.4495
15	618.0373	633.2102	33	893.4495	908.8179
16	633.2102	648.4052	34	908.8179	924.1844
17	648.4052	663.6208	35	924.1844	939.5481
18	663.6208	678.8560	36	939.5481	954.9074

Tabla 2. Bandas espectrales del sensor aerotransportado CASI.

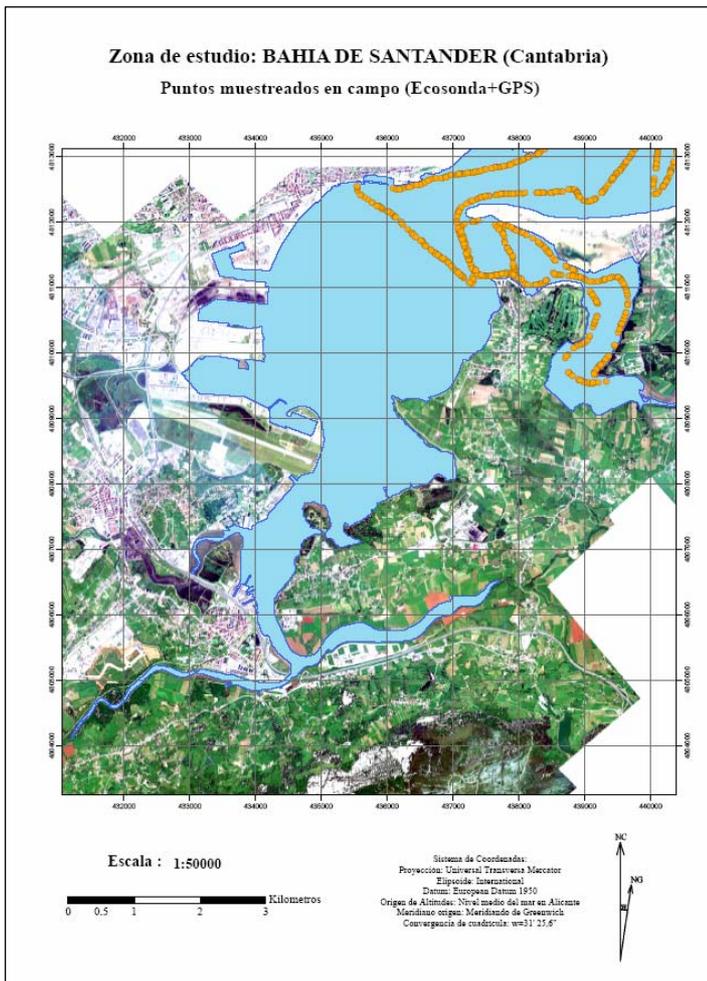
El proceso seguido para realizar el muestreo de profundidades, ha consistido en tomar valores de profundidad con la eco-sonda fijando simultáneamente la posición absoluta de ésta con el GPS. De este modo, se pueden integrar ambos datos para calcular un valor de profundidad respecto a un origen, que en nuestro caso es el nivel medio del mar en Alicante; si bien podría haber sido también el cero del puerto o bien el nivel instantáneo del mar.

Para completar el muestreo de profundidades se necesitaron 5 horas, la velocidad máxima de navegación fue de 6 km/h para conseguir una sincronización de los datos GPS-ecosonda precisa y así transformar las alturas a profundidades en el origen de referencia elegido.

El proceso de medición tenía la limitación de no poder medir profundidades menores a 1m, debido a que el barco necesitaba al menos esa profundidad para poder navegar, por lo que se circuló por los canales de navegación principalmente. Los puntos muestreados se indican en la Figura 2.



Figura 1. Material utilizado en el trabajo de campo.



Se observa como se han tomado puntos cubriendo toda la zona de estudio, con un total de 400 puntos muestreados, de los cuales se han utilizado 100 para la fase de calibración y 100 para la validación de los resultados. Se ha decidido excluir todos aquellos puntos que pertenecían a zona intermareal ya que las imágenes fueron tomadas en marea baja y los que quedaban fuera de la imagen capturada. El vuelo se realizó sobre las 12 del medio día (hora de reloj), una vez transcurridos 60 minutos de la hora prevista para la bajamar, de ahí que se observe suelo y no agua en dicha zona.

Figura 2.- Puntos muestreados *in situ*.

4.- METODOLOGY

Una vez corregidas las imágenes facilitadas por el sensor CASI, eliminando todos los factores que hubieran introducido ruido y, por tanto, perturbado la señal captada por el sensor, se procedió a la obtención de la componentes principales clásicas y robustas.

El análisis de componentes principales tiene por objetivo reducir la dimensionalidad de los datos, reuniendo en las dos o tres primeras componentes la mayor parte de la información producida por las variables originales (36 bandas capturadas por el sensor). Se han considerado las 5 primeras componentes de cada uno de los procedimientos estudiados ya que recogen más del 95% de la variación total.

El análisis de componentes principales clásico (ACP clásico) es un método estadístico ampliamente utilizada en teledetección multispectral que fue propuesto por Pearson en el año 1901. Los datos originales se transforman en un conjunto de nuevas bandas no correlacionadas entre si (componentes principales ortogonales) ordenadas de acuerdo a la cantidad de varianza que cada una explica.

El análisis de componentes principales robusto (ACP robusto) tiene por objetivo robustizar los resultados clásicos obtenidos en el análisis multivariante detectando la presencia de datos anómalos. Para ello es necesario robustizar la matriz de covarianzas ya que en ella están basados la mayoría de estos resultados. En el ACP robusto se ha utilizado como matriz de covarianza el estimador elipsoide de mínimo volumen, el cual alcanza un punto de ruptura casi igual a 0,5, ya que tiene menos eficacia que el del mínimo determinante de la matriz de covarianza pero tiene mayor punto de ruptura.

Ambos procedimientos han sido aplicados a las imágenes facilitadas por el sensor CASI mediante el programa “R: A Language and Environment for Statistical Computing” (versión 2.0.0) al cual se le ha añadido un módulo de estadística robusta implementado por Alfonso García Pérez (UNED). Todas las rutinas necesarias para la transformación de dichas imágenes en matrices así como su posterior inversión se han implementado en IDL 6.0.

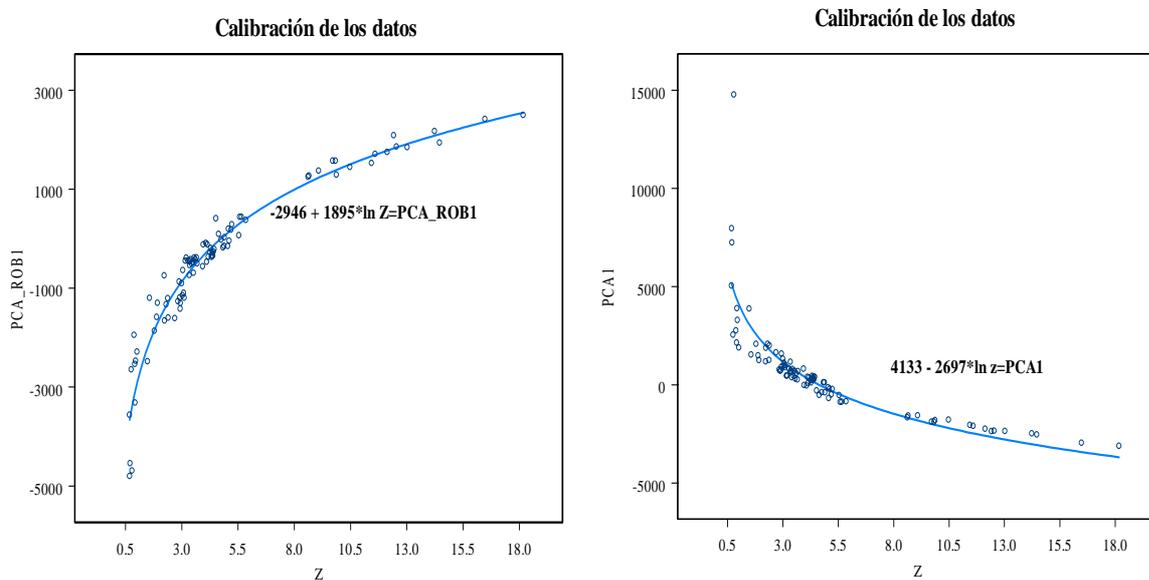


Figura 3.- Ajuste establecido entre los datos tomados *in situ* y la primera componente principal robusta y clásica de los datos recogidos por el sensor.

Una vez determinadas las componentes principales clásicas y robustas se calibraron dichos valores con los de las muestras tomadas *in situ* obteniéndose las correlaciones que se indican (Tabla 3).

	pcarob1	pcarob2	pcarob3	pcarob4	pcarob5
Z	0.8692	-0.0101	-0.2746	-0.4673	-0.1953
	pca1	pca2	pca3	pca4	pca5
Z	-0.7104	-0.1938	-0.2657	0.0732	-0.0555

Tabla 3. Correlación entre las componentes principales clásicas y robustas y el muestreo de campo.

Se observa que la primera componente es la que tiene mayor correlación con la muestras tomadas con la ecosonda-GPS (en ambos casos) y que los valores obtenidos en el caso robusto son considerablemente mejores que los proporcionados por el método clásico. La calibración de los datos proporciona las expresiones de ajuste que se indican a continuación (Figura 3).

Una vez invertidas dichas ecuaciones y generada la imagen correspondiente (Figura4) los errores que se observan se ajustan a las gráficas exponenciales, al igual que las ecuaciones de ajuste (Figura 5).

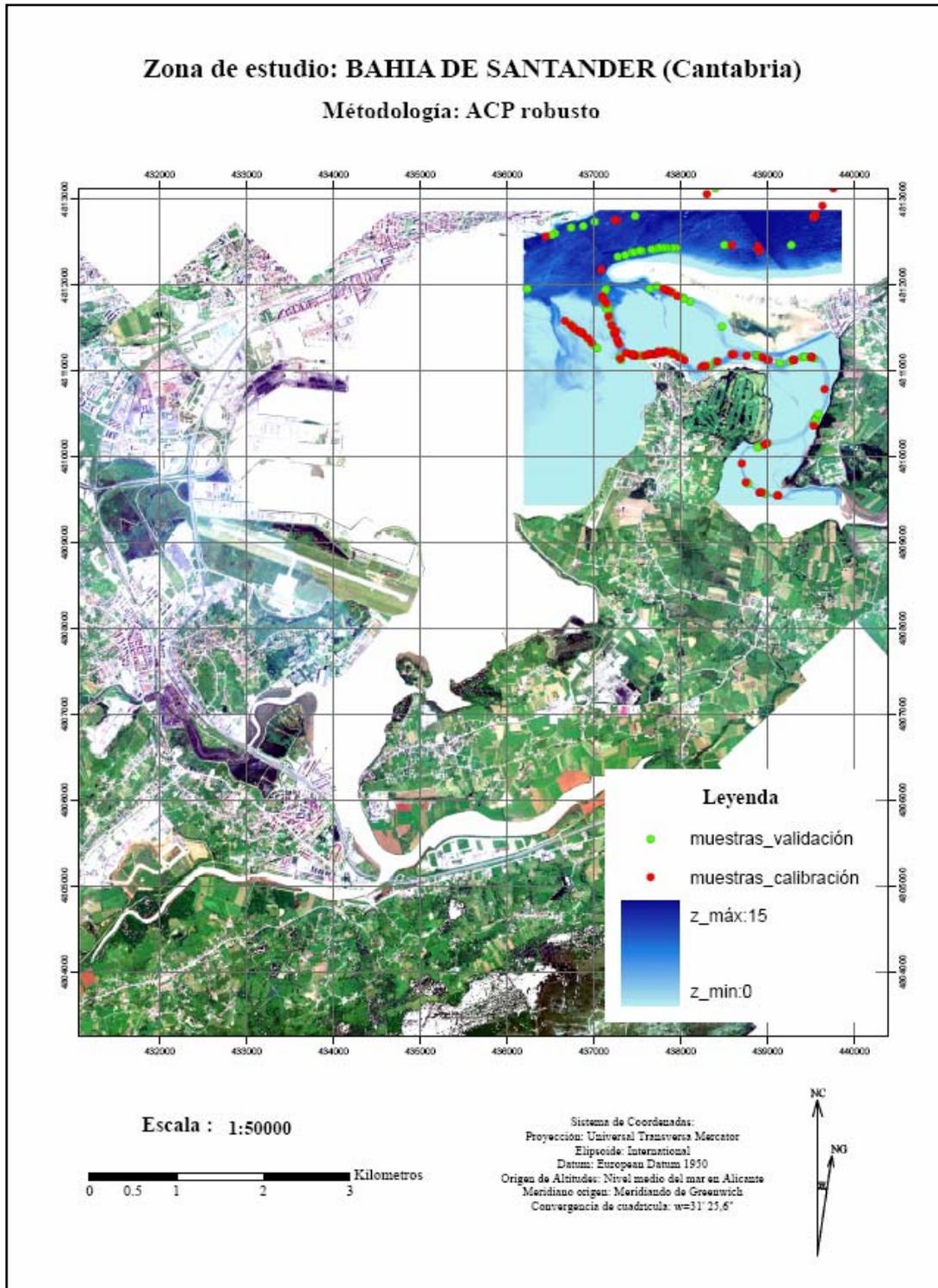


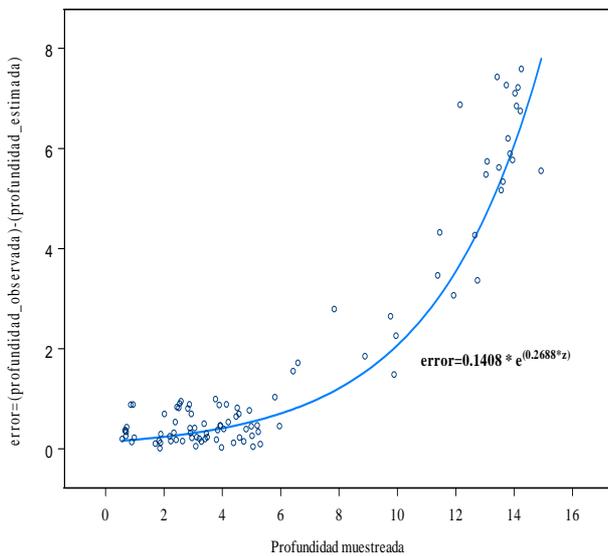
Figura 4. Plano de profundidades.

Se ha detectado que para profundidades inferiores a los 10 metros la media de error en valor absoluto, cometida en las muestras utilizadas para la validación de los resultados y correspondientes a dicho rango de cotas, es de 58.8 cm utilizando el ajuste con la primera componente principal robusta y de 76 cm utilizando la componente clásica. Si se consideran todas las profundidades (rango comprendido entre los 0 y los 15 m. aproximadamente) el error cometido en valor absoluto es del orden de 1,756 m en el caso robusto y de 2,088 en el clásico (Tabla 4).

	Media	Desviación típica	
Promedio (100 muestras de calibración)	1.756	2.317	PCA_robusto
Promedio de las profundidades $0 < z < 10$ m	0.588	0.585	
Promedio (100 muestras de calibración)	2.088	2.640	PCA_clásico
Promedio de las profundidades $0 < z < 10$ m	0.760	0.727	

Tabla 4. Estadísticas obtenidas tras realizar el ajuste.

Errores cometidos al realizar el ajuste con la PCA_ROB1



Errores cometidos al realizar el ajuste con la PCA1

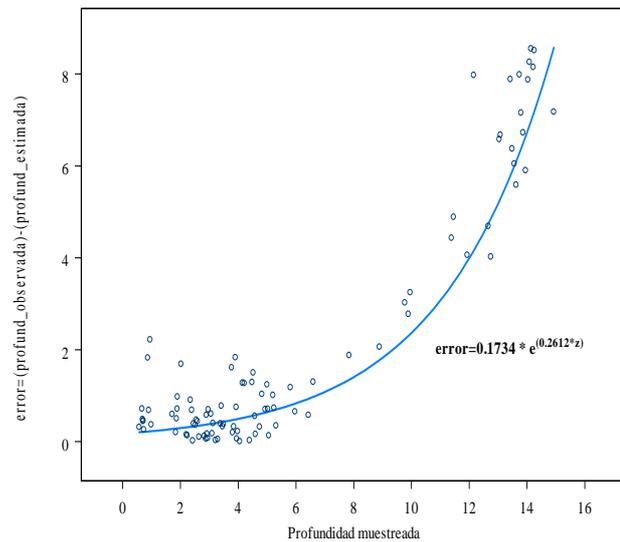


Figura 5. Errores cometidos entre las profundidades estimadas y las observadas al realizar el ajuste con la primera componente principal robusta y clásica.

5.- CONCLUSIONES

El análisis de componentes principales es un procedimiento válido y efectivo en la determinación de profundidades a partir de sensores satelitales o aerotransportados, obteniéndose precisiones aceptables para muchos trabajos de ingeniería civil.

El análisis de componentes principales robusto proporciona mejores resultados que el análisis clásico pero se observa que en ambas metodologías los errores aumentan considerablemente a partir de los 10 metros de profundidad.

Por otra parte, tanto en la fase de calibración de los datos como en la de validación de los resultados las ecuaciones de ajuste obtenidas siguen una forma exponencial respecto a la profundidad que se ajusta muy bien a los datos.

6.- BIBLIOGRAFIA

Bierwirth, P.N., Lee, T.J. and Burne, R.V. "Shallow sea-floor reflectance and water depth derived by unmixing multispectral imagery". PE&RS, Vol.59, N°3, pp.331-338, Mar.1993.

Lyzenga, D.R., "Passive remote sensing techniques for mapping water depth and bottom features", *Applied Optics*, Vol.17, N° 3, pp. 379-383, Feb.1978.

Mobley, C.D. "Estimation of the remote sensing reflectance from above-surface measurements". *Applied Optics* 38 (36), pp. 7442-7454. 1999.

Polcyn, F.C., Brown, W.L. and Sattinger, I.J. "The measurement of water depth by remote sensing techniques". Report 8973-26-F, Willow Run Laboratories, University of Michigan, Ann Arbor, 1970.

Sandidger, J.C. and Holyer, R.J. "Coastal bathymetry from hyperspectral observations of water radiance", *Remote Sensing of Environment*, Vol. 65, pp. 341-352, 1998.

Yarbrought, L.D. and Easson, G. "Comparison of Techniques for Deriving Bathymetry from Remotely Sensed Data" AMRS-Conference 2003.