Characterising Regional Sea Level Variability On The Basis Of Quality Controlled Tide Gauge Records

A.G.P Shaw[1], M.N. Tsimplis[1], K. Bergant [2], M. J. García[3], V. German[4], N. Kjær[5], M. Pasaric[6], H-P. Plag[7,8], J. J. Martinez Benjamin[9], M. Martinez-Garcia[10], M. Sušnik[11], E. Tel[3], F. Raicich[12] and H. Yildiz[13]

1. National Oceanography Centre, Southampton, James Rennell Division, Empress Dock, Southampton, S014 3ZH, United Kingdom.
2. Nova Gorica Polytechnic, Centre for Atmospheric Research, Vipavska 13, SI-5000 Nova Gorica, Slovenia.
3. Instituto Español de Oceanografía, (MEC) C/Corazón de María, 8, 28002, Madrid, Spain.
4. Israel Oceanographic and Limnological Research, National Institute of Oceanography, Department of Marine Geology and Coastal Processes, Tel Shikmona, POB 8030, Haifa 31080, Israel.
5. Danish National Space Center, Department of Geodesy, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark.
6. Andrija Mohorovicic Geophysical Institute, Faculty of Science, University of Zagreb, Horvatovac bb, 10000 Zagreb, Croatia.
7. Norwegian Mapping Authority (Statens kartverk), Geodetic Institute, Kartverksveien 21 Service box 15, N-3504, Hønefoss, Norway.
8. Nevada Bureau for Mines and Geology, University of Nevada, Reno, Nevada, USA
9. Universidad Politécnica de Cataluña (UPC), Department of Applied Physics, C/ Jordi Girona, 31. 08034. Barcelona, Spain.
10. Universidad Politécnica de Cataluña (UPC), Department of Applied Mathematics and Telematics, Jordi Girona, 1-3, Mòd. C-3 Campus Nord, 08034 Barcelona, Spain.
11. Environmental Agency of the Republic of Slovenia, Vojkova 1b, SI-1000 Ljubljana, Slovenia.
12. CNR Istituto di Scienze Marine, Sezione di Oceanografia Chimica e Fisica, viale Romolo Gessi, 2, Trieste, Italy.
13. General Command of Mapping, Department of Geodesy Dikimevi, Ankara, TR-06100, Turkey.

Email address : Andrew.G.P.Shaw@noc.soton.ac.uk

Abstract

Tide-gauges provide records of relative sea level that are of value for anumber of climate and global change studies. Based on mean monthly values of sea level contained within the Permanent Service for Mean Sea Level (PSMSL) database three data sets of particular interest to satellite altimetry comparisons and to studes of inter-annual to decadal variability. This paper details the methodology used to produce three global tide gauge datasets. The first a dataset contained quality controlled, detrended and deseasoned tide gauges that spanned 20 years or more, the second dataset contained all tide gauges overlapping the period of TOPEX/POSEIDON altimetric observations, i.e. from 1993 onwards. The third dataset consisted of 45 regionally coherent sea level signals extracted from the local sea level records by use of Empirical Orthogonal Functions (EOF). Within each of the regions defined, the first EOF described between 59 and 95% of the variability in the mid to high latitudes, but only 43-68% in the tropical regions. Thirty five regions had over 80% of the total variance explained within the first two EOFs. In spite of the high percentage of the variance explained in most of the regions, the fact that many distinct regions were identified demonstrates the significance of regional sea level variability.

Keywords: PSMSL; tide gauge; interannual; decadal; EOF analysis

1. Introduction

Regional sea level indices have been used in the past to describe long term variability in sea level. In its simplest form, the regional sea level indices have been expressed as a straight average of detrended tide gauge data with respect to time (Woodworth, 1990, Shennan and Woodworth, 1992 and Woodworth, et al., 1999). However, to extract coherent sea level variability, empirical orthogonal functions (EOFs) also referred to as principal component analysis (PCA) (Krzanowski, 1990, Jolliffe, 2002) may be used. Woodworth, et al., (1999) continued by comparing straight averaged sea level records with the first EOF from four regions within the British Isles and found that they were remarkably similar. Recent work by Papadopoulos and Tsimplis (in press) used EOF analyses to describe high coherency of decadal sea level variability between different oceanic regions. They also found that in most cases coherent regional sea level variability is explained in the first EOF of the tide gauge data and in a few cases in the second EOF. Church, et al. (2004), estimated global EOFs gained from satellite altimetry data and combined them with historical tide gauge data to estimate large scale sea level variability over the period 1950-2000.

To determine the inter-annual to decadal variability, it was necessary to develop a dataset which contained the best quality long tide gauge series. Thus extensive quality control and comparison with nearby stations took place in parallel by different institutions providing cross checks and complementarity into the data-quality process. The quality controlled dataset was then the basis for determining and extracting the regionally coherent sea level variability.

The spatial distribution of tide-gauges is not uniform around the globe (Woodworth and Player, 2003) a fact which restricts the study of the decadal variability to the areas where tide gauge measurements exist. The areas populated with large numbers of tide-gauges pose the opposite problem: How does one separate local from regional variability? In addition many tide-gauge records exist are fragmented in time. Excluding those with gaps from the analysis would be a waste of valuable information. The use of EOFs provides answers to these problems because, provided that we focus on regional rather than local sea level variability, and we chose areas where sea level is coherent, combining gappy records to formulate longer EOFs would not bias the signal. Therefore, rather than using the EOFs solely as a statistical tool we use the geographical location of the tide-gauges and any pre-existing knowledge in order to define regions within which we consider permissible to combine the available incomplete data series.

Knowledge of the history of sea level variability from tide gauges cannot be translated to variability of oceanic sea level without knowing the spatial variability of sea level in the open ocean and without knowing the transfer function between sea level variability in the open ocean and that at the tide-gauge sites or at least at the regions of coherent variability we have identified. Satellite altimetry has revolutionalised sea level research by providing the spatial variability of sea level in the open ocean. If one assumes that the observed spatial structures are invariable in time and could also show that the tide gauges are significantly correlated with the nearshore parts of such structures then one could reconstruct sea level variability globally by exploiting the correlations of the present spatial structures with the extracted regional structures. The first step in this process, that is the correlation of the spatial structures with the tide-gauge values as well as the the assimilation of the tide gauge data with the satellite altimetry data from the TOPEX/POSEIDON mission necessitated the development of a third product, a subset of quality controlled tide gauges spanning the period of the TOPEX/POSEIDON mission.

This paper gives a detailed account of the methodology used to develop the regional EOFs and the two other data sets. In a second paper in this volume the stability, in time of these EOFs is explored (Shaw et al. this issue) while the usefulness of the derived EOFs is demonstrated in Tsimplis and Shaw (this issue). In the following two sections, the data selection process and the methodology are explained. In section four some new results on the accuracy of the trends, and the seasonal cycle based on the various processing strategies followed by each partner are discussed. Finally, the regions with apparent sea level coherency are described.

2. PSMSL Tide Gauge Data

Most of the data used are mean monthly values included in the Permanent Service for Mean Sea Level (PSMSL) dataset. A number of additional stations were also provided by some partners.

There are two types of data within the PSMSL database, Revised Local Reference (RLR) and METRIC data. RLR data are those for which information allowing linkage of the sea level data to a benchmark on land is provided. Data for which such information is not provided are considered as METRIC. In principle, METRIC data cannot be used for the estimation of relative sea level trends whilst RLR data can be used. The spatial distribution and the number of tide gauges are shown in Figure 1a for RLR and Figure 1b for METRIC data. It is worth noting that the length of the time-series for both the METRIC and the RLR tide gauges is far from uniform and while a few of them span more than a century a considerable number of them contain short time series (less than 10 years) or long periods of missing data.

The METRIC data contain valuable information and, although they cannot be relied upon for trend estimation due to the possibility of benchmark changes, they can be used for the estimation of decadal, inter-annual and seasonal variability provided that care has been taken to avoid records in which evident shifts of the reference level has taken place. For this reason the METRIC data were subjected to a careful analysis to identify those records that had valid information especially in the areas with sparse RLR sea level time series. Thus, although the PSMSL data are quality checked by the responsible National Authorities and further quality checked and flagged if necessary by the PSMSL (Woodworth and Player, 2003) the combination of RLR with Metric data in order to reduce both the spatial gaps and the interruptions in time of the dataset necessitated additional quality checking. The monthly tide-gauge data was downloaded from the PSMSL database on 08/05/2003.

3. Methodology

3.1 Quality control of the tide gauge dataset

Only tide gauges spanning at least 20 years of data with less than 25% missing values during this period were included in the analysis. Sixteen stations where long gaps were present and/or a datum shift was suspected each part of the time-series were considered independently as separate tide-gauges. Overall 596 tide gauges of which 503 RLR and 93 Metric stations passed the requirements and were included in the analysis. Figure 2 (a & b) show the geographic distribution RLR and METRIC PSMSL data, respectively. Figure 2c represents the number of station-years collected each year. The decrease in the 1990's is due to data backlogs (Woodworth and Player, 2003).

The quality checks of the data sets were done on the mean monthly values, on groups of stations in spatial proximity and against 9 control stations and not on each station independently (see Section 4.1, and Fig. 3). Thus each station was visually compared with nearby stations and examined for datum shifts and for periods of marked different behaviour from its neighbours. The difference of each station from each neighbour and the control stations was also examined for values or timing errors. Where datum shifts were identified, the time series was not necessarily excluded from the analysis but was split into shorter series consisting of data without datum shifts. If the length of these segments was longer than 20 years then they were included in the analysis. Thus many of the datum shifts documented in the PSMSL database did not invalidate the use of the whole record but permitted the use of the segments of the record before and after the shift as independent records of shorter duration. The output of this analysis therefore was the production of a quality control dataset suitable for the analysis of interannual and decadal variability but not for the study of sea level trends. In regions with long records of established good quality, these stations were used, in addition to the control stations, as the basis for quality control of all shorter stations.

Linear trends and the mean seasonal cycle was removed from each quality controlled tide-gauge record (or segment of record) using linear regression (E.q. 1). The linear trends are provided with the dataset but as these include trends from non-RLR records they should not be considered as estimates of sea level rise.

$$TG = TG' + at + constant \tag{1}$$

Where TG = Quality Controlled Tide Gauge time series

TG' = Detrended Tide Gauge time series

t = Time (in years)

a = Slope

The mean seasonal cycle (Eq. 2, $\overline{\mathrm{TG}'}_{(m)}$) was estimated by calculating the mean monthly value for each month on the basis of full years only. The mean value of each month was then removed from all corresponding months in all years.

$$\overline{\mathrm{TG}'}_{(m)} = \sum_{i=1}^{i=nyears} \mathrm{TG}'_{(m,i)} \Big/ nyears$$
$$\mathrm{TG}''_{(m,i)} = \mathrm{TG}'_{(m,i)} - \overline{\mathrm{TG}'}_{(m)} \qquad (2)$$

Where $\mathrm{TG}''_{(m,\,i)}$ = Detrended and Deseasoned Tide Gauge Data for month (m) and year (i)

$\mathrm{TG}'_{(m,\,i)}$ = Detrended Tide Gauge Data for month (m) and year (i)

$nyears$ = Number of complete years

$\overline{\mathrm{TG}'}_{(m)}$ = Seasonal cycle

For the dataset developed for comparison with TOPEX/POSEIDON altimetry data all time series covering the period 1993 onwards were quality checked.

3.2 Defining regional sea level variability.

Empirical orthogonal functions (EOFs) were used to extract the coherent regional signal and to check for further errors or local signals. This technique is not new for analysing tide gauge (e.g. Barnett, 1983). Standard EOF analysis (Krzanowski, 1990, Jolliffe, 2002) does not provide for the existence of gaps. There are several possible ways of dealing with this problem. Four methods were implemented to extract the coherent regional signal, three of which were variants of the EOF time series analysis while the fourth method was used as a check.

The first method was the standard EOF without any interpolation involved. This standard Principal Component Analysis (PCA) is the most robust statistically. Unfortunately, this method is not very useful with regional sea level values as it necessarily limits the analysis to the minimal number of common monthly values. The second method interpolates the gaps in the time series using maximum likelihood (Tipping and Bishop, 1999; Houseago-Stokes and Challenor, 2004) prior to running an EOF analysis. This method is referred to as probabilistic principal component analysis (PPCA).

The third technique (which will be called LR) produces an EOF dataset where the gaps are filled in using linear regression from those nearby stations which were best correlated with the station that contained gaps. In this case, a minimum correlation coefficient of 0.65 was required between stations before filling the gaps. This procedure can produce long EOFs but the significance of the values is very much reduced when there are values missing from several stations, which implies that the different data values are not independent. Nevertheless, to the extent that tide gauges sample sea level variability in the same region it is arguable that the measurements are not independent anyway. In addition, small gaps of 1 to 2 months were filled using linear interpolation if there were no significantly correlated neighbouring stations.

Finally, the fourth method provided an independent check of the other three methods mentioned above. The quality checked tide gauge stations within the each region were straight averaged with respect to time. This method was referred to as Regional Indices (RI).

The correlation of the first EOF with the averaging technique indicates that in many cases the RI can be relied upon to assess the sea level variability during periods when few tide gauges were operational. This, of course depends on whether the regions are defined in such a way that the first EOF represents a significant part of the variability. Both the EOFs and the regional averages provide the empirical determination of decadal sea level variability.

3.3 Regional Empirical Orthogonal Functions Determination

After the dataset was finalised and the gap filling technique developed we proceeded in extracting regional Empirical Orthogonal Functions across the globe. The purpose of using EOFs was to create longer time series containing the spatially coherent signals and to reduce dimensionality within the dataset, where possible. The way this was done was to an extent subjective and iterative: geographical regions were selected and the tide gauge records within these regions with sufficient spatial distribution to cover the particular region were used to extract the EOFs. The resulting spatial weights of the EOFs were then examined to determine whether the first EOF was representative of all the selected tide gauges, i.e. whether it explained an important part of the variability in all the tide gauge records within the selected region. If only a subset of stations were represented in the first EOF then the EOF extraction was performed again within this subset or by excluding any tide gauge which was not contributing to the first EOF. As a check, the tide gauges in each region were reconstructed based on the first three modes of EOFs to verify no hidden outliers. In cases where an outlier (i.e., a particular tide gauge record) was clearly incoherent with nearby records then that particular record was excluded and the EOF calculation was repeated. This procedure eliminates weighting of the EOFs on particular "temperamental" records like those close to river outflow.

3.4 Parallel processing of the data.

The processing of the data set was done in parallel by various institutions. Although the development of a common processing tool was discussed as a possibility, it became apparent that as various computer and software platforms were in use and this would necessarily have been a time consuming process hindered in some cases by the hardware and software options available at each participating institution. Thus statistical and mathematical tools were left to the choice of the individual researcher and the methodology focused on establishing consistent quality control amongst the partners. To maximize the gain from the parallel processing of the data sets we developed suitable quality tests and procedures of the various parts of the data. Thus nine of the longest available tide gauge records were analysed by all the partners (Fig. 3) and results of these analyses were compared. Moreover, where tide-gauges were analysed by more than one institution the resulting quality controlled time series were cross checked.

4. Results

Overall, most records are of good quality. The creation of a good quality dataset is based on consistency with nearby records and does not necessarily mean that the coastal data are consistent with the open water sea level observations. The only

European region within the PSMSL database that does not have enough good quality records is the Ionian Sea and the Mediterranean regions east of it (see SELF project Zerbini et al. (1996), Tsimplis and Spencer, 1997). For these regions the approach used by Tsimplis and Josey (2001) is probably more appropriate. This method is highly subjective and involves selection of parts of the available records that appear to be consistent with nearby records for even short periods of time. This approach has resulted in an index for the Eastern Mediterranean which was well correlated with the Black Sea the Adriatic and the Western Mediterranean (Tsimplis and Josey, 2001).

4.1 Comparison of results from Control stations

4.1.1 Time domain

The trend estimates by the various partners using different numerical techniques and computer programs by partners for the control stations that required no gap filling (i.e., Honolulu, San Francisco and Stockholm) agreed better than 0.05 mm/year. Where gaps were present the trend difference between partners was of the order of 0.10 mm/year between partners, mainly due to differences in expert judgement of what is acceptable data within the data set. Small differences of approximately 0.03 mm/year in the slope between partners was associated with some partners having used the annual mean to calculate the trend while other used the monthly values. The annual mean was defined here as having at least 10 out of the 12 months present, provided that the 2 missing months were not consecutive.

4.1.2 Frequency domain

The analysis of the frequency domain aimed in identifying any dominant frequencies that could then be used as the basis of describing the sea level variability. Thus power spectra of the control stations were produced by all the partners. These dependent on the length of record, the filter used and the interpolation of gaps.

4.2 Cross- Checking of Final Product

In addition to the control stations, there were 48 tide gauges that had been duplicated by different partners. A comparison between these tide gauge stations was carried out before the compilation of the dataset took place. In general the quality-checked data appeared consistent having an mean rms value of 17mm and median rms value of

13mm. Examples of records for which agreement or disagreement were noticed are shown in Figures 4a and 4b respectively.

In Figure 4b (top) the discrepancy between the two partners was due to the difference in the way of calculating the mean seasonal cycle. This is somewhat surprising, but it is demonstrated in Figure 5 that such differences can be as high as 25 mm for a month if there are gaps in the data and these do not occur randomly but within particular months. Figure 4b (middle) shows differences for the Cadiz III station where national data have been used and the reference level is different between these data and the PSMSL data. Finally, Figure 4b (bottom) demonstrates the differences where different lengths of the data have been used to estimate the trend.

The discrepancy for the Cadiz III station was investigated further. Two different data sets were available for this station, one from the Instituto Español de Oceanografia (IEO), and the other from the PSMSL. A careful cross-checking procedure was carried out, and as a result the data from the PSMSL dataset were preferred even though there was an apparent data shift as they compared better to nearby stations than in the IEO tide gauge series.

The use of harmonics for the description of the seasonal cycle would not have been as sensitive as the present method to the use of gaps. However, it would not have been able to describe non-harmonic cycles as successfully as the method, used here.

4.3 The Data Products

The quality checked data set of mean monthly values used in the development of the EOFs is available at (http://www.eseas.org/eseas-ri/-). This dataset contains the values of the trend and seasonal cycle along with a information about each station. This information includes the station name, PSMSL code name, the geographic coordinates, type of data, (i.e. RLR, METRIC or in a few cases, METRIC data that has combined RLR quality & METRIC) , the number of complete years of data per station. The first zero-crossing and the times/lags of (first) 50% auto correlation are also provided.

The second dataset that was quality controlled only for the period of 1993 onwards contained 507 tide gauge records, of which 457 stations were RLR and 50 were METRIC data. The spatial distribution of this dataset is displayed in Figure 2. For consistency, both datasets had no interpolation of monthly values.

4.4 Regional Sea level indices

Forty five regions were found around the globe that were considered to have a coherent signal, these regions are shown in Figure 6. Although the selection process of each region was initially subjective however, the iterative procedure of resampling each region to maximise the coherency in the signal was very effective

A time series dataset combining the first two EOFs was created for these 45 regions,. These data can also be found on the ESEAS-RI web pages (see Section 4.3). All the four methods (i.e., PCA, PPCA, LR and RI) are available. The total variance explained for each EOF using PCA and LR method were in general very similar, where the PPCA method was between 5-10% less than the other two methods. The best method that maximises the length of the time series was found to be the LR method. Thus, the total variance explained using the LR method of the first EOF ranged between 59-95% in the mid to high latitude regions, with three exceptions, Sea of Japan, the eastern side of South America, and Greenland having the total variance explained of 48%, 56% and56%, respectively. In the tropical regions the first EOF ranged between 43-68%. However, the combined first two EOFs produced the minimum total variance explained of 65%. Thirty five regions had over 80% of the total variance explained within the first two EOFs.

As examples, we present the results of the regional sea level in the Southern North Sea in Figures 7 and 8 and the Western Central Pacific in Figures 9 and 10. The resulting spatial weights from the three EOF methods are very similar (Fig. 7). The stations are ranked in terms of overall length and thus the PPCA method appears to have spatial weights affected by the number of valid data available in each station. Therefore, when comparing the RI to the PPCA and LR methods, it became evident that the PPCA method was not representing the time series well. This can be seen in Figure 8b where the amplitude in signal is reduced between 1850 -1890. The time variation representing the sum of the first two EOFs, the regional index of the average of the stations and their power spectra are shown in Figure 8. The correlation coefficients “r” between the three estimates of the EOF and the RI is also shown in Figure 8, in this particular case were higher than 0.9.

Figure 8e shows the power spectra for Southern North Sea of the same period of time produce the same frequencies. Nevertheless, when the whole period of time is considered different peaks appear (Fig. 8f). This is a result of the different lengths used and the variation of energy for example at the 14 month period of oscillation. Thus, most of the stations have data in the late 1800, a period for which there was reduced energy in the 14m oscillation signal while the common period starts after 1930 a period with increased energy at the 14 month oscillation (Plag, 1988). In the North Sea and the Baltic, sea level variations in the 14 to 16 month period band are mainly driven by meteorological forcing (Tsimplis et al.,1994; Plag, 1997; O'Connor et al. 2000; Aoyama, et al. 2003). Therefore variations of the dominant peak in that band can be expected.

In the tropical Western Central Pacific (Figs. 9 and 10) the spatial weights from the three EOF methods appear to be very similar. No interpolation was applied to Station 5, (Rabaul, Fig. 10g) because the correction between neighbouring stations was less than 0.65. However, this station was selected because the spatial weight of Rabaul (Fig.9h) was close to one as well as the EOF reconstruction of the tide gauge station (not shown) explained 61% of the total variance from the first 2 EOFs. The most dominating signal in the time series (Fig. 10a-d) and power spectra (Fig. 10e) appears to be the El Niño Southern Oscillation (ENSO). This is consistent will Papadopoulos and Tsimplis, (in press) EOF analysis, showing that sea level forcing in the Pacific region is dominated by the ENSO. Also shown is that PPCA technique in (Fig. 10b) did not perform well interpolating the time series prior to 1975 showing a reduction in the amplitude of the signal.

5. Conclusion

The development of three sea level datasets suitable for the analysis of inter-annual to decadal variability and for comparison and assimilation with the TOPEX/POSEIDON altimetric dataset is discussed. The datasets are available through the ESEAS-RI webpage.

The exploration of the produced dataset has led to the development of regions which demonstrate significant coherency in sea level variability. However, and somehow surprisingly a number of open ocean tide gauges in the tropical zone do not demonstrate coherent variability with nearby stations and thus pose questions on their suitability in monitoring sea level change. In developing the technique for identifying regional coherency, several statistical techniques were used in order to tackle the problem of gaps in the data. The linear regression (LR) method using neighbouring stations for interpolation was found as the most useful in practical terms. The EOFs and alternatively the regional indices (RI) reproduce regionally coherent sea level variability. These two methods were in general in good agreement with each other.

One special aspect was that data processing was done in parallel by various European Institutions. A detailed evaluation of the discrepancies introduced by the different methodologies is available from ESEAS-RI website under Work Package Task 3.1.

The regional EOFs identified here can only be used to interpolate sea level in time and space if they are stable in time and if their cross correlation is stable in time. The first question is addressed by Shaw et al. (this issue). The sea level spectrum examined show dominant peaks. The question on whether some of the peaks dominate in space or/in time is are discussed elsewhere in this issue (Pasaric et al., 2005).

Acknowledgements

The work was carried out in the framework of the European Sea Level Service Research Infrastructure project (ESEAS_RI, contract EVR1-CT -2002-40024). The authors will like to thank the PSMSL data centre and the National Authorities for their superb work over the years in maintaining the database as well as carrying out their own quality control of the tide gauge data. Special thanks to Prof. Philip Woodworth for providing advice during the development of the task.

References

Aoyama Y, Naito I, Iwabuchi T, Yamazaki N. 2003. Atmospheric quasi-14 month fluctuation and excitation of the Chandler wobble. Earth Planets and Space 55 (12): E25-E28.

Barnett, T.P. Recent Changes in Sea Level and Their Possible Causes. Climatic Change, 5: 15-38.

Church, J.A, White, N.J., Coleman, R., Lambeck, K. and Mitrovica, J.X. 2004, Estimates of the Regional Distribution of Sea Level over the 1950-2000 period. Journal of Climate. 17 (13): 2609-2625.

Douglas, B.C., Kearney, M.S and Leatherman, S.P. (Editors), 2001. Sea level rise: history and consequences. Academic Press. (International Geophysics Series; 75).

Houseago-Stokes, R.E. and P.G. Challenor, 2004. Using PPCA to estimate EOFs in the Presence of Missing Values. Journal Atmospheric and Oceanic Technology. 21 (9): 1471-1480.

Jolliffe, I.T., 2002. Principal Component Analysis. Spring-Verlag 2nd Edition, New York.

Krzanowski, W.J. 1990. Principles of Multivariate Analysis, Oxford University Press, New York.

O'Connor WP, Chao BF, Zheng DW, Au AY. 2000 Wind stress forcing of the North Sea 'pole tide' Geophysical Journal International 142 (2): 620-630

Papadopoulos, A. and Tsimplis, M.N. Coherent Coastal Sea Level Variability at Inter-Decadal and Inter-annual Scales from Tide Gauges.(in press). Journal of Coastal Research.

Plag H.-P., 1988. A regional study of Norwegian coastal long-period sea-level variations and their causes with special emphasis on the pole-tide . Berl. Geowiss. Abhandl. Reihe A 14 1-175.

Plag H.-P., 1997. Chandler wobble and pole tide in relation to interannual atmosphere-ocean dynamics. In: Wilhelm H., Zurn, W. Wenzel, H.-G. (Eds). Tidal Phenomena, number 66 in Lecture Notes in Earth Sciences, Springer, pp. 183-218.

Pasaric M. Long-term variability of sea level in the North Sea. Chemistry and Physics of the Earth (submitted)

Tipping, M.E. and , C.M. Bishop (1999) Probabilistic principal component analysis *J* Journal of the Royal Statistical Society Series B-Statistical Methodology 61: 611-622 Part 3.

Tsimplis M.N. Flather R.A., Vassie, J.M., 1994. The North Sea Pole Tide described through a tide-surge numerical model. Geophysical Research Letters, 21: 449-452.

Tsimplis, M.N. and. Josey, S.A, 2001. Forcing of the Mediterranean Sea by Atmospheric Oscillations over the North Atlantic. Geophysical Research Letters, 28(5) : 803-806.

Shaw, A.G.P.,Tsimplis, M.N, Challenor, P.G. and Houseago-Stokes, R.E. ESEAS Development of Regional Indicies of Sea Level. Chemistry and Physics of the Earth (submitted)

Shennan, I. and Woodworth, P.L. 1992. A comparison of late Holocene and twentieth-century sea-level trends from the U.K and North Sea region. Geophysical Journal International, 109 (1) : 96-105.

Woodworth, P.L. 1990. Towards Routine Global and Regional Sea Level Indices , in *Towards an Integrated System for Measuring Long term Charges in Global Sea Level*, pp 121-132, edition Eden H.F. (Report of a workshop held at Woods Hole Oceanographic Institution, May 1990), Joint Oceanographic Institutions Inc., Washingtion, DC.

Woodworth P.L. and R. Player, 2003. The Permanent Service for Mean Sea Level: an update to the 21st century. Journal of Coastal Research. 19(2): 287-295.

Woodworth P.L., Tsimplis, M.N. Flather, R.A. and Shennan, I. 1999. A review of the Trends observed in British Isles mean sea level data measured by tide gauges. Geophysical Journal International, 136 (3); 651-670.

Zerbini S., Plag H.-P., Baker T., Becker M., Billiris H., Burki B., Kahle H.-G., Marson I., Pezzoli L., Richter B., Romagnoli C., Sztobryn M., Tomasi P., Tsimplis M., Veis G. and G.Verrone, 1996. Sea Level in the Mediterranean: a first step towards separating crustal movements and absolute sea-level variations. Global and Planetary Change, 14, 1-48.

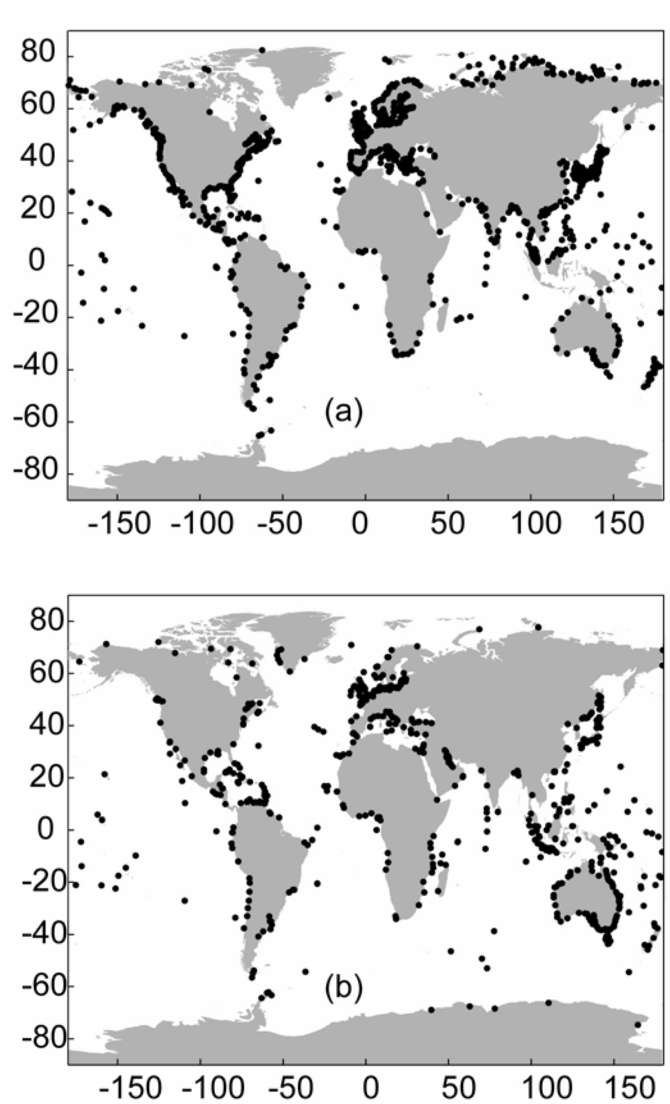

Figure 1 Positions of all tide gauges from Permanent Service for Mean Sea Level (PSMSL) dataset containing only RLR data (a) and METRIC (b) as of 24 January 2005. There are 1265 RLR and 893 METRIC tide gauges.

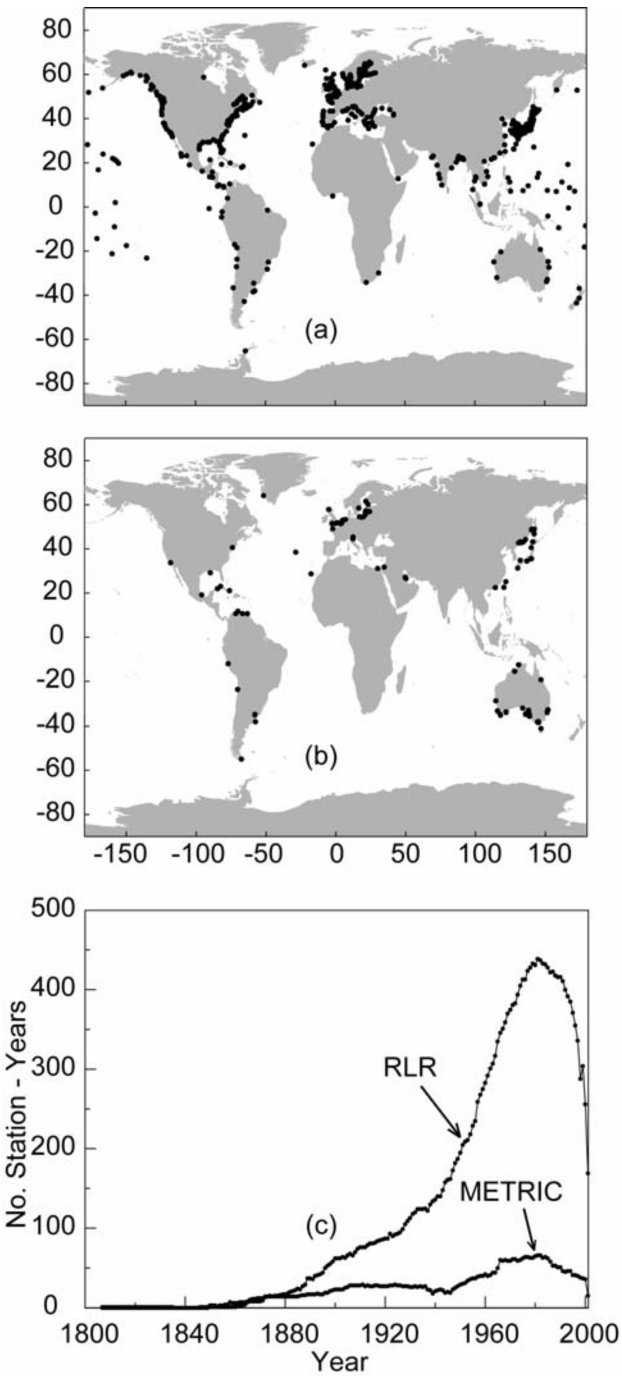

Figure 2 The locations of tide gauges with records from the PSMSL data base (as for 8/5/2003) that are longer than 20 years which were quality checked. RLR (a) contains 503 stations and METRIC (b) contains 93 stations. Within the METRIC dataset there were 12 stations that contain RLR data. The number of station-years (that are longer than 20 years) of data collected each year is shown in (c). However, due to data backlogs, the 1990's appear only half complete from a PSMSL perspective (Woodworth and Player, 2003).

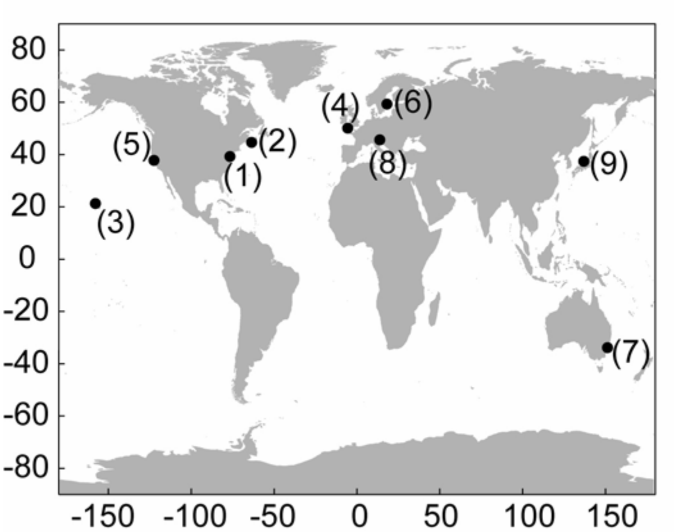

Figure 3 The location of the nine control stations used. (1) Baltimore, (2) Halifax, (3) Honolulu, (4) Newlyn, (5) San Franciso, (6) Stockholm, (7) Sydney, (8) Trieste and (9) Wajima

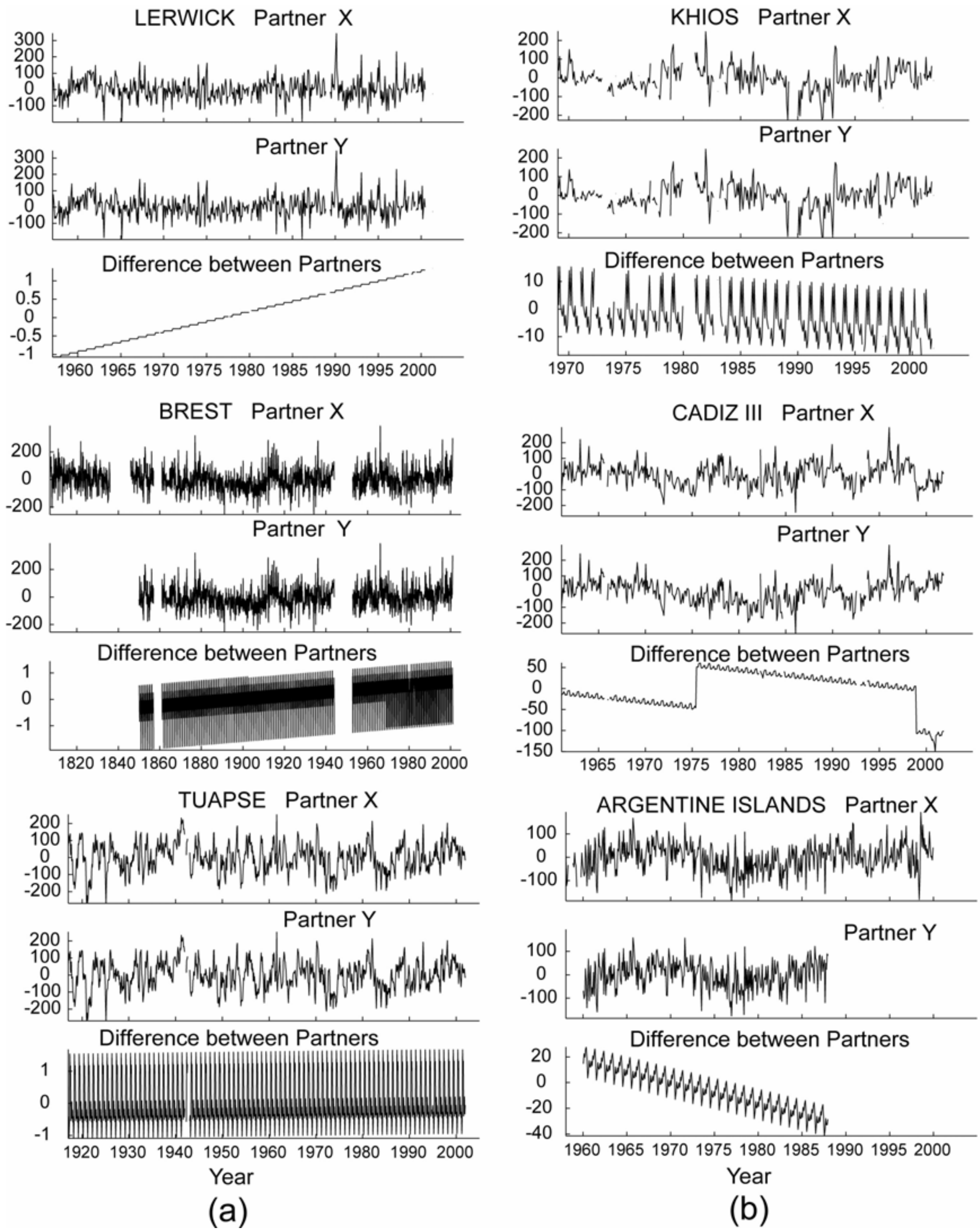

Figure 4 (a) shows examples of good comparison of the final results between partners. The y-axis is in mm. The differences in the analysis techniques are less than 3 mm where the same period of analysis is used. These results are consistent with those for the analysis of the control stations. Column (b) represents less good results due to different ways of estimating the seasonal cycle (top), or use of national dataset (middle) or use of different periods (bottom) resulting in significant differences which needed to be resolved.

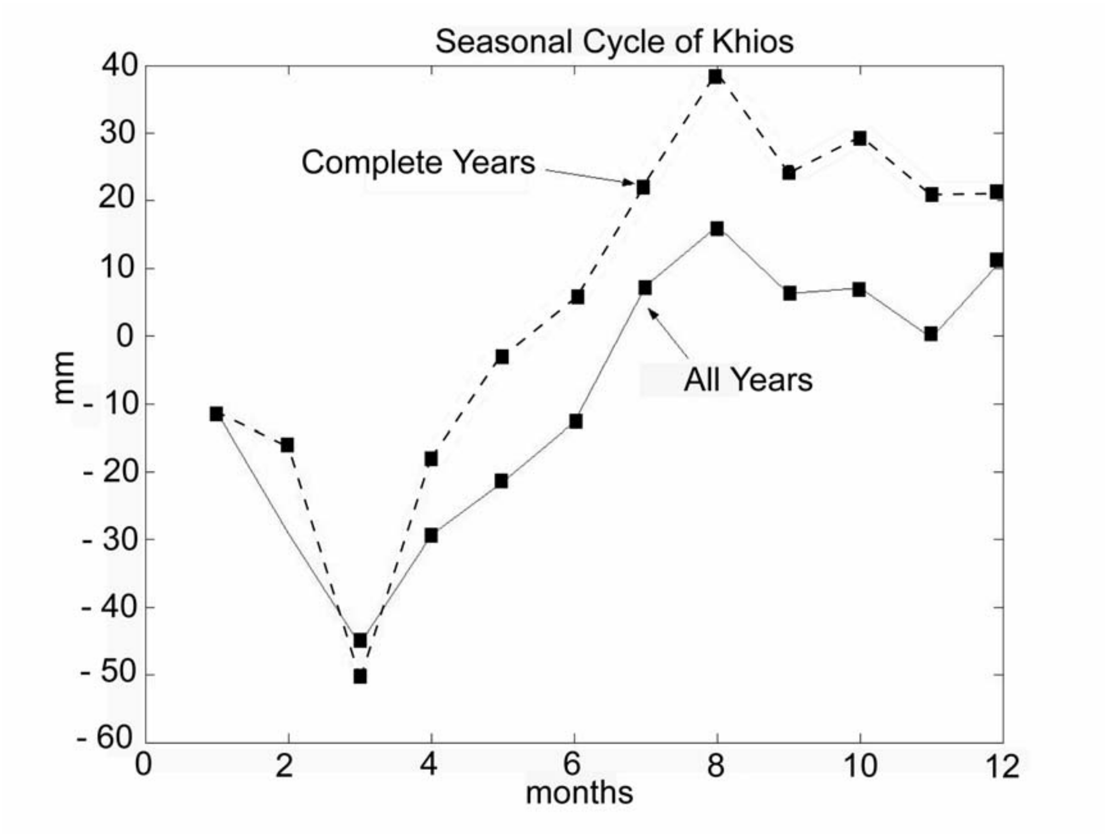

Figure 5. Mean seasonal cycle for Khios estimated from complete years only (dashed line) and for all available values (solid line). Discrepancies of up to 25 mm can occur in particular months.

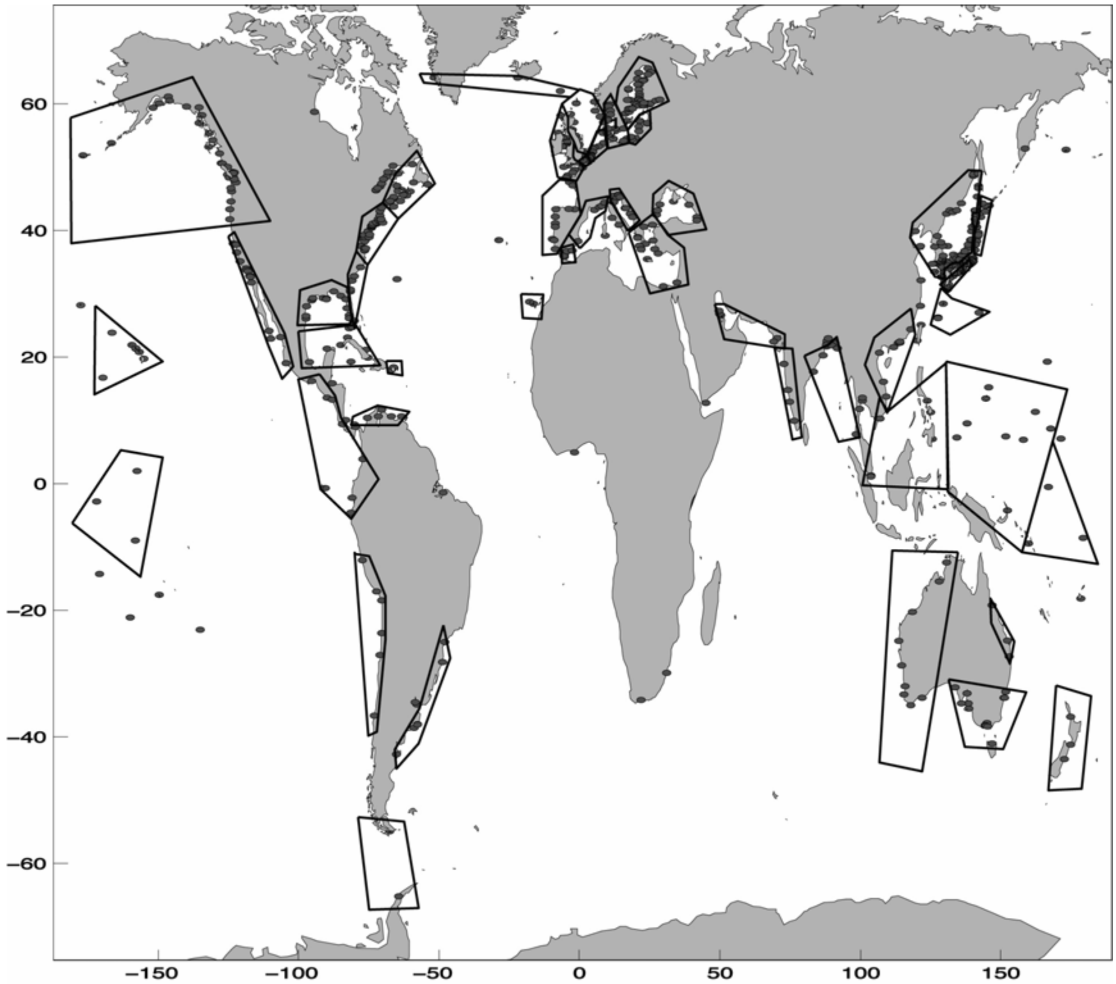

Figure 6 The tide gauges with records longer than 20 years which were quality checked and analysed at SOC. The regions enclosed dote the areas for which coherency in the sea level signal was found and for which the EOF's were extracted.

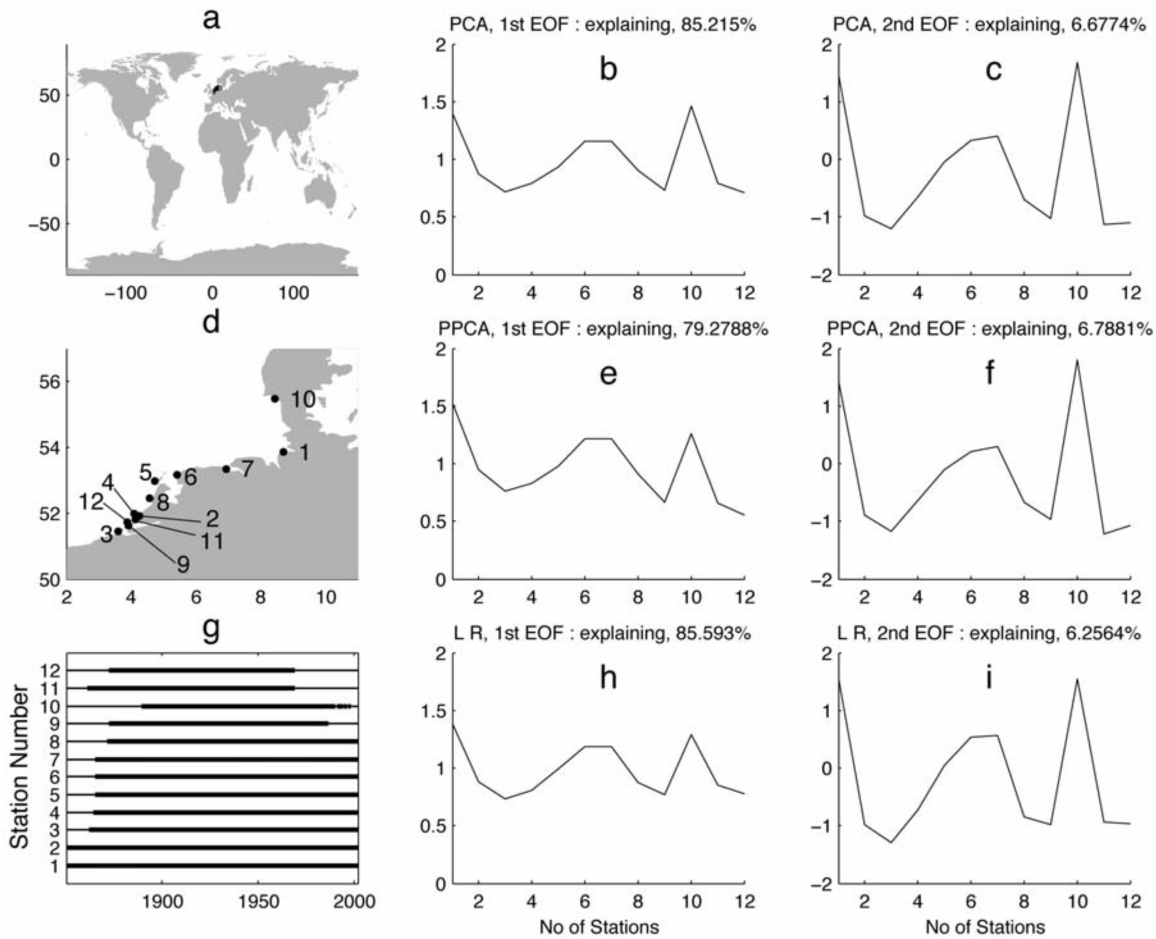

Figure 7 An example of Regional Empirical Orthogonal Functions (Southern North Sea) where there is a coherent sea level signal. The location of the tide gauge stations selected are shown a) and d). (g) The names of the tide gauge stations are as follows: Cuxhaven 2 (1), Masssluis (2), Vlissingen (3), Hoek Van Holland (4), Den Helder (5), Harlingen (6) Delfzijl (7), Ijmuiden (8), Zierikzee, (9), Esbjerg (10), Hellevoetsluis (11), Brouwershaven (12). The periods for which the selected stations have data (thin lines are interpolated data). The spatial weights of the first and second EOFs are shown in the other plots. (b) and (c) are principal component analysis (PCA) estimates; (e) and (f) are probabilistic principal component analysis (PPCA) estimates; (h) and (i) are LR estimates.

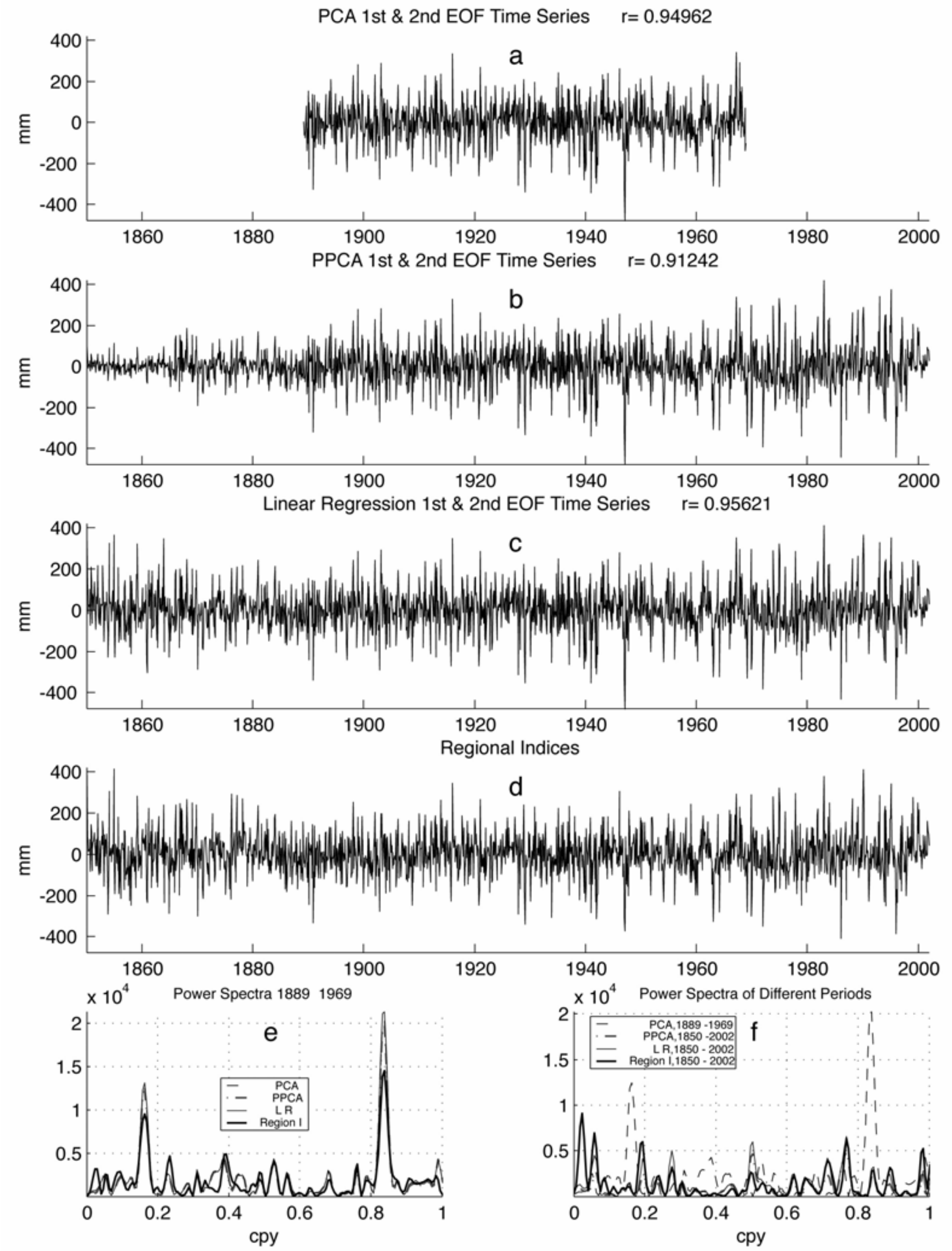

Figure 8 The time variability of the EOFs (Southern North Sea region) from (a) principal component analysis (PCA); (b) probabilistic principal component analysis (PPCA); (c) linear regression (LR). The regional index (RI) is shown at (d). The correlation coefficients are between each type of EOF and the RI. (e) shows the power spectra of (a-d) for the common period, while (f) the power spectra for the maximum period available for each of the EOF types.

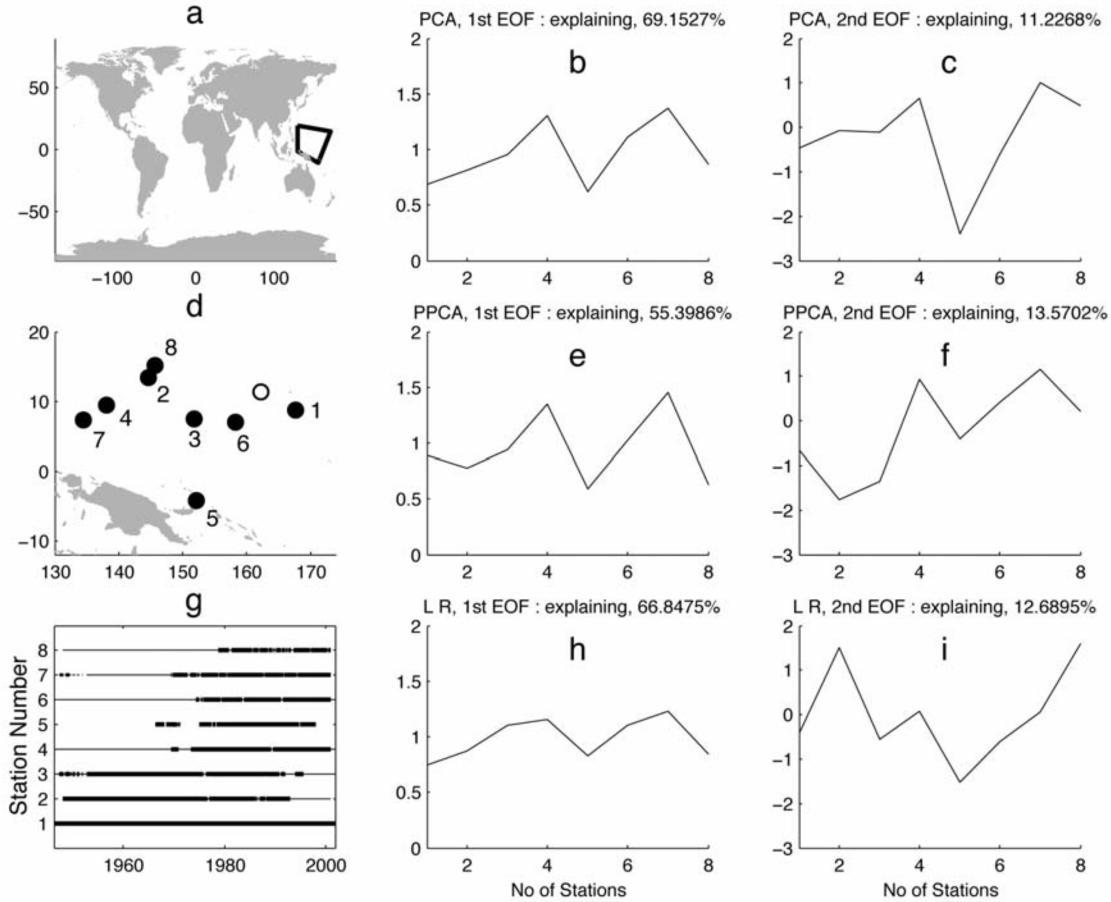

Figure 9 An example of Regional Empirical Orthogonal Functions (Western Central Pacific) where there is a coherent sea level signal. The location of the tide gauge stations selected are shown a) and d). The filled in circles are the stations that were selected which are also numbered. The circles are the stations that were not selected. (g) The names of the tide gauge stations are as follows: Kwajalein (1), Guam Chuuk (2), Moen Island (3) Yap B (4), Rabaul (5), Pohnpei-B (6) Malakal (7), Saipan (8). The periods for which the selected stations have data (thin lines are interpolated data). The spatial weights of the first and second EOFs are shown in the other plots. (b) and (c) are principal component analysis (PCA) estimates; (e) and (f) are probabilistic principal component analysis (PPCA) estimates; (h) and (i) are LR estimates.

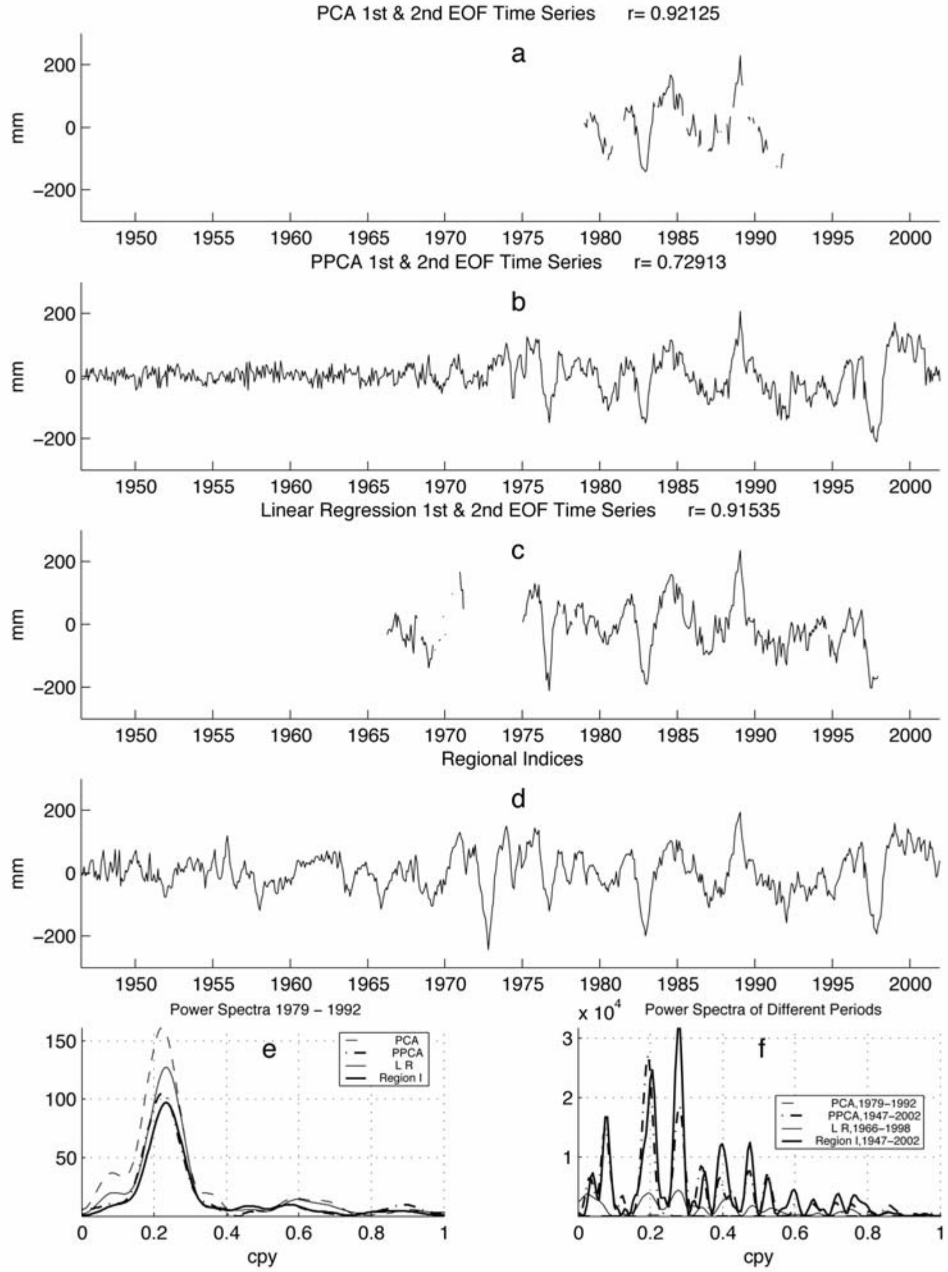

Figure 10 The time variability of the EOFs (Western Central Pacific region) from (a) principal component analysis (PCA); (b) probabilistic principal component analysis (PPCA); (c) linear regression (LR). The regional index (RI) is shown at (d). The correlation coefficients are between each type of EOF and the RI. (e) shows the power spectra of (a-d) for the common period, while (f) the power spectra for the maximum period available for each of the EOF types.