CHAPTER 172

COMPLEX PRINCIPAL COMPONENT ANALYSIS OF SEASONAL VARIATION IN NEARSHORE BATHYMETRY

Guoxiong Liang[1], Thomas E. White[2], Member, ASCE
and Richard J. Seymour[1], Member, ASCE

ABSTRACT Both the conventional Principal Component Analysis (PCA) and the Complex Principal Component Analysis (CPCA) were applied to analyze six nearshore profiles in Siuslaw, Oregon. Results indicated that the first two components derived from CPCA always outperformed those derived from PCA. This suggests that CPCA is a better method of describing the seasonal variation in nearshore bathymetry. The relative performance of CPCA on different profiles depended upon the coherence of the variation within those profiles. Furthermore, the concept of an absolute amount of variance was used in explaining the spatial variation of the predictability of principal components.

INTRODUCTION

Complex principal component analysis (CPCA), developed for meteorological application (e.g., Wallace and Dickson, 1972; Barnett, 1983), has been successfully used to describe an event of a fast-moving sand bar (Liang and Seymour, 1991). In comparison with conventional principal component analysis (PCA), also known as the Empirical Orthogonal Function (EOF) technique (e.g. Aubrey *et al.*, 1980; Seymour, 1989), CPCA offers significant advantages. Besides being able to give a more compact description for the variation of the data set (fewer functions required), it can also detect propagating waves.

However, most nearshore bathymetric data sets available consist only of seasonal surveys. One may argue that the seasonal variation in nearshore profiles is more like a standing wave than a propagating feature, and thus it might be unnecessary to apply CPCA to the seasonal bathymetric data. Therefore, a test using

1 Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0222, USA.

2 USAE Waterways Experiment Station, Coastal Engineering Research Center, 3909 Halls Ferry Road, Vicksburg, MS 39180-6199, USA.

seasonal data appears to be a rigorous method of evaluation. In order to compare the efficiencies of CPCA and PCA in analyzing the seasonal variation in nearshore bathymetry, a set of multiple profile data from Siuslaw, Oregon, was analyzed.

SEASONAL SURVEY DATA FROM SIUSLAW, OREGON

The bathymetric data were collected from an area that is located in the mouth of Siuslaw River, Oregon (Figure 1). The Helicopter-borne Nearshore Survey System (HBNSS) was applied to measure the seabed elevation (Pollock, in press). The survey helicopter is fitted with a 26-meter weighted cable, graduated like a surveyor's rod. The elevations are read by a shore-based surveyor's level. The horizontal positioning is obtained using a shore-based electronic total distance station (TDS) aiming at a cluster of prisms mounted on the helicopter. The HBNSS surveys were carried out every winter and summer from 1981 to 1990.

Six cross-shore profile sets, which are within an alongshore segment extending 762 m (2500 feet) from the North Jetty, were chosen for analysis. Each of them, with a length of 354 m, includes 30 grid points and 20 time steps. Some profiles show pronounced seasonal cycles (Figure 2).

ANALYSIS TECHNIQUE AND RESULTS

To apply CPCA, a profile is required to be transformed into a complex data set such as:

$$U_j(t) = u_j(t) + i\hat{u}_j(t)$$

The real part is simply the original scalar field. The imaginary part is the Hilbert transform of the real part. On the basis of the complex data, the complex cross-correlation matrix can be computed, consisting of: the eigenvectors (functional decompositions of the data) and eigenvalues (portions of data's variation represented by each eigenvector). It is customary to evaluate the performance of these analysis tools by the percentage of the variation about the mean, which is represented by each of the principal components. A comparison of PCA and CPCA is shown in Table 1.

Table 1 Percentage of Variation Explained by PCA and CPCA

Profile	1st Component		1st & 2nd Components	
	PC	CPC	PC	CPC
1	86.57	85.65	95.43	95.47
2	77.00	75.28	88.21	89.41
3	60.35	58.65	75.36	78.33
4	56.01	55.50	71.95	73.80
5	38.83	40.71	58.77	64.22
6	36.91	41.63	55.98	64.46

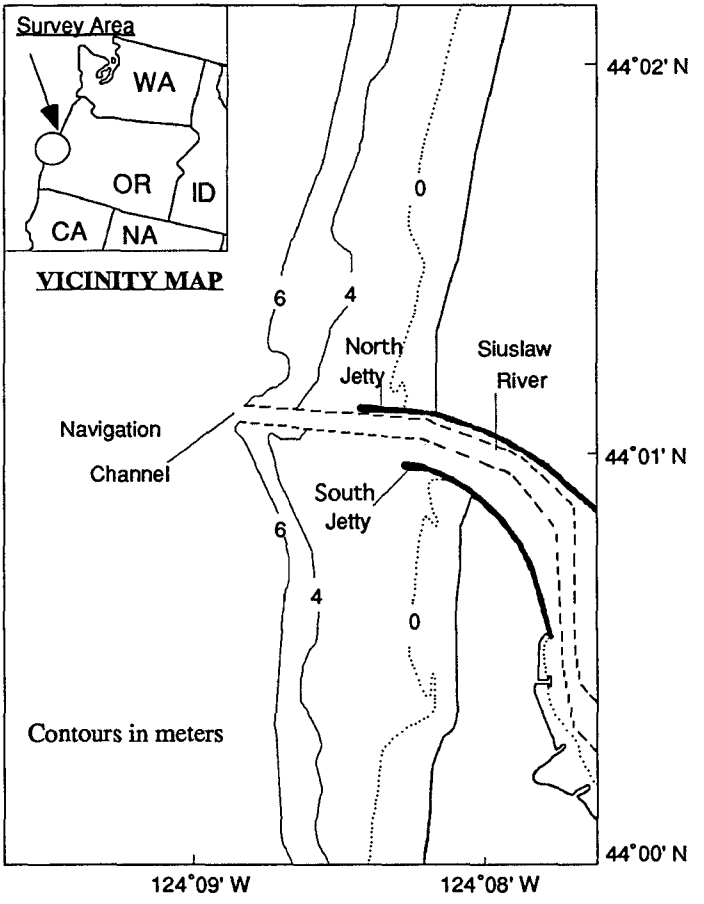

Figure 1. Location map of Siuslaw, Oregon.

The first and second complex components (CPC) can explain more variation than the corresponding conventional components (PC) in every case. In some

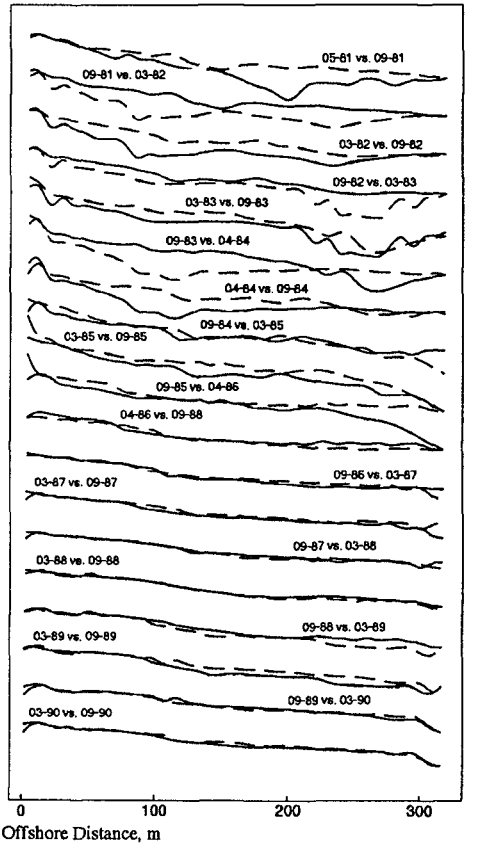

Figure 2. An example of seasonal variation in nearshore profile (Profile 1).

profiles such as profile 6, the CPC is about 8.5% better. The data in Table 1 are shown graphically in Figures 3 and 4.

DISSCUSSION

However, it must be noted that, in some profiles (e.g. Profile 1 and Profile 2), the first conventional component can explain 1% to 2% more variation than the first complex component does. Also, the less variation contained in the first component, the better the performance of the first complex component compared to the conventional one. One possibility is that the complex analysis requires more degrees of freedom, and when artificially constrained to a very low number of functional modes (e.g., only one component), it will not behave reliably statistically. From the data in Table 1, it can be seen that once the number of components exceeds the bare minimum necessary for computation (one), the complex method performs better.

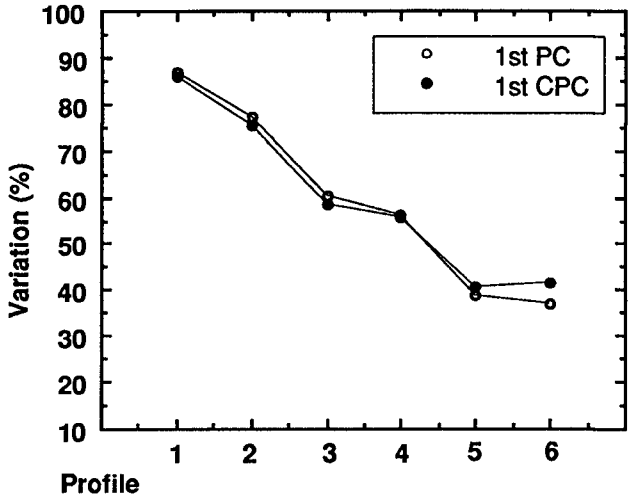

Figure 3. Variation explained by the first component.

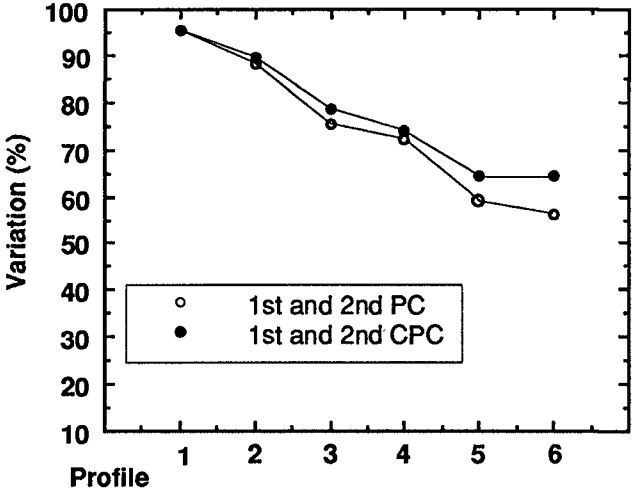

Figure 4. Variation explained by the first and second components

However, for the case of a fast-moving bar, the first component always outperformed the conventional component (Liang and Seymour, 1991).

The data suggest that the relative performance depends upon the coherence of the variation within a profile as described in the following paragraphs. For simple variation (coherent changes), most of the variation can be explained by the first component and the complex method has no particular advantage for a single component. For more complicated variations (non-coherent changes), there is less variation explained by the first component, and CPCA is always more effective.

Figure 5 shows the conventional correlation (left) and complex correlation (right) between each grid point and grid points 5, 10, 15, 20 and 25 in Profile 1.

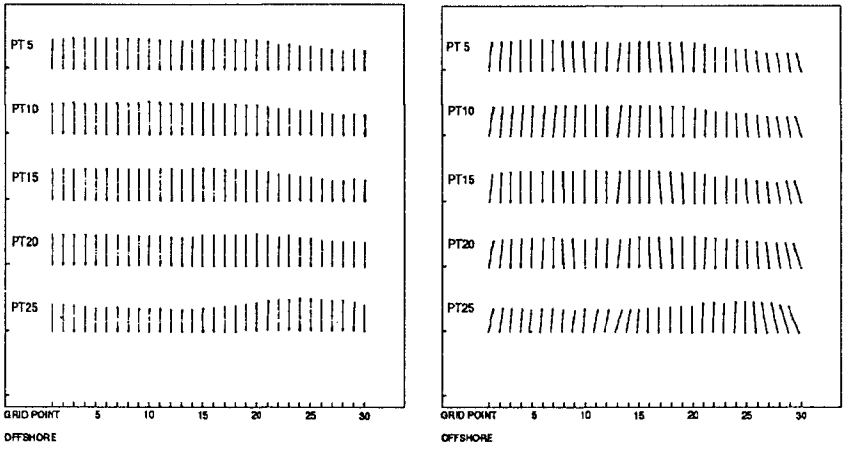

Figure 5. Comparison between conventional correlation (left) and complex correlation (right) in Profile 1.

Each correlation is plotted in a vectorial format where the magnitude is indicated by the length of the vector. Full scale (correlation equals 1.0) is indicated by the correlation between point 5 and itself. For the complex correlation, the phase is arranged like the hour hand of a clock. A vector pointing upwards (downwards) indicates that the two time series are in-phase (out-of-phase); one pointing to the right (left) indicates that the grid point time series lags (leads) the time series indicated in the left margin by 90°, etc. It is obvious that the complex correlation between the grid points within this profile is rather similiar to the conventional correlation. Every grid point shows positive correlation with each of the other ones. The entire profile appears to be involved in highly coherent motion. In this case, the results of CPCA and conventional PCA are almost identical.

A different result is revealed in Profile 6 (Figure 6). The difference between complex correlation and conventional correlation is significant. The correlation between grid points are much poorer than those in Profile 1. Also, correlation between certain points show an out-of-phase relationship. It suggests that the motion in this profile is not as coherent as that in Profile 1. In this case, CPCA can explain more variation than conventional PCA does.

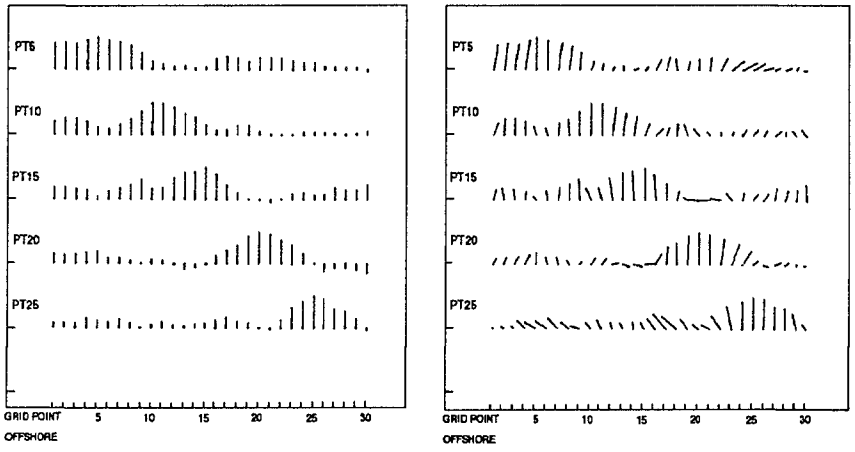

Figure 6. Comparison between conventional correlation (left) and complex correlation (right) in Profile 6.

In many instances, the first CPC may indicate a strong in-phase relationship between the time series at grid points while the second one indicates a weaker, out-of-phase relationship (Horel, 1984). In the conventional PCA, the correlation is simply a scalar. Therefore, it is not possible for it to reveal the phase relationship between different grid points. The data and results described above indicate that, within some standing-wave-like variation in nearshore profiles, both in-phase and out-of-phase correlations can exist. Therefore, since this cannot be determined *a priori*, use of the complex technique appears to be the prudent approach.

One of the other interesting observations is that the percent of the variation explained by either PC or CPC decreases monotonically with distance from the jetty (see Figures 3 and 4), a reduction in predictive capability of about one third. To understand this change better, the absolute value of the variability was calculated. Figure 7 shows the mean value (over time) of the standard deviations (in space) of the six profiles. This shows clearly that there was much greater variability close to the jetty (nearly three times that of the minimum). Figures 8 and 9 show the data of Figures 3 and 4 normalized by the standard deviations shown in Figure 7. These show that, in both PC and CPC, profile 4 is predicted best in terms of the absolute amount of variance, and Figure 7 shows that profile 4 exhibits very close to the minimum amount of absolute variance. Therefore, the performance on profile 4 could be considered to represent the basic capability of the principal component method to extract signals from the noise in the measurements and the physical processes. If this noise is uniform in magnitude, then as the absolute values of the variance increases (toward the jetty) the relative predictability ought to increase – and it does. On the other hand, if noise is somehow proportional to the signal, the relative predictability ought to remain more or less constant – and it does not. The

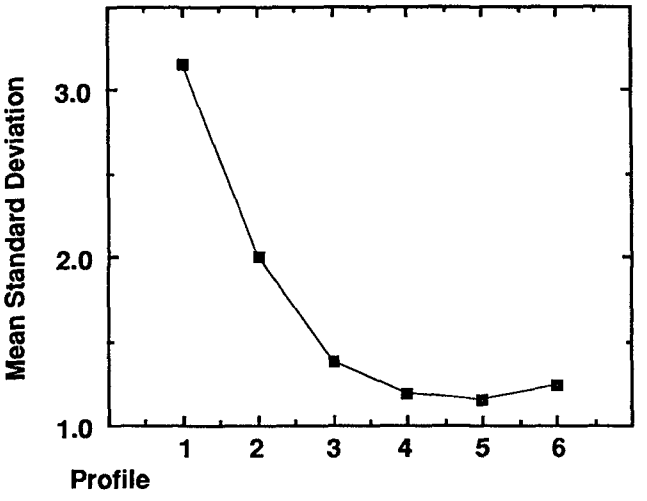

Figure 7. Mean standard deviations of profiles.

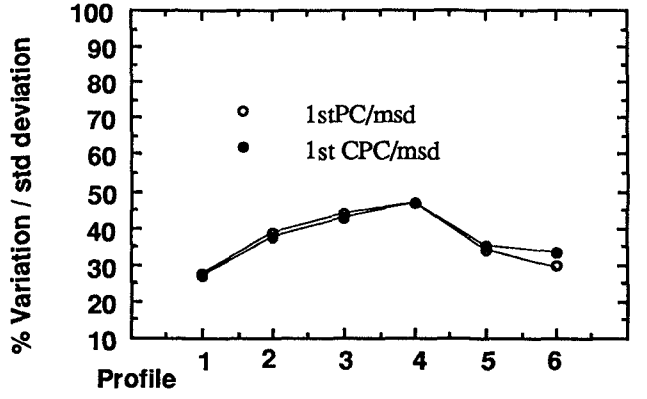

Figure 8. Variation explained by the first component normalized by the standard deviations.

data set is too sparse to make conclusive statements on this hypothesis, but it might be of interest to test this on richer data.

ACKNOWLEDGMENTS

This research was supported by US Army Engineer Waterways Experiment Station, contract DACW39-90-K-0007. The CERC MCCP program kindly provided the observed data which made this work possible.

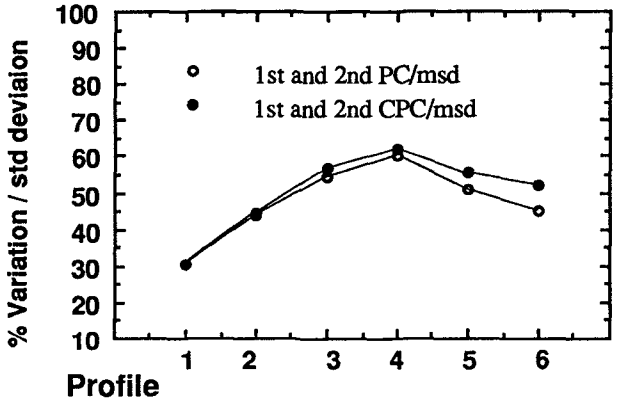

Figure 9. Variation explained by the first and second components normalized by the standard deviations.

REFERENCES

Aubrey, D.G., D.L. Inman and C.D. Winant (1980). The statistical prediction of beach changes in Southern California. *J. Geophys. Res.* 85(C6), 3264-3276.

Barnett, T.P. (1983). Interaction of the monsoon and Pacific trade wind systems at Interannual time scale. Part I: The equatorial zone. *Mon. Wea. Rev.* 111, 756-773.

Horel, J.D. (1984). Complex principal analysis: Theory and examples. *J. Climate Appl. Meteor.*, 23, 1660-1673.

Liang, G. and R.J. Seymour (1991). Complex principal component analysis of wave-like sand motions. In: *Coastal Sediments '91*, N.C. Kraus, K.J. Gingerich and D.L. Kriebel, Eds. ASCE. pp. 2175-2186.

Pollock, C.E., (1992). Helicopter-borne Nearshore Survey System, a valuable tool in difficult survey area. In press.

Seymour, R.J. (1989). Cross-shore transport. In R.J. Seymour, Ed., *Nearshore Sediment Transport*, Plenum press, New York. pp. 273-286.

Wallace, J.M. and R.E. Dickson (1972). Empirical orthogonal representation of time series in the frequency domain, Part I. Theoretical considerations. *J. Appl. Meteor.*, 11, 887-892.