

STATISTICAL PREDICTION OF COASTAL AND ESTUARINE EVOLUTION

Vanessa Magar¹, Markus S. Gross², George Probert³, Dominic E. Reeve⁴, Yuzhi Cai⁴

This paper presents a novel data-driven methodology based on empirical orthogonal teleconnections (EOTs) to analyse and forecast the evolution of coastal navigational channels near the mouth of the Exe estuary, UK. This is the first time EOTs are used in coastal morphodynamics. Therefore, particular emphasis is placed on the comparison of EOTs with the well established empirical orthogonal functions (EOFs) method. EOTs and EOFs are used with a series of 14 surveys, taken approximately every 8 months, covering the period between January 2001 and February 2010. The skill of the methods in producing accurate bathymetric one-step forecasts for February 2010 is analyzed and compared with one-step forecasts based on the raw data. It is found that, provided the order of the autoregressive forecast method is chosen appropriately, EOTs and EOFs are better than the raw data and EOTs outperforms than EOFs. This is attributed to the fact that EOTs, without the orthonormality restriction for the temporal eigenfunctions required in EOFs, capture the temporal patterns within the data more accurately than EOFs.

Keywords: empirical orthogonal functions; empirical orthogonal teleconnections; forecasting skill.

INTRODUCTION

The understanding and forecasting of nearshore sandbank and channel evolution is of crucial importance for marine and coastal users, including: port and harbor developers, leisure and commercial mariners, and sand dredging companies. However, prediction of the long-term evolution of sandbanks and navigational channels remains elusive, particularly in environments where both waves and tides are of importance. River deltas, estuaries and barrier island systems are examples of such environments. This paper focuses on a new, data-driven method to characterise and forecast regional geomorphic evolution at monthly to yearly timescales.

Data-driven methods are statistical techniques that have been extensively used to analyse the evolution of nearshore bedforms. Some of these techniques are based on bulk statistics; others rely on some data decomposition method. The data decomposition method most commonly applied in coastal morphodynamics is the empirical orthogonal function (EOF) analysis, first introduced to this field by Winant et al. (1975). The key advantages of data-driven techniques over process-based models include stability and robustness, straight forward implementation, and ability to assess uncertainty in forecasts; their key disadvantages are their lack of generality (as they are always applied to particular case study sites) and their inability to forecast behaviours not contained in the observations (Reeve et al., 2008).

Different linear and nonlinear statistical techniques applied to different coastal locations are discussed by Larson et al. (2003) and Southgate et al. (2003), respectively. The methods focused, generally, on the assessment of spatial and temporal characteristics of nearshore bars, shoreline sand waves, and nesses. More recently, comparisons of nonlinear techniques have been performed on nearshore sandbar evolution (Rattan et al., 2005; Pape and Ruessink, 2008). Although nonlinear techniques, such as Singular Spectrum Analysis (SSA) or wavelets, are better at capturing non-stationary features, they can be quite difficult to implement and interpret (Reeve et al., 2007; Magar et al., 2012).

This paper focuses on a new methodology that has not been used before for pattern analysis in coastal morphodynamics, namely the empirical orthogonal teleconnections (EOT) method (Van Den Dool et al., 2000). Given that there are some similarities between EOFs and EOTs, a comparison of the forecasting skill of EOFs and EOTs is made to assess whether EOTs are superior to EOFs in their ability to produce accurate forecasts. The rationale for testing this new method is that EOTs are more flexible than EOFs: the temporal eigenfunctions in EOTs are not required to satisfy the orthonormality condition. Hence the hypothesis is that EOTs will capture the temporal patterns within the data more accurately than EOFs and hence will give superior forecasts. To test this hypothesis, EOTs and EOFs are applied to the evolution of a coastal region located eastwards of the Exe estuary. This region has been surveyed roughly every 8 months since 2001 and the methods are assessed against the last survey available, carried out in February 2010. The forecasting method used is Burg's maximum entropy method (Burg, 1967, 1968) and a one-step

¹Plymouth University, PL4 8AA, UK; ²University of Cambridge, UK; ³Bridgezone Ltd, UK; ⁴Swansea University, UK.

forecast is performed in order to assess the forecasting skill of EOT and EOF reconstructions. The survey data and the methodology are described in detail in the next section.

METHODS

Starting with an outline of the survey data this section introduces the Empirical Orthogonal functions (EOFs) and Empirical Orthogonal Teleconnections (EOTs) utilised in this study.

Survey data

The Exe estuary (or the Exe) is located in the Southern coast of the Devon County, in the UK. Its position within the UK is indicated with the red dot shown on the map on the left of Figure 1. The Exe is a site of special scientific interest (SSSI) because of its large migrant and wading bird populations visiting the estuary and its saltmarshes every winter. The circulation around the ebb delta is driven by waves and tides. The mean spring tide range is 3.80m and the mean neap tide range is 1.48 m. Tidal current speeds along the coast are relatively weak, but in the approach channel between Pole Sands and Exmouth Beach, as well as in the channel between Exmouth and Dawlish Warren spit, they may reach speeds of 3 ms^{-1} during spring tides and 1 ms^{-1} during neap cycles. These are high velocities, capable of moving large quantities of sediment, up to medium sized sand. A rose diagram summarising 20 years of wave height data shows that the majority of the waves come from the South and that most of them have a height of less than 3m. High-energy SE storms are most likely causes of significant damage to Exmouth Beach and to Dawlish Warren (Halcrow, 2008). Pole Sands, the ebb-tidal shoal south of the mouth of the estuary, moves over time and this motion shifts the position of the navigational channel located between Pole Sands and Exmouth Beach. The different features of the estuary are shown on the right of Fig. 1, with mudflats and sandflats in brown. Although it appears that Exmouth Beach is reliant on sediment supply from the tidal deltas during storm events, conclusive research focusing on regional, detailed morphodynamical evolution of the area at timescales of 0-20 years is still needed. This, however, would not be achieved with the data available and is, thus, beyond the scope of the present work.

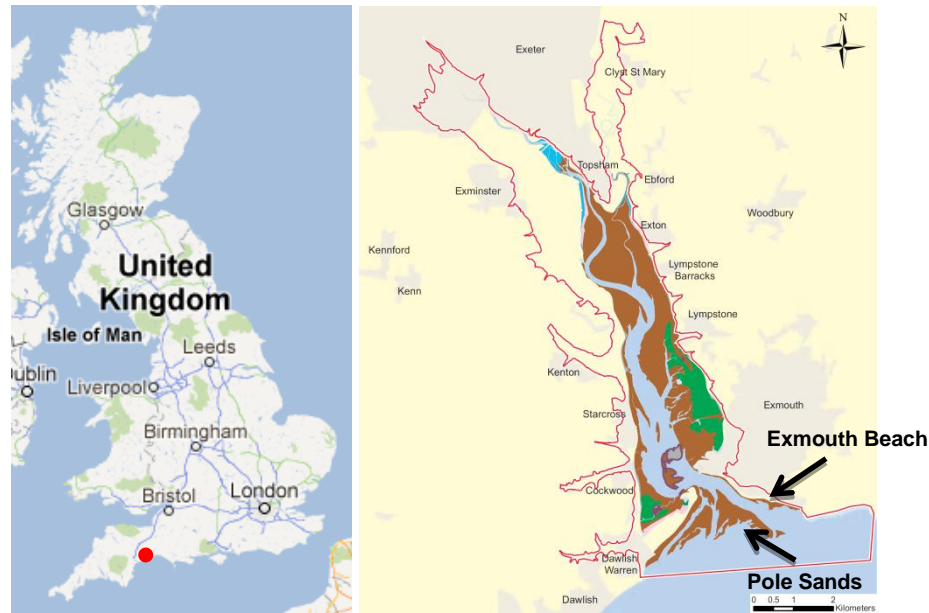


Figure 1. Location of Exe (left) and main features (right), with mudflats and sandflats in brown (Crown, 2006).

Given that the estuary is heavily used for recreational navigation, the Exe Harbour Authority (Exeter County Council) monitors regularly the position of the navigation channel north of Pole Sands. Surveys carried out are sent to the UKHO for archiving. Figure 2 shows the location and details of the first digital, geo-referenced bathymetry made available; the survey was taken in January 2001. A total of 14 surveys

were obtained from the UKHO, spanning the period between January 2001 and February 2010. The surveys were taken approximately every 8 months, because in 2005 no surveys were taken, a time interpolation had to be performed to provide surveys throughout the period of interest that were evenly spaced in time. The timestep was adjusted slightly and was of 7.58 months to ensure that the first and last surveys coincided exactly with the UKHO surveys. Similarly, a spatial interpolation was performed and the area analysed corresponded only to the area covered by all the surveys. Although this reduced the size of the area significantly, it was large enough to include the northern part of Pole Sands and part of the navigational channel.

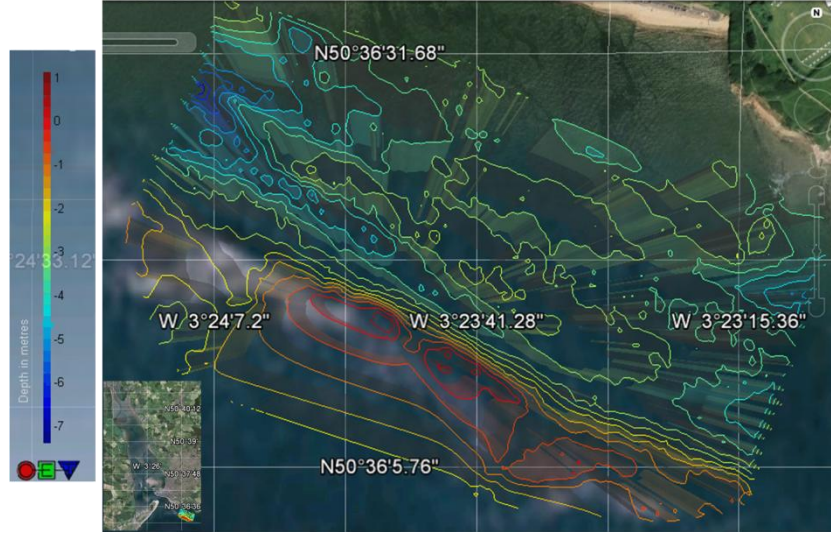


Figure 2. Location and Bathymetric contours of the January 2001 survey.

Empirical Orthogonal functions (EOFs)

The purpose of most data-driven methods is to separate noise and signal and to identify some statistical patterns within the data. Empirical orthogonal functions, or EOFs, is a well-established technique used first in the geosciences and coastal engineering by Winant et al. (1975), who applied it to long-term measurements of beach profiles at Torrey Pines Beach, California. The term was first coined by Lorenz (1956), for meteorological time series analysis. Because in EOFs the basis functions are orthogonal in time and space simultaneously, the temporal and spatial eigenfunctions are basis functions and projection coefficients at the same time (Van Den Dool, 2007).

The first step is to separate the time mean and trend at a given location, and apply the EOF methodology to the anomalies, that is, the remainder of the signal. With $f(s_l, t_k)$ representing the space-time data set at discrete locations s_l and times t_k , and $\langle f(s_l, t_k) \rangle$ denoting the mean and trend component, then the EOF analysis is applied to the anomaly $f'(s_l, t_k)$, defined as

$$f'(s_l, t_k) = f(s_l, t_k) - \langle f(s_l, t_k) \rangle. \quad (1)$$

Once the anomaly has been identified, it can be expanded as

$$f'(s_l, t_k) = \sum_{p=1}^P \alpha_p c_p(t_k) e_p(s_l), \quad (2)$$

with $e_p(s_l)$ denoting the eigenvectors of the data's correlation matrix, $c_p(t_k)$ the temporal eigenfunctions, and α_p the eigenvalues. The p^{th} eigenvalue is a measure of the variance resolved by the p^{th} eigenfunction. The eigenvalues are usually organised in decreasing order; this permits to separate noise from signal. Generally, the eigenvalues will decrease slowly at first, then will suffer a sharp decrease - this is called a "shoulder" or "knee", and then will decrease slowly again. The eigenvalues found before the sharp decay describe parts of the signal (with the composition of all of them describing the totality of the signal), those after the decay parts of the noise.

Empirical Orthogonal Teleconnections (EOTs)

Empirical Orthogonal Teleconnections (EOTs), proposed by Huug Van Den Dool (2000), is gaining popularity within the meteorological data assimilation community. It is a statistical technique that is very similar to EOFs. In EOTs, the anomaly is expanded as

$$f'(s_l, t_k) = \sum_{p=1}^P c_p(t_k) e_p(s_l), \quad (3)$$

with the spatial and temporal eigenfunctions defined as

$$e_p(s_l) = \text{cor}(s_l, sb_p) [\text{std dev}(s_l) / \text{std dev}(sb_p)] \quad (4)$$

and

$$c_p(t_k) = f'(sb_p, t_k), \quad (5)$$

respectively. $e_p(s_l)$ is a regression coefficient as defined in ‘‘regression through the origin’’. The functions $\text{cor}(x, y)$ and $\text{std dev}(x)$ are the correlation function and the standard deviation, respectively. In the EOT method, sb_p is a point in space that captures most of the spatial variance within the domain at a given time. The temporal eigenfunctions are simply the anomaly evaluated at sb_p , and hence they are not necessarily orthonormal. This freedom makes EOTs more advantageous than EOFs, at least for some applications. This is because the patterns captured with EOTs are likely to have a clearer physical meaning than those captured with EOFs.

RESULTS AND DISCUSSION

Starting with demonstrating the resolved variance of both methods EOT and EOF reconstructions of survey data are presented and differences discussed qualitatively. Then the results of a forecast are presented, including the resulting errors, which are compared in magnitude and in resulting reconstruction (qualitative comparison).

Resolved variance and signal reconstruction

Figure 3 shows the cumulative variance resolved as the number of patterns (or modes) included in the reconstruction is increased. This figure shows it suffices to have 2 patterns to resolve more than 65% of the variance, but it is necessary to include at least 6 patterns to resolve more than 90% of the variance. Some studies have claimed that for systems with red noise, such as this one, it is only necessary to keep enough eigenmodes so as to resolve more than 50% of the variance (Vautard and Ghil, 1989). Here this hypothesis is tested by analysing the reconstructions obtained with 2 and 8 eigenmodes.

A comparison of the temporal eigenfunctions obtained in the EOF (T-EOF) and EOT (T-EOT) decomposition is shown in Fig. 4. This shows that the eigenfunctions are very similar in regard to the oscillatory patterns. The T-EOTs have, however, a slightly larger amplitude.

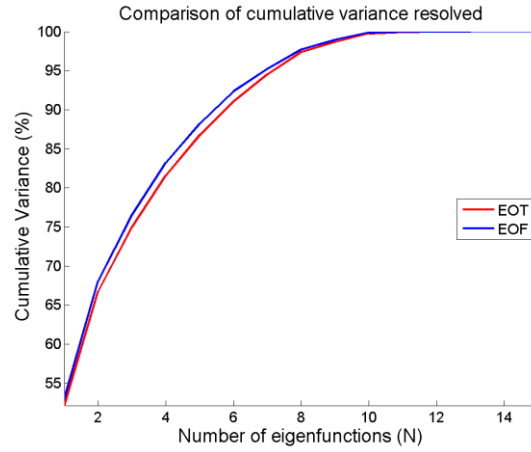


Figure 3. Cumulative variance explained by EOT (red) and EOF (blue) eigenvalues.

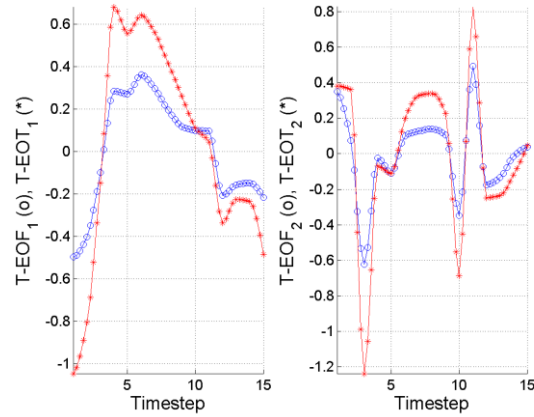


Figure 4. Comparison of the first two temporal eigenfunctions for EOFs (o) and EOTs (*)

A comparison of the reconstructions using 8 modes is shown in Figure 5. A number of encircled small features along the western side of the domain, and a characteristic width for the sandbank running from the North West to the South East of the domain is shown as a red line. By comparing the three panels it can be readily noted that the highlighted small features and the characteristic lengthscale of the sandbank are well reproduced both with the EOTs and the EOFs. The same is true for other small features in the area. However the EOF failed to reproduce the second peak (middle circle).

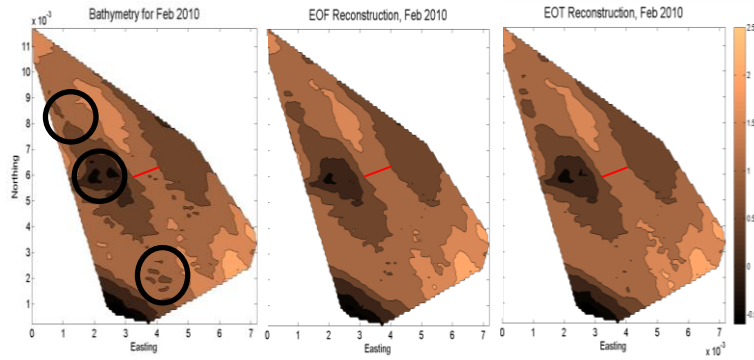


Figure 5. Survey and 8-modal EOF and EOT reconstructions

Performing the reconstructions with only two modes (see Fig. 6) it is clear that the small bedforms are smoothed out and that the width of the sandbank is smaller than in the surveys for both EOT and EOF reconstructions. Figures 5 and 6, hence, show that if important geometrical properties of the bedforms have to be conserved, a reconstruction has to use sufficient modes such that at least 90% of the variance is resolved.

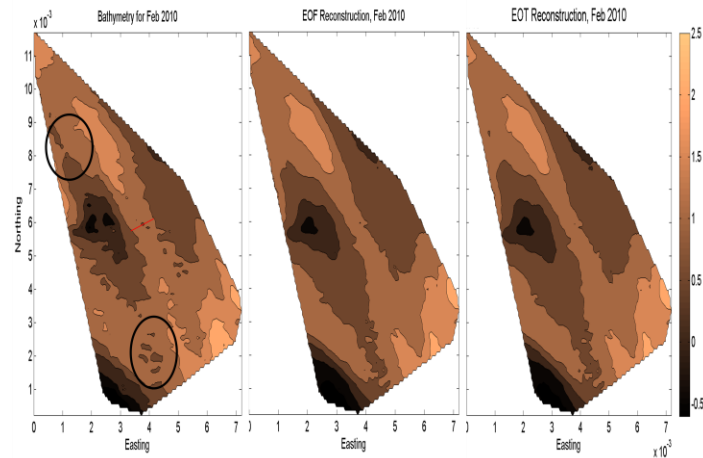
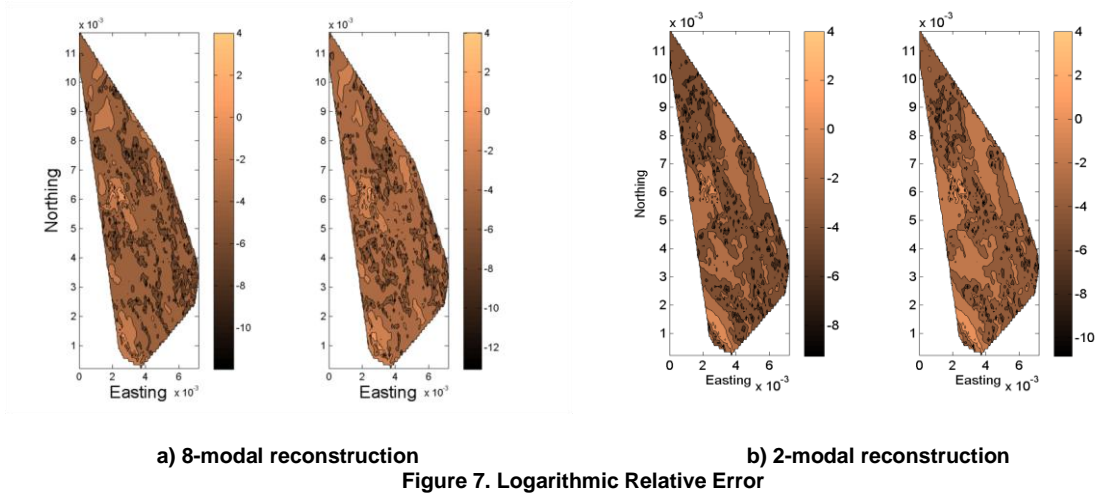


Figure 6. Survey and 2-modal EOF and EOT reconstructions

In the following the relative errors obtained with each method for the two types of reconstructions are compared. The panels on the left in each of the two figures in Fig. 7 correspond to the relative errors for the EOFs, and the panels on the right for the relative errors for the EOTs. The errors are small in all cases, although they are a bit smaller with the 8-modal reconstruction. Hence, the relative errors cannot be used as a criterion for choosing neither a type of reconstruction nor a type of decomposition. All that can be said is that the relative errors with the 8-modal reconstruction appear to be random, but the relative errors obtained with the 2-modal reconstruction seem to be correlated with the position of the sandbank.



One-step forecast analysis

The next stage of the analysis consisted of performing a one-step forecast using Burg's maximum entropy algorithm, as explained in the introduction. In order to be able to assess the skill of the forecast, the last survey (February 2010) was removed from the dataset. This enabled the comparison of the forecast obtained and the data using the root-mean-square error (RMSE). Burg's algorithm is of the form:

$$c_p(t_k) = -b_1 c_p(t_{k-1}) - b_2 c_p(t_{k-2}) - \dots - b_r c_p(t_{k-r}) + \varepsilon_r(t_k),$$

with r being the order of the algorithm. Performing a one-step forecast means the forecast is based purely on surveyed data. Figure 8 shows the results for the forecasts from a) EOF reconstruction and the b) EOT reconstruction, with different modal reconstructions (ranging from 1 to 10 eigenmodes), as well as with forecasts performed directly from the data. For the reconstruction based on 2 eigenmodes, it appears that the best order for the forecast would be an order 5, with orders larger than 5 performing worse than direct forecasts. The same holds with the EOT reconstruction. However, an algorithm of order 2 seems to give good results too, and is actually the order with lowest RMSE for the 8-modal reconstruction.

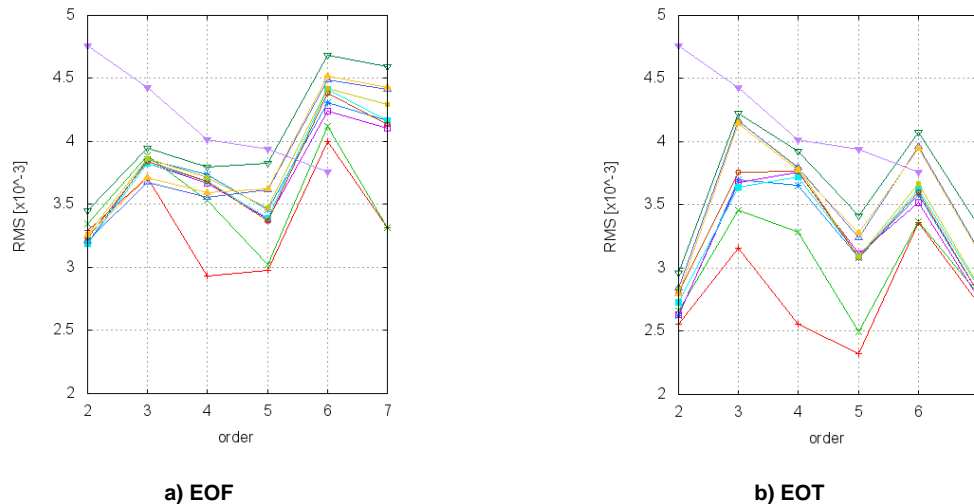


Figure 8. Forecasts' RMSEs (mode max: 1 - red cross, 2 - green cross, 3 - blue asterisk, 4 - magenta square, 5 - cyan square, 6 - red circle, 7 - light green circle, 8 - blue triangle, 9 - yellow triangle, 10 - green triangle. Purple triangle forecast from surveys)

Figure 9 shows a comparison of the forecasts obtained using the EOT, when using a 8-modal reconstruction. The figure in the middle corresponds to a forecast of order 5, the figure on the right to a

forecast of order 2. It appears that a forecast of order 2 smoothes out the small features, which is consistent with it achieving a relatively low RMSE.

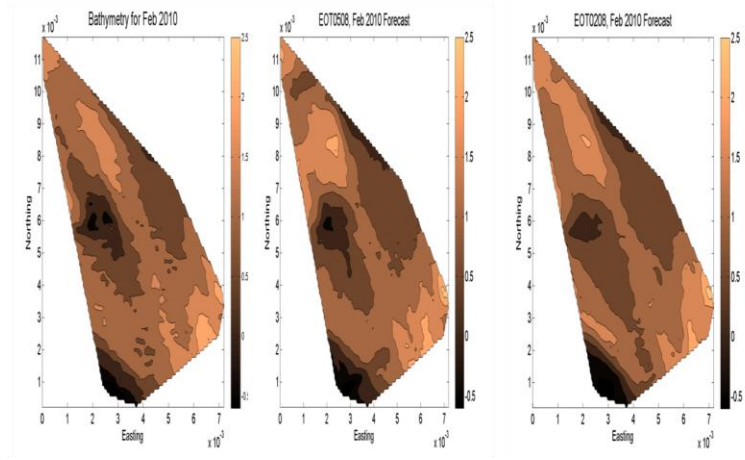


Figure 9. Comparison of forecasts based on an 8-modal EOT reconstruction

Figure 10 shows a comparison of the forecasts obtained using the EOT, but this time using a 2-modal reconstruction. The figure in the middle corresponds to a forecast of order 5, the figure on the right to a forecast of order 2. With a 2-modal reconstruction, the differences between reconstructions are not very significant.

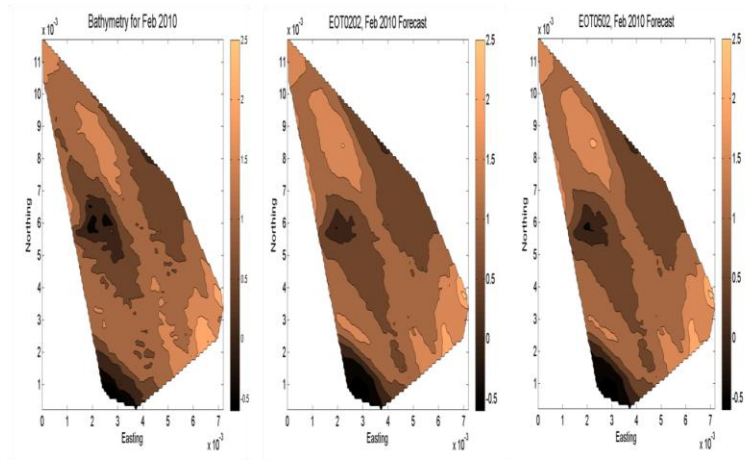


Figure 10. Comparison of forecasts, 2-modal EOT reconstruction

CONCLUSION

The present paper focused on a morphodynamic analysis and forecasts produced using EOFs and EOTs at a small region near the Exe estuary, UK. The study showed that the number of modes of EOFs and EOTs chosen for the reconstruction should capture geometric characteristics of bedforms, in this case nearshore sandbanks. The study showed that small features are more accurately represented by the EOTs, and hence in this respect the study has demonstrated that EOTs are more advantageous than EOFs. Both techniques lead to similar, and small, Logarithmic Relative Errors when comparing the bathymetry for February 2010 with the reconstructions. The one-step forecasts were produced with different orders, with the best choice of order being 2 or 5. Finally, the results also show that a Burg's algorithm of order 2 produces small RMSEs because the fields are smooth, but not necessarily better forecasts. Finally, it is

worth noting that this paper focuses on the bedlevel evolution without specific links to the external environmental conditions, however the impact of the forcings will be reflected in the coefficients of the decomposition. Also, any future changes in environmental forcings cannot be captured by the method, which will impact on long-term forecast validity.

ACKNOWLEDGEMENTS

The authors would like to thank the UKHO for their assistance, and for kindly providing the digital bathymetries used for this study. George Probert, who performed the data curation and preliminary analyses, was sponsored by a pump-priming award from the School of Marine Science and Engineering, Plymouth University, UK. Support from The Software Sustainability Institute (University of Edinburgh, UK) for attendance to ICCE2012 is acknowledged and was greatly appreciated.

REFERENCES

- Burg, J. P., 1967. Maximum entropy spectral analysis, paper presented at the 37th Annual International Meeting, Soc. Of Explor. Geophys., Oklahoma City, Okla., Oct. 31, 1967.
- Burg, J. P., 1968. A new analysis technique for time series data, paper presented at Advanced Study Institute on Signal Processing, NATO, Enschede, Netherlands, 1968.
- International Hydrographic Organisation – IHO, 2008. IHO Standards for Hydrographic Surveys. Publication S-44, third ed. International Hydrographic Bureau, Monaco. Available from: http://www.iho.shom.fr/publicat/free/files/S-44_5E.pdf, 36 pp. (last accessed: 19.03.08).
- Larson, M., Capobianco, M., Jansen, H., Różyński, G., Southgate, H. N., Stive, M., Wijnberg, K. M., Hulscher, S., 2003. Analysis and modeling of field data on coastal morphological evolution over yearly and decadal time scales. Part 1: background and linear techniques. *Journal of Coastal Research*, 19 (4), 760–775.
- Magar, V., Lefranc, M., Hoyle, R. B., Reeve, D. E., 2012. Spectral quantification of nonlinear behaviour of the nearshore seabed and correlations with potential forcings at Duck, N.C., U.S.A. *PLoS ONE* 7(6): e39196. doi:10.1371/journal.pone.0039196
- Pape, L., Ruessink, B.G., 2008. Multivariate analysis of nonlinearity in sandbar behavior. *Nonlinear Processes in Geophysics* 15 (1), 145–158.
- Rattan S. S. P., Ruessink B. G., Hsieh W. W., 2005. Non-linear complex principal component analysis of nearshore bathymetry. *Nonlinear Processes in Geophysics* 12 (5), 661–670.
- Reeve, D. E., Horrillo-Caraballo, J., Magar, V., 2008. Statistical analysis and forecasts of long-term sandbank evolution at Great Yarmouth, UK. *Estuarine, Coastal and Shelf Science* 79 (3), 387–399.
- Reeve D. E., Li, Y., Lark, M., Simmonds D. S., 2007. An investigation of the multiscale temporal variability of beach profiles at Duck using wavelet packet transforms. *Coastal Engineering* 54 (5), 401–415.
- Southgate, H.N., Wijnberg, K.M., Larson, M., Capobianco, M., Jansen, H., 2003. Analysis of field data of coastal morphological evolution over yearly to decadal timescales. Part 2: non-linear techniques. *Journal of Coastal Research*, 19 (4), 776–789.
- Van Den Dool, H., Saha, S. Johansson, A., 2000. Empirical Orthogonal Teleconnections. *Journal of Climate* 13, 1421–1435.
- Van Den Dool, H., 2007. *Empirical Methods in Short-Term Climate Prediction*. Oxford University Press.
- Winant, C. D., Inman, D. L., Nordstrom, C. E., 1975. Description of seasonal beach changes using empirical eigenfunctions. *Journal of Geophysical Research* 80(15), 1979–1986.