“Contribution of classical and robust principal component analysis to depth estimation in coastal environments”

E. Castillo[1], V. Bayarri[1], R. Ferrer[1], J. Viguri[2], J. Sánchez[1], J.M. de Luis Ruíz[1], R. Pereda[1] and M. Ruíz[1].

[1] University of Cantabria, E.T.S.I. Caminos, Canales y Puertos, Avda. Los Castros s/n, Santander 39005, Cantabria, SPAIN.
Tlf: +34942201795. Fax: +34942201703
[2] University of Cantabria, E.T.S.I. Industriales y de Telecomunicaciones, Avda. Los Castros s/n, Santander 39005, Cantabria, SPAIN

Abstract:

The Santander Bay is located in the north of the Cantabrian Community (Spain). It has been suffering morphodynamic changes from the eighties due to anthropogenic activity. This fact can become dangerous in safety navigation within navigational channel terms; so accurate information is needed.

For more than 30 years, accurate bathymetric algorithms from multispectral imagery have been searched and Principal Components Analysis (PCA) has been traditionally considered as a reliable method. This method implies an advance respect to other traditional methods such as spectral rationing since water’s physical parameters are not needed.

This article are shows the experience obtained after applying the new PCA robust method developed by Huber's team to CASI (Compact Airborne Spectrograpic Imagery) imagery with 36 spectral bands within 400 to 950 nm range. Two different correction levels have been considered aimed to assess whether the final product has been improved. On the other side a GPS-Echo sounder bifrecuency performing in RTK mode has been employed to calibrate and validate the data taking measures on site.

1.- Introduction

Traditionally bathymetric methods have been based on sampling data by positioning planimetrically the point and a technique for measuring the thick of the water sheet. Nevertheless, sometimes the working area is not accessible or it is very hard to sample directly since they are deathtrap.

Nowadays bathymetries are made from GPS and Echo-sounder data, reaching higher accuracies than previously; but bathymetric works in shallow water with traditional sounding techniques are slow, expensive and even dangerous. This is why bathymetries from airborne or satellite imagery are provoking so much interest. Bathymetries obtained from this imagery have the advantage of availability and reasonable prices but unfortunately they are not so accurate.

Advances in technology in the last few years have generated new methodologies. High spatial, spectral and radiometric resolution sensors have generated news proceedings in depth estimation with very good results from the obtained accuracies point of view.

This information is required in multiple aspects of civil engineering (docks construction, dikes, submarine emissaries, trench control, etc…) and according to the final objective a different accuracy will be required.

2.- Study area location

The Bay of Santander is located in a privileged place within Cantabrian community (Spain). It is a depression whose ecosystem is very rich both from biological and from the socio-economical point of view. More than 250.000

inhabitants are concentrated in this area, which means more than 50% of the community population. The bay has been conceiving from the Mesozoic and it is very important for a population which develops several activities, all of them creating residually spills.

This ecosystem is very sensitive and with high ecological fragility. It presents sediments such as limes, clays and other particles from Miera River. There are lots of points which collect the municipalities' spills that surround the bay (Santander, Camargo, Astillero, Marina de Cudeyo, Ribamontan al Mar y Villaescusa) and that flow into it, either directly either nearby estuary. Fluvial, fecal and even residual waters from industries are collected.

The consequence of the coastal dynamics and mouth of the Miera River, also known as Cubas Estuary, has produced that a singular structure called "el puntal de Somo" had appeared. It is a sand tongue 2.5 km length and 250 m width approximately.

This tongue is located at the end of navigational channel and it is periodically measured by spatial geodesy, by using double frequency GPS receivers in RTK mode due to the growing rhythm in length that is experimenting, 5 m per year approximately. Nowadays Real Time photographs are being taken with semi-metric sensors, but the metric determination of the advance is more precise if double frequency echo-sounder & GPS bathymetries are used for the survey.

Classical and Robust Principal Components Analysis algorithms are presented as an alternative in depth sampling. The following sections content the result from assessing the accuracies in each case in order to being able to validate such proceedings showing if they can be presented as a viable alternative for monitoring this process.

3.- Material

The imagery used in the study has been taken by CASI sensor, owned by Catalonian Cartographic Institute (ICC). CASI is a pushbroom imaging spectrograph with a two-dimensional CCD array of 512 spatial pixels and 288 spectral pixels scanning the scene in the VNIR (405-950nm).It lets the user to set up the number and width of the bands in which the sensor will record data.

The sensor was installed on board of the plane "Cessna Citation I" belonging to ICC. The flight was made the 3rd of June and to capture the whole bay, 10 tracks were needed. The imagery was atmospherically (6s model), radiometrically and geometrically corrected and then orthorectified.

The spectral configuration of the images was configured by the ICC, with 36 channels and a spatial resolution of 4 x 4 m^2. Typical operating specifications and configuration are summarised in Table 1.

The images were recorded in *enhanced spectral* mode. This mode lets the users obtain a continuous channels configuration between a spectral range from 408 nm to 953 nm and minimum spectral width of up to 1.8 nm. The designed configuration has consisted in 15nm width per channels, and 36 spectral channels as it can be seen in the table 2 in the appendices.

The data for the bathymetries was obtained by using sonar installed on a ship. The central instrument was the sounder Atlas Elektronic Deso 15 which needed several accessories as a sounder foot with subjection for GPS antenna, batteries and AC generator to feed the instruments and the PC needed to store the data measured as it can be appreciated in figure 1. This sounder is capable to measure a depth range between 0.2 and 650 m by using two different frequency channels (33 Khz and 210 Khz). It is also possible to adjust the sound speed in water from 1400 to 1600 m/s.

Depth values are obtained when GPS and echo-sounder data are integrated. On one hand, two GPS receiver are needed to perform in RTK mode. The base station is sit on a point with known coordinates and the second receiver on the top of the sounder foot, differential corrections are transmitted with a radio-modem and X,Y and Z coordinates recorded each second. On the other hand, echo-sounder records each three seconds a depth value in both frequency; to be able to obtain the depth values, first of all the recorded data have to be cleared of noise. A value can be considered as noise when the difference between depths in both frequencies is not very close to zero or when there is a big difference between the previous and following point. Zero values are received when the power of the signal is not enough to estimate the depth, and it have to be tuned. Once depth values are corrected, they are linked with the GPS data which have the same GPS time and X, Y and Z coordinates for all the points calculated.

Table 1.- Technical specifications of CASI sensor.

Parameter	**Description**
IFOV (Instantaneous Field Of View)	
Across Tack	54.4 ° (custom lens)
Along track	0.1151 °
Aperture	f/2.8 - f/11 (Automated iris control)
Spectral range	405 - 950 nm
Spatial samples	512 spatial pixels
Spectral samples	288 at 1.8nm intervals (2.2nm FWHM @ 650nm)
Dynamic range	16-bits (65536 levels)
Operating Modes	
Spatial Mode	512 pixels across swath, up to 18 spectral bands (fully programmable).
Spectral Mode	full spectrum (288 channels) for up to 39 look directions spread across swath (4, 8, 12, or 16 pixel spacing between look directions). Includes a monochromatic image at full spatial resolution (Scene Recovery Channel).
Enhanced Spectral Mode	full spectrum (288 channels) in a block of 101 adjacent spatial pixels.
Full Frame	512 pixels across swath x 288 spectral pixels (~1-2 sec. Integration time limits use to laboratory calibration or ground-based field use).

A statistical process was develop to extract the data used to calibrate and to validate the methods; depth values were extracted according the frequency they appeared in the samples taken.

Depth measurements were taken during 5 hours, the maximum speed was 6 km/h to be able to synchronize the sounder and GPS data and hence correct both measurements to obtain depth values in the Spanish official reference system for heights. Nevertheless any other height reference system could have been chosen.

The sampling process was limitated to depth higher than 1m, since the vessel needs at least that depth to be able to sail, so sampling was made along navigational channels. Sample points can be appreciated in figure 2.

Points covering all the study area have been taken. 400 points have been statistically extracted and all the points within intertidal region removed since the flight was made in a very low tide. Two sets of 100 points were taken to calibrate and to validate the results. The flight was made at 12 a.m. (clock time), 60 minutes after low tide; this is why soil instead of water is observed in that area.

Table 2.- CASI Spectral configuration of the flight.

Channel number	From (nm)	To (nm)	Channel number	From (nm)	To (nm)
1	408.67349	423.3988	**19**	678.856	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	755.277
6	482.6924	497.6072	**24**	755.277	770.6011
7	497.6072	512.556	**25**	770.6011	785.9353
8	512.556	527.5374	**26**	785.9353	801.2784
9	527.5374	542.5502	**27**	801.2784	816.629
10	542.5502	557.5929	**28**	816.629	831.986
11	557.5929	572.6643	**29**	831.986	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	678.856	**36**	939.5481	954.9074

Figura 1.- Material for field work.

4.- Methodology

Once imagery taken by CASI sensor was corrected, and all the factors which could have inserted noise and disturbed the signal received had been eliminated, the classical and robust Principal Components Algorithms were performed.

Principal Components analysis is aimed to reduce the dimensionality of data, collecting in the two or three first components the majority of the produced information by the original variables (36 bands). The first 5 principal components have been considered of each of the studied proceedings since they collect more than 95 % of the total variation.

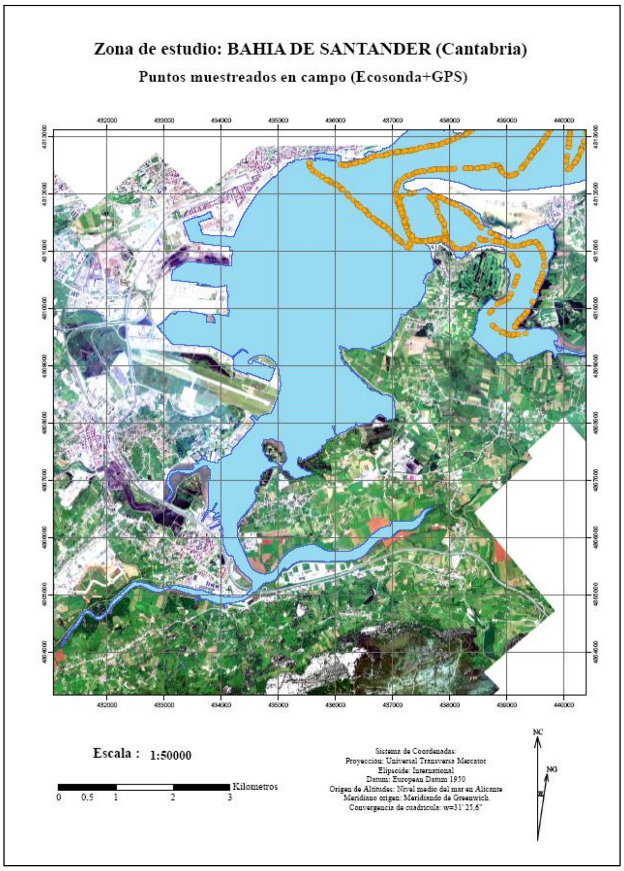

Figure 2.- Sample points taken on site.

Classical Principal Components Analysis (Classical PCA) is a very common statistical process in multispectral remote sensing that was proposed by Pearson in 1901. Original data is transformed into a new set of uncorrelated bands (orthogonal principal components) sorted by the amount of variance that any of them explains.

Robust Principal components Analysis (Robust PCA) has the objective of “robusting” the classical data obtained in the multivariate analysis detecting the presence of anomalous data. To be able to do this It is necessary to robust the covariance matrix since the most of results are based on it. The minimum volume ellipsoid estimator has been used as a covariance matrix. This estimator reaches a breaking point close to 0.5, since it has less efficacy than the one of least determinant of the covariance matrix but it has higher breaking point

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.

Both proceedings have been applied to CASI imagery by using the Program “R: A Language and Environment for Statistical Computing” v. 2.0.0 and a robust statistics module implemented by Professor Alfonso García Pérez (UNED). All the routines needed to convert imagery into matrices and undo such process. This procedure has been implemented in IDL 6.0.

Once Classical and Robust PCA have been determined, the calibration process was developed with the on-site samples, obtaining the correlation coefficients that can be appreciated in table 3.

Table 3.- Correlation between classical - robust PCA ando n site simples.

	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

As it can be appreciated, the first component is the one with highest correlation with the Echosounder-GPS samples (in both cases) and the obtained values in the robust case are considerably better that the obtained with the classical method. The data calibration give the adjust expressions which are indicated in figure 3.

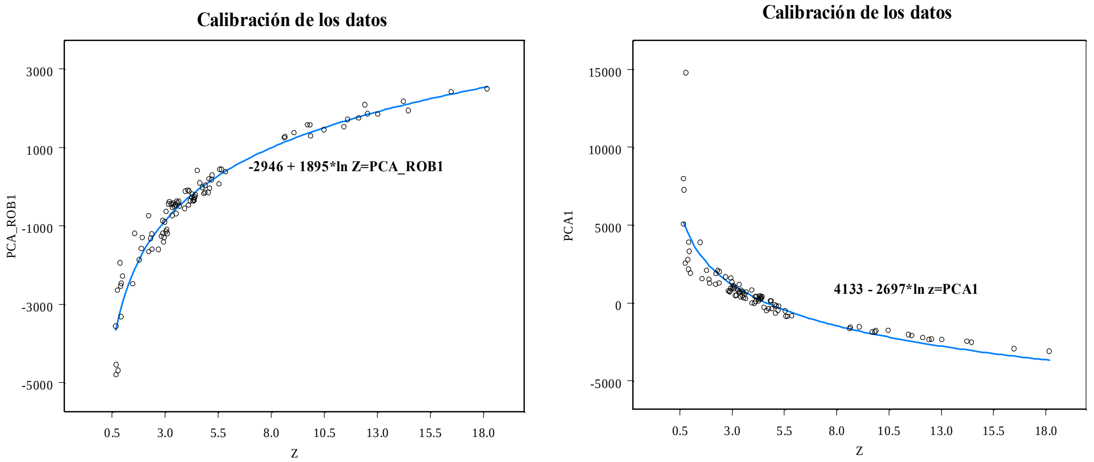

Figure 3.- Adjust between on site data and the first robust and classical principal components of the imagery

Once the previous equations are inverted and the corresponding image generated (Figure 4) the observed errors can be approximated by exponential functions, in the same way than the adjusted equations. (Figure 5)

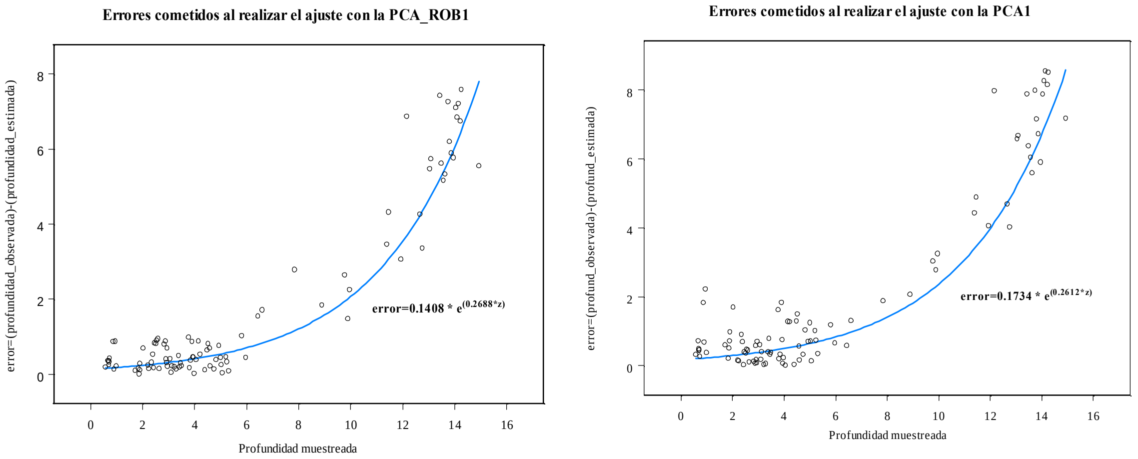

Figure 5.- Errors when an adjust with the first classical and robust principal component is performed.

It has been detected that when depths are lower than 10 m, the mean absolute error value for the samples in the validation of the results is 58.8 cm when it is adjusted to the first robust principal component and 76 m. when the classical component is used. If the whole depth range is considered (between 0 and 15 m approximately) el absolute mean error is 1.75 in the robust case and 2.08 m in the classical case (Table 4)

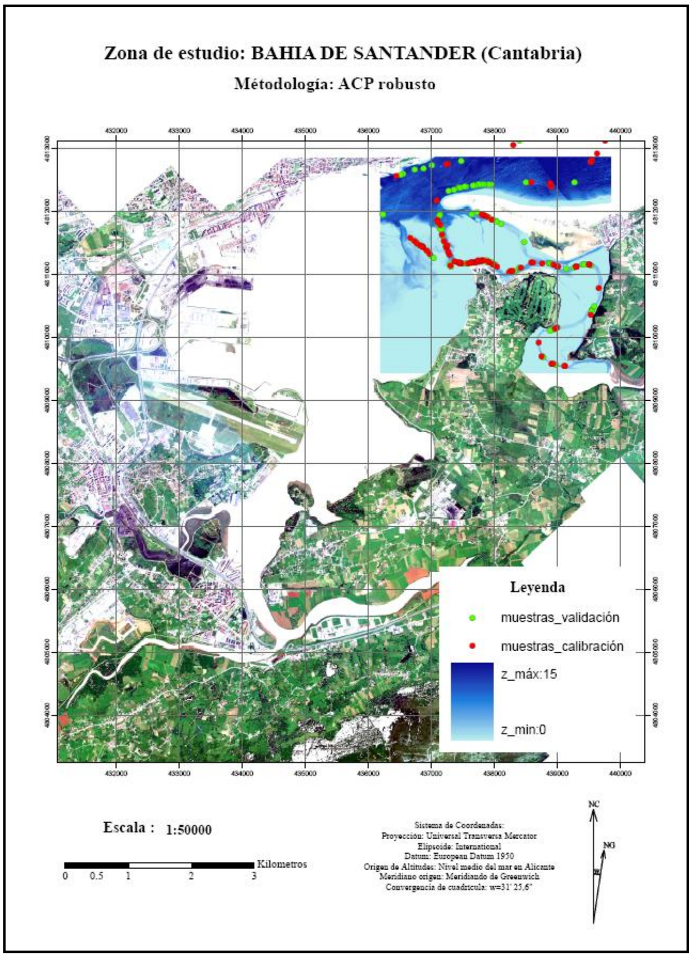

Figure 4.- Depth plane containinng the generated bathymetry and calibration and validation samples.

Table 4.- Statistics obtained for the adjust.

	Mean	Standard Deviation	
Average (100 calibration samples)	1.756	2.317	PCA_robusto
Depth Average 0<z<10 m	0.588	0.585	
Average (100 calibration samples)	2.088	2.640	PCA_clásico
Depth Average 0<z<10 m	0.760	0.727	

5.- Conclusions

Principal components analysis is a valid and effective process for depth determination from both satellite and airborne sensors, and it is possible to obtain acceptable accuracies for some works in civil engineering.

Robust Principal Components Analysis provides better results than classical but it is observed that both methodologies provide high errors when depth is higher than 10 m.

It has been also observed that the best adjust equations for calibration and validation data are those that follow an exponential model to depth values.

6.- Bibliography

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, Nº3, pp.331-338, Mar.1993.

Lyzenga, D.R., "Passive remote sensing techniques for mapping water depth and bottom features", Applied Optics, Vol.17, Nº 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.