The use of ERS-1 crossovers and TOPEX/Poseidon repeat pass data for global sea surface variability studies.

Matthew Dominic Reynolds
Doctor of Philosophy

Aston University
September 1998

Aston University

The use of ERS-1 crossovers and TOPEX/Poseidon repeat pass data for global sea surface variability studies.

Matthew Dominic Reynolds

Doctor of Philosophy

September 1998

Thesis Summary

In previous sea-surface variability studies, researchers have failed to utilise the full ERS-1 mission due to the varying orbital characteristics in each mission phase, and most have simply ignored the Ice and Geodetic phases. This project aims to introduce a technique which will allow the straightforward use of all orbital phases, regardless of orbit type. This technique is based upon single satellite crossovers.

Unfortunately the ERS-1 orbital height is still poorly resolved (due to higher air drag and stronger gravitational effects) when compared with that of TOPEX/Poseidon (T/P), so to make best use of the ERS-1 crossover data corrections to the ERS-1 orbital heights are calculated by fitting a cubic-spline to dual-crossover residuals with T/P. This correction is validated by comparison of dual satellite crossovers with tide gauge data.

The crossover processing technique is validated by comparing the extracted sea-surface variability information with that from T/P repeat pass data. The two data sets are then combined into a single consistent data set for analysis of sea-surface variability patterns. These patterns are simplified by the use of an empirical orthogonal function decomposition which breaks the signals into spatial modes which are then discussed separately. Further studies carried out on these data include an analysis of the characteristics of the annual signal, discussion of evidence for Rossby wave propagation on a global basis, and finally analysis of the evidence for global mean sea level rise.

Keywords:

- single & dual satellite crossovers
- ERS-1
- TOPEX/Poseidon
- sea surface variability
- Rossby wave
- orbit error correction

Acknowledgements

Thanks to Dr Philip Moore at Aston and Dr Helen Snaith at Southampton for supervisory duties, and for pointing me in the right direction when necessary.

Thanks go to my parents for their support throughout life.

Special thanks to Laura Doherty for being a first-class best-friend, and for believing I could do it.

Thanks also for advice, support and good humour from my fellow group members (past and present): Simon, Chris L., Gordon, Stuart, Ruaraidh, Chris M., Henno, Rob, Russ, Grant and Paul, and to everyone else I've met (however briefly) during my time at Aston and elsewhere who's been a friend.

Grateful acknowledgement is made to the Natural Environment Research Council and the James Rennel Division at Southampton Oceanography Centre for financial support.

Contents

List of Figures

List of Tables

Glossary

ACC	Antarctic Circumpolar Current
AMI	Active Microwave Instrument
AO	Atlantic Ocean
ATSR	Along Track Scanning Radiometer
CNES	Centre Nationale d'Etude Spatiales
DGM4	Delft Gravity Model 4
DORIS	Doppler Orbitography and Radiolocation Integrated by Satellite
DUT	Delft University of Technology
DXOs	Dual satellite crossovers
ECMWF	European Centre for Medium-range Weather Forecasting
EOF	Empirical Orthogonal Function
ERS-1	European Remote Sensing satellite-1
ERS-2	European Remote Sensing satellite-2
ESA	European Space Agency
GDR	Geophysical Data Record
GFO	Geosat Follow On
GM	Geodetic Mission
GMT	Generic Mapping Tools
GOME	Global Ozone Monitoring Experiment
GPS	Global Positioning System
IO	Indian Ocean
JGM3	Joint Gravity Model 3
JPL	Jet Propulsion Laboratory (at NASA)
LRR	Laser Retro-Reflector
MJD	Modified Julian Date
MR	Microwave Radiometer
NASA	National Aeronautics and Space Administration
OI	Optimally Interpolated
OPR	off-line Ocean Product
PO	Pacific Ocean
PRARE	Precise Range-And-Range-rate Equipment
SAR	Synthetic Aperture Radar

SATAN	Satellite Analysis software
SLR	Satellite Laser Ranging
SPTR	Scanning Point Target Response
SWH	Significant Wave Height
SXOs	Single satellite crossovers
T/P	TOPEX-Poseidon
USO	Ultra Stable Oscillator
WOCE	World Ocean Circulation Experiment

1. Introduction

1.1 Overview.

There is now a hitherto unimagined amount of data available to satellite oceanographers. The parallel running of the European Remote Sensing satellite (ERS) mission and the TOPEX/Poseidon (T/P) mission provides consistent, concurrent data, collected over a period of 6 years at the time of writing. This is set to continue with the ongoing running of ERS-2 and T/P, the upcoming launch of Topex Follow-On (JASON-1), and the attempts to resolve the problems currently affecting Geosat Follow-On (GFO) which was launched on Feb 10th 1998.

This data set can provide enormous amounts of information about the dynamics of the ocean, as well as previously unseen information about errors present in the mean sea-surface models and the gravity field that are presently used. (Discussion of these errors is beyond the scope of this project).

This thesis aims to present a method of extracting information on sea-surface variability patterns from the entire ERS-1 satellite altimetry data set. The results observed on applying this method and the feasibility and effectiveness of combining this data set with that provided by T/P are discussed. This work is one of the first attempts to use a data set of this size for these purposes. In previous cases, problems have been encountered in extracting variability information from ERS-1 data. The long 35-day repeat pattern gives poor temporal resolution compared with the 10-day repeat of the T/P satellite, whilst the Geodetic phase (see Section 1.2) has a 168 day period, which therefore cannot be used with straightforward techniques such as repeat pass.

A process using single satellite crossovers (SXOs) has been devised in an effort to overcome this problem. Crossovers provide a data set which inherently contains variability information, and eliminates the problem of finding a fixed repeat period on which to base any time series. The crossover data set also has accuracy advantages over some other schemes, due to the cancelling of geographically fixed (correlated) errors in

the data. Repeat pass data allows cancellation of geographically variable (anti-correlated) as well as fixed errors, but this is not a feasible method when using data from a satellite which has taken in many orbital phases.

SXOs are relatively free from systematic error sources. They also preserve variability patterns, are straightforward to produce, and are easy to work with. They are not suitable for mean sea surface work or geoid determination, since any mean offset in an altimeter reading is lost when the crossover difference is calculated, but this is not a problem in a variability study.

In this thesis a method will be shown for extracting sea surface variability from a data set of corrected SXOs. This method involves:

- Construction of a time series of sea-level anomalies in each geographical location ('bin') from the SXOs using a least squares technique.
- Construction of similar time series from T/P Repeat Pass data.
- Combination of these two data sets, temporal smoothing and re-gridding onto a regular 5 day grid by optimal interpolation.
- Splitting this signal into spatial modes using an Empirical Orthogonal Function (EOF) decomposition which reveals individual signals as separate modes of variability, enabling more straightforward interpretation.

The majority of the data used in this thesis will come from the two satellites which are described in Sections 1.2 and 1.3.

1.2 The ERS Satellite Missions.

ERS-1 was designed and built almost entirely within the European Space Agency (ESA). The satellite carries a single Radar Altimeter, a Laser Retro-Reflector (LRR) array, Precise Range and Range-Rate Equipment (PRARE), an Active Microwave Instrument (AMI) which can behave as both a Synthetic Aperture Radar (SAR)

and a Wind Scatterometer, and an Along Track Scanning Radiometer (ATSR) (*Carnochan*, 1996).

The altimeter is a single frequency device which uses a radar pulse in the Ku band at 13.8 GHz and gives continuous information from 81.5° North to 81.5° South, including information about ice sheet coverage with a dedicated 'ice-mode'. The altimeter also provides information about significant wave heights and an estimate of the sea surface wind speed. The LRR is used as a precision and operational orbit determination tool in conjunction with the network of ground based Laser Ranging Stations. PRARE was designed to be a precision orbit determination tool but shortly after launch the equipment failed irrecoverably and the full abilities of the system were never explored. The ATSR equipment functioned successfully, however, and combined an infrared radiometer and a microwave sounder, these providing measurements of sea surface and cloud-top temperature, cloud cover, and atmospheric water vapour content. The SAR can operate in both image mode and wave mode. In image mode it acquires wide-angle, all weather images over the oceans, polar regions, coastal areas and land. Wave mode involves producing small images (of approx. 5km by 5km) at regular intervals for the determination of the length and direction of ocean waves. The Wind Scatterometer uses three antennae for finding sea surface wind speed and direction. Also on board are two sun acquisition sensors and the solar array which maintain the satellite's battery charge over the long term.

ERS-1 was launched on the 16th July 1991 into a sun synchronous, near polar orbit with inclination of 98.5°, a mean altitude of 785 km and a period of 100.6 minutes. The ERS-1 mission took in several different orbital phases in an attempt to provide the widest possible range of information about the Earth's surface.

The intended life time of ERS-1 was extended when it became clear that the originally planned shutdown time would be significantly before the launch of ERS-2, and so ERS-1 was maintained active in its second multi-disciplinary phase until the 22nd June 1996, which was as long as budgetary constraints allowed.

Table 1 gives a summary of the orbital phases ERS-1 completed during its mission. Dates are also given in Modified Julian Date (MJD), or the number of days since 18th October 1860. The three-day-repeat orbit pattern during the commissioning phase was maintained but shifted by approximately 3° longitude for the first Ice Phase (phase B), where the emphasis was on Arctic and Antarctic regions and inland ice or snow. The altimeter includes a specialist 'ice-mode' specifically for this part of the mission. There was then a short experimental 'roll tilt mode' campaign between the 2nd and 14th of April 1992, which allowed the use of the SAR at an incident angle of 35 degrees to the Earth's surface in image mode. The first Multi-Disciplinary phase (C) was designed to be used to determine the mean sea surface and to aid the study of sea surface variability. The second Ice Phase (D) had identical characteristics to phase B. The Geodetic Phase is divided into two: Phases E and F. Each is a 168 day repeat with identical orbit parameters except for a westward shift of the ground track by half of the across track spacing at the end of the first cycle. This provided high density coverage of the Earth's surface, with ground track spacing at the equator of only 8 km on combining the two sub-cycles. This phase was intended to enable the determination of a high resolution geoid using the Altimeter. ERS-1 returned to the 35-day Multi-Disciplinary orbit pattern for the remainder of it's nominal mission (phase G).

ERS-1 Phase	Start Date	End Date	Repeat Period	Ground Track Spacing
A - Commissioning	16 Jul 91 *(48453)*	27 Dec 91 *(48617)*	3	837
B - Ice 1	28 Dec 91 *(48618)*	31 Mar 92 *(48712)*	3	837
C - Multi 1	14 Apr 92 *(48726)*	23 Dec 93 *(49344)*	35	72
D - Ice 2	1 Jan 94 *(49353)*	7 Apr 94 *(49449)*	3	837
E - Geodetic 1	8 Apr 94 *(49454)*	27 Sep 94 *(49622)*	168	16
F - Geodetic 2	29 Sep 94 *(49624)*	21 Mar 95 *(49797)*	168	16
G - Multi 2	21 Mar 95 *(49797)*	22 Jun 96 *(50256)*	35	72

Table 1. ERS-1 Mission Phases. Ground track spacings are quoted at the equator in km.

Once the second ERS satellite (ERS-2) was launched in April 1995, a tandem mission from both satellites took place in the standard 35 day Multi-Disciplinary pattern. The orbits were identical, but the satellites were separated by one day. ERS-1 was kept operational for approximately 15 months into the start of the ERS-2 35-day mis-

sion, but was eventually shut off with the satellite being put into safe mode and left on standby. In this configuration, the satellite is turned to face the Sun and orbit maintenance manoeuvres are stopped, maintaining the possibility of a future return to activity since the satellite's batteries are continually charged. Any switch on would depend on further funding - once a satellite is launched, the major costs are involved in maintaining the orbit and processing the data which the satellite returns. The satellite is still tracked routinely (as is every other Earth satellite in an effort avoid potential collisions with other space craft) so a switch on could be easily arranged.

ERS-2 has a similar design to ERS-1, only differing by the inclusion of the full PRARE equipment, rather than a test device, and the Global Ozone Monitoring Experiment (GOME) which is an absorption spectrometer tuned to measure the presence of Ozone, aerosols and trace gases in the stratosphere and troposphere. The ERS-2 PRARE device is fully operational and is proving to be a useful orbit determination tool. ERS-2 will be operational throughout 1998 and is expected to be kept running until the year 2000.

1.3 The TOPEX/Poseidon Satellite.

The TOPEX/Poseidon (T/P) satellite is a joint project between the U.S. National Aeronautics and Space Administration (NASA) agency and the French space agency Centre National d'Etudes Spatiales (CNES). The satellite is project managed by NASA's Jet Propulsion Laboratory (JPL).

The payload of the satellite consists of two Altimeters, a Microwave Radiometer, a Laser Retro-Reflector array, a DORIS (Doppler Orbitography and Radiolocation Integrated by Satellite) receiver and a demonstration Global Positioning System (GPS) receiver (*JPL*, No date). The two altimeters are a dual frequency device developed by NASA (called TOPEX) and a single frequency altimeter from CNES (Poseidon). The altimeters share an antenna and thus cannot be active simultaneously - TOPEX is active about 90% of the time and Poseidon the remainder. The Microwave Radiometer provides information about the total water content of the atmosphere, which can be used to

correct the altimeter readings. The LRR is used with the same network of satellite laser ranging stations as used for ERS-1 orbit determination. The DORIS equipment provides an alternative tracking system which uses Doppler techniques. A network of 40-50 ground transmitters in conjunction with the receiver on board the satellite give almost continuous all-weather global tracking of the satellite. This information is used for precise orbit determination. The on-board GPS receiver is for demonstration purposes only, showing a possible future alternative tracking method which makes use of the network of GPS satellites already launched.

T/P was launched in August 1992 from the European Space Agency's Space Centre in Kourou, French Guiana. Late into the month after launch, the satellite had completed its deployment and was manoeuvred into an orbit approximately 1336 km above the Earth's surface on a 9.9 day repeat track with an inclination of 66.059° and a period of 112.43 minutes. The low inclination of the satellite means that its ground tracks do not extend further north or south than 66° from the equator. This is not a major hindrance since it only prevents the observation of areas of the ocean which are covered with ice for at least some parts of year, and there is no ice component in the satellite's design objectives, unlike the ERS satellites. T/P was initially intended as a three year mission, but is in continued use in an 'extended observational phase'. It is likely that T/P will remain active until JASON-1 goes on line, expected in 2000: the condition of the satellite's batteries reveals that T/P could stay active until 2000 or later, but once JASON-1 is active and fully calibrated, T/P will no longer be needed. There will, however, be a tandem mission for a period of several weeks or months with JASON-1 following the T/P ground track only a few minutes behind T/P. This will enable a cross-calibration of the new satellite from the well known characteristics of the T/P satellite. Thus JASON-1 will be a direct follow-on mission from T/P. It will carry a very similar payload, and will be placed in the same orbit configuration as T/P. It is designed to be part of an on-going series of JASON satellites (hence the '-1' designation) which it is hoped will aid in the resolution of very long term sea-level changes.

T/P is designed primarily as an altimetric mission, unlike the more general Earth observation purposes of the ERS satellites. The satellite is placed in a high orbit to facilitate improved orbit determination accuracy (air drag and gravitational errors are

lower at higher altitude). The TOPEX altimeter is a dual frequency device, allowing ionospheric corrections to be applied from exact measurements, rather than from models or third-party observations. As a result the T/P satellite is a more accurate tool for this specific task than ERS-1.

1.4 Precision Orbit Determination.

A satellite's orbits are generated using available observations of its position. These observations, a mathematical model of many of the forces which are affecting the satellite as it moves, and a numerical integration system are all used to produce what is known as an 'ephemeris'. This is simply a log of the satellite's position above the Earth's surface at typically 30 second intervals. From this ephemeris it is possible to interpolate to the satellite's exact position at any time. More details about this procedure can be found in Chapter 3.

The calculated orbits for a satellite can be improved using many techniques (e.g. *Carnochan*, 1996 or *Ehlers*, 1993). The most obvious is to include more observation data in the determination procedure, but, in order to produce improvements, these data have to be of sufficient accuracy. SLR data are used in most procedures because they are the most versatile form of tracking data. All altimeter satellites carry a LRR, and with ERS-1 (as with most satellites) it is the only system to offer any information in the along and across-track directions.

Another potential observation data set is the altimetry itself. This, however, must be used carefully to avoid absorbing the sea-surface height signals (i.e. the desired information) as if it were orbit error, and also because altimetry only provides information about the radial orbit error. As a result it is generally only included in the orbit determination system as a guideline, since the altimeter residuals will, on average, drop as the orbit improves.

Also used are Single Satellite Crossovers (SXOs), Dual Satellite Crossovers (DXOs) and, if the satellite carries the appropriate hardware, PRARE, DORIS, or GPS

observations. SXOs and DXOs are discussed in subsequent parts of this thesis. More information on the PRARE, DORIS and GPS systems is given in Chapter 3.

1.5 The Altimetry Data.

A Radar Altimeter can produce readings of high precision. Readings are continuously taken and stored on board the satellite for most of an orbit. They are then transmitted to the ground in a burst of data as the satellite passes over one of its ground stations.

The data are supplied to the user community as Geophysical Data Records (GDRs) on CD-ROM by the appropriate agency (AVISO for T/P and CERSAT for ERS-1).

With the supplied corrections to the raw measurements the GDRs are a useful and accurate data set, but any actual inference of the sea surface height from these readings is corrupted by the relative inaccuracy to which the satellite's position is known. Therefore much effort is channelled into minimising the errors in the orbit determination procedure because any error in the calculated orbit height will be directly absorbed into the retrieved sea-surface signal. Further details of the altimetry processing are given in Chapter 2.

1.6 Single Satellite Crossovers.

The single satellite crossover (SXO) is defined as the difference in observed sea surface height at two epochs when an ascending (northward moving) arc passes the same point on the Earth's surface as a descending (southward) arc.

Since the altimeter readings at the two epochs are taken at the same physical location, there are several error sources which are no longer important. Any error which is

geographically fixed occurs in both readings, and when the difference is taken therefore cancels out (*Rosborough*, 1986).

With a sufficiently short time between the epochs it is possible to assume that there has been no significant change in the local sea surface height, and therefore the SXO gives a direct measurement of the orbit errors from the two passes of the satellite. Short time difference SXOs can therefore be minimised during the orbit determination procedure as a means of directly reducing the orbit error.

Conversely, over a sufficiently long period of time, the SXO represents a combination of the orbit errors and the sea surface variability signal. When the orbit error is removed through the use of short time difference SXOs (as described above), the sea surface variability signal can be obtained to a much improved accuracy from SXOs with longer time differences.

1.7 Dual Satellite Crossovers.

The dual satellite crossover (DXO) is defined similarly to SXOs, except that two satellites provide each crossover residual: the residual is the difference between the sea surface height at epoch t_1 and location **x** observed by satellite 1 and that observed by satellite 2 at the same location **x** at epoch t_2.

As for SXOs, sea surface variability can be assumed static over short periods of time, and a crossover represents the difference in orbit height between the two satellites at epochs t_1 and t_2. Since T/P orbits are significantly more accurate than those of ERS-1, it becomes clear that a DXO residual between these two satellites is mainly a function of ERS-1 orbit error, and therefore it is possible to solve for this orbit error if it is assumed that T/P orbits are error-free. Whilst this is not actually true, ERS-1 orbits are sufficiently improved using some form of dual crossover correction to justify its use.

There are several methods for carrying out the DXO correction in practice (e.g. *Le Traon* 1995 and 1996 or *Jolly and Moore* 1995) , and the one used in this project is

shown in detail in Chapter 4. It involves using a cubic spline fitting technique to produce a time series of radial orbit error, which can then be removed from the observation data before further processing.

1.8 Repeat Pass Data.

A satellite will pass over a particular part of its ground track once per cycle (i.e. once every 9.9 days for T/P). Therefore an altimetric observation of sea surface height can be obtained for every cycle, and these observations can be combined to produce a time series of sea surface height for this particular location. This technique is simple to apply in practice and has accuracy advantages over other techniques.

As is the case for SXOs above, repeat pass observations are taken at the same geographical location, and so certain sources of error in the underlying altimetric observations are cancelled out. Repeat pass data have the additional advantage over SXOs that more error terms are cancelled because the altimetric observation is always taken in the same *sense*. The observations at a certain point are always from ascending (or descending) arcs, and this allows the geographically anti-correlated components of various errors to be cancelled. An SXO, in contrast, is the difference between observations from an ascending and a descending arc and the anti-correlated effects become additive.

The main drawback with repeat pass data is its unsuitability for use with data sets such as the ERS-1 mission, in which no fixed repeat period exists so that any repeat pass analysis is forced to ignore large parts of the mission. Typically a 35-day repeat is used, thereby ignoring the ice and geodetic phases. This 35-day repeat period is also rather too long for most applications, and ideally an observation every 10 days or less is required.

T/P provides an ideal data set for repeat pass analysis. A long time series can be generated from the entire T/P mission with observations every 9.9 days as a minimum (less if a coarser grid size can be used).

1.9 Sea Surface Variability.

'Sea surface variability' is any change in the sea surface height which is not part of the normal tidal movements of the sea (which are removed from the altimetry at an early stage).

Many factors affect the sea surface height:

- the presence of a current will cause a change in surface level, even if the current flows deep below the surface,
- movement, strengthening, and weakening of currents are observable on the surface,
- the surface rises and falls approximately linearly with temperature, bringing about variability patterns which depend upon the seasons,
- wind stress promotes currents in surface waters, and thus strong winds indirectly affect the sea surface height (as well as producing surface waves).

In addition, each source of sea surface variability depends on the others. Any change at the sea surface will tend to feed back as a change in the climatic conditions that caused it, and the net result is a highly dynamic and complex system of causes and effects.

A sea height anomaly is defined as the actual height difference between the observed sea height and the mean sea surface at that point. The height of the mean sea surface at a point includes any effects from constant current flows, as well as the basic geoid height above the reference ellipsoid. By convention, a positive anomaly is an increase in surface height with respect to the mean.

2. The Altimetry Data

The altimetry data as supplied contains various ready-applied corrections to the raw altimeter range measurement. The models used for these corrections are detailed below. Updates to several of these models and some additional corrections are made prior to using the altimeter measurements for either orbit determination or oceanographic work. These updates are listed in Section 2.2, whilst the standard corrections are listed below in Section 2.1.

2.1 The Altimeter Measurement And Standard Corrections.

The sea-surface-to-satellite height measurement must be corrected for the various geophysical effects listed below. These corrections are derived by the agency which distributes the data set to the user community. The value of each correction is stored in the data set along with the corrected final value to enable removal of a default correction if an alternative is to be applied.

If the raw measurement of satellite to sea surface distance is h_{raw}, the corrected height h_{obs} is related to this raw measurement by the following formula:

$$h_{obs} = h_{raw} + h_{geoid} + h_{SSH} + h_{SE} + h_{tide} + h_{ib} + h_{ion} + h_{wet} + h_{dry} + h_{inst} - h_{SSB}$$

where (in brief),

h_{raw} = raw altimeter height measurement,

h_{geoid} = geoid height w.r.t. reference ellipsoid,

h_{SSH} = sea surface height w.r.t. geoid,

h_{SE} = solid Earth tide correction,

h_{tide} = ocean tide correction,

h_{ib} = barotropic correction (inverse barometric effect),

h_{ion} = ionospheric correction,

h_{wet} = wet tropospheric correction,

h_{dry} = dry tropospheric correction,

h_{inst} = instrument correction,

h_{SSB} = sea state bias correction.

These terms warrant further explanation:

h_{raw} is the distance the altimeter measures between the satellite and the sea-surface. This is simply calculated from the measured time taken for a radar pulse to travel to and from the sea surface directly beneath the satellite.

h_{geoid} is the geoid height above the reference ellipsoid at this particular location. The geoid is the level which the sea surface would fall to in the absence of any forces other than gravity. This is not a perfectly flat surface, since local features such as sub-oceanic ridges apply local variations to the gravitational force (as well as the obvious surface curvature of the Earth). This parameter is cancelled out when calculating cross-over differences since it is geographically correlated and time invariant.

h_{SSH} is the observed sea surface height at a particular geographical location. As for geoid height a quasi-stationary sea surface height cancels in a crossover difference.

h_{SE} compensates for the effect on the sea of the shifting of the fluid mass within the earth due to gravitational attraction from other celestial bodies.

h_{tide} is a combination of the ocean tide due to the Moon's and Sun's gravitational pulls, the pole tide correction (due to movement of the Earth's rotational axis), and the ocean loading tide (due to the effect of water mass compressing the ocean floor). All are taken from the CSR3.0 model from *Eanes and Bettapur,* 1995).

h_{ib} takes the effect of air pressure on the sea surface into account. An increase of 1 mBar in air pressure results in a local drop in sea level of approximately 1cm.

h_{ion}, h_{wet} and h_{dry} remove the errors which result from the radar pulse's differing propagation rates through the atmosphere - any propagation rate different to that through

a vacuum translates directly into a distance mis-reading at the altimeter. The atmosphere results in the radar pulse being slowed down, and therefore the satellite interprets this as an *increase* in distance. Greater water content in the atmosphere further slows the radar pulse.

For ERS-1 data, the Bent model (*Llewellyn and Bent,* 1973) is used to calculate h_{ion} whilst the ATSR microwave instrument on board the satellite provides information on the total water content of the atmosphere for the wet troposphere correction (h_{wet}). The dry tropospheric correction (h_{dry}) is taken from French meteorological data (*CERSAT*, 1994).

h_{inst} is a correction for the instrument's range bias due to internal timing inconsistencies. The currently observed range bias for the ERS-1 satellite is an under-reading of approximately 42cm. This includes a time varying compensation due to Ultra Stable Oscillator (USO) drift and bias jumps. The TOPEX satellite has an effective range bias of zero, since the bias is removed in the initial processing of the data before distribution takes place. As with ERS-1 however, there are time varying corrections to be applied due to similar problems. (See Section 2.2)

h_{SSB} is in response to the way that the radar pulse is returned from the sea surface. The sea state bias is a combination of the electromagnetic bias as a function of the significant wave height and the tracker algorithm bias. Significant wave height (SWH) is defined as the average of the highest one-third of the waves. From a perfectly flat ocean, the returned radar pulse would be very simple to interpret, but in the case of an ocean with waves on its surface, this causes changes in the return pulse shape with the wave-tips returning a small proportion of the pulse *before* the wave-troughs return the rest. This makes it difficult to determine the actual sea-surface height. Currently this correction includes the assumption that the sea surface height is miscalculated by 5.95% of the significant wave height (*Carnochan*, 1996).

2.2 Updates And Additional Corrections Applied At Aston.

2.2.1 ERS-1 Altimetry.

The updates and additional corrections applied to the ERS-1 altimetry (as supplied by CERSAT - see *CERSAT,* 1994) are:

- Ultra Stable Oscillator (USO) drift
- Scanning point target response (SPTR) bias jumps
- Sea state bias: 0.0595 of SWH (*Carnochan,* 1996)
- Radiometric wet tropospheric correction with linear scaling given by: $\Delta h_{corr} = 0.8182\Delta h - 1.9256$
- CSR 3.0 ocean tide model (*Eanes and Bettapur*, 1995)
- Aston, Delft University of Technology (DUT) JGM3 and DGM4, and DXO corrected (see Chapter 4) orbit heights used in preference to the supplied DPAF orbit heights

The ERS-1 altimetry has been released in several versions. Phases C and D of the ERS-1 mission are version 3.1 - 3.3; phase E and F version 3.5 and phase G version 5 as far as MJD 49922 and version 6 thereafter. Versions 3 to 5 can be considered to be consistent, but version 6 includes changes which may be significant. As a result only versions 3 to 5 altimetry are used in this project. The USO drift and SPTR bias jump corrections were downloaded from the Internet (*ESA*, No date).

2.2.2 ERS-1 sea state bias modelling issues.

Previous work has shown that the sea state bias used in satellite altimetry can be poorly modelled. The global distribution of wind speed and SWH show marked variations (Figure 1 - the equivalent wind speed distribution is very similar and is not presented here), and therefore it would seem to make sense to produce a regional value for SSB, as opposed to the single global, empirical scaling-factor as used here.

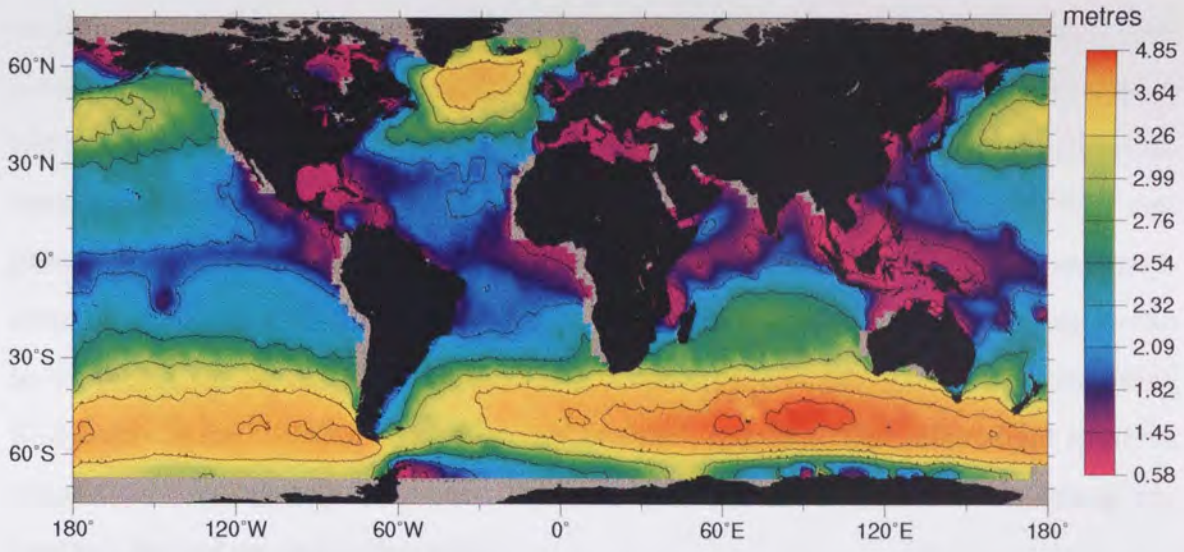

Figure 1. Global SWH distribution from 100 cycles of T/P altimetry. Contour lines are at 0.5m intervals.

However, in addition to the almost zonal distribution of the mean SWH and wind speed, these parameters also exhibit seasonal variations. To illustrate this, northern hemisphere summer and winter distributions of SWH are shown in Figure 2, each taken from two cycles of T/P altimetry data (cycles 103 and 104 from June 1995 and cycles 84 and 85 from December 1994/January 1995).

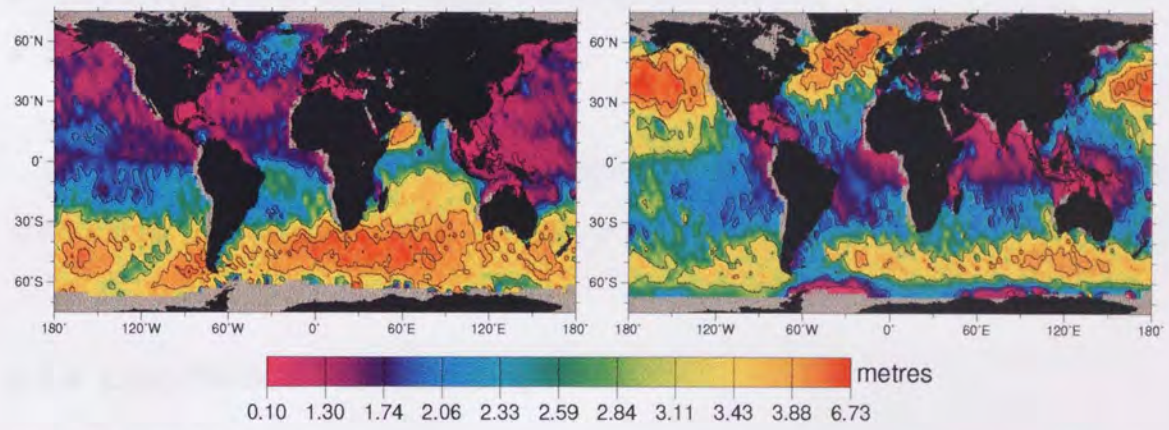

Figure 2. Global distributions of SWH in a Northern hemisphere summer (left) and a Southern hemisphere summer (right). Contour lines appear at 1m intervals.

Glazman et al., (1994) have shown that regional scaling factors or higher-order polynomials used to describe SSB (as a function of SWH, wind-speed or both) yield significant improvements in accuracy of up to 7% for T/P data. At the same time, however, the author acknowledged that empirical terms such as these do not address the

main problem, which is that of efficiently and accurately modelling the physical phenomena involved. As a result, the author felt that further investigations should be carried out into fully understanding the physical mechanisms of SSB before making radical alterations to the current altimetry dataset. As can be seen from the Figures, the SSB amplitude distribution is broadly zonal and long-wavelength in nature. Therefore any modelling problems in ERS-1 data arising from the use of a single global SSB value would be corrected for by the DXO correction procedure to T/P. Obviously, any remaining SSB errors in the T/P altimetry will go uncorrected, and this could bring about spurious zonal signals in the output data set. This must be borne in mind whilst discussing any satellite-derived oceanographic data.

2.2.3 TOPEX/Poseidon Altimetry.

All TOPEX/Poseidon altimetry used in this project is version 3 as supplied by AVISO (*AVISO,* 1995), with further changes as follows:

- USO algorithm errors
- *Gaspar et al* (1994) four parameter model for sea state bias
- CSR 3.0 ocean tide model

The supplied NASA JGM3 orbit heights are of sufficient quality to be used uncorrected for generating crossover residuals or for repeat pass calculations.

2.2.4 Long Periodic Tidal Modelling Errors.

At this point it is also worth discussing the implications of the potential modelling errors in long periodic solid Earth and ocean tides. These errors give rise to two distinct problems: a perturbation to the predicted orbit of the satellite, and an error in the measured altimetric height.

Shorter wavelength errors in the geophysical tidal corrections within the altimetric signal, due to an inadequate temporal sampling frequency by the satellite, are aliased to much longer wavelengths. These signals can then appear and be interpreted as if they were actual oceanic variability. For example the K_1 tide when sampled every 9.9156 days by T/P is aliased to a 173-day signal (Table 2). This is very close to half a year (the period of the semi-annual signal) and so any error in the modelling of the K_1 tide can result in spurious semi-annual sea surface variability signals being observed.

These problems also occur in ERS-1 altimetry, and different aliasing effects are observed for the geodetic and multi-disciplinary phases due to the differing orbital geometries. In particular, the sun-synchronous nature of the ERS-1 orbit means certain terms cannot be resolved at all, either because they are always sampled at the same phase (e.g. the S_2 tide) or they are aliased exactly into the annual or semi-annual signal.

Tidal Constituent	Q_1	O_1	P_1	K_1	N_2	M_2	S_2	K_2
Tidal period on Earth's surface	1.12	1.08	1.003	0.997	0.527	0.517	0.5	0.499
T/P alias period	69.4	45.7	88.9	173.2	49.5	62.1	58.7	86.6
ERS-1 alias period	132	75	365	365	97	95	∞	183

Table 2. Original tidal periods in days on the earth's surface and the altimetric aliasing period for T/P and ERS-1 per tidal constituent. ERS-1 data pertains to multi-disciplinary orbit pattern. All values in days.

Errors in the tidal model cause orbital perturbations because the movement of water mass alters the local gravity field and this effect is modelled by the orbit prediction procedure. ERS-1 and T/P have different resonance wavelengths for each tidal constituent due to their differing orbits. These wavelengths are tabulated in Table 3.

Fortunately ERS-1 orbit error perturbations can be removed by the correction of ERS-1 orbits against T/P by DXO, since the errors have wavelength much greater than one orbital cycle. Aliased tidal errors in the altimetric signal will not necessarily be removed by this correction, however, since whilst they are observed as long periodic signals at *each geographical location*, they appear as short wavelength variations in the DXO dataset which is used to perform the correction. Additionally, in the process of this

correction any errors in the T/P data (orbital- or altimetric-based) will be assumed to be true oceanographic signal, and as such will be passed uncorrected directly to the output data set.

Tidal Constituent	Q_1	O_1	P_1	K_1	N_2	M_2	S_2	K_2
T/P orbit resonance	8.68	12.7	89.0	174.0	8.26	11.8	58.9	86.9
ERS-1 orbit resonance	9.4	14.2	369	360	9.6	14.8	∞	180

Table 3. Orbital perturbation resonance periods (days) per tidal constituent for T/P and the ERS-1 multi-disciplinary phase.

Lerch et al., (1994) showed that there are 12 dominant tidal frequencies giving rise to altimetric errors. Of these, the ones where the altimetric, aliased signal and the orbit error signal are of similar wavelengths give rise to the largest errors. *Marshall et al.*, 1995, (using an older tidal model than that used here) have shown by comparison with GPS data that of these 12 dominant frequencies, only those at 58.7 (the M_2 tide) and 62.1 (S_2) days wavelength produce any significant observable errors. (It should be noted that this analysis stopped short of investigating errors at wavelengths of greater than 110 days.) *Schum et al.* (1997) have compared the CSR3.0 tide model with several others, and concluded that this model provides the best fit to all currently modelled tidal terms (as is expected since it is also the most recent). An extract from their results of a comparison against tide gauge data at 102 sites globally for the here-used CSR3.0 tide is included in Table 4, showing that the M_2 and S_2 terms have RMS errors of 1.64 and 1.01 cm respectively. (Note that these are likely to be pessimistic estimations of the true accuracy as tide gauges are frequently sited in shallow seas where the global ocean models are degraded.)

Tidal Component	Q_1	O_1	P_1	K_1	N_2	M_2	S_2	K_2
T/P-gauge RMS diff. (cm)	0.30	0.95	0.40	1.12	0.67	1.64	1.01	0.52

Table 4. CSR3.0 tide model comparison against 102 tide gauge stations per tidal constituent for T/P data.

Note that for ERS-1 phases C and G altimetry, the P_1 and K_1 tides are aliased exactly into the annual signal, the K_2 into the semi-annual, and the S_2 is totally unob-

servable and is aliased into the quasi-stationary sea surface. Since there is effectively no repeat period during the ERS-1 geodetic mission, errors such as these will simply appear as uncorrelated noise in the altimetry.

As a result of this contamination, efforts were made to remove the effect of the T/P M_2 and S_2 altimetric aliasing errors by filtering out a sinusoidal signal at both 58.7 and 62.1 days wavelength by least squares from the time series in each bin. This successfully removes tidal errors at these particular wavelengths, but also removes any sea-surface variability signal of comparable wavelength.

As can be seen from Table 4, other terms with large RMS errors are the O_1 and K_1 terms. The O_1 has a small impact on the overall accuracy of the system (*Marshall et al.*, 1995), and its removal is unlikely to offer significant benefit. The K_1 term has a larger influence, but cannot easily be distinguished from the semi-annual sea-surface oscillation, and therefore was left unchanged. This means, however, that semi-annual signals in the sea-surface variability data may in fact be spurious tide-model errors, and any interpretation should be made with this in mind. Additionally, residual aliased tidal errors from the ERS-1 altimetry may be retained in the final data set, and may give rise to spurious signals at several wavelengths, though the amplitude of these errors will be small in comparison to the true oceanographic signals due to the accuracy of the CSR3.0 tide model parameters.

3. Precise Orbit Determination

In order to extract meaningful information from satellite altimetric data, it is necessary to find the exact location of the satellite, and most importantly, its exact radial height above the reference ellipsoid. Without this information, it is impossible to know whether any signals observed in the data are sea surface variability, or if they are simply artefacts of orbit errors.

A radial height is supplied in the altimetry data, calculated using a similar procedure to that shown below by whichever organisation supplies the altimetry data to the user community (NASA or CNES for T/P or ERS-1 respectively). Heights for ERS-1 are then replaced with an improved orbit generated at Aston or DUT before any further processing takes place. This is required before attempting to calculate any single or dual satellite crossover residuals. A description is given in the following section of the data used in this procedure, and the orbit determination procedure itself.

3.1 The Tracking Data.

Satellites are tracked by various methods. The most common method is Satellite Laser Ranging (SLR). An LRR is mounted on the earth-facing side of the satellite. It has the property that any light entering the device is bounced internally such that it leaves the device in exactly the opposite direction. The returned pulse is used to record the overall journey time. This can be translated directly into a range measurement which, in conjunction with the known co-ordinates of the laser station, provides the location of the satellite.

A LRR is simple to place on most satellites, as it is a passive device which requires no electrical hardware on board and therefore places no power demands on the satellite's batteries. It is light, simple, and cheap. Ground stations are widespread, though concentrated in the northern hemisphere due to there being more land mass North of the equator, and are generally well maintained and produce accurate data. The

SLR system does have problems, however, notably that laser light does not penetrate dense cloud, so many stations cannot provide information for long periods of time and also that it is difficult for a station to track more than one satellite at a time.

Other systems in use currently include DORIS, PRARE, and satellite to satellite GPS. These all rely on electronic hardware local to the satellite. DORIS and PRARE both rely on electronic hardware at ground stations, similar to that required for laser ranging.

DORIS uses a passive ground station system, which transmits a constant twin frequency signal. The signals are at 2036.25 MHz for accurate Doppler measurement and 401.25 MHz for ionospheric correction determination. The on-satellite equipment recognises the station identity from the signals and calculates an approach speed towards the station by the Doppler shift. Using such range-rate data as the satellite passes over the ground station, the satellite position can be post-processed or can be used in real time for on-board calculation of position. DORIS is in use on board the T/P satellite.

PRARE (Precise Range And Range-rate Equipment) uses active ground stations. A ground station will lock onto a particular satellite passing overhead on one of a selection of frequency bands. A series of timed signal-response messages are passed and the satellite location can be calculated from this information. PRARE was initially put on ERS-1 as a trial system, but failed shortly after switch-on due to a hardware fault. This was corrected on the ERS-2 satellite, which carries a fully functional PRARE tracking system.

GPS (the Global Positioning System) uses a different concept for tracking the satellite. GPS is a U.S. military sourced system which provides a system for global navigation. It utilises a network of satellites with orbits determined for data processing purposes arranged such that there is maximum availability of data at all points on the globe. These satellites transmit information about their location and identity to a receiver, which range from hand held receivers to more bulky ones with external aerials, enabling the software in the receiver to find its location to a high degree of accuracy. The small size of the receivers allows them to be mounted on board a satellite which

permits tracking to take place as for a ground based receiver. The receiver mounted on T/P is a demonstration instrument (i.e. the data is not currently used for orbit determination), but this system showed potential for future use on board satellites due to the well known and reliable nature of the hardware and software. GPS will be used on Jason-1.

Altimetry height measurement data are also sometimes used as tracking information. Mostly they are heavily down-weighted in the solution procedure, since they contain the ocean variability signals which could otherwise be lost in the course of the correction technique. Altimetry data only have an impact on the radial positioning, providing no along or across track information. More commonly they are used as an indicator of the quality of the orbit, rather than as an independent tracking method.

The radial orbit accuracy can be improved through the use of short time difference SXOs. A crossover difference represents the radial orbit error difference between two epochs. If this crossover difference is minimised in the calculation of the orbit then the radial orbit error is also minimised.

Another method (which will be discussed later) is the use of dual satellite crossovers (DXOs). In this case we can improve the ERS-1 orbit accuracy by including DXOs with T/P. Low-orbiting satellites such as ERS-1 (~760 km orbit height) suffer greater problems with gravitational and air drag models than the higher T/P (~1340 km) due to the increased atmospheric density at lower altitudes and attenuation of the gravity field with height. Therefore they cannot provide orbits of precision comparable with those of T/P. The application of DXOs allows an improvement in the radial precision of the ERS orbits which in turn improves the accuracy of any oceanographic data extracted from the ERS data. As with SXOs, a short time difference between the two epochs must be applied in this technique to avoid sea-surface variability signals being absorbed by the orbit correction procedure. This technique is discussed in Chapter 4.

The accuracy of these tracking methods varies with several problems impacting each. SLR is the most commonly used method, and typically has an accuracy of 2-3 cm RMS. The main drawback of SLR data is that the laser does not penetrate dense cloud layers, causing data gaps. This affects laser stations in higher northern and southern

latitudes more seriously since they suffer from cloud cover for greater proportions of the year. DORIS range rate measurements are not as precise as SLR, but are more plentiful which compensates. PRARE can provide similar accuracy to SLR, but only if a range bias is recovered per orbit. GPS is not currently used in precise orbit determination systems, but it is hoped that this will change with Geosat Follow-On (GFO) which uses GPS as its primary tracking system. Precision of better than 3cm RMS is hoped for.

3.2 Orbit Prediction/Correction Procedure.

In order to produce precise orbits, a satellite's motion is modelled using Newton's equations of motion and models which estimate the forces acting on the satellite. Various parameters are recovered as part of the orbit determination procedure including the actual position and velocity, scale factors for surface forces, and empirical 1 cyc/rev accelerations. These parameters are then refined by a differential correction procedure using the least squares method, which minimises residuals between observed data (such as SLR, DORIS or PRARE observations) and the calculated equivalents. This corrects the parameters, which can then be used to regenerate the arc more accurately.

This system comprises several parts:

- a force model,
- an integrator,
- the least squares, differential correction procedure.

Together these make up the SATAN-A suite of software, which is an acronym for 'Satellite Analysis at Aston'. The SATAN-A suite was developed from the core SATAN programs written by Graham Appleby and Andrew Sinclair (*Sinclair and Appleby*, 1986).

3.2.1 The Force Model.

a) The Earth's gravitational potential.

This is the greatest force that acts on an Earth satellite.

From Newton's law of gravitational attraction between two bodies of mass m and M a distance r apart, the force F is given by

$$F = \frac{GmM}{r^2},$$

where G is the universal gravitational constant, which acts between those two bodies. From this, using Newton's second law, it can be deduced that the mass m experiences an acceleration of magnitude

$$a = \frac{GM}{r^2}$$

due to this attracting force. This acceleration vector can be written as a potential $\mathbf{a} = \nabla V$ where

$$V = \frac{GM}{r}$$

The above potential will satisfy Laplace's equation and is thus harmonic.

A more complex mass, such as the Earth, has a potential which can be expanded as

$$V = G\int_x \int_y \int_z \frac{\rho(x,y,z)}{l(x,y,z)} dxdydz$$

where $\rho(x,y,z)$ is the density at the point (x,y,z) within the Earth's mass, a distance $l(x,y,z)$ from the external mass m. This potential also satisfies Laplace's equation (*Kaula*, 1966). Also note that due to the nature of the problem it is more usual to work in spherical co-ordinates.

Solving Laplace's equation in spherical co-ordinates for the potential V by the method of separation of the variables yields the expression

$$V = \frac{GM_E}{r}\left(\sum_{l=0}^{\infty}\sum_{m=0}^{l}\left[\frac{R_E}{r}\right]^l \mathrm{P}_{l,m}(\sin\phi)\left(C_{l,m}\cos m\lambda + S_{l,m}\sin m\lambda\right)\right) \qquad 1.$$

where M_E is the Earth's mass and R_E its mean radius. The functions $\mathrm{P}_{l,m}$ are the associated Legendre polynomials and $C_{l,m}$ and $S_{l,m}$ are the harmonic coefficients which collectively define a given gravity field. ϕ and λ are latitude and longitude respectively in a geocentric Earth-fixed co-ordinate system, and r is the distance between the attracted point mass m and the Earth's centre.

Equation 1 is the most commonly used form of the potential V for expressing the gravitational force $\mathbf{F}_G$ on an Earth satellite,

$$\mathbf{F}_G = m\nabla V . \qquad 2.$$

b) Solid Earth and ocean tides.

The Earth is not a rigid body and is deformed by gravitational attraction of other masses in the solar system. Under these effects both the solid Earth and the oceans are perturbed. These changes affect the gravitational attraction $\mathbf{F}_G$ as modelled in 3.2.1a.

The ocean tides and the frequency dependant portion of the solid Earth tides are modelled by variations in the gravity field harmonic coefficients $C_{l,m}$ and $S_{l,m}$. The frequency independent part of the solid Earth tides is modelled by the potential

$$U_2 = \sum_{j=\text{sun, moon}} \frac{GM_j}{r_j^3} \frac{R_E^5}{r^3} k_2 \left(\tfrac{3}{2} \cos^2 s_j - \tfrac{1}{2} \right)$$

where s_j is the angle subtended by the sun/moon and satellite at the Earth's centre, k_2 is the second degree Love number; r_j is the geocentric Earth-body distance; and r is the geocentric Earth-satellite distance. M_j is the mass of the jth body (*Lambeck,* 1980). R_E and G are constants as defined in 3.2.1a.

$\mathbf{F}_G$ as used in the force model therefore also incorporates the solid earth and ocean tidal effects ($\mathbf{F}_{SET}$).

c) Third body attraction.

Other bodies in the solar system will also exert a gravitational attraction on an Earth satellite. These bodies can be modelled as simple point masses due to the large distances involved. The celestial bodies currently used are the Sun, Moon, Venus, Mars, Jupiter and Saturn.

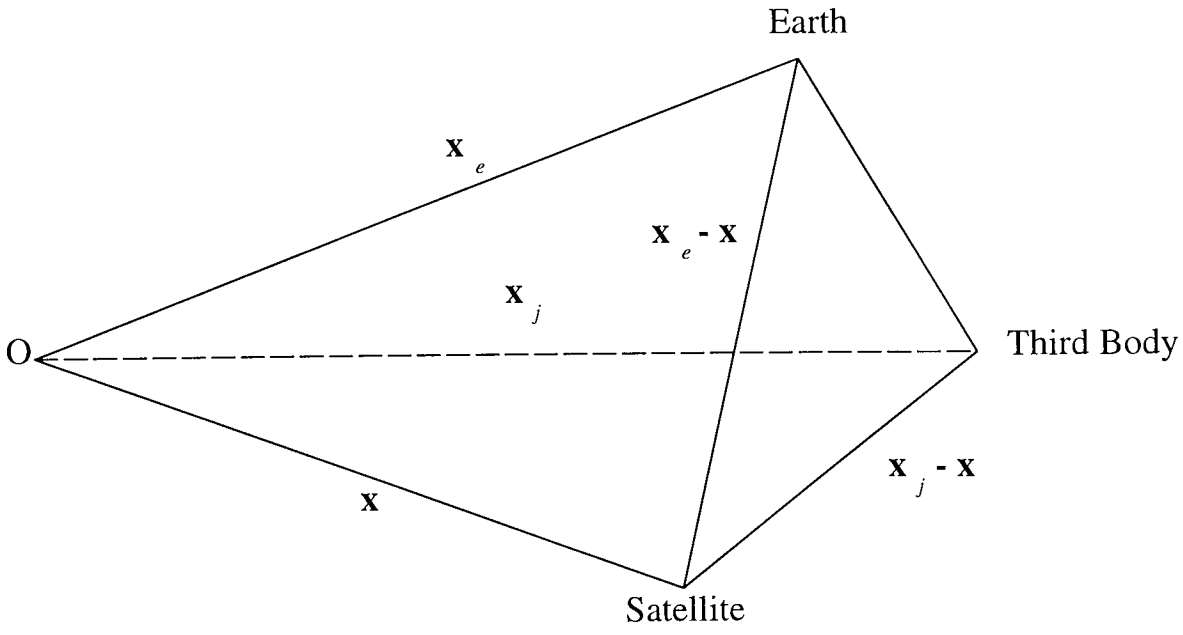

Figure 3. The system of vectors describing the locations of the three bodies relative to the origin O.

It can be shown that the force applied at the satellite by these bodies is

$$\mathbf{F}_{TB} = G\sum_{j} M_j \left(\frac{(\mathbf{x}_j - \mathbf{x}_e) - (\mathbf{x} - \mathbf{x}_e)}{\left|(\mathbf{x}_j - \mathbf{x}_e) - (\mathbf{x} - \mathbf{x}_e)\right|^3} - \frac{(\mathbf{x}_j - \mathbf{x}_e)}{\left|(\mathbf{x}_j - \mathbf{x}_e)\right|^3} \right) \qquad 3$$

where j is each of the bodies used in turn (*Brouwer and Clemence*, 1961). The vectors $\mathbf{x}_j$, $\mathbf{x}_e$ and $\mathbf{x}$ relative to the origin O are shown in Figure 3.

d) Atmospheric drag and lift.

The satellite experiences a retarding force (in the opposite direction to its movement) due to atmospheric particles impacting and rebounding from its surfaces. This effect is known as air drag and acts predominantly in the along track direction but also has an important influence on the radial height of the satellite. Other forces applied by the atmosphere are due to atmospheric rotation, affecting the across track velocity of the satellite, and the effect of imperfect collisions between particles and the satellite's surface (which results in atmospheric lift). These forces are, however, relatively unimportant compared with drag.

Drag is modelled by

$$\mathbf{F}_{\mathrm{D}} = -\tfrac{1}{2}\rho C_{\mathrm{D}} A_{\mathrm{D}} \mathrm{v}_r \underline{\mathrm{v}}_r, \qquad 4$$

where ρ is the atmospheric density, A_{D} is the effective area of the satellite in the direction of travel (which is calculated using a model of the satellite's major external surfaces), $\underline{\mathrm{v}}_r$ is the velocity vector and $\mathrm{v}_r = |\underline{\mathrm{v}}_r|$. The minus sign indicates that the force applies in the direction opposite to the satellite's travel. C_{D} is known as the drag coefficient but, in practice (due to poor modelling of ρ) it is reduced to being used as a scaling factor which is solved for empirically at several points during a long arc orbit. Currently, C_{D} is solved for at quarter daily intervals in order to absorb atmospheric mis-modelling problems (*Ehlers,* 1993).

e) Direct solar radiation pressure.

The flux of solar radiation hitting the surface of a satellite gives rise to a significant orbit perturbation. Direct solar radiation from the sun is modelled simply as (*Aksnes*, 1976)

$$\mathbf{F}_R = -s_E P C_R A_R \mathbf{e}_s \qquad 5$$

where P is the solar flux per unit area at the satellite (which is calculated from observed solar flux data and the fact that the flux follows an inverse square relationship to distance), A_R is the effective satellite area (calculated similarly to the A_D term used in Equation 4) perpendicular to the unit satellite-to-sun vector $\mathbf{e}_s$. C_R is the solar radiation scale factor which allows for interactions between the incident photons and the satellite surface (i.e. non-perfect collisions). s_E is the shadow factor, introduced to take account of the effect of the Earth blocking the radiation from reaching the satellite. Its value ranges from zero in total shadow to one in full sunlight, and smoothly changes from one to the other based on the penumbra from an approximated (spherical) Earth.

f) Earth reflected albedo and infra-red forces.

The final model concerns the radiation reflected from the Earth to the satellite and the radiation that is absorbed and re-emitted later as infra-red radiation.

The infra-red radiation is always present, even during the hours of darkness, but the reflected radiation comes only from the regions that are visible from both the sun and the satellite simultaneously. Reflected radiation is modelled by splitting the Earth into 13 regions and calculating the effects of each one in turn - taking into account the effective area, the reflectivity of that area, the distance and other factors. The force is written as

$$\mathbf{F}_{\text{AlbIR}} = \sum_{i=1}^{13} \mathbf{dF}_i \qquad 6$$

where

$$\mathbf{dF}_i = -\left(\gamma_i C_{\text{R}} P \cos\theta + P_{\text{IR}}\right)\left(\frac{A_R dA_i \cos\alpha}{\pi c d^2}\right)\mathbf{e}_d$$

where dA_i is each respective region's area, P and C_R are as defined in 3.2.1e for F_R (Equation 5), γ_i is the albedo of the ith surface element; P_{IR} is the emitted infra red radiation (interpolated from tables as in *Vonder Haar and Suomi*, 1971); θ is the angle between the surface normal and the satellite geocentric position vector; A_R the area of the satellite perpendicular to the satellite-to-Earth unit vector $\mathbf{e}_d$ and d is the distance between the satellite and the Earth surface-element.

Thus the force model as used in the prediction program is the sum of these forces, i.e.

$$\mathbf{F} = \mathbf{F}_{\text{G}} + \mathbf{F}_{\text{TB}} + \mathbf{F}_{\text{D}} + \mathbf{F}_{\text{R}} + \mathbf{F}_{\text{AlbIR}} \qquad 7$$

where $\mathbf{F}_G$ incorporates $\mathbf{F}_{SET}$.

3.2.2 The Integrator.

The integrator is used to compute the satellite ephemeris (which contains the position and velocity, **x** and **x'**, at times t_i) and the required partial derivatives with respect to the solution parameters ($\frac{\partial}{\partial P_k}\mathbf{x}(t_i)$) for use in the correction program.

The SATAN software uses a standard eighth-order Gauss-Jackson numerical integrator. This, along with the force model **F** (detailed above) and a set of a priori parameters P (including initial position and velocity, values of C_D and several others) is used to produce an ephemeris. Further details of the parameters P are given in section 3.3.

The integrator's step length is chosen by the user, along with the gravity field model (i.e. which set of $C_{l,m}$ and $S_{l,m}$ are used in the force model), the atmospheric model, and the tide model. All such configurations are placed in a text input file which controls the integration process.

3.2.3 Differential Orbit Correction.

The basic premise is that of minimising weighted squared residuals. A residual is the difference between the satellite's position as calculated by the integration procedure C_i and its position as found from some physical observation O_i (from observations such as SLR, altimetry, SXOs or PRARE). The sum of such residuals is given by

$$\sum_i \omega_i (O_i - C_i)^2 \qquad 8$$

where ω_i are the weights for each individual observation.

Suppose that an orbit can be modelled using a set of true parameters P^*. These values are not known, but we have an estimate P^0, such that

$$\mathbf{P}^* = \mathbf{P}^0 + \Delta\mathbf{P}\,. \qquad 9$$

Given that the C_i are functions of P^0, i.e. $C_i = C_i(P^0)$, it is possible to refine ΔP to improve our estimate P^0 to P^1.

Minimisation of expression 8 is undertaken in the usual least squares method by differentiating with respect to each model parameter and setting all the resultant equations equal to zero, i.e.

$$\sum_i \omega_i \left(O_i - C_i\left(\mathbf{P}^*\right)\right)\frac{\partial C_i\left(\mathbf{P}^*\right)}{\partial p_k} = 0, \forall p_k \in \mathbf{P}\,.$$

Since P^* is not known, and equation 9 holds, we can write

$$C_i(P^*) = C_i(P^0 + \Delta P). \qquad 10$$

Equation 10 can then be expanded about P^0 by use of a Taylor series. This series is truncated to a linear series, enabling an approximate solution to ΔP to be found by simple means. A text input file controls the operation of the correction procedure. It includes the ability to fix certain parameters, set rejection levels on a per-data-type basis and define the weighting for each of the various observation data types. This program then writes a results summary file and writes a set of corrected parameters P^1. The orbit prediction is subsequently re-run using the improved parameters and these two procedures iterated until the system converges to near P^*.

3.3 The Aston Orbits.

On a practical level, the SATAN-A suite requires various text input files to be completed to fully configure a particular run. There are separate input files controlling the integrator and the correction program. The parameters P^0 are stored in another file and these initial parameters are referred to as a 'start vector'. The start vector defines an approximate location and velocity vector for the satellite ($\mathbf{x}(t_0)$, $\mathbf{x}'(t_0)$), a set of drag coefficient C_D values (four per day for ERS-1, one per day for T/P), a solar radiation scale factor C_R, and a set of along and across track empirical accelerations. This vector is used to initialise the appropriate components of the force model, and an arc is integrated. Arc length is generally less than 6 days for ERS-1 and 10 days for T/P (a whole cycle) due to the differences in modelling accuracy for the two satellites. This arc is then compared with the observation data for the time period in question, and residuals calculated for each observation. These are filtered with a simple thresholding system to remove erroneous values. Since these residuals are functions of the parameters in the state vector, refinements are solved for by least squares in order to minimise the remaining residuals.

Different observation types are given different weights in the least squares procedure. Laser observations are weighted according to the historical accuracy of the par-

ticular station that recorded that pass. Single satellite crossovers are given a similar weighting to some of the better stations, as are PRARE data. DORIS data are weighted differently since they are less accurate than SLR, and consist of range-rate observations instead of absolute range. Altimetry data, if included, are given a low weight in the solution, to leave the orbit unaltered but enable certain additional parameters such as the altimeter bias to be recovered.

For ERS orbits, where a six day arc is produced (for example), the first and last half days are effectively discarded by overlapping the start of the next and the end of the previous arcs by a full day. This avoids using the least precise parts of the orbit, which are effectively unconstrained by a lack of data before or after. In many cases a discontinuity in the satellite's position at the switch over point between arcs can be observed. The overlapping of arcs minimises this effect, though cannot remove it entirely.

The methods employed in this section are standard orbit determination techniques which have evolved over many years and with enhancements devised by many authors.

For generation of single or dual satellite crossovers a basic orbit must be available, since the generation procedure uses the ephemeris files to find the exact locations of the crossovers. To this end, the currently available Aston produced ERS orbits were utilised. These covered the multi-disciplinary phase (C), second ice phase (D) and second multi-disciplinary phase (G). Precise orbits were required, however, for the Geodetic phase (E and F). A summary of the generation of these orbits is given in section 3.4.

3.4 The Orbits Produced.

Using the procedure outlined in section 3.3 and the software described in sections 3.1 and 3.2, seventy-two 6-day ERS-1 arcs were produced covering the Geodetic Mission (GM).

All available SLR data were used in the determination procedure, along with altimetry and SXOs. Initial laser only orbits are required in order to generate the SXOs, but then further iterations of the prediction-correction procedure can be carried out including the SXO information to produce enhanced orbits.

Figure 4 is a plot of RMS residual values (in metres) for the ERS-1 mission from MJD 48726, the start of phase C, to the end of phase G at MJD 50232. The arc with very low RMS (0.0014 m at MJD 49622) is a very short arc, less than 1 day, with few tracking data which joins two large manoeuvres at the link between the sub phases of the Geodetic mission. Short arcs were often necessary near to manoeuvres (since it is not possible to integrate through a manoeuvre), and they can easily achieve RMS residuals of less than 0.06m. This fact, however, has little meaning in terms of the quality of the orbit produced, as it is easy to fit an orbit to so few data points and it is likely that these arcs are in fact *less* accurate. Very short arcs such as these are discarded before any further processing.

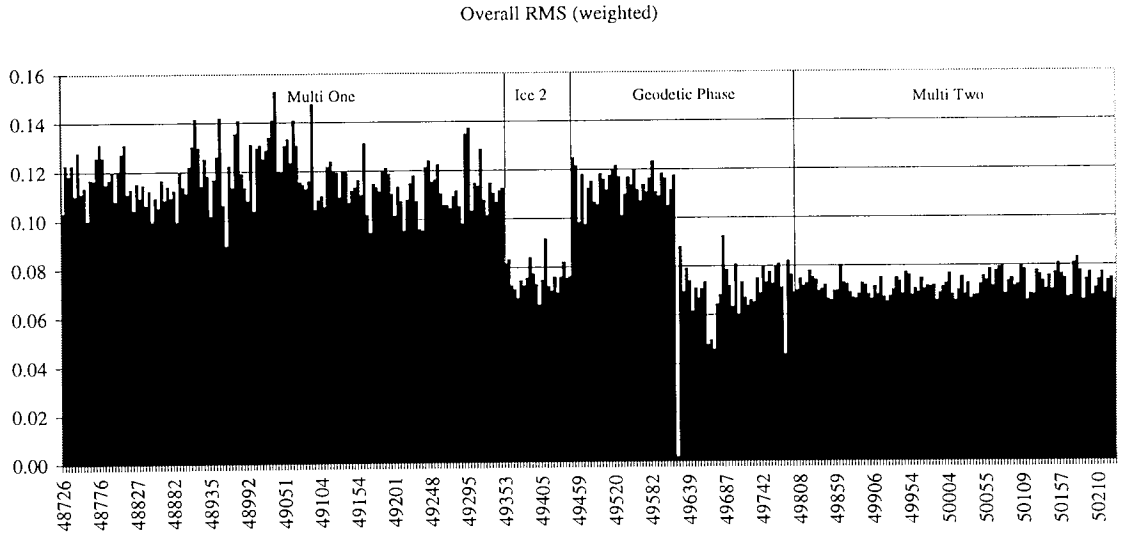

Figure 4. Overall weighted RMS residuals (metres) for the ERS-1 mission.

Figure 4 shows that there was a general improvement in RMS residuals in the second half of the GM compared to the first (average values are shown in Table 5). This is due to an improvement in determination technique whereby a greater number of empirical along and across track accelerations are solved for per arc. Orbits generated ear-

lier used a single set per arc which was subsequently increased to one set per day, thereby improving the fit to the observations. It is not clear without independent verification whether this has any significant effect on the errors within the orbit determination procedure, or simply makes a better fit to imperfect data. It was therefore decided that to re-run the early orbits was unnecessary as the statistical benefits were outweighed by the extra time required, and more importantly because the final data set will be corrected against ERS-1–T/P DXOs before any further processing is carried out, much reducing the impact of this type of difference.

Phase	Overall RMS (m)
Multi One (C)	0.1159
Ice Two (D)	0.0752
Geodetic Phase E	0.1130
Geodetic Phase F	0.0698
Multi Two (G)	0.0717

Table 5. RMS residuals in metres for each ERS-1 mission phase.

Table 5 shows the RMS residual values averaged over each phase of the ERS-1 mission. It should be noted that the phase C orbits were produced using the inferior JGM2 gravity field model, which has since been replaced with the JGM3. The fact that some arcs were produced with the JGM2 gravity field model is problematic as it introduces a systematic error and inconsistency to the system. As a result, the Aston orbits were subsequently used only for the derivation of crossover locations. Orbital heights for further analysis were taken from the DUT-supplied JGM3 or DGM4 orbits package (*Scharroo*, 1999). These packages were constructed using more consistent modelling than the then-available Aston orbits, and as such were more suited for further analysis.

The DGM4 orbits were generated using the DUT gravity model DGM4 (*Scharroo and Visser*, 1998), which is produced by using ERS-1 SXOs to solve for a correction to the JGM3 gravity model parameters. The JGM3 orbits are based on the standard *Tapley et al.* (1996) model.

Though the DUT orbits were more consistent, there was still a need for the use of a correction scheme and to this end a dual crossover-based correction scheme was used, which is discussed in Chapter 4.

4. The Dual Crossover Orbit Correction System

4.1 The Dual Crossover

A dual crossover residual ΔDXO is the difference between sea-surface height residuals measured by two satellites at different epochs over the same point on the ocean's surface. i.e.

$$\Delta DXO = \Delta h_{sat\ 1} - \Delta h_{sat\ 2}.$$

The height residual for each satellite is calculated as $\Delta h = h_{orb} - h_{corr}$ where h_{corr} is the corrected observed altimeter measurement (i.e. the distance from the sea surface to the satellite) and h_{orb} is the satellite's orbital height above the reference ellipsoid. Thus Δh represents the observed sea surface height above the reference ellipsoid.

If the true height of the satellite above the reference ellipsoid is h_{true}, and the orbit error in the radial direction is given by Δr, h_{orb} can be written as $h_{orb} = h_{true} + \Delta r$. Δr can then be broken into the following main components:

$$\Delta r \cong \Delta r_{grav} + \Delta r_{ipv} + \Delta r_{srp} + \Delta r_{drag} \qquad 11.$$

where each error is caused by modelling problems in the orbit determination procedure. Equation 11 represents errors caused by gravity field mis-modelling (Δr_{grav}), the initial position (Δr_{ipv}) of the satellite and the surface forces solar radiation pressure (Δr_{srp}) and atmospheric drag (Δr_{drag}). The Δr_{grav} term is a combination of the geographically correlated (Δg_{corr}) and the geographically anti-correlated ($\Delta g_{anti\text{-}corr}$) components (*Rosborough and Tapley*, 1985). The geographically correlated component is constant for a particular type of orbit, whereas the anti-correlated part gives results of different sign depending on whether the satellite is on an ascending or a descending pass when the altimetry measurement is taken.

The mean sea surface and the geoid height used in each altimetric measurement are geographically correlated and therefore any errors in these values do not affect the accuracy of a crossover residual. It is therefore an accurate guide to the orbit errors present in the two altimeter readings.

The ERS-1 satellite suffers substantial air drag and gravitational modelling problems. The resulting large orbit errors can severely affect the accuracy of the altimeter data set. The relatively poor quality of ERS-1 orbits (5-10cm RMS) compared with T/P orbits (3-4cm RMS [*JPL*, 1993]) means that a short-time-difference DXO (where the two satellite passes making up the crossover difference occurred less than 10 days apart) can be considered to be a direct observation of the ERS-1 orbit error. Two assumptions are required: firstly, that there is *no* error in the T/P orbit and, secondly, that there is no sea-surface signal (after the removal of tidal effects) over such a short period of time. The first assumption is reasonable considering the much greater relative accuracy of the T/P orbits. The second depends on the knowledge that very little sea-surface variability occurs over time spans of less than about 20-30 days. The ERS-1 orbits can therefore be corrected in order to minimise such DXOs, thereby reducing the ERS-1 orbit error to a level similar to T/P.

This method must be used with caution because although T/P orbits are substantially better than those for ERS-1, problems exist with T/P orbits. In particular there are several discrete wavelengths of tide modelling errors in the observed T/P data at around 60 days wavelength (*Marshall et al*, 1995). Given that the ERS-1 orbit will be corrected to these data, the majority of these errors will be transferred to the ERS-1 data (*Jolly and Moore*, 1995). It is thus important to filter out such errors before further processing.

Le Traon et al. (1995) demonstrated that a cubic spline fitting procedure using single and dual satellite crossovers can be an effective means of computing a continuous approximation to the radial orbit error. They achieved a typical fit of approximately 7cm to the ERS-1 single satellite crossover data set. This is clearly superior to alternative methods which involve fitting a sinusoid wave to dual crossover residuals. *Jolly and Moore* (1995) showed a fit of 11.8 cm RMS for a 35 day cycle while *Cazenave et al.* (1996) reported a range of 7.5 to 9 cm for ten 37 day sub-cycles from the geodetic

phase. *Le Traon and Ogor* (1996) have recently updated their methodology and claim reductions to as little as 2cm RMS difference between the ERS and T/P orbits.

A modified version of the *Le Traon et al.* (1995) algorithm is used in this project which does not use ERS-1 single satellite crossovers in the cubic spline determination. This significantly reduces the complexity of the spline fitting procedure without seriously impeding the effectiveness of the system. This is particularly true since ERS-1 SXOs can only realistically influence the accuracy of the orbits outside the T/P latitudinal limits (±66°) and these areas are generally unsuitable for long term variability studies due to ice sheet extent during each hemisphere's winter causing large gaps between observations.

4.2 The Procedure.

A data set of dual crossover residuals with time difference between the two epochs of five days or less (to minimise the effects of sea surface variability) was produced for each of the ERS-1 phases used in this study.

An initial relative bias of 40 cm was removed from all residuals before a rejection threshold of 1 metre was applied to avoid skewing the data. The data were split into ERS-Topex and ERS-Poseidon residuals and the RMS of fit and an estimate of the relative bias for each was calculated. At this stage a three-sigma rejection criterion was applied to each data set (i.e. statistics were calculated and points outside 3σ rejected) to remove the effect of bad arcs. Remaining data points were then collated into a single discrete ERS-1 referenced residual time series.

Any errors in the time series are assumed to be randomly distributed. Following *Hayes* (1974), $N_{n,i}(t)$ is defined as the ith normalised B-Spline of degree n - 1 through a set of h knots $\lambda_1, \ldots, \lambda_h$ with the restriction that $N_{n,i}(t) = 0$ for t outside $[\lambda_{i-n}, \lambda_i]$. Within this range, $N_{n,i}(t)$ can be calculated from recurrence relation 12.

$$N_{n,i}(t) = (t - \lambda_{i-n})N_{n-1,i-1}(t) + (\lambda_i - t)N_{n-1,i}(t) \qquad 12.$$

For cubic splines, degree 3 is required, i.e., $n = 4$. This recurrence is initialised with

$$N_{1,i}(t) = \begin{cases} 1, & \lambda_{i-1} \le t < \lambda_i \\ 0, & \text{otherwise} \end{cases}$$

In order to create a full set of splines, the system requires a further n knots at each end of the time series, i.e.

$$\lambda_{-n} < \lambda_{-(n-1)} < \ldots < \lambda_{-1} < \lambda_0 \le t_{\min} < t_{\max} \le \lambda_{h+1} < \ldots < \lambda_{h+(n-1)} < \lambda_{h+n}$$

Assuming $n = 4$ and dropping the subscript for clarity, the full set of $h + 8$ knots therefore produces $h + 4$ B-Splines, $N_i(t)$, $i = 1,2, \ldots, h+4$, and thus an estimate of radial orbit error, Δr can be found by evaluating

$$\Delta r(t) = \sum_{i=1}^{h+4} c_i N_i(t)$$

The coefficients c_i can be determined via a least-squares technique by the generation of the observation equations f_j for each data point j over a set of m crossover differences (Equation 13).

$$\sum_{i=1}^{h+4} c_i N_i(t_j) = f_j, \qquad j = 1, \ldots, m \qquad 13.$$

This can be written in matrix form as

$$\mathbf{Ac} = \mathbf{f}. \qquad 14.$$

A is a m by $h+4$ matrix and **c** and **f** are column matrices with elements c_i and f_i respectively. The observation equations are converted to normal equations and equation 14 becomes

$$\mathbf{A}^T\mathbf{A}\mathbf{c} = \mathbf{A}^T\mathbf{f} \qquad 15.$$

and the coefficients c_i can be retrieved via a suitable matrix inversion of $\mathbf{A}^T\mathbf{A}$ and pre-multiplying through Equation 15.

Knot locations were chosen via improved rules compared to those used by *Le Traon et al.* (1995). Locations were set according to the number of residuals present in each arc. Where an arc had less than 10 residuals, a knot was placed at each end. Any arcs with more than ten residuals had an extra knot placed at the central point. Initially, a method was implemented which placed the extra knot at the epoch of the observation nearest to the centre of the arc, but this was seen to cause a problem due to the large number of arcs not containing residuals near to the equator, compromising the spline. Therefore the physical centre of the arc (i.e. the epoch of the satellite's crossing of the equator) was used, regardless of the spread of observations.

In addition, an iterative procedure was used to remove bad residuals which could not be picked out by simple thresholding. Knots were selected for the data set, the spline fitted, statistics calculated and residuals that differ from the fitted spline by more than 3σ were removed. The spline was then re-calculated using the remaining residuals. This procedure was iterated until no further residuals are rejected.

A further amendment to the technique was made after observing that spline fits in the equatorial region were worse than those near the latitudinal extremes. This was probably due to the much greater number of residuals at the latitudinal extremes forcing the spline fit. This effect was reduced by applying a weighting scheme giving greater weight to equatorial residuals. The weightings applied were:

$$W_j(t) = \cos^2 \phi \qquad 16.$$

i.e. the residuals were weighted according to the square of the cosine of the latitude.

An example of a section of cubic spline is shown in Figure 5, showing the locations of the knots chosen for each arc and an example of the spread of the residuals on each arc. The segment shown is typical. Note that where large variability exists in the residuals (e.g. around MJD 49571.86) the spline is heavily smoothed by comparison. Post correction fits were then calculated from the corrected DXOs. The results are shown in Table 6.

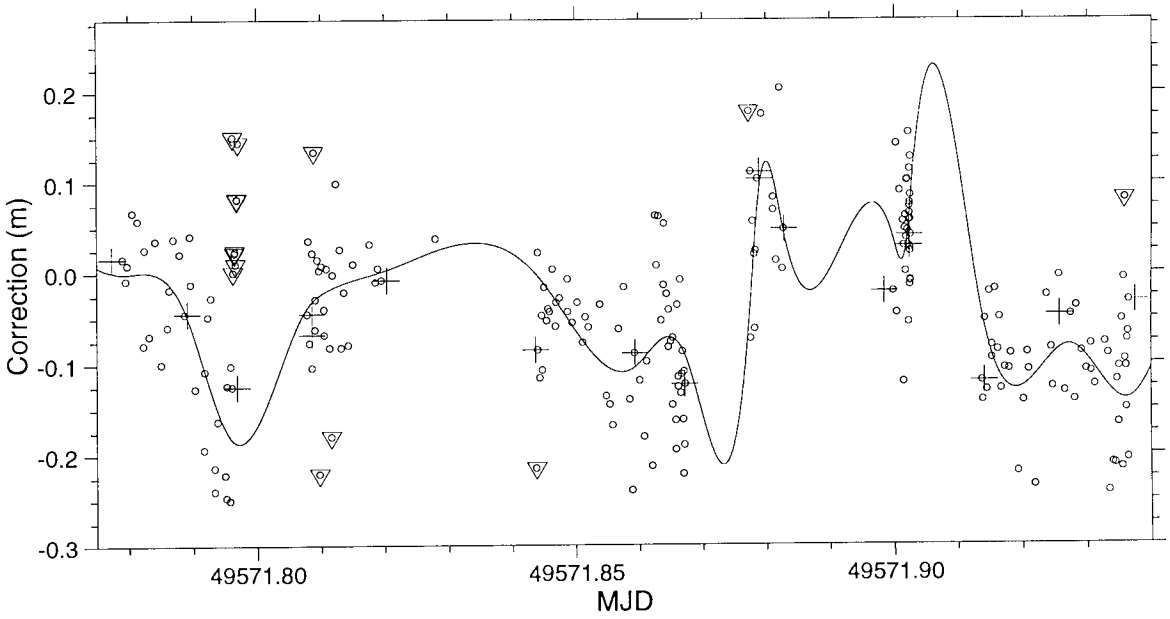

Figure 5. Example cubic spline segment. Plot shows residuals (open circles), rejected residuals (triangles) and knots (crosses).

As can be seen from Table 6 the procedure results in a significant improvement in fit to the dual crossover residuals, with all ERS-1 phases having final RMS fits of lower than 6 cm. Note that both Geodetic sub-phases E and F were processed as a single block.

Phase	Period (MJD)	No of XOs	Pre-corr fit T	Pre-corr fit P	No of Arcs	No of Knots	No of Rejs	Fit σ (cm)
C	48725-49350	616960	13.1	13.1	16508	45851	26365	5.92
D	49350-49450	157760	10.7	10.8	3532	9942	9950	5.91
E + F	49450-49790	460120	12.1	11.4	12152	34102	17315	5.80
G	49790-49960	210393	10.3	11.1	5580	15659	6904	5.75

Table 6. Results of fitting cubic splines to dual T/P–ERS-1 crossover residuals for the ERS-1 phases C, D, E, F and G. Initial ERS-1 orbits used Aston orbit heights.

Procedures relating to the determination of the cubic spline correction were carried out in collaboration with *Cullen* (1998).

4.3 Validating The Correction Technique By Comparison With Tide Gauge Data.

As a further validation of the DXO correction method in Section 4.2, comparisons were made between tide gauge observations and local ERS-1 single satellite crossovers (with unlimited time difference between the epochs). This technique allows us to ascertain whether the cubic spline procedure is absorbing any significant oceanographic signals along with the orbit error as well as whether it absorbs a useful quantity of orbit error.

When comparing satellite observations and tide gauge data, it should be noted that the Inverse Barometer correction must not be applied to the crossover observations since this effect is observed directly by the tide gauge. It is normal practice to apply a correction to the altimeter measurements to remove this effect for other uses of altimetry (See Chapter 2).

4.3.1 Tide Gauge Selection Process

Tide gauges were chosen from the global database of over a hundred in the World Ocean Circulation Experiment (WOCE) fast-release database. A subset of these (53) were chosen because they are based on remote islands, which tend to be in areas where altimetry does not suffer from poor shallow-water tidal models (e.g. on the continental shelf). These gauges give global coverage, though they are more concentrated in the Pacific and at low latitudes than elsewhere.

A study (*Moore et al*, 1999a) was carried out to validate these gauges. In this study, correlations and linear dependencies between TOPEX repeat pass time series,

ERS-1 repeat pass data and crossover data, and the tide gauge derived time series itself were calculated.

ERS-1 repeat pass data and crossover data including all but the second ice phase from 1992 onwards were used. The TOPEX short repeat time (9.9 days) meant that the use of a time series at just four points (from two ascending and two descending passes) near to each tide gauge gave sufficient temporal density. The longer ERS-1 repeat required the use of seven ascending and descending passes near to the gauge for the construction of the time series.

ERS-1 altimetric data were interpolated to every 0.25° latitude, and the nearest 7 ascending and descending points were chosen at the quarter degree of latitude adjacent to the tide gauge. This repeat pass procedure gave at most 22 points for the two 35-day Multi-Disciplinary phases, so no correlation analysis could be carried out. Similarly, temporal spacing is a problem for the ERS-1 crossover data, and as a result it was necessary to restrict the latitude of crossovers to within $\Delta\phi$ of the tide gauge latitude. $\Delta\phi$ was calculated as half the span between consecutive circles of latitude containing the cluster points of crossovers for the 35 day repeat cycle at the latitude of the gauge. Similarly to the repeat pass procedure, the longitude span was taken to be a factor of 3.5 of the corresponding span between consecutive ground tracks for the 35 day repeat data, thus including the nearest 7 SXO cluster points.

Though the latitude and longitude limits above are based around the 35-day repeat/501-orbit pattern of the ERS-1 multi-disciplinary phase, it is also an acceptable generalisation to the Geodetic mission's parameters. The Geodetic phase does not have 'cluster points' as such, due to the very long repeat period, but it does have a similar orbit pattern with 37 day sub-cycles. The ice phase is not well suited to this method due to the short 3-day repeat and wide ground-track spacing.

WOCE tide gauge data are distributed as hourly, daily or monthly values. The daily values used for this study were derived from the hourly data after removing spurious values and short period tides. The tidal filtering involves smoothing 119 hourly values to give a single value at midday on the central day.

Any gauge which was found to have a correlation exceeding 0.3 and a small or negligible slope when compared with both ERS-1 and TOPEX-derived time series was taken to be representative of the signals present in the open ocean. This eliminated eight gauges, including two which are known to suffer from subsidence or post-glacial rebound (where an area is rising due to the slow recovery of the Earth's crust previously compacted by ice sheets), making the reference tide gauge position unreliable. A further gauge was rejected when inspection of the time series revealed that there were negligible differences between the time series observed at gauges 3 and 30 since they are geographically close together (both in the Galapagos Islands) and are highly correlated. Gauge 3 was rejected since gauge 30 showed slightly better fits with TOPEX observations.

This procedure left 44 tide gauges, which are considered to be representative of open ocean variability signatures and can therefore be used to validate the DXO orbit correction scheme. The geographical locations of these gauges is plotted in Figure 6.

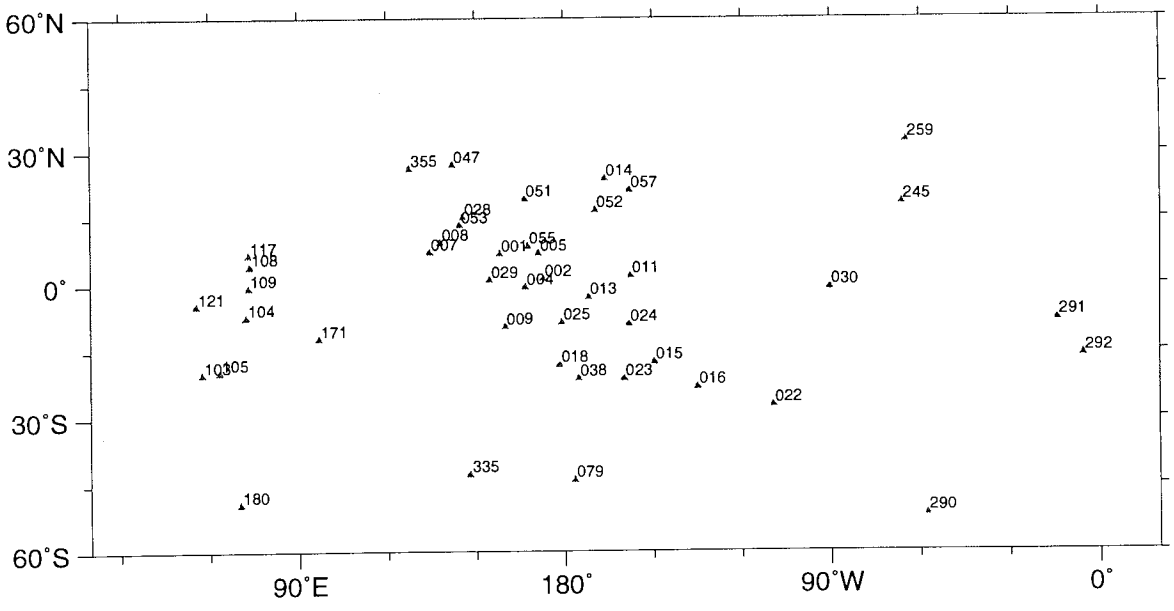

Figure 6. Locations and IDs of the 44 tide gauges used in the study analysing the DXO orbit correction procedure.

4.3.2 Assessment Of Correction Technique Against Tide Gauges

The 9,163,249 ERS-1 single satellite crossovers for the phases C through F were filtered to 41,622 crossovers which were within the required distance (as for the analysis in Section 4.3.1 above) from one of the tide gauges on Figure 6.

Gauge ID and name	# Obs	Latitude	Longitude	RMS Difference (m)		Correlations	
				Uncorr	Corrected	Uncorr	Corrected
1 Pohnpei	787	6.98	158.25	0.09352	0.08792	0.50532	0.62138
2 Betio	733	1.37	172.93	0.10554	0.08053	0.47228	0.59626
4 Nauru	1208	-0.53	166.90	0.11619	0.10582	0.39190	0.42771
5 Majuro	806	7.10	171.37	0.11031	0.11192	0.40456	0.47211
7 Malakal	792	7.33	134.47	0.10994	0.11340	0.55079	0.52421
8 Yap	932	9.52	138.13	0.08476	0.08794	0.72028	0.69937
9 Honiara	422	-9.43	159.95	0.09506	0.09345	0.59071	0.59852
11 Christmas	1110	1.98	202.52	0.08356	0.07316	0.64178	0.72313
13 Kanton	1284	-2.82	188.28	0.08141	0.07225	0.56191	0.65421
14 French Fr Shall	1240	23.87	193.72	0.10928	0.10439	0.49638	0.55560
15 Papeete	1127	-17.53	210.43	0.10763	0.09794	0.24069	0.33517
16 Rikitea	1270	-23.13	225.05	0.12381	0.10700	0.28240	0.37662
18 Suva	1139	-18.13	178.43	0.14025	0.12830	0.16695	0.17620
22 Easter	1173	-27.15	250.55	0.10771	0.10469	0.46656	0.45243
23 Rarotonga	1124	-21.20	200.22	0.12230	0.11835	0.33962	0.31817
24 Penrhyn	1224	-8.98	201.95	0.08692	0.08145	0.52863	0.61305
25 Funafuti	924	-8.53	179.22	0.14000	0.08079	0.15944	0.65713
28 Saipan	685	15.23	145.75	0.12250	0.09973	0.57889	0.73697
29 Kapingamarangi	745	1.10	154.78	0.08620	0.07602	0.56584	0.64389
30 Santa Cruz	1408	-0.75	269.68	0.09193	0.09329	0.54678	0.45292
38 Nuku'alofa	963	-21.13	184.80	0.13128	0.11227	0.29633	0.49624
47 Chichijima	728	27.10	142.18	0.16061	0.15671	0.46419	0.50204
51 Wake Island	1812	19.28	166.62	0.13718	0.12805	0.17461	0.20598
52 Johnston Island	995	16.75	190.48	0.13322	0.12687	0.48827	0.53853
53 Guam	984	13.43	144.65	0.10739	0.09942	0.64774	0.70741
55 Kwajalein	931	8.73	167.73	0.08141	0.07844	0.63673	0.64845
57 Honolulu	779	21.30	202.13	0.11212	0.10779	0.31377	0.26735
79 Chatham Island	93	-43.95	183.43	0.13043	0.07980	0.52196	0.83373
103 Port Louis	886	-20.15	57.50	0.11414	0.10896	0.50356	0.56183
104 Diego Garcia	772	-7.28	72.40	0.11411	0.08222	0.44185	0.66501
105 Rodrigues	1140	-19.67	63.42	0.11031	0.10972	0.54643	0.53380
108 Hulhule	627	4.18	73.53	0.09740	0.08932	0.34652	0.41308
109 Gan	1832	-0.68	73.15	0.10952	0.07870	0.37440	0.56395
117 Hanimaadhoo	968	6.77	73.17	0.10212	0.10362	0.44240	0.48721
121 Point La Rue	1030	-4.67	55.53	0.09754	0.08402	0.49799	0.58931
171 Cocos Island	1024	-12.12	96.90	0.12917	0.12028	0.53449	0.60137
180 Kerguelen	360	-49.35	70.22	0.12767	0.10593	0.46072	0.62905
245 San Juan	1675	18.47	293.88	0.08725	0.08820	0.56727	0.51396
259 Bermuda	600	32.37	295.30	0.12462	0.12437	0.48250	0.46453
290 Port Stanley	451	-51.70	302.15	0.11531	0.11075	0.59389	0.59063
291 Ascension	953	-7.90	345.62	0.08347	0.06127	0.41239	0.53758
292 St. Helena	781	-15.97	354.30	0.07724	0.05741	0.58976	0.67042
335 Spring Bay	277	-42.55	147.93	0.15674	0.13985	0.33575	0.50633
355 Naha	828	26.22	127.67	0.13860	0.13862	0.39188	0.42084
Mean Values				**0.11131**	**0.10025**	**0.46084**	**0.53599**

Table 7. RMS differences (in metres) and correlations between tide gauge data and SXOs, before and after orbit error corrections were applied. Tide gauges were a subset of the WOCE global set, and were specifically chosen as being representative of sea level change in the open ocean.

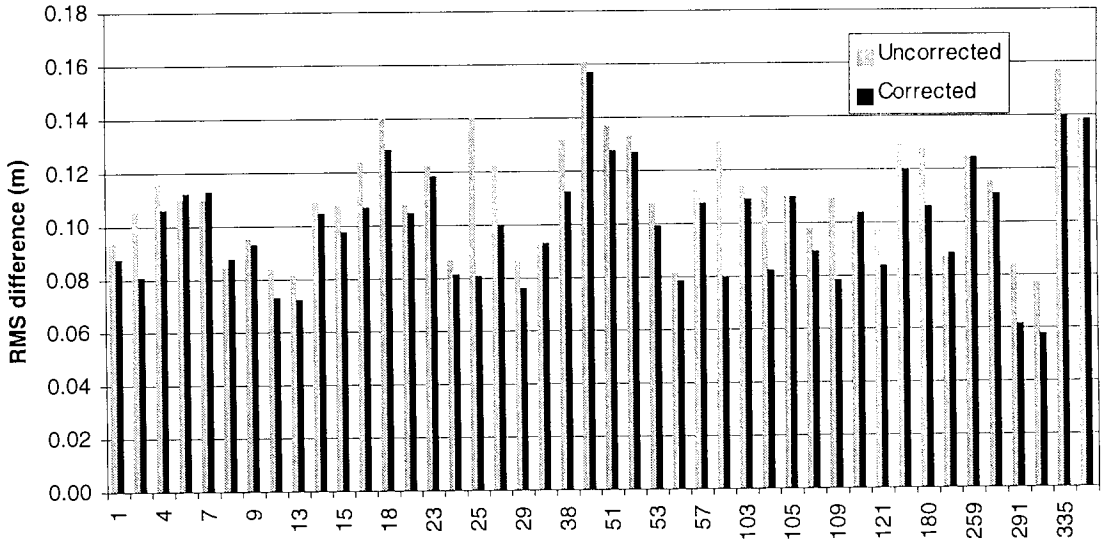

Figure 7. RMS difference (metres) between tide gauge data and ERS-1 SXOs before and after orbit correction (listed by gauge ID).

For each of the 44 tide gauges verified in section 4.3.1 all local crossover observations were compared with the difference between the tide gauge readings at the two epochs of the crossover. The difference between these two measurements was calculated. The mean and RMS differences for each tide gauge are shown in Table 7 for both the corrected and the uncorrected orbits. The RMS difference before and after correction is also plotted in Figure 7 and the correlations before and after correction are plotted in Figure 8.

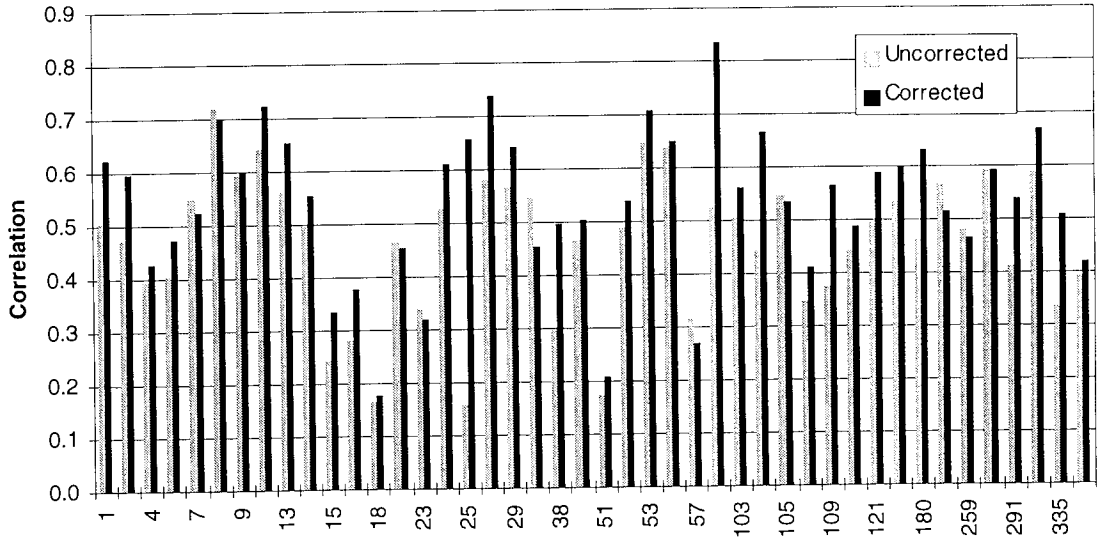

Figure 8. Correlation between tide gauge data and ERS-1 SXOs before and after orbit correction (listed by gauge ID).

If oceanographic signals are being absorbed by the DXO correction procedure, the RMS differences between the crossover data and the tide gauge data will increase, since theoretically the tide gauge observes such signals with total accuracy.

The data show that there are improvements in both the correlation with the tide gauges and the RMS differences, because of the DXO correction. The RMS difference between the tide gauge derived crossover and the satellite crossover drops from 11.13cm to 10.03cm, which represents an improvement of fit by around 5 cm, and the mean correlation increases from 0.46 to 0.54.

This DXO cubic-spline-based correction can be equally easily calculated for orbits which are derived from the Aston or other orbital heights, making it a simple and effective scheme.

4.4 Derivation Of ERS-1 Range Bias Drift

At this point an opportunity arose to derive an estimate of the relative bias drift between ERS-1 and TOPEX using the data set of ERS-1–TOPEX dual crossovers. The TOPEX bias drift is known accurately from other sources and it is therefore possible to isolate the ERS-1 bias drift to a good level of accuracy. Note that no similar comparisons are made for the Poseidon altimeter due to its infrequent operation.

The cubic spline fitting procedure is carried out as described previously in Section 4.2, and the overall relative bias is noted between ERS-1 and TOPEX. The cubic spline then represents the difference from this mean value at any particular time. By averaging these differences over some short period of time, it is possible to obtain an instantaneous estimate of the relative bias.

In practice it is necessary to choose a suitable time period over which to average the data. A period which does not represent an integer number of orbital cycles will result in contamination of the data with some once-per-revolution orbit modelling errors.

Therefore an interval of 1435.8 minutes was chosen, being 14 full orbits (a single rotation of ERS-1 about the Earth takes approximately 100.58 minutes) and almost a single day in length (one day is 1440 minutes).

Within each interval, a certain number of evaluations of the relative bias (i.e. evaluations of the cubic spline) were carried out. These were then summed and a mean and standard deviation was calculated in the normal way. In this case the spline was evaluated once every 300 seconds, giving 281 values from which to calculate statistics. Since each cubic spline has had the overall mean relative bias (Table 8) removed, this must be added back on to the daily mean values before further processing. At this stage, any values with standard deviation greater than 20 cm were rejected as erroneous, along with a further two points which were observed to be problematic due to satellite manoeuvres taking place near to those epochs.

Phase	C - Multi-One	D - Ice Two	E+F - Geodetic	G - Multi-Two
Dates (MJD)	48900-49353	49353-49454	49454-49790	49790-50000
Overall Relative Bias	0.41907	0.42437	0.42679	0.42730

Table 8. Overall relative biases per ERS-1 phase between ERS-1 and TOPEX derived From DXOs.

918 values of approximately daily relative range biases were yield from this data set. Upon performing a linear regression on these data, the relative bias drift is found to be -2.871 ± 0.682 mm/year (Figure 9). *Moore et al* (1999a) found an ERS-1 drift of -2.6 ± 0.4 mm/year after applying dual crossover orbit corrections to ERS-1 crossover data (thereby aliasing the TOPEX bias drift into the ERS-1 data), suggesting that there has been little change in the ERS-1 range bias.

Equivalent studies using longer time intervals were carried out, which resulted in increases in the variances of the estimated drift, but the actual estimate did not vary significantly from the above quoted -2.87 mm/year value. Specifically, two and five day intervals yielded relative drifts of -2.839 ± 1.07 and -2.917 ± 1.10 mm/year.

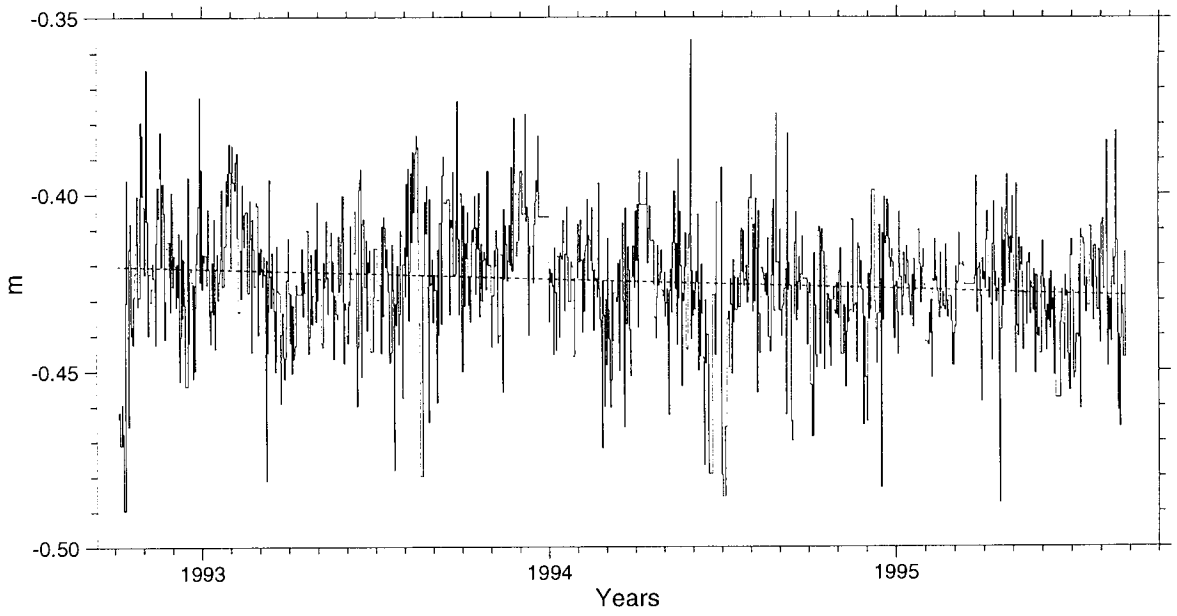

Figure 9. Relative Bias in metres between ERS-1 and T/P. Daily values from Cubic Spline technique.

The reported value of -2.87 mm/year is similar to reported values of the TOPEX bias drift. Current studies of TOPEX bias drift for this period have revealed drifts of -3.25 ± 0.53 mm/year (*Moore at al,* 1999a) and -2.3 ± 1.2 mm/year (*Nerem et al.* 1997a) from comparisons between TOPEX derived sea heights and tide gauge data. Much of this is thought to be due to a drift in the microwave radiometer (MR), which is used to calculate corrections incorporated into altimeter readings. *Kubitschek et al.* (1997) found a MR drift of -1.9 ± 1.2 mm/year by comparisons with an upward-looking water vapour radiometer and *Haines and Bar-Sever* (1998) found a drift of -1.2 ± 0.4 mm/year using terrestrial GPS receivers.

4.5 Conclusions.

The application of a DXO correction to ERS-1 orbits was shown to improve radial orbit error by around 5cm RMS. The fit to tide gauge data was also shown to improve upon applying this correction. The full impact of the cubic-spline DXO correction technique detailed in this chapter is explored further in Chapter 5.

The analysis in section 4.3 enabled the recovery of a relative bias drift between ERS-1 and TOPEX. This drift was found to be -2.87 ± 0.68 mm/year of which between -2 and -3 mm/year is due to TOPEX bias drift, suggesting that ERS-1 suffers from little or no such drift.

5. Analysis of the errors present in the ERS-1 orbits.

An analysis procedure was carried out using ERS-1 – T/P DXO data in an effort to better understand and quantify the errors present in the uncorrected ERS-1 orbits (using both the DUT JGM3 and DGM4 orbital heights). This analysis was repeated after applying the DXO correction detailed in Chapter 4 and the results compared, giving an indication of the magnitude of improvement possible through the use of this technique.

This analysis is made possible by assuming that any height difference in a short-time-difference DXO is entirely due to ERS-1 orbit error (as discussed in Section 4.1). Though this assumption should not be used carelessly, the short time difference applied to these DXOs removes the impact of a long-term effect such as the bias drift observed in Section 4.4.

5.1 Data Processing

The DXO data were split into blocks of time. Each block was sized in order to achieve as close as possible a full global coverage of data. During the first and second Multi-Disciplinary phases, these blocks are 35 days long and coincide with the normal cycles. The Ice phase is split into 36-day blocks with a single 34-day block. The Geodetic phase is split into nine 37-day blocks (to coincide with the 37-day sub-cycle which is a property of this mission) and a final 49-day block. The number of dual crossovers in each block is shown in Table 9.

An initial bias of 40cm was applied to the observations to remove most of the ERS-1–T/P relative bias and a 30cm threshold rejection criterion was applied. For the purposes of this study the ERS-1 orbit error (Δr) is modelled as:

$$\Delta r = A_T + A_P + A_3 \cos\phi \cos\lambda + A_4 \cos\phi \sin\lambda + A_5 \sin\phi + A_6 r_\tau + A_7 (\cos^2\phi - 0.5) \qquad 17.$$

where ϕ is latitude and λ is longitude. A_T is the relative ERS-1–TOPEX altimeter bias, and A_P the relative ERS-1 – Poseidon altimeter bias. A_3, A_4 and A_5 are the coefficients of the first order spatial errors, A_6 is the altimeter time tag error and A_7 the coefficient of a latitude-symmetric term. Note that for some cycles no Poseidon data exist, and the relative bias cannot be solved for.

Block #	Start MJD	Length	Mid MJD	Date in Years	# DXOs
1	48901	35	48918.5	1992.8101092	45580
2	48936	35	48953.5	1992.9057377	45305
3	48971	35	48988.5	1993.0013698	50464
4	49006	35	49023.5	1993.0972602	58445
5	49041	35	49058.5	1993.1931506	58076
6	49076	35	49093.5	1993.2890410	53037
7	49111	35	49128.5	1993.3849315	50133
8	49146	35	49163.5	1993.4808219	44440
9	49181	35	49198.5	1993.5767123	37575
10	49216	35	49233.5	1993.6726027	40651
11	49251	35	49268.5	1993.7684931	43024
12	49286	57	49316.0	1993.8986301	73141
13	49346	36	49364.0	1994.0301369	46750
14	49382	36	49400.0	1994.1287671	59838
15	49418	34	49435.0	1994.2246575	50410
16	49452	37	49470.5	1994.3219178	51289
17	49489	37	49507.5	1994.4232876	50627
18	49526	37	49544.5	1994.5246575	45146
19	49563	37	49581.5	1994.6260273	44577
20	49600	37	49618.5	1994.7273972	44177
21	49637	37	49655.5	1994.8287671	44933
22	49674	37	49692.5	1994.9301369	50829
23	49711	37	49729.5	1995.0315068	58445
24	49748	49	49774.0	1995.1534246	66700
25	49800	35	49817.5	1995.2726027	54116
26	49835	35	49852.5	1995.3684931	47935
27	49870	35	49887.5	1995.4643835	43381
28	49905	52	49931.0	1995.5835616	63438

Table 9. DXO data used for ERS-1 orbit accuracy analysis.

A_3, A_4 and A_5 represent spatial errors caused by the first-order geopotential coefficients or by the displacement of the geocentre from the reference centre used by the tracking network. A_6 is the time tag bias (the difference between the recorded time at which an altimetric observation is made and the actual or effective time) which affects crossovers as an anti-correlated term with radial error (*Wagner and Klocknick*, 1994) given by

$$r_\tau = - n R_E \sin^2 I (f + C_{2,0} R_E / 2a) \sin 2u \qquad 18.$$

where n, a, I are the standard Keplerian elements, R_E the Earth's equatorial radius, f the flattening coefficient, u the argument of latitude and $C_{2,0}$ the unnormalised geopotential

second-degree harmonic. A_7 is an empirical term that was required to absorb discrepancies in the representation of the altimetric corrections such as the ionosphere, sea-state bias and so on, which tend to have a simple geographical distribution. It is also intended to absorb any small modifications to the corrections such as the change in dry tropospheric field correction source in the ERS-1 version 5 altimetry. The terms in Equation 17 are solved for using a standard least-squares technique.

5.2 Results

5.2.1 DUT JGM3 orbits

The results obtained after solving for the terms in Equation 17 are plotted in Figure 10. The results show that the relative bias between ERS-1 and TOPEX or Poseidon respectively are mostly constant throughout the time span in this study, although Figure 10a shows that the TOPEX data are less noisy than the Poseidon data. This is mainly due to the brief periods of operation of the Poseidon altimeter. Apart from the very early section, the TOPEX time series is consistently decreasing, representing a consistent drift in either the ERS-1 or the TOPEX bias. The first four data points appear to be noisy in all parts of these data.

The A_3 term in Figure 10b is consistently substantially negative (greater than 6cm) which shows that the JGM3 gravity model can be problematic when used for ERS satellite applications. The JGM3 model was generated by incorporating T/P data and data from the French SPOT-2 satellite (*Tapley et al*, 1996). No ERS data were used and therefore any gravity field terms which are more significant to the geometry of the ERS orbits than for T/P or SPOT, will be poorly resolved. A_3 represents an effective shift in the co-ordinates of the reference frame ($a \cos \phi \cos \lambda$ term is most positive at $[\phi,\lambda] = [0,0]$ and most negative at $[0,\pi]$). The A_4 and A_5 terms represent similar shifts on the other two of the three axes, which are less pronounced, and any coherent signal is difficult to separate from the noise in these data.

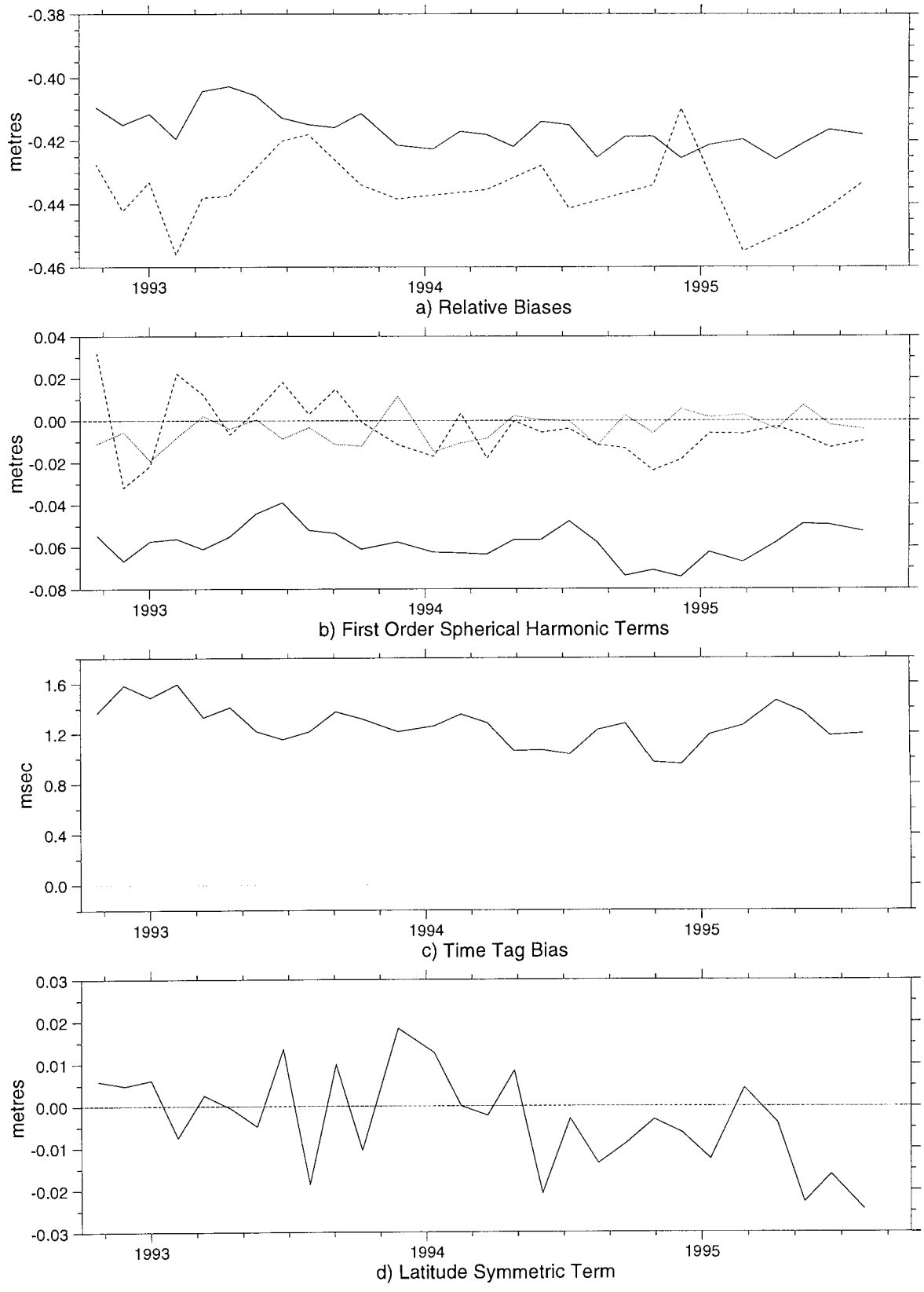

Figure 10. Results of DXO based orbit analysis procedure for DUT JGM3 orbits. **a)** Relative biases ERS-1 – TOPEX (solid line) and ERS-1 – Poseidon (dashed). **b)** First order geopotential errors A_3 (solid) A_4 (dotted) and A_5 (dashed). **c)** ERS-1 time tag bias. **d)** Latitude symmetric term A_7.

Figure 10c shows time tag biases exactly as expected for ERS-1 altimetry data - consistently around 1-1.5 ms. A slight downward trend is evident, which may indicate some systematic change, though the value rises again towards the end of the time series.

Figure 10d shows a possible jump at the start of 1994 with a zero mean previously and a mean of around -1.0 cm after this point. This date coincides with the start of the Geodetic mission. The last few values are noisy, but seem to be low after the introduction of the version 5 altimetry at the start of the second Multi-Disciplinary phase.

5.2.2 DUT DGM4 orbits

To provide an independent orbit source, the same process was carried out using the DUT DGM4 orbits and results compared. These plots differed by less than 1cm from the JGM3 data set, with the exception of the A_3, A_4 and A_5 terms, for which the *difference* between the two results is plotted in Figure 11.

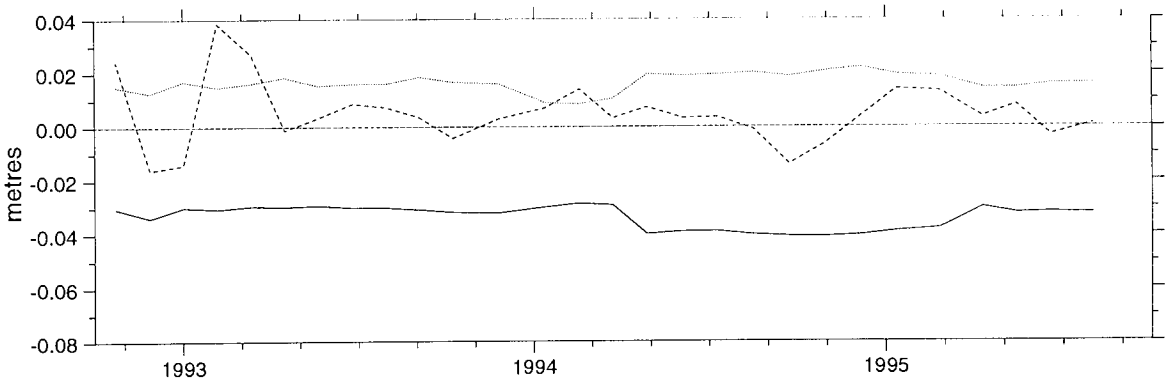

Figure 11. *Change* in first order spatial terms (A_3 {solid} A_4 {dotted} and A_5 {dashed}) from ERS-1 – T/P DXOs when using DUT DGM4 orbits in place of JGM3.

From Figure 11, the A_3 term appears to have increased consistently by around 3cm with a group of observations having a jump of an additional centimetre which coincides with the Geodetic phase. This suggests that the DGM4 orbits provide better fits to one or other of the Multi-Disciplinary phases or the Geodetic phase. It is also possible to see a slight drop (about 0.25cm) during the Ice phase. These trends seem to be repeated in reverse for the A_4 term which appears to be reduced by approximately 1.5 cm during

the multi-disciplinary phases, less during the ice phase (approx. 1.0 cm) and by an average of 2 cm during the geodetic phase. A_5 appears to suffer from noise problems at many stages in the data. As has been seen in the first four data points in the other plots, strong variations were observed for this parameter, but in this case variations were observed throughout the time period i.e. there is no systematic element.

These results show that the A_3, A_4 and A_5 terms may represent some aspects of ERS-1 gravity-modelling-based errors. The fact that trends observed in the other terms remain mostly unchanged suggests that these trends are likely to be due to some other cause (e.g. systematic changes to geophysical corrections).

5.2.3 Cubic Spline Corrected JGM3 Orbits

As a further means of assessing the effectiveness of the cubic-spline DXO correction procedure, the analysis detailed in Section 5.1 was also carried out using the same DUT JGM3-based DXO data set as in section 5.2.1, with the addition of the cubic-spline correction as described in Chapter 4.

The results (plotted in Figure 12) give strong confirmation that the correction procedure removes the major gravity modelling errors from the orbit. The TOPEX bias (A_T) is noticeably smoother (all comparisons are made with the time series shown in Figure 10), but exhibits a linear trend of approx. -0.5 cm/year (Figure 12a). As before, the Poseidon bias (A_P) is noisy, with a particularly high value just before the end of 1994, but its variability is reduced. All three spatial terms (A_3, A_4 and A_5) are reduced to insignificance (Figure 12b), being consistently less than 1cm, with only two epochs in the A_5 time series having value in excess of $\pm$0.5cm. Time tag biases (Figure 12c) are reduced to almost zero. The A_7 term is reduced in amplitude, and its mean is shifted to slightly above zero (Figure 12d). The cubic-spline correction procedure may not be as powerful as the use of an improved gravity field or better geophysical corrections for removing the effects of the A_7 parameter since DXOs are concentrated at latitudinal extremes, leading to poor resolution of signals near the equator, and the A_7 term is symmetrical about the equator.

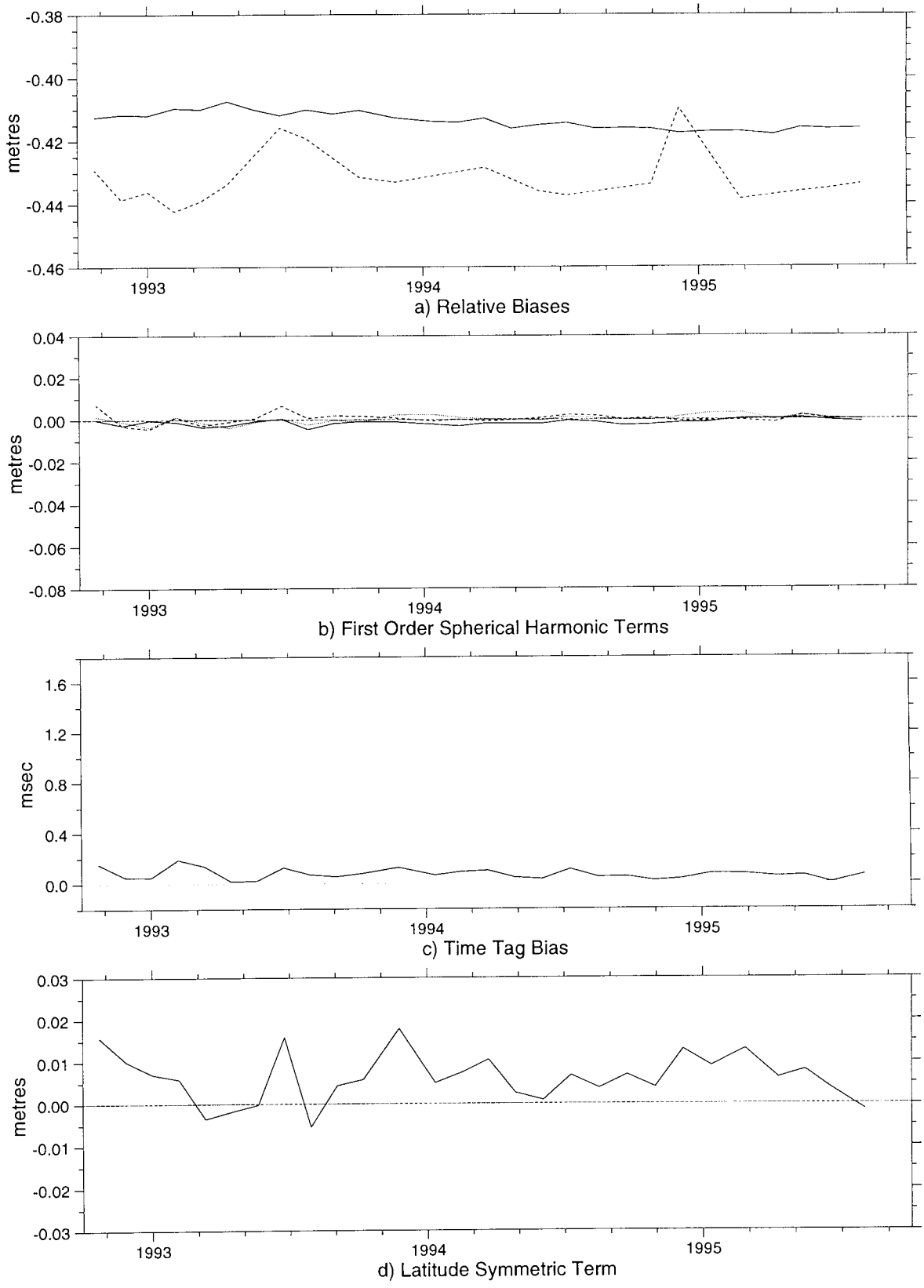

Figure 12. ERS-1 orbit error from the DUT JGM3 orbits after DXO correction. **a)** Relative biases ERS-1 – TOPEX (solid line) and ERS-1 – Poseidon (dashed). **b)** First order geopotential errors A_3 (solid) A_4 (dotted) and A_5 (dashed). **c)** ERS-1 time tag bias. **d)** Latitude symmetric term A_7.

This information gives further evidence that the cubic-spline orbit correction procedure is capable of removing many of the problems inherent to the ERS-1 orbits. The major gravity modelling errors are reduced, and there appear to be no systematic problems introduced by the use of this procedure, in that none of the solved parameters worsen.

5.3 Conclusions.

ERS-1 orbits produced using the JGM3 gravity model reveal the weaknesses of this model for this satellite. Strong first-order harmonic terms are not accurately resolved in the JGM3 model when applied to the ERS orbital configuration, as expected, since the JGM3 model included no data from an ERS-type satellite in its original make-up.

These data also show that it is worthwhile refining this gravity field model by the inclusion of data from the ERS missions. DUT have shown improvement in the low order gravity terms under this type of analysis by the generation of the DGM4 gravity model. Other researchers are reporting good results in carrying out similar work: *Boomkamp and Moore* (1997) have reports of orbit error reductions for ERS-1 from 7.61 cm RMS to 5.52 cm RMS for SLR orbits, along with reductions in SXO residuals from 10.18 cm RMS to 8.88 cm RMS upon the use of a refined gravity model.

6. World Ocean Circulation.

This chapter is a brief summary of the dominant circulation patterns in the world's oceans. This information can be found in many texts on introductory oceanography and sources such as *Pond and Pickard* (1983), *Pickard and Emery* (1990) and *Tomczak and Godfrey* (1994) are typical examples. Figure 13 shows a schematic of circulation patterns on a global basis.

Figure 13. Schematic map of climatological average surface ocean currents. From *Summerhayes and Thorpe* (1996).

6.1 Overall Circulation

The surface ocean circulation is mostly wind driven. Alterations from the wind flow patterns are due to physical interruption to this flow, either as a result of land masses or bottom topography.

Due to the Earth's rotation and the temperature gradients formed by the Sun's heating of the land and oceans, the wind patterns either side of the equator are split into three latitude bands. Nearest to the equator are the strong easterly trade winds (blowing from east to west). A high pressure band at about 30° north (or south) separates these

from the westerlies. These continue north (or south) until a band of low pressure at about 65° north (or south) separates these from the polar easterlies. Ocean circulation broadly follows these patterns with change forced by the presence of the continents.

6.2 Antarctic (or Southern) Ocean.

The term 'Southern Ocean' is generally considered to refer to the entire ocean south of the tip of South Africa.

This Ocean is unique in that there is an almost completely unhindered path for circulation around the entire globe, so that surface water circulation in the Antarctic Ocean closely resembles the atmospheric circulation. At no other latitude can this occur. The Antarctic Circumpolar Current (ACC) dominates variability in this ocean. Most variability is a result of shifts of the ACC, production of eddies, or changes in the overall flow rate. Since the Southern Ocean has total connectivity with other oceans, it has the possibility of unlimited water mass transport between itself and these other oceans, and this makes it important to understand the activity in this area before attempting to gain any knowledge of the other oceans. The ACC has many effects on the southern areas of the other oceans - for example the propagation of eddies northwards from the tip of South America and its contribution to the strength of the Peru current.

The ACC enjoys almost complete freedom to circulate about the Antarctic continent, with the exception of a few specific locations where its flow is restricted by local physical boundaries. Drake's passage is the most obvious example. Here the gap between land masses is less than 850km. Combined with the effects of the Scotia Ridge 2000km to the east of the passage this produces a strong effect - the current accelerates to go through the passage and then becomes highly turbulent upon hitting the ridge and swings sharply northward. The Scotia Ridge is the joint between the tip of the African continent and the Antarctic continent. It is not complete, but does block most flow below about 3000 metres. Other ridges between the major ocean basins have similar effects on the ACC. The Kerguelen plateau is about 70° east and between about 60° south and 48° south. This plateau carries several islands and blocks most ACC flow below

about 2000m. There is, however, a small gap between this plateau and the continental shelf at the edge of the Antarctic. This permits the passage of very deep water, and the net result is no major change in flow direction of the ACC over this area. The final major flow restriction is provided by a combination of the Macquarie Ridge, the Pacific-Antarctic Ridge and the South-East Indian Ridge which spread around the south of Australia. These three ridges combine to make a significant alteration to the path of the ACC in this region. Due to Coriolis effects and the complex topography, this results in a northward shift of the ACC.

6.3 Pacific Ocean.

The basic flow pattern in the Pacific consists of two large circulations (the sub-tropical gyres), one north of the equator and one south of the equator. In keeping with the atmospheric rotation in these areas, the flow in both hemispheres is westward at the equator and eastward at higher latitudes. There also exists a smaller sub-polar gyre in the extreme north of the Pacific. This is a result of the wind direction changing at the boundary between the westerly and the polar easterly winds. In this ocean an equatorial countercurrent occurs at approximately 5° north of the equator, flowing against the weak westerly winds present in this region.

The northern limb of the southern sub-tropical gyre is the westward flowing South Equatorial current. The bulk of its volume transport is transferred to the East Australian current, which runs south along the east coast of Australia. It then turns east just north of New Zealand to form the South Pacific Current, formerly known as the Sub Tropical Front (STF). This loops back on the eastern side of the basin and returns to the South Equatorial current.

The northern hemisphere between 8-10° and 35° N repeats this general pattern. The northern sub-tropical gyre consists of the westward flowing North Equatorial current, which turns north near the Philippines to form the Philippines Current. This then becomes the Kuroshio current off Japan and is deflected eastwards upon meeting the southward component of the sub-polar gyre (the Oyashio current). This then follows

what is known as the Polar Front eastwards to form the North Pacific Current and splits into two components. One heads south in the California Current to complete the northern sub tropical gyre, and the other heads north in the Alaskan Current to form the start of the sub-polar gyre. This then heads westwards as the Alaskan Stream to be diverted south by the continental shelf to form the Oyashio current, thereby completing the flow pattern in the north Pacific ocean.

Between December and March, the region is affected by the north-west monsoon, which has effects on some of the minor features of the flow patterns.

6.4 Atlantic Ocean.

The Atlantic Ocean has a North/South basin split giving a pair of gyres, as with the Pacific. The fundamental differences between the Atlantic and the Pacific are mainly a result of the geographical layout of the Atlantic basin.

A major influence is the Mid-Atlantic Ridge, which runs down almost the entire ocean. It rises to within 1000m of the surface in some places, and is rarely deeper than 2000m. The latitudinal extent of the ocean basin is some 22000 km, whilst at its narrowest point it is only 8300 km across. This severely affects the standard flow patterns which would otherwise be evident in such an area of the globe.

Sub-tropical gyres exist in both hemispheres in the Atlantic Ocean. The North Atlantic gyre is intensified on the western boundary of the North American Basin, with the strong Gulf Stream forming the western boundary current of this gyre. This current meanders extensively along this area, and runs along the edge of the continental shelf. It carries a great deal of warm Caribbean water north towards the western tip of Europe, significantly influencing the weather conditions in the UK, turning what would otherwise be a cold area of the world at 50° North into a temperate zone.

At the Eastern side of the North Atlantic the remainder of the Gulf Stream flows back south via the Portugal Current and the Canary Current (named due to their respective nearest land masses).

The southern Atlantic exhibits a sub-tropical gyre between about 20° and 45° south which rotates anti-clockwise. The western boundary current (the Brazil Current) is intensified as for the north Atlantic gyre. This current meets the Falkland current at about 40° south and both are forced to turn eastwards. The Falkland current is a branch from the eastward motion of the ACC as it is deflected around the southern tip of South America.

The eastern parts of this basin have relatively little current activity. The sub-tropical gyre turns northwards as the more leisurely Benguela Current and returns to the west of the ocean as part of the south equatorial current. The far southern Atlantic (greater than 40° south) is dominated by the ACC.

6.5 Indian Ocean.

The Indian Ocean has complex variability patterns, partly because its northern boundary is a large land mass near the equator. This area is unique in that there is a seasonal switch between two almost unrelated ocean circulation patterns, so much so that it is not meaningful to define a single 'mean' ocean circulation. Instead two such means are defined, one for the Southwest Monsoon pattern which peaks in August and another for the Northeast Monsoon pattern which peaks in February.

During the Northeast or Winter monsoon (peaking in February) the pattern resembles the Pacific circulation pattern with the exception that the equatorial counter current occurs below the equator (at about 5° S) rather than above as in the Pacific due to differing land balances. Wind stress essentially comes from the north east and forces a generally south westward surface flow pattern. This forms a gyre in the north of the ocean, an equatorial counter current and a mirror-image circulation to the South. During

this period, the region's weather is dominated by the dry air coming from the land: this is the dry season.

Conversely in the Southwest or Summer monsoon the wind changes direction and the strong Somali current is set up along the western boundary of the Ocean. This invokes a large circulation in the whole Indian Ocean with the south equatorial current flowing westward and the Southwest Monsoon current flowing eastwards across the Southern tip of the Indian sub-continent. The winds driving these currents carry a great deal of moisture, and as this air hits the Indian sub-continent and rises, the result is the characteristic heavy monsoon rains.

7. Construction Of A Sea Surface Anomaly Time Series From SXOs.

Sea surface anomalies cannot be observed directly from crossover data. A crossover represents the difference in sea surface height at two epochs, and therefore it is less straightforward to obtain a time series from a crossover data set than from a set of along-track altimetry (for example). However, as described in Section 4.1, along-track altimetry contains several error terms which cancel out upon taking the crossover difference making crossovers a more accurate data set, and crossover data are flexible enough to work with even across the boundaries of mission phases. Thus they have the potential to form a very useful data set.

In this chapter a method of obtaining a time series of sea surface heights from a database of SXOs is shown in detail.

7.1 Single Satellite Crossovers

The crossover difference is defined as the difference in observed sea surface height at this point at epochs t_1 and t_2 (the epochs of the two arcs at the crossover point). The crossover residual (ΔSXO) is formulated as:

$$\Delta SXO = \Delta h_{t1} - \Delta h_{t2}$$

where each sea-surface height residual Δh is given by $\Delta h = h_{obs} - h_{orb}$ where h_{obs} is the corrected observed altimeter measurement, and h_{orb} is the orbital height of the satellite above the reference ellipsoid. The crossover can therefore be written as:

$$\Delta SXO = (h_{obs1} - h_{obs2}) - (h_{orb1} - h_{orb2}) \text{ or}$$
$$\Delta SXO = \Delta h_{obs} - \Delta h_{orb}$$

i.e. it is possible to calculate the observation difference and the orbit height difference separately. With this form of the residual any orbital height (e.g. DUT or Aston) can be easily applied.

Heights are calculated from the off-line Ocean Product (OPR) Geophysical Data Records (GDRs) which were edited to remove land or ice data. Shallow water observations are retained in order to provide as much information as possible at the boundaries of the data set, even though tidal models are notoriously poor in shallow water and coastal areas. Further details on altimetry processing is presented in Chapter 2.

7.2 Generation Of The ERS-1 SXOs.

A database of ERS-1 SXO residuals was generated for phases C, D, E, F and G of the ERS-1 mission (see Introduction, Table 1 for mission phase details).

The orbits for each phase are stored as a number of ephemeris files. The first Multi-Disciplinary phase (C) was stored as 145 overlapped 5-day arcs and Phase D as 22 overlapped 5-day arcs. The geodetic phases (E and F) consisted of 70 overlapped 6-day arcs and phase G 39 overlapped 5-day arcs. A short period only from the second multi-disciplinary phase (G) was used due to fundamental processing differences between the altimetry data for the early part of phase G and that released to cover the latter part. A version change from 5 and 6 means it is potentially difficult to combine the two, whereas versions 3 and 5 (which make up the remainder of the ERS-1 mission) are internally consistent and can be used together with few reservations.

An arc is the path of the satellite, which is calculated using the Satan-A suite of software (Chapter 3). An ephemeris file contains the position vectors of the satellite at 30 second intervals in an Earth fixed reference frame, allowing the calculation of an exact position at any epoch by interpolation. The ephemerides are used to calculate the exact crossover locations as shown in *Rothwell* (1989). The residuals are then calculated by reading the altimetry data and interpolating to the exact location from the nearest data points.

In order to shorten the computation time for generating such a large data set, the task was split into sections using the convenient phase boundaries. In an effort to further reduce the amount of data produced, and to decrease the likelihood of bad arcs contaminating the data set as a side effect, not all arcs in each phase were used to generate residuals. The first arcs to be deleted were those near to orbital manoeuvres or that had sparse data problems. Any further reduction was carried out by selecting a proportion of remaining arcs or by removing alternate arcs. Details of the number of arcs used from each phase, and the number of crossovers produced is given in Table 10. Note that for simplicity, no distinction is made between the first and second Geodetic phases and both are listed as phase 'E'.

Phase One			Phase Two			Crossovers	RMS residual
#	Total Arcs	Used	#	Total Arcs	Used	Produced	(cm)
C	145	54	C	145	54	1244869	18.6
C	145	7*	D	22	22	238314	17.3
C	145	39	E	70	29	1378459	20.4
C	145	7*	G	39	39	319945	17.6
D	22	22	D	22	22	302991	13.7
D	22	22	E	70	29	851435	15.7
E	70	70	E	70	70	3249845	15.5
E	70	28	G	39	29	903676	15.4
G	39	39	G	39	39	673715	13.8
					Total:	9163249	

Table 10. Number of arcs used in creating the crossover database for each pair of phases in the ERS-1 mission. RMS values calculated after a 3σ rejection criterion has been applied and using Aston orbit heights. * - Only cycle 9 of the Multi-One phase was used in these cases.

The SXO residuals were calculated and the radial orbital heights (which come from the ephemeris file itself and were therefore uncorrected Aston heights) were replaced with those from the DUT JGM3 orbit files (*Scharroo*, 1999). The JGM3 data set was chosen because this gravity model is commonly used within the Earth observation scientific community and was the gravity model used by NASA while generating the T/P orbits. Ideally, since the aim of this project was to compare and combine T/P and ERS-1 data, the generation of both data sets should be as consistent as possible.

Though the stored SXO residuals include uncorrected orbits from the DUT JGM3 package, it is a trivial matter to replace these with orbital heights from another

source or add the cubic spline DXO correction detailed in Chapter 3 during the reading of the SXOs from the database. Hence multiple data sets using different orbit types can be generated from the same basic set of SXO residuals.

7.3 Processing of the SXOs.

The crossover data set was gridded into geographical bins. The grid size chosen was a compromise between spatial density (i.e. 1° x 1° or better) and temporal density, since with a larger grid size, more crossover points are included per bin, and hence more data points can be generated for each time series. Using a 1° x 1° grid resulted in less than one observation every 13.4 days on average in each bin and many bins with no data at all due to the orbit configurations in use, which is unacceptable. It is possible to use even a 1° x 1° grid through the use of a spatial interpolation technique but in the present study this would have placed an unnecessary load on computational resources and offered little benefit, so a coarser initial grid was chosen.

Where TOPEX/Poseidon data are to be included during later processing, the grid spacing is 2.83464 (or 360 / 127)° longitude by 2° latitude to coincide with the T/P longitudinal ground track spacing (there are 127 orbits per T/P 9.9-day cycle). Some runs were also completed without including T/P data and these were gridded to various dimensions, including 2.835° x 2° as above, 2° x 2° and 1.5° x 2°.

During the gridding process, any obviously erroneous crossover residuals were rejected by use of a 1 metre threshold, followed by an iterative 3σ rejection system: i.e. σ is calculated, values in excess of 3σ rejected and σ is calculated again. This procedure was repeated until no further points were rejected, and in practice the system converged in 3 or 4 iterations. RMSs for each section are shown in Table 10, along with the number of accepted crossovers.

The time series of sea surface variability is extracted from the crossover differences by a least-squares procedure. Each crossover is the sea surface height difference between that observed during an ascending and a descending arc (where an arc in this

case is the satellite's path between its southern and northern extreme). If there are M ascending arcs and N descending arcs in the current bin, each residual can be written as

$$r_{mn} = a_m - d_n$$

where a_m is the mth ascending arc height and d_n is the nth descending arc height with $1 \leq n \leq N$ and $1 \leq m \leq M$. It is then possible to solve for each a_m and d_n (since the r_{mn} are known) by minimising the sum:

$$Z = \sum_{m=1}^{M} \sum_{n=1}^{N} (a_m - d_n - r_{mn})^2 \quad .$$

An ideal data set (i.e. a complete set of crossovers between every one of M ascending arcs and N descending arcs) would yield the following set of equations (*Fu and Chelton,* 1985)

$$N a_m - \sum_{j=1}^{N} d_j = \sum_{j=1}^{N} r_{mj} \quad (m = 1, \ldots, M) \qquad 19.$$

$$\sum_{i=1}^{M} a_i - M d_n = \sum_{i=1}^{M} r_{in} \quad (n = 1, \ldots, N) \qquad 20.$$

These equations can then be written in matrix form:

$$\mathbf{Z\,l = R}$$

where $\mathbf{Z} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$ in which

A is a MxM matrix whose elements are equal to N, **B** an NxM matrix with all elements set to -1, **C** an MxN unit matrix and **D** an NxN matrix with all elements set to $-M$. **l** is a column vector with elements given by:

$$\mathbf{l} = \begin{bmatrix} a_1 \\ . \\ a_M \\ d_1 \\ . \\ d_N \end{bmatrix}$$

and R is a column vector with each element equal to the sum of all residuals which involve the i^{th} arc, either ascending or descending.

Under real conditions, the altimetry for some arcs may be missing for certain periods of time. This results in a more complex system, where not all crossovers are present for each arc. Therefore the construction of the matrix must be carried out in a slightly different manner.

First, it is necessary to determine how many ascending and descending arcs are represented by the crossovers in a particular bin (remembering that a single ascending arc can cross several descending arcs in the same bin and vice versa, resulting in several crossover residuals). This was done by calculating the time it takes for the satellite to pass over the entire latitudinal extent of the bin and assuming that any epoch further away temporally than this from the current arc must represent another arc. As each new arc is numbered, each crossover can be labelled with its ascending and descending arc numbers (n and m). The crossover can then be inserted into the same matrix representation as before:

$$\text{i.e. } \mathbf{Zl} = \mathbf{R} \qquad 21.$$

by adding entries in the nth and mth rows and columns.

Note that there is a degree of freedom inherent in this system. There are $M + N$ unknown sea height values and $M + N$ equations, but the sum of Equations 19 is equal to the sum of Equations 20 and thus a singularity occurs on attempting to invert the re-

sulting matrix. In order to better condition the matrix a constraint must be applied. The method used by *Fu and Chelton* (1985) was to set $a_1 = 0$ but this technique relies heavily upon the accuracy of arc a_1 so the method chosen here was to set the mean of the arcs to zero i.e. to apply:

$$\sum_i a_i + \sum_j d_j = 0. \qquad 22.$$

The arc with most crossover values is found. Treating this as a 'reliable' sea level, the equation representing it in the matrix is eliminated by subtracting it from all the other equations and replacing it with the representation of the constraint (Equation 22). The resulting set of simultaneous equations is solved with a standard least squares technique: the inverse of matrix **Z** is calculated and then pre-multiplied through Equation 21 to produce the equation

$$[\mathbf{Z}^{-1}\,\mathbf{Z}]\,\mathbf{l} = \mathbf{Z}^{-1}\,\mathbf{R} \qquad \text{i.e.}$$

$$\mathbf{l} = \mathbf{Z}^{-1}\,\mathbf{R}. \qquad 23.$$

Thus the values of a_n and d_m which make up the column vector **l** can be directly calculated from Equation 23.

The system used caused some problems. On some occasions, it was noted that the mean of the output arcs was substantially non zero (rounding error causing means approaching $1*10^{-6}$ was not uncommon, however). This indicated a badly conditioned system, and such bins were rejected. This phenomenon mainly occurred near coastlines, and is probably a result of poor altimetry data contaminating the result. Tidal models are not as reliable in shallow water areas on the continental shelf as in deeper waters and may cause this effect. Alternatively, data over land may be erroneously flagged as being over shallow water. The altimeter's footprint radius is in excess of 7km and if any land appears within this, then it can produce spurious values.

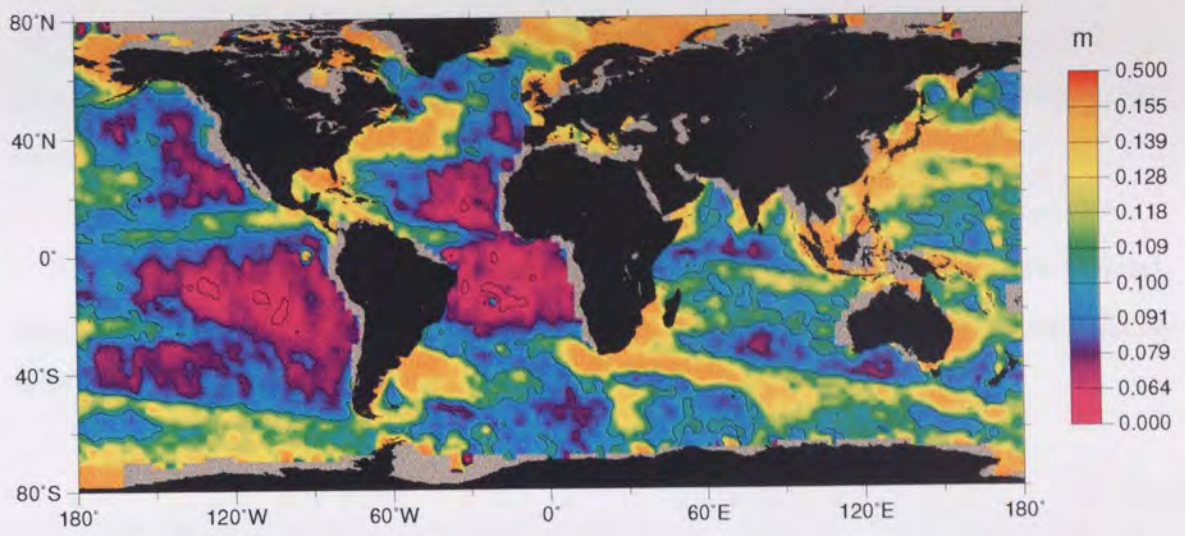

Figure 14. RMS sea level variability from the single crossover derived time series. Contours at 0.02 m intervals, plot is histogram-equalised for colour balance.

Determination of the sea surface heights for each bin produced a noisy time series at irregularly spaced intervals (i.e. at the epochs when the satellite passed over the bin). This information (even without further processing) clearly reflects the broad variability patterns present in the ocean. Figure 14 shows the RMS of the arcs produced by this procedure using the standard DUT JGM3 orbits which are corrected via DXOs with T/P transferred onto a world map. The original 2.83° by 2° gridded information is smoothed and re-sampled onto a finer grid by the GMT (Generic Mapping Tools) visualisation package (*Wessel and Smith*, 1991). The figure clearly shows the low RMS (and hence low variability) in the Tropics of the eastern Pacific and the low variability in the southern Atlantic. Also present is high variability in the vicinity of the major ocean currents: the Gulf Stream in the north-western Atlantic; the Agulhas Retroflection off the south-western tip of Africa; the Antarctic Circumpolar Current visible right around the south-polar region and then tracing a path up towards the south-eastern tip of Africa from west of Australasia.

7.4 Conclusions.

This SXO based procedure (Section 7.3) can be used to extract a noisy time series of sea surface anomalies from any satellite mission, even missions which have included varied orbital phases (such as ERS-1). Previously it has proved difficult to use

different satellite phases together in a variability study due to the limitations of the repeat-pass procedure, and unusual orbital phases (such as the ERS-1 Geodetic phase) have been ignored.

The data produced by this technique will be used to analyse sea surface variability patterns on a global basis and validated by comparison and ultimately combination with T/P repeat pass data.

8. Processing Of The Sea Surface Anomaly Time Series.

The noisy, irregularly spaced time series produced either by processing of ERS-1 SXOs, or by T/P repeat pass analysis (or a combination of both) must be smoothed and re-sampled onto a regular temporal grid before it is possible to carry out an EOF de-composition. This is done by the use of an optimal interpolation procedure.

8.1 The Optimal Interpolation Procedure.

This is a mathematical procedure that outputs a regular grid of data regardless of the input data spacing. It also attempts to improve the signal to noise ratio of the data set by modelling the signals involved and filtering out other signals. Optimal interpolation is also known as 'least squares prediction' or 'objective analysis'.

8.1.1 Theory.

After *Davis* (1976), consider the covariance of the q measurements $\mathbf{l}$ and the m components of the signal $\mathbf{s}$ and an optimal estimate $\mathbf{s}^*$. If

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ \cdot \\ l_{q-1} \\ l_q \end{bmatrix}$$

is a vector of q measurements of sea surface heights, assume a similar vector exists for the actual signal $\mathbf{s}$:

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ . \\ s_{m-1} \\ s_m \end{bmatrix}$$

The assumption is made that **l** and **s** have an expected value of zero, i.e. E(**l**) = E(**s**) = 0, where the expected value is the average or mean.

Covariance matrices are introduced:

$\left.\begin{matrix} \mathbf{C}_{ll}(q\text{x}q) \\ \mathbf{C}_{ss}(m\text{x}m) \end{matrix}\right\}$ auto-covariance matrices

$\mathbf{C}_{sl}(m\text{x}q)$ cross-covariance matrix.

The elements of these matrices are given by:

$$\mathbf{C}_{ll} = \text{E}\{l_i, l_j\} \quad 1 \le i,j \le n$$
$$\mathbf{C}_{ss} = \text{E}\{s_k, s_h\} \quad 1 \le k,h \le m$$

and $\mathbf{C}_{sl} = \text{E}\{s_k, l_i\}$ $1 \le i \le n$

$1 \le k \le m$.

The above can also be written as:

$$\mathbf{C}_{ll} = \text{E}\{\mathbf{l}\,\mathbf{l}^{\text{T}}\}, \quad \mathbf{C}_{ss} = \text{E}\{\mathbf{s}\,\mathbf{s}^{\text{T}}\} \text{ and } \quad \mathbf{C}_{sl} = \text{E}\{\mathbf{s}\,\mathbf{l}^{\text{T}}\}$$

where $^{\text{T}}$ denotes the matrix transpose.

A linear estimate for **s** can be written

$$\mathbf{s}^* = \mathbf{H}\mathbf{l}$$

where $\mathbf{H}$ is a mxq matrix. Introduce the error vector $\varepsilon = \mathbf{s}^* - \mathbf{s}$ with covariance matrix

$$\mathbf{C}_{\varepsilon\varepsilon} = E\{\varepsilon\,\varepsilon^T\} = E\{(\mathbf{s}^* - \mathbf{s})\,(\mathbf{s}^* - \mathbf{s})^T\}$$

The diagonal elements of $\mathbf{C}_{\varepsilon\varepsilon}$ are the error variances σ^2_k of the estimated signal s^*_k, i.e.

$$\sigma^2_k = E\{\varepsilon^2_k\} = E\{(s^*_k - s_k)^2\}.$$

As the best linear estimate of $\mathbf{s}$ in terms of $\mathbf{l}$ we seek the minimum variance unbiased estimate.

Now, $E\{\mathbf{s}^*\} = E\{\mathbf{H}\mathbf{l}\}$

$$= \mathbf{H}E\{\mathbf{l}\}$$

$$= 0$$

i.e. the condition for an unbiased estimate.

The matrix H is arbitrary, and

$$\varepsilon\,\varepsilon^T = (\mathbf{s}^* - \mathbf{s})(\mathbf{s}^* - \mathbf{s})^T$$

$$= (\mathbf{H}\mathbf{l} - \mathbf{s})(\mathbf{H}\mathbf{l} - \mathbf{s})^T$$

$$= (\mathbf{H}\mathbf{l} - \mathbf{s})(\mathbf{l}^T\mathbf{H}^T - \mathbf{s}^T)$$

$$= \mathbf{H}\mathbf{l}\mathbf{l}^T\mathbf{H}^T - \mathbf{s}\mathbf{l}^T\mathbf{H}^T - \mathbf{H}\mathbf{l}\mathbf{s}^T + \mathbf{s}\mathbf{s}^T$$

the average of which gives $\mathbf{C}_{\varepsilon\varepsilon}$:

$$\mathbf{C}_{\varepsilon\varepsilon} = E\{\varepsilon\,\varepsilon^T\}$$

$$= \mathbf{H}E\{\mathbf{l}\,\mathbf{l}^T\}\mathbf{H}^T - E\{\mathbf{s}\,\mathbf{l}^T\}\mathbf{H}^T - \mathbf{H}E\{\mathbf{l}\,\mathbf{s}^T\} + E\{\mathbf{s}\,\mathbf{s}^T\}$$

$$= \mathbf{H}\mathbf{C}_{ll}\mathbf{H}^T - \mathbf{C}_{sl}\mathbf{H}^T - \mathbf{H}\mathbf{C}_{ls} + \mathbf{C}_{ss}\,. \qquad 24.$$

But as $\mathbf{C}_{ls} = \mathbf{C}_{sl}^T = E\{\mathbf{l}\,\mathbf{s}^T\}$, Equation 24 gives

$$\mathbf{C}_{\varepsilon\varepsilon} = \mathbf{C}_{ss} - \mathbf{C}_{sl}\mathbf{C}_{ll}^{-1}\mathbf{C}_{ls} + (\mathbf{H} - \mathbf{C}_{sl}\mathbf{C}_{ll}^{-1})\,\mathbf{C}_{ll}\,(\mathbf{H} - \mathbf{C}_{sl}\mathbf{C}_{ll}^{-1})^T \qquad 25.$$

(by $\mathbf{C}_{ll}^T = \mathbf{C}_{ll}$ due to the symmetry of an auto-covariance matrix)

Equation 25 is split into two components

$$\mathbf{A} = \mathbf{C}_{ss} - \mathbf{C}_{sl}\,\mathbf{C}_{ll}^{-1}\,\mathbf{C}_{ls}$$

$$\mathbf{B} = (\mathbf{H} - \mathbf{C}_{sl}\mathbf{C}_{ll}^{-1})\,\mathbf{C}_{ll}\,(\mathbf{H} - \mathbf{C}_{sl}\mathbf{C}_{ll}^{-1})^T\,.$$

For the minimum variance solution the minimum diagonal elements of $\mathbf{A} + \mathbf{B}$ must be found. Since $\mathbf{A}$ is dependent upon the signal and the measured values, it cannot be altered. $\mathbf{B}$, however, depends on $\mathbf{H}$, which is arbitrary so can be chosen to provide this minimum.

Let the kth row of $(\mathbf{H} - \mathbf{C}_{sl}\mathbf{C}_{ll}^{-1})$ be given by γ_k^T then

$$\mathbf{B} = \begin{bmatrix} \gamma_1^T \\ \gamma_2^T \\ \cdot \\ \gamma_q^T \end{bmatrix} \mathbf{C}_{ll} \begin{bmatrix} \gamma_1 & \gamma_2 & \cdot & \gamma_q \end{bmatrix}$$

and kth diagonal term of $\mathbf{B}$ is $(\gamma_k^T \mathbf{C}_{ll} \gamma_k)$, but covariance matrices are positive definite, i.e.

$$\mathbf{x}^T \mathbf{C}_{ll} \mathbf{x} \geq 0$$

for any $\mathbf{x}$ and with equality only if $\mathbf{x} = 0$. Thus with $\mathbf{C}_{\varepsilon\varepsilon} = \mathbf{A} + \mathbf{B}$, the diagonal terms of $\mathbf{C}_{\varepsilon\varepsilon}$ are minimised if diagonal terms of $\mathbf{B}$ are zero, that is if

$$\mathbf{H} - \mathbf{C}_{sl} \mathbf{C}_{ll}^{-1} = 0 \text{ , or}$$

$$\mathbf{H} = \mathbf{C}_{sl} \mathbf{C}_{ll}^{-1} .$$

Therefore the solution of minimum variance is given by

$$\mathbf{s}^* = \mathbf{Hl} = \mathbf{C}_{sl} \mathbf{C}_{ll}^{-1} \mathbf{l} \qquad 26.$$

with error covariance

$$\mathbf{C}_{\varepsilon\varepsilon} = \mathbf{C}_{ss} - \mathbf{C}_{sl} \mathbf{C}_{ll}^{-1} \mathbf{C}_{ls}$$

Equation 26 therefore represents the required solution.

Optimal interpolation allows any system of observations to be compared with any signal system. In addition, any output grid can be used independently of observation data spacing.

8.1.2 Applying Optimal Interpolation To A Sea Surface Variability System.

If optimal interpolation is applied to a system of altimeter observations, then a regularly spaced data set containing the appropriate signal can be obtained from an irregularly spaced noisy observation data set. In order to construct the covariance matrix

$\mathbf{C}_{ll}$, we note that the observation data exhibits variation due to the altimeter signal, orbit errors and the actual variability signal (*Blanc et al*, 1995), i.e.

$$\{\mathbf{C}_{ll}\}_{i,j} = A\delta(\tau) + Be^{-\tau^2/d^2}\cos M\tau + De^{-\tau^2/g^2}e^{-\psi^2/R^2} \qquad 27.$$

where $\tau = t_i - t_j$ i.e. the temporal separation, and $\psi = |\mathbf{x}_i - \mathbf{x}_j|$ i.e. the spatial separation. The first term is the variance of the altimeter noise, the second represents a once per revolution orbit error (damped) and the third is the variance in the ocean signal. The values chosen for A, B and D are dependent on the altimeter noise, an estimation of orbit error and ocean surface variability respectively. d and g are user specified decorrelation times representing the time in which the variance of the parameter is reduced by a half, i.e. $d \log_e 2$ for the orbit error term and $g \log_e 2$ for the variability term.

Typical values are d = 3 days, g = 20 days and A = 4cm^2 (i.e. 2cm noise in altimeter signal), B = 100cm^2 (i.e. 10cm orbit error), D = 100cm^2 (i.e. 10 cm variability). These values are chosen on the basis of prior knowledge of the effects involved. Orbit errors are known to be strongly decorrelated within 3 days, and ocean variability is known to be small over less than about 20 to 30 days except in regions of very high activity (e.g. the Gulf Stream). The altimeter noise is known from studies of the electronic accuracy of the device itself and is not the same as the accuracy of the observations. Orbit error has been extensively studied and is generally taken as between 5 and 10 cm for ERS-1 data. Variability is set at 10 cm as a global average. These parameters are the 'guidelines' which the optimal interpolation technique uses to filter the data set.

8.1.3 The Practicalities of Optimal Interpolation.

The raw time series data were read from the data files. Where both T/P and ERS-1 data were used together, both time series were read and combined into one new time series. Any mean offset in each set of observations was removed.

The raw time series was then filtered to remove erroneous values. Any bin with an RMS variability in excess of 1 m was taken to be in error, and removed from further analysis. In practice this succeeded in removing a very small number of bins (generally less than 20 out of over 5000), a few of which had RMS variabilities of around 15-20 metres.

A further rejection criterion was included to prevent use of bins with few arcs, for example near the polar regions where there are long gaps in altimeter data due to ice coverage of the Ocean's surface. Any bin where there was a time gap exceeding 80 days between subsequent arcs was rejected. This was necessary because of the low spatial resolution obtainable from the ERS-1 ice phase. Less than a third of the total number of bins contain any crossovers in the ice phase. This creates a problem in that two thirds of the bins therefore have almost a three month gap between subsequent arcs. Therefore any time limit of less than 80 days would have eliminated two thirds of the bins. The 80 day limit is nonetheless still effective in removing bins which are worst affected by seasonal ice contamination.

It was assumed that orbit error had been removed, so the final parameters used in the construction of the covariance matrix (from Equation 27) are:

- decorrelation time g, effectively controlling how 'smoothed' the data set is,
- maximum variability D, which applies a constraint to the amplitude of the variance,
- altimeter noise A, representing the random noise in any single observation.

It is this optimally interpolated data set which can be used for an empirical orthogonal function decomposition.

A comparison of Figure 14 (unsmoothed data) and Figure 15 (after smoothing with optimal interpolation), bearing in mind the altered z-scale, reveals that the global RMS variability pattern is mostly unchanged and accurately follows the patterns expected from a knowledge of ocean circulation. This fact is reassurance that the parameters chosen above were acceptable, and that there is no corruption or loss of oceanographic signal by the optimal interpolation procedure. Note that the increase in areas ex-

cluded from the data set (near land masses) is due to the bin rejection criteria applied during the OI procedure.

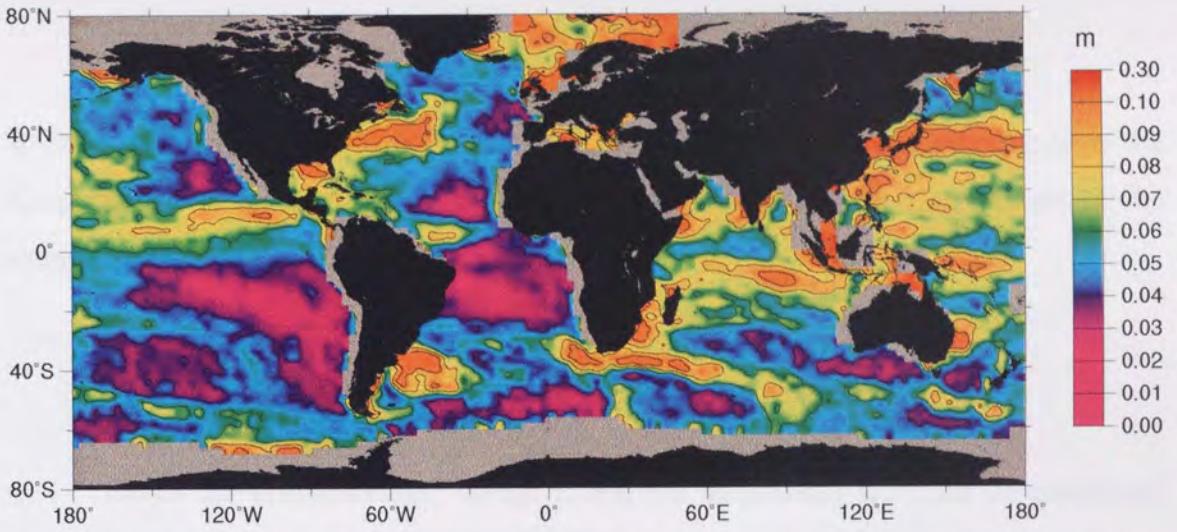

Figure 15. Global RMS variability from the Combined ERS-1–T/P data set after optimal interpolation. Contour lines are placed at 0.02m intervals.

8.2 Empirical Orthogonal Functions (EOFs)

EOFs can be used to represent any given signal in terms of spatial modes. As a by-product, EOFs provide the most effective means of compressing data, can be regarded as uncorrelated modes of variability, and simplify the understanding of the procedures of minimum mean square error estimation. A brief summary of the EOF system is presented in the following section. Further details can be obtained from *Davis* (1976).

8.2.1 Theory (summary).

If $n_{\mathbf{x}}(t)$ is the derived sea level at time t at location $\mathbf{x}$, $n_{\mathbf{x}}(t)$ can be expressed at the locations $\mathbf{x}_m$, m=1, …, M, in terms of M orthogonal functions $F_k(\mathbf{x}_m)$ as

$$n_m(t) = \sum_{k=1}^{M} a_k(t) F_k(\mathbf{x}_m) \qquad m = 1, \ldots, M$$

The EOFs are uniquely defined within the space of all orthonormal function sets by the constraint that the corresponding amplitudes a_k satisfy

$$\langle a_i(t).a_j(t)\rangle = \mathbf{A}_{ij}\delta_{ij}$$

where $\langle \ldots \rangle$ denotes the sample mean value over time, A_{ij} is constant, and δ_{ij} is the Kronecker delta function. It can be shown that the EOFs are the eigenvectors of the square mean cross product matrix **C** with elements

$$\mathbf{C}_{ij} = \langle n_i(t).n_j(t)\rangle$$

Given the orthonormality of the EOFs, the amplitudes $a_k(t)$ can be calculated from the observed time series and the k^{th} EOF by

$$a_k(t) = \sum_{m=1}^{M} n_m(t)F_k(\mathbf{x}_m).$$

The mean square variability summed across all locations **x** is spread amongst the M modes such that the fraction of total mean square variability accounted for by the k^{th} EOF is

$$\langle a_k(t).a_k(t)\rangle \Big/ \sum_{m=1}^{M} \langle n_m(t).n_m(t)\rangle.$$

The EOFs give an extremely high efficiency of presentation, as opposed to other orthonormal expansions. If the EOFs are ordered by decreasing fraction of total mean square variability, no other expansion such as

$$n_m{}^K(t) = \sum_{k=1}^{K} b_k(t)G_k(\mathbf{x}_m)$$

using the $K < M$ functions $G_k(\mathbf{x}_m)$ gives a lower total mean square error

$$E = \sum_{m=1}^{M} \langle [n_m(t) - n_m^K(t)].[n_m(t) - n_m^K(t)] \rangle$$

than can be obtained when the $G_k(\mathbf{x}_m)$ are EOFs. Hence when the signals are dominated by a few large scale coherent signals (such as the annual cycle in the case of sea surface variability), fewer modes are required to represent the majority of the signal than for any other set of basis functions.

8.2.2 The Practicalities Of EOFs.

The optimally interpolated data set is filtered before the EOF decomposition is performed. Included in this data set is information about the density of 'real' arcs (as opposed to arcs generated by the optimal interpolation procedure) in the original time series. Thus a filter is applied which rejects bins for which the average time gap between real arcs is greater than a certain value. This value was set at 10 days (i.e. half the normal value given for the temporal decorrelation during the optimal interpolation procedure). Thus no bins with sparse data will be accepted, as this may corrupt the EOF modes produced by incorporating optimal interpolation artefacts into the data - even in the absence of actual observation data, optimal interpolation still produces a 'signal'.

In practice it was necessary to compromise between the number of bins for which EOF data were produced and the length of the time series. In order to carry out a successful EOF decomposition, every bin must contain a full set of observations at the correct epochs as otherwise the bin will be rejected. Combining this with the fact that the optimal interpolation procedure will not produce data before the first observation or beyond the last, it becomes necessary to set appropriate start and end dates according to the 'worst' bins it is wished to retain in the data set, i.e. the bins with least time span.

8.3 Which Data Set To Use?

In this project, there are two data sets. The T/P data set consists of a long term mission in a single orbit type, and benefits from accurate knowledge of the satellite's orbit, thereby making the associated altimetry highly reliable. The mission compromises a spatial density for the sake of temporal density, so that every arc is overflown once every 9.9 days, but there is a spacing of 2.83° between each ground track. From Nyquist sampling theory, this means it is impossible to resolve an across track signal with wavelength of less than twice these values (560 km across track at the equator or 19.8 days temporally globally). The ERS-1 mission, on the other hand has been split into several distinct orbital phases (see Table 1 in the Introduction) each with significantly different characteristics. Most of the mission was carried out in the Multi-Disciplinary format which provides greater spatial density (0.72° ground track spacings) but lower temporal density (overflown once every 35 days). The Geodetic mission (making up the majority of the remainder) places the emphasis entirely upon spatial accuracy. It has a high density of ground tracks but has no repeat cycle, thereby making any kind of repeat pass study impossible whilst using this phase. Thus ERS-1 offers unrivalled spatial properties as compared to T/P but does not provide a simple variability extraction method due to the differing phases and long repeat cycles. In addition, ERS-1 suffers from being in a lower orbit thus increasing atmospheric drag and gravity field modelling difficulties during the orbit determination procedure. This reduces the accuracy of the orbits compared with T/P. The lack of a single repeat pattern for the ERS-1 mission is overcome through the use of crossover data. This also theoretically allows the use of any grid density. In practice, however, this is restricted by the type of data which are included in the procedure. In cases where T/P data are included, a grid spacing which places bins between the T/P ground tracks (and which thus have no T/P data) will cause problems during the later EOF decomposition since these bins will contain a different set of signals from the others. A similar problem occurs upon including the ERS-1 ice phase in an EOF decomposition. Several EOFs are produced which simply represent the difference in behaviour between bins which contain data in the ice phase and bins which do not, and other EOFs are contaminated with similar effects.

Where only ERS-1 data are used in a study, a longitudinal grid spacing of 1.5° or less can be used, since each bin will contain at least 2 ground track passes during the Multi-Disciplinary phases (and consequently also during the geodetic - see the Introduction). The only limiting factor is computational: a global EOF decomposition for a grid spacing of 1.5° x 1.5° requires determination of the eigenvalues and eigenvectors of a square matrix of approximately 15,000 x 15,000 elements. This requires 1.8 GB of memory for the duration of the process. A reduction in resolution is thus required for the production of global EOFs, as the required amount of memory was unavailable.

On the inclusion of T/P data to the procedure, the longitude grid spacing is forced to 2.835° or more as a finer grid results in bins containing no T/P data. This does, however, also result in a drop in the number of bins required. Using a 2° latitude gridding requires a more manageable 5800 bins, allowing the EOF decomposition to be carried out more easily in around 256 MB of memory. On a grid size such as this, there are more ERS-1 arcs per bin than T/P arcs, due to the repeat pass technique in use for T/P and single satellite crossovers for ERS-1 (even though this is only a subset of the theoretical maximum number of SXOs which could have been produced from the entire ERS-1 mission).

9. Sea Surface Variability

This chapter is a discussion of the relative merits of inclusion of both ERS-1 and T/P data in a sea surface variability study. This discussion is supported by analysis of EOFs generated as the difference between the two types of data. The results of the global EOF decomposition are then presented, along with some discussion of the ocean variability implications from the EOFs. Supporting evidence for these implications is drawn from the optimally interpolated time series and an analysis of the nature of the annual signal is carried out.

9.1 The Data Sets.

The ERS-1 SXO data set used in this project is as detailed in Section 6.2. The T/P repeat pass data includes information from cycles 1 to 108 (MJD 48901 to 49959) of the mission.

The data sets available for analysis are as follows:

- 'The Dual Crossover (DXO) data set'. A set of DXO residuals between ERS-1 and T/P during the entire period spanned by the ERS-1 mission. The composition of this data set is detailed in Chapter 3.

- Optimally Interpolated (OI) time series:
    1. 'The DXO Corrected ERS-1 Data'. Derived from ERS-1 SXOs, using DUT JGM3 orbits, DXO corrected using the cubic spline technique in Chapter 4.
    2. 'The T/P Repeat Pass Data'. Derived from T/P repeat pass data. The generation of this data set is described in Section 4.3.
    3. 'The Combined Data Set'. Derived from a combination of the ERS-1 DXO corrected SXOs (as in 1) and the T/P repeat pass data (as in 2).

- Empirical Orthogonal Function decompositions:
    4. 'The ERS-1 EOFs'. Derived from OI time series 1 above.
    5. 'The T/P EOFs'. Derived from OI time series 2 above.
    6. 'The Combined EOFs'. Derived from OI time series 3 above.

A data set which includes T/P observations is gridded to the default 2.83° by 2° mesh. ERS-1 data sets may be gridded to finer dimensions, which is indicated if so.

Note that for all of the EOF figures, the colours in the spatial mode plot are histogram equalised by the GMT package in order to clarify the signal present. Contour lines are placed at 0.01m intervals. The amplitude time series has no units - it is simply a factor by which the EOF is multiplied. Note also that any oceanographic interpretation of the EOFs is qualitative and based on prior knowledge, and is in part an effort to support the validity of the data set rather than an analysis of the data for its own sake.

9.2 Combination Of ERS-1 SXO Data And T/P Repeat Pass Data.

In a global EOF decomposition, what are the benefits of using both ERS-1 and T/P data? No finer grid can be used than with T/P alone (without the use of a spatial interpolation scheme) and therefore the spatial resolution is still restricted to the same level. The system becomes more problem-tolerant upon the addition of extra data, but is this at the expense of consistency?

To answer these questions, a global EOF decomposition was carried out on a data set constructed as the arithmetic difference between that produced by the combined ERS-1–T/P data set and that from T/P alone. In performing this differencing operation, care was taken so that only equivalent values (at the same location and at the same epoch) are differenced.

This data set provided 5438 bins, and of these 5386 contained sufficient data to perform an EOF decomposition. This produced 83 EOFs which on output were sorted

according to how much of the overall signal they represented, thereby making the first few the most significant - the first 11 represented over 50% of the total signal present.

9.2.1 The Combined Minus T/P Difference EOFs

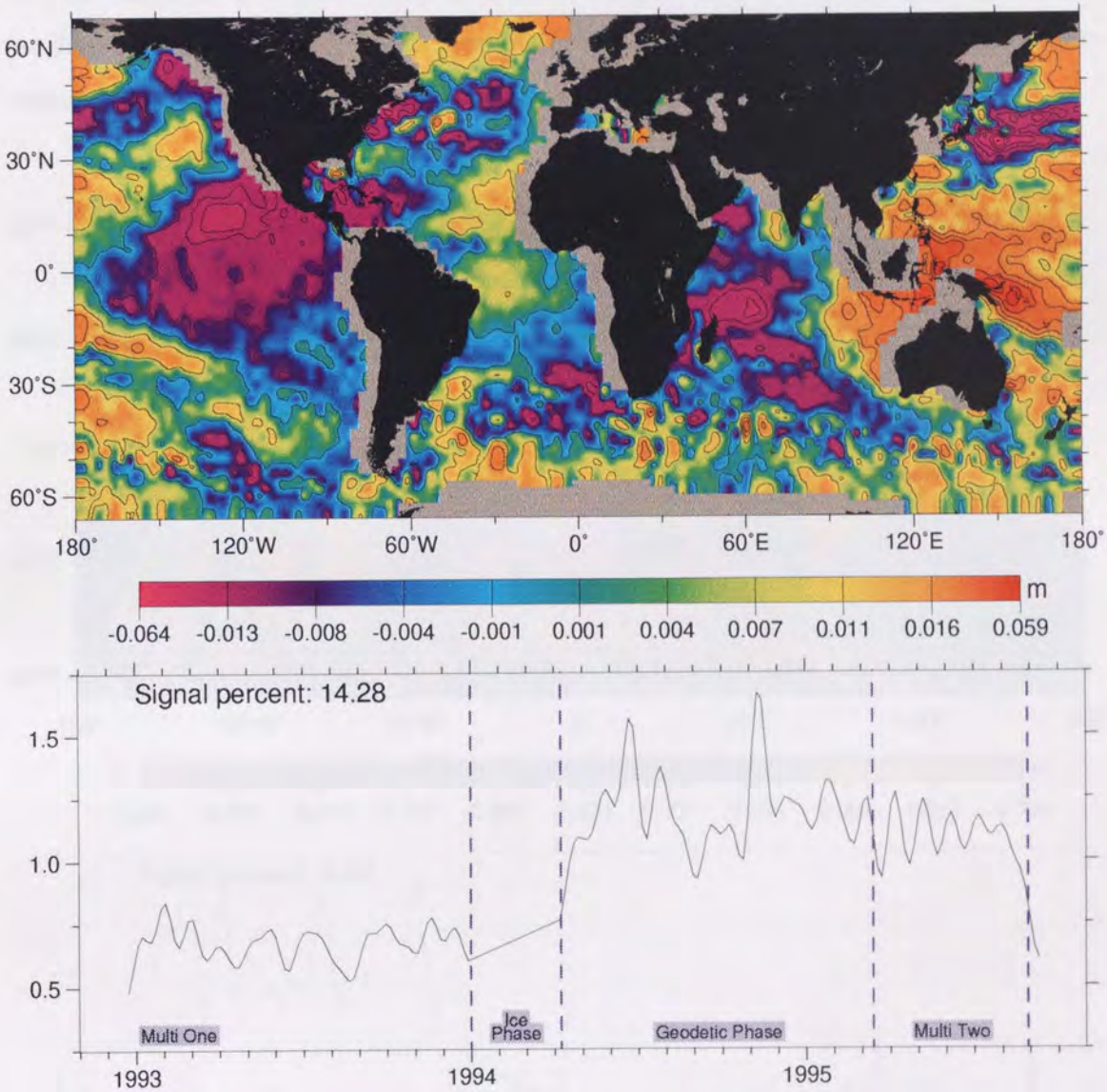

Figure 16. EOF 1 from the Combined-minus-T/P difference data set. Contour lines at 0.01m intervals.

EOF 1 from the combined-minus-T/P difference data set represents 14.28% of the total signal. The EOF (Figure 16) shows opposite peaks in the west and east Pacific and strong signals in the Indian Ocean. The mid-Atlantic appears mostly unaffected, but the extreme north and south show high values. This resembles a first-order spherical

harmonic, and could be a result of deficiencies in gravity modelling (the JGM3 gravity model of *Tapley et al.* [1996] does not include any ERS-derived information) or the reference frames used by the tracking systems for each satellite. The amplitude time series is consistently positive, suggesting a fundamental change between the parameters for the two missions. This is unexpected, since it was presumed that the DXO orbit error correction would remove most of these long-wavelength differences. The major weakness (as discussed in Chapter 4) in this type of scheme is the distribution of the DXO observations themselves: it is thus possible for some signals to remain uncorrected.

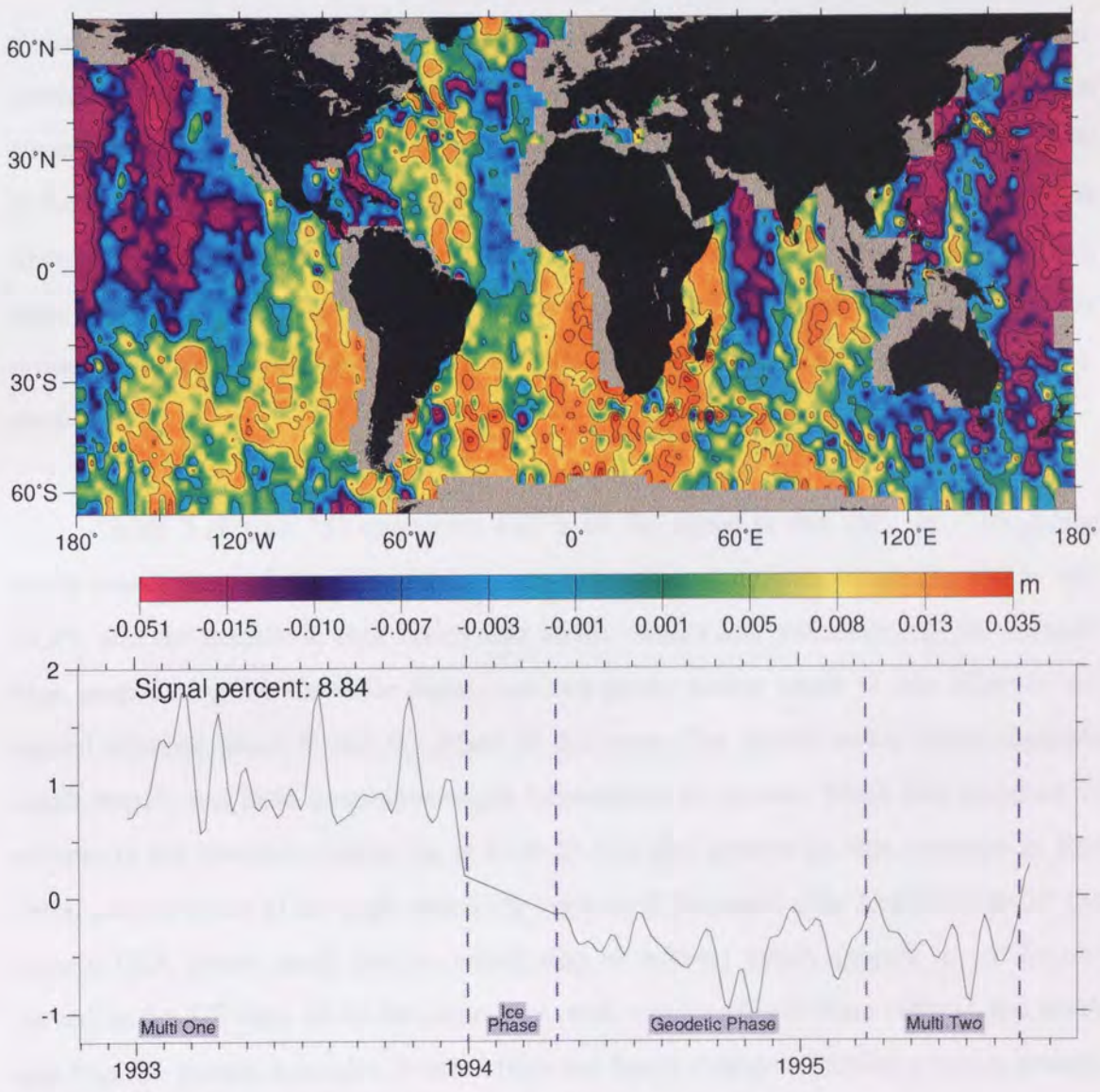

Figure 17. EOF 2 from the Combined-T/P difference data set.

The time series also shows an oscillation during the first Multi-Disciplinary phase (C) with wavelength of around one month (close to the repeat cycle of 35 days), clearly a remnant of uncorrected ERS-1 orbit error. A similar signal occurs during the second Multi-Disciplinary phase (G), and even during the Geodetic phase (which has a sub-cycle of 37 days). An offset appears to be present between phase C and the other two phases (this has been observed in the ERS-1-only EOFs also). The ice phase (phase D) has not been included in any of the EOF decompositions.

The long wavelength information in the spatial mode of **EOF 2** (Figure 17) shows opposite values at approximately opposite sides of the globe - lows in the mid-north Pacific and highs in the south Atlantic, centred on around 0° longitude (a similar cause to the signal observed in the first EOF is suspected here). Also of note in this EOF is the fine structure in the Indian and Atlantic oceans at around 35° south. The amplitude time series for this EOF shows an offset between phase C and the other ERS-1 phases, with a weak oscillation at about one month wavelength as in EOF 1. Strong peaks are present at several epochs during phase C. This EOF represents 8.84% of the total difference signal.

EOF 3 (Figure 18) represents 4.07% of the signal in this data set. The spatial mode shows most of the globe having very low values (i.e. little variability due to this EOF), and the amplitude time series also shows mainly low values except for a single high peak during the Geodetic phase, and two peaks during phase G. No offset is observed between phase C and the others in this case. The spatial mode shows multiple small details and little long-wavelength information is present. More fine structure is evident in the Southern Ocean (as in EOF 2), but also present on this occasion is fine detail around many of the high variability regions of the world. The Gulf Stream off the eastern USA shows small details (which may be eddies) which appears to not be observed in the T/P data set in the same way, with similar effects observable in the north east Pacific, around Australia, South Africa and South America. Similar detail is present in some of the less significant EOFs produced from T/P data only (not shown here), suggesting that T/P *is* capable of observing these signals and that their presence in this difference data set is simply an artefact of the differing observation times and rates for the two satellites. Areas such as these change rapidly and the signals may be observed at

different stages of their cycle by the two satellites simply because they do not take measurements at the same time.

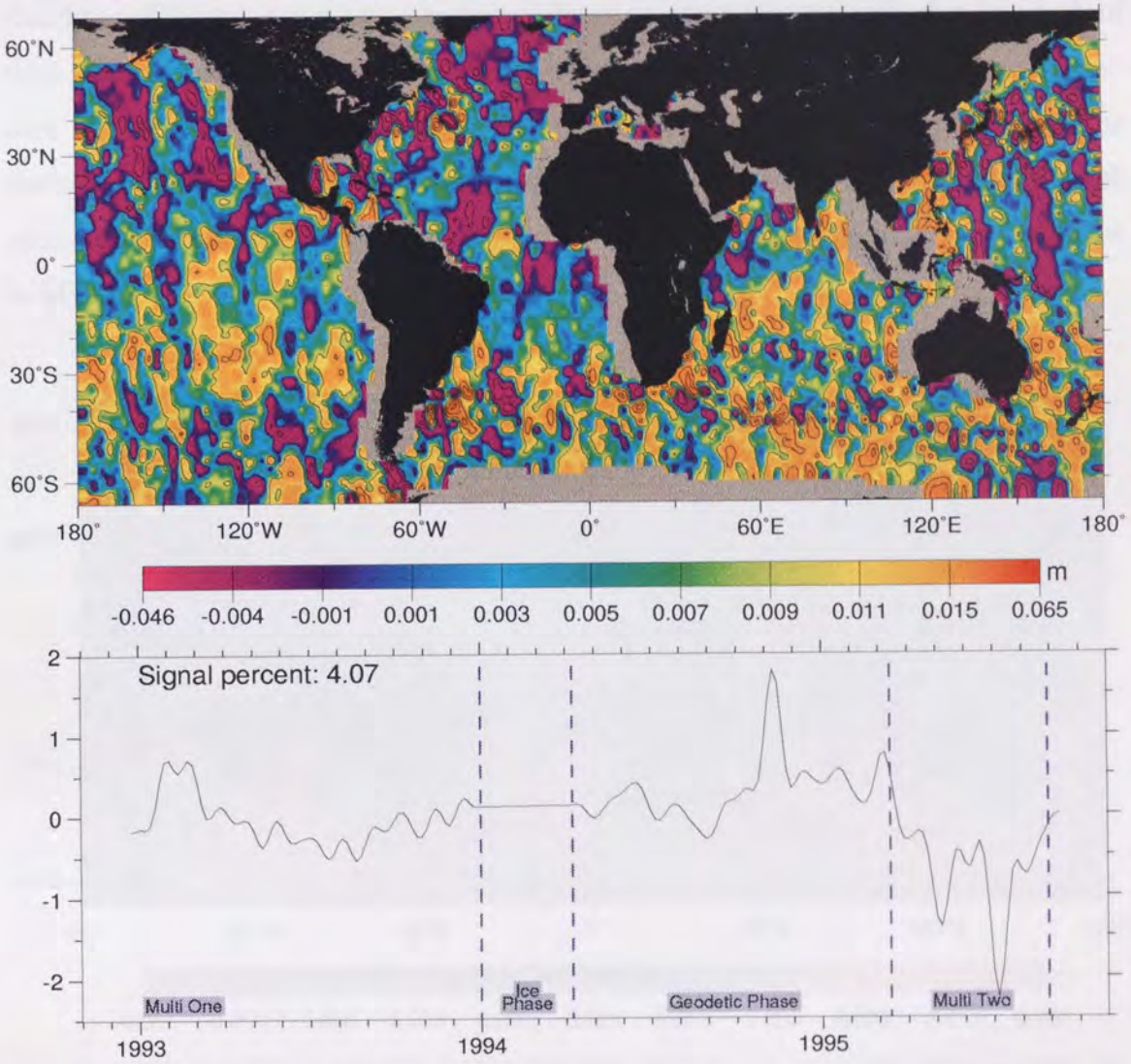

Figure 18. EOF 3 from the Combined-T/P difference data set.

EOF 4 (3.84% of the signal) (plotted in Figure 19) has similar effects to those observed in EOF 3 above in the amplitude time series - mainly low values, with several sharp peaks and an overall, small-amplitude oscillation with a period of about 35 days throughout the time span. The peaks in the time series for each EOF appear to be random, but on observation of the whole data set it is clear that they occur at certain fixed epochs throughout the time span. A strong peak occurs at the end of October 1993 for both the 2nd, 4th, 5th and 7th EOFs (and possibly more EOFs which have not been closely analysed). These epochs coincide with Poseidon cycles of the T/P satellite (the

Poseidon altimeter is used instead of TOPEX for several cycles; since the two altimeters share the same receiver dish they cannot be active simultaneously). During the repeat pass data processing, Poseidon data are rejected (for consistency purposes), since Poseidon has a different range bias which is difficult to accurately quantify due to a lack of data. The optimal interpolation system will therefore be unconstrained during such cycles. All the large peaks observed in these EOFs (right through the first 10 EOFs) are during Poseidon cycles. The spatial mode for this EOF shows a broadly north-south split, and this could be a further indication of low-order gravity modelling problems (as in EOF 2).

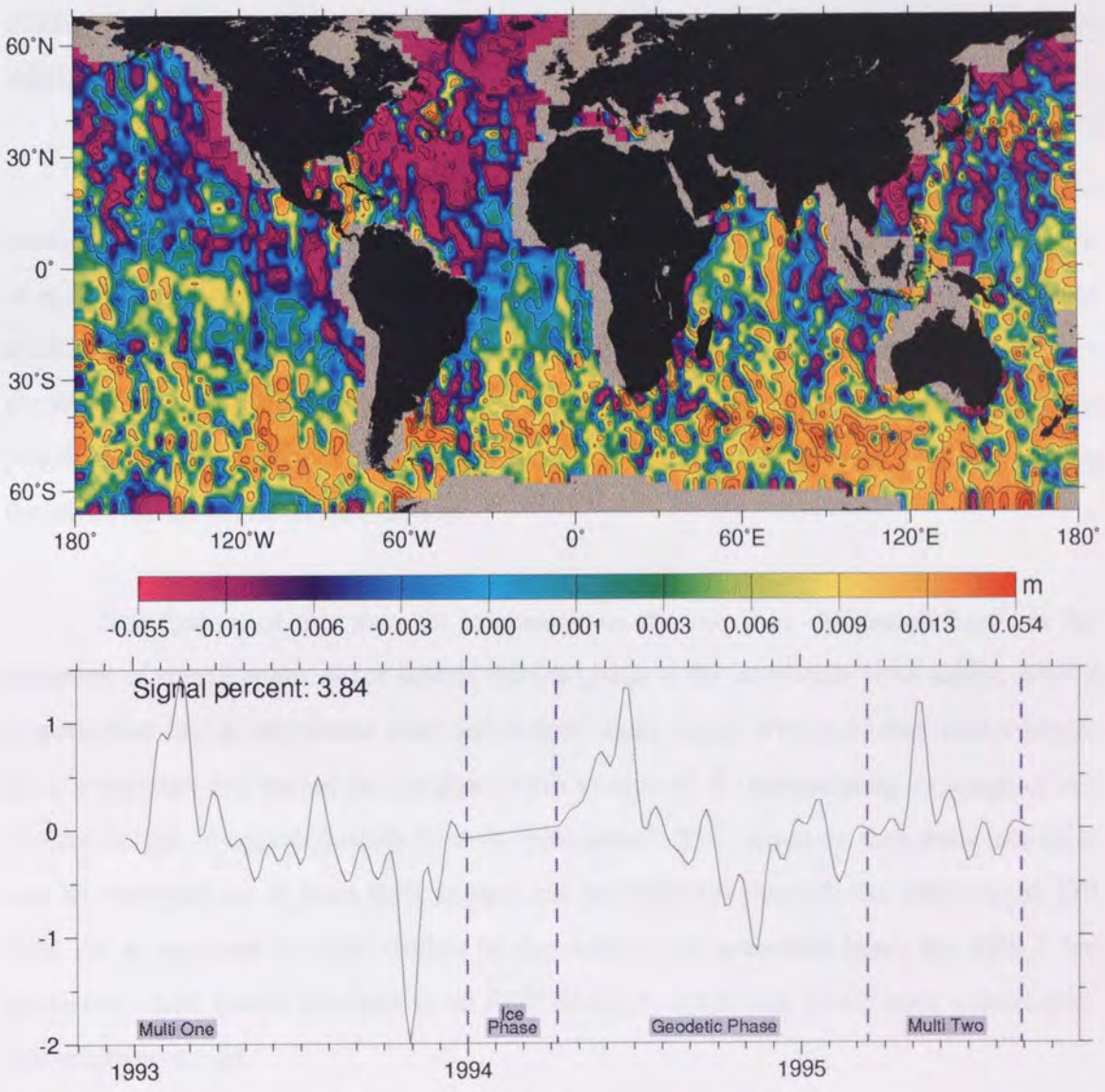

Figure 19. EOF 4 from the Combined-T/P difference data set.

EOFs 5-10 (collectively 17.69% of the signal) all show mainly low values in the amplitude time series (less than ±0.5), with several peaks coinciding with Poseidon cycles and an approximately 35 day oscillation throughout. The spatial modes are difficult to interpret as there is very little coherent structure to the data. As in previous EOFs, areas of high variability show up as having complex fine structure.

9.2.2 The Combined Minus ERS-1 Difference EOFs.

A further decomposition was carried out on a data set of combined ERS-1–T/P minus ERS-1 optimally interpolated values. This produced a set of 89 EOFs of diminishing importance.

As expected, the EOFs from this data set reveal similar patterns to those in the previous section. EOF 1 (8.86% of the signal) shows peaks in the amplitude time series at epochs coinciding with Poseidon cycles and a clear 35-day ERS-1 orbit-error related oscillation. There is also a significant offset between the first ERS-1 Multi-Disciplinary phase (C) and the other two phases (E+F and G) used in this study. The spatial mode is practically identical to that observed as EOF 2 in the previous section, thus supporting the observations made in that section.

One feature of this data set in contrast to the previous difference EOFs is the presence of strong peaks either side of the ice phase in the amplitude time series. EOF 2 in particular has an amplitude time series with most values within ±1 but with a single peak at the start and end of the ice phase with a value of -6 (representing a change of sea surface height of approximately 8cm in most areas). This suggests that these artefacts can be removed (or at least their impact can be reduced) through the addition of T/P data. As is observed in other studies by the author (not presented here) the ERS-1 ice phase can cause severe artefacts in an EOF decomposition, due to the poor spatial density of observations.

Other significant points are the presence of, in some cases, a strong 35-day oscillation in the amplitude time series, particularly during ERS-1 phase C. This oscillation is much weaker during phases E and F.

9.2.3 Conclusions

As expected, there is very little signal remaining in these two data sets. All spatial modes have extreme values of sea surface height of less than ±2cm, with an RMS of around 1.3cm. Thus even the highest peaks observed in the amplitude time series represent sea surface height changes of only about 4cm maximum (with one exception noted above). The RMS change is 1.5cm upon addition of the ERS-1 data to T/P alone and this 1.5cm difference appears to be due to ERS-1 'filling the gaps' left by Poseidon cycles in the T/P data, rather than adding any new problems to the data. On the other hand, adding T/P data to ERS-1 data is capable of making improvements to the system by reducing the 35-day ERS-1 orbit error signal and problems related to ERS-1 phase D. Therefore this shows that ERS-1 and T/P are complementary systems, and there are no problems with using both data sets together. In fact the system benefits from the extra data density only achievable through use of data from both satellites.

9.3 The EOF Data Sets.

9.3.1 The T/P Repeat Pass EOFs.

Repeat pass data produced between 170 and 700 data points in each bin from TOPEX cycles, depending on the number of T/P arcs passing through each bin. Latitude bands near the equator had the minimum number of passes and at higher latitudes bins frequently had twice this number, up to the maximum ~700 data points near the 66° extremes.

An EOF decomposition on this data set produces 72 EOFs, though only the first 22 contribute more than 1% to the overall signal. It should be noted that the first 8 EOFs

describe 50.23% of the total variability in this data, and the first two 26.44%. These values will be compared with those from the ERS-1 and combined data sets below.

9.3.2 The ERS-1 EOFs.

The EOF decomposition based on the DUT JGM3 ERS-1 SXO derived time series gave rise to 85 EOFs, of which the first 27 were each responsible for in excess of 1% of the total signal.

In this data set, the first 9 EOFs represent 50.45% of the total variability and the first two represent 25.09%. Note that one more EOF is required to represent 50% of the signal, and the first two EOFs represent slightly less of the signal than was the case for the T/P data in Section 9.3.1. Thus the variability signal from the ERS-1 data set is spread over more EOFs than is the case with T/P alone. This suggests that ERS-1 data is either more noisy than T/P or that ERS-1 is resolving finer signals than T/P can. It is likely, however, that the different ERS-1 phases are causing the addition of distinct modes of orbit error and hence distinct EOFs. Were the ice phase (D) to be included in this data set, more EOFs would be produced.

9.3.3 The Combined Data Set EOFs.

Upon using both ERS-1 and T/P data in the Optimal Interpolation / Empirical Orthogonal Function Decomposition process, the EOF data are changed slightly. The EOFs are broadly similar, but subtle changes are observable.

This EOF decomposition resulted in 87 Empirical Orthogonal Functions. Of these, only the first 23 were each responsible for more than 1% of the total sea surface variability signal. These are together responsible for over 78% of the observed signals. Table 1 shows the first 25 eigenvalues, along with the percentage of total signal that each one represents, and the cumulative total percentage. The remaining values are not shown here because they are negligible. Note that because EOFs 1 and 2 represent the

two spatial parts of the dominant annual signal they will always represent a large proportion of the total variability. (In this case 29.66%.)

As shown in section 9.1 this combined data set has improved characteristics over either of the individual satellite data sets in many aspects. This improvement justifies the extra computational time and resources necessary to carry out the procedure.

EOF Number	Eigenvalue	Percentage	Cumulative %
1	4.26417615192	18.54	18.54
2	2.55596769179	11.12	29.66
3	1.82937095796	7.96	37.61
4	1.07429753428	4.67	42.29
5	0.88133731249	3.83	46.12
6	0.71037848128	3.09	49.21
7	0.67970867359	2.96	52.16
8	0.63196330431	2.75	54.91
9	0.58153505560	2.53	57.44
10	0.49701608719	2.16	59.60
11	0.41714917221	1.81	61.42
12	0.40851817135	1.78	63.19
13	0.38382851590	1.67	64.86
14	0.37072327030	1.61	66.47
15	0.36386539001	1.58	68.06
16	0.34801725442	1.51	69.57
17	0.33641635715	1.46	71.03
18	0.31893720782	1.39	72.42
19	0.29441402335	1.28	73.70
20	0.27881142022	1.21	74.91
21	0.27269761759	1.19	76.10
22	0.25064333446	1.09	77.19
23	0.23452575255	1.02	78.21
24	0.22285150386	0.97	79.18
25	0.21537485875	0.94	80.11

Table 11. Eigenvalues and percentages of total signal for each EOF in the combined ERS-1 - T/P data set.

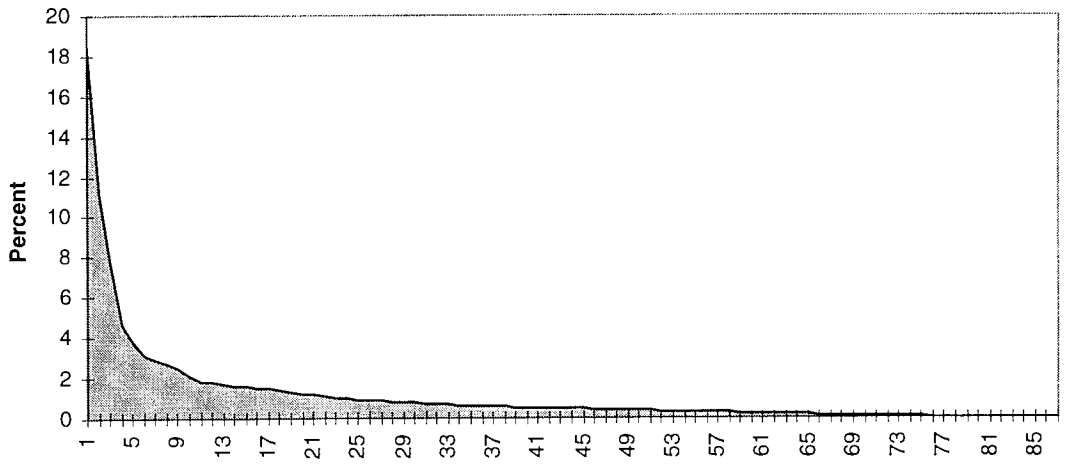

Figure 20. Percentage of total variability represented by each EOF for the combined ERS-1 - T/P data set.

As can be seen from Figure 20 and Table 11, the EOFs are efficient in representing the total variability signal. The percentage of signal per EOF drops off sharply in the first ten EOFs, and slowly drops to zero thereon: EOFs 83 to 87 inclusive represent only 0.01% or less of the total signal.

9.4 The Spatial Modes

In this section, the full detail of the first few Combined EOFs is shown, along with some thoughts about the oceanographic implications of each.

EOF 1 is plotted in Figure 21. This EOF is the largest component of the annual signal, displaying its peak amplitudes in February or March, and its minima each autumn. This EOF represents 18.54% of the total global variability. The spatial mode shows an approximately north/south split, representing a 'tilt' of the sea surface during the course of the year. This is due to surface heating by solar radiation which causes the surface to rise along with temperature. The maximum amplitudes occur approximately three months after the peak of summer in each hemisphere, which is a result of the thermal capacity of the water in the ocean: the heating effect continues to build up until well after the peak of the summer since the air is still warmer than the water, and vice versa in winter (*Levitus*, 1984).

The north/south phase change occurs at slightly different latitudes in each Ocean basin. The equatorial counter current in the Pacific (i.e. its effective equator) occurs about 5-10° north of the physical equator, whilst in the Atlantic it is again slightly north of the equator, but does not run along a line of latitude because of the encroachment of the two land masses on either side. The Indian Ocean displays the same north/south change and its effective equator is slightly south of 0° latitude.

Signals in the southern Indian Ocean are noisy due to the complexity of currents in the area. Variability due to the Antarctic Circumpolar Current (ACC) makes the fine detail complex.

A region of opposite phase to that in the remainder of the hemisphere occurs either side of the equator. This is a result of wind driven effects in the tropics. These type of signals are more evident in EOF 2.

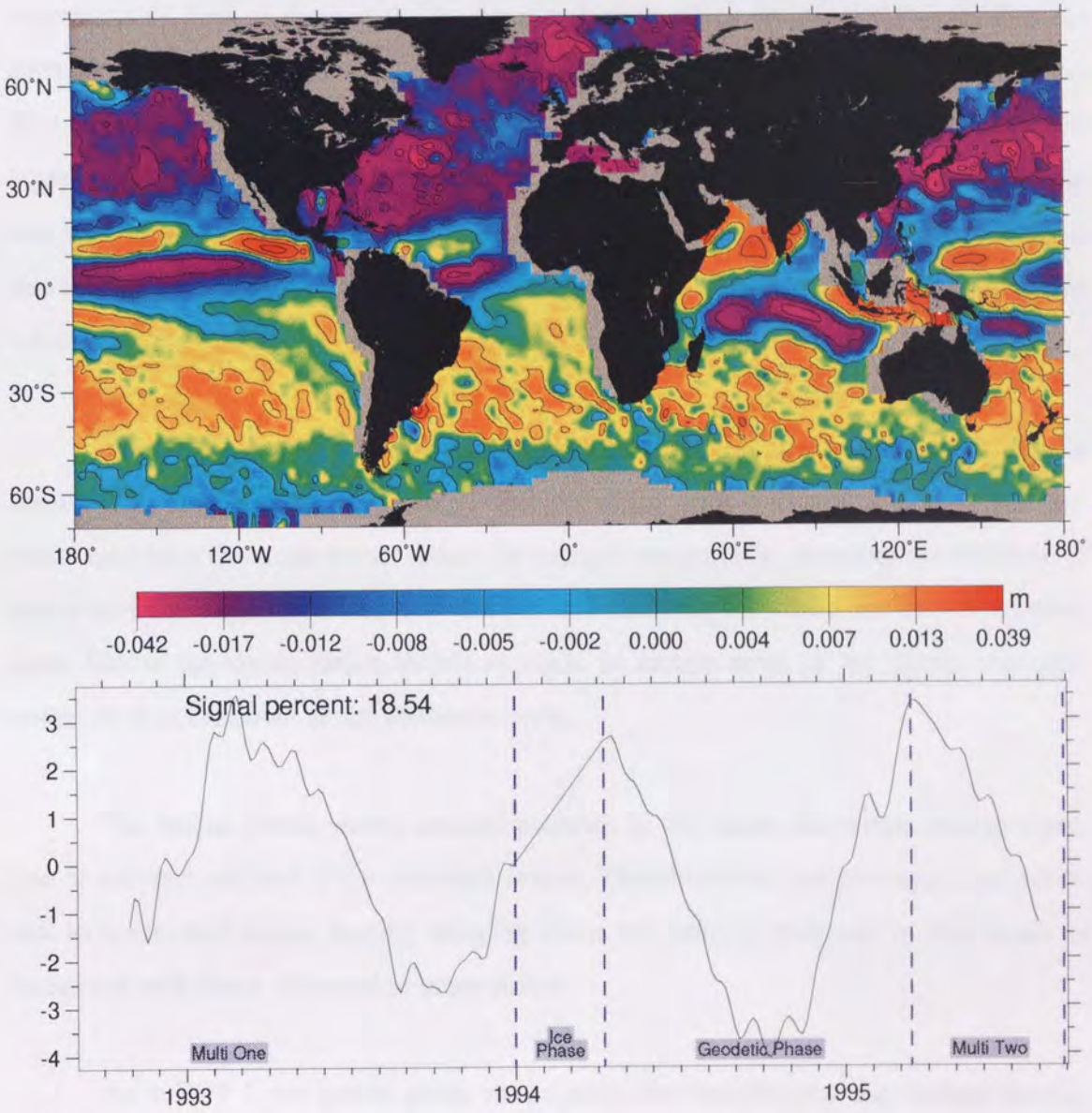

Figure 21. EOF 1 from the Combined ERS-1–T/P data set. Contour lines at 0.01m intervals.

The spatial mode shows similar features (short spatial wavelength, eddy-like) in the vicinity of both the Gulf Stream and the Kuroshio currents. These are both intense western boundary currents with high variability and thus have similar altimetric signatures. These display high variability in this annual cycle because of the large quantities

of heat transfer they provoke by carrying warm water away from the equator towards cold air sources.

EOF 2.

This EOF (Figure 22) is the second component of the overall annual signal and represents 11.12% of the signal. The amplitude time series reveals that this EOF is approximately 3-4 months out of phase with EOF 1. The spatial mode for this EOF is unlike the first in that the bulk of the signal present in this EOF is concentrated on the equatorial regions. The northern and southern Pacific and Atlantic parts reveal little signal, with heights rarely in excess of ±1.2cm, compared with over 10 cm of variability in the highest areas. The extreme north Atlantic part of EOF 2 contains strong signals, which are thought to be an artefact of the DXO correction scheme.

This EOF displays a strong signal (up to 10 cm RMS) in the equatorial region of the Pacific. This is due to the annual changes in the trade winds which drive the currents: each year the trade winds reduce in strength temporarily, stopping the build up of warm surface waters in the eastern Pacific and allowing the ocean surface to equalise again before the winds return to full strength. In certain areas of the globe, this phenomenon is also known as the Monsoon cycle.

The Indian Ocean shows unusual patterns. In this basin, the winds change direction at the start and end of the monsoon season. This direction change occurs in a north-east to south-west sense, thereby bringing about the patterns observed in this ocean as compared with those observed in other oceans.

As in EOF 1, the spatial mode shows eddy-like features near the western boundary currents of the Atlantic and Pacific oceans.

EOF 3.

This EOF (Figure 23) represents just under 8% of global variability. The time series reveals an apparently linear trend, but is in fact a partial cycle of a longer signal and is an example of how improper sampling of long-term trends can give rise to spurious estimates of sea-level rise (as also seen by *Hendricks et al.* [1996] and discussed

further in Chapter 11). A longer sampling period reveals more of the signal, such as can be achieved by using the full ERS-1 data set, or the full T/P data set.

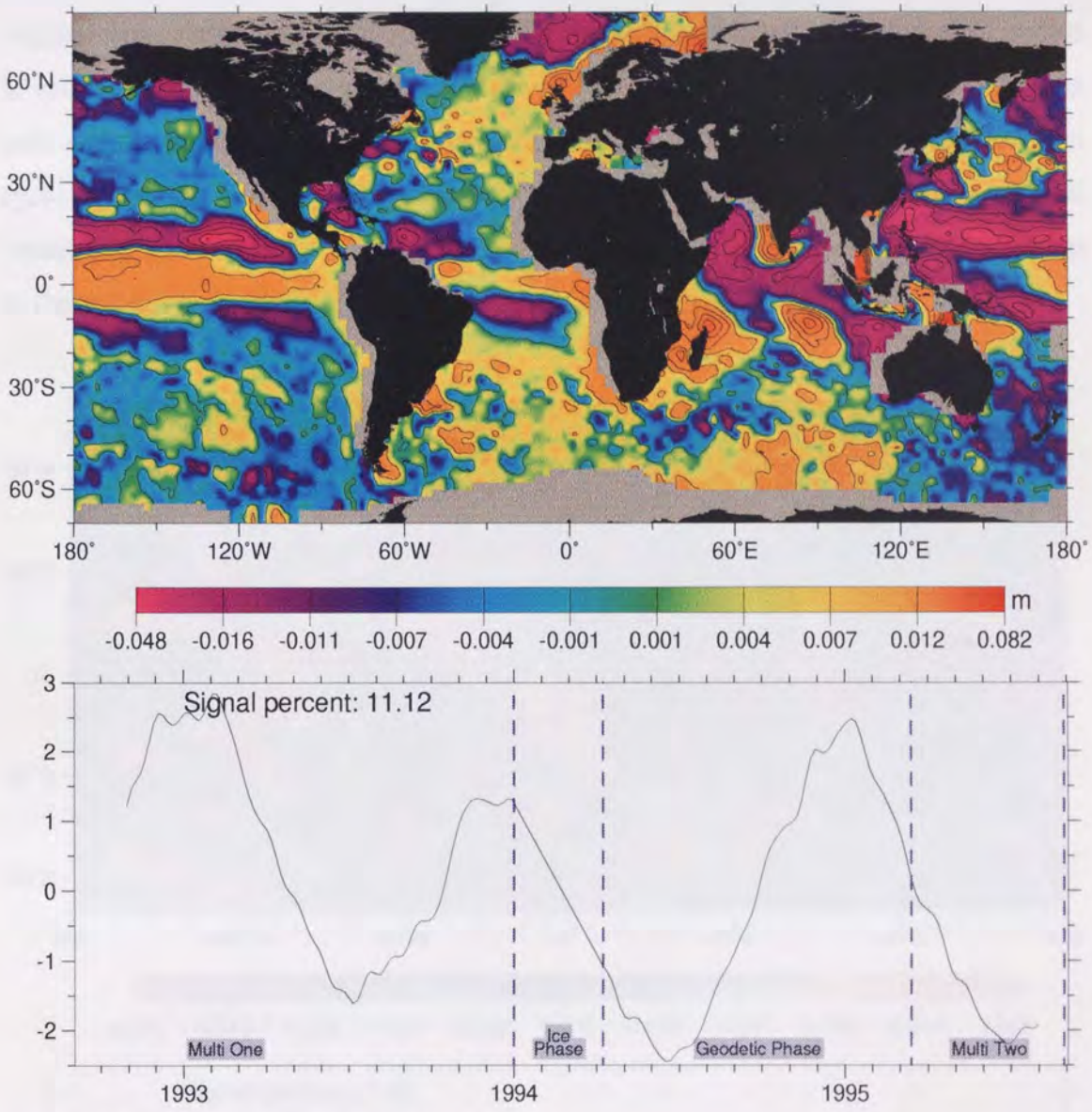

Figure 22. EOF 2 from the Combined data set.

This signal is consistent with the effects of the El Niño phenomenon, which changes the normal yearly pattern of winds in the equatorial regions. The easterly trade winds result in warm surface waters being pushed across the Pacific and building up on the western boundary. An associated upwelling in the east causes cooler surface waters to be prevalent here. In a normal year the trade winds weaken but remain easterly. During an El Niño, the trade winds weaken further, the upwelling stops and the warm pool in the west spreads eastwards as a Kelvin Wave. The trade winds then switch to wester-

lies and continue to feed the progress of this warm surface water eastwards. This water splits either side of the central eastward flow when it reaches the coast of America, and returns westward (as so called Rossby waves) in bands to the north and south of the original flow. At the end of the El Niño, the trade winds switch back to easterlies, return to full strength, and the warm water in the eastern pacific is forced westwards again as cold water upwells at the eastern margin (*Tomczak and Godfrey*, 1994). The El Niño cycle upsets the normal balance in the Pacific weather systems, and causes unusual weather on a global basis. Another signal evident in this plot is the east-west oscillation in the Indian Ocean (once again an El-Niño-related effect).

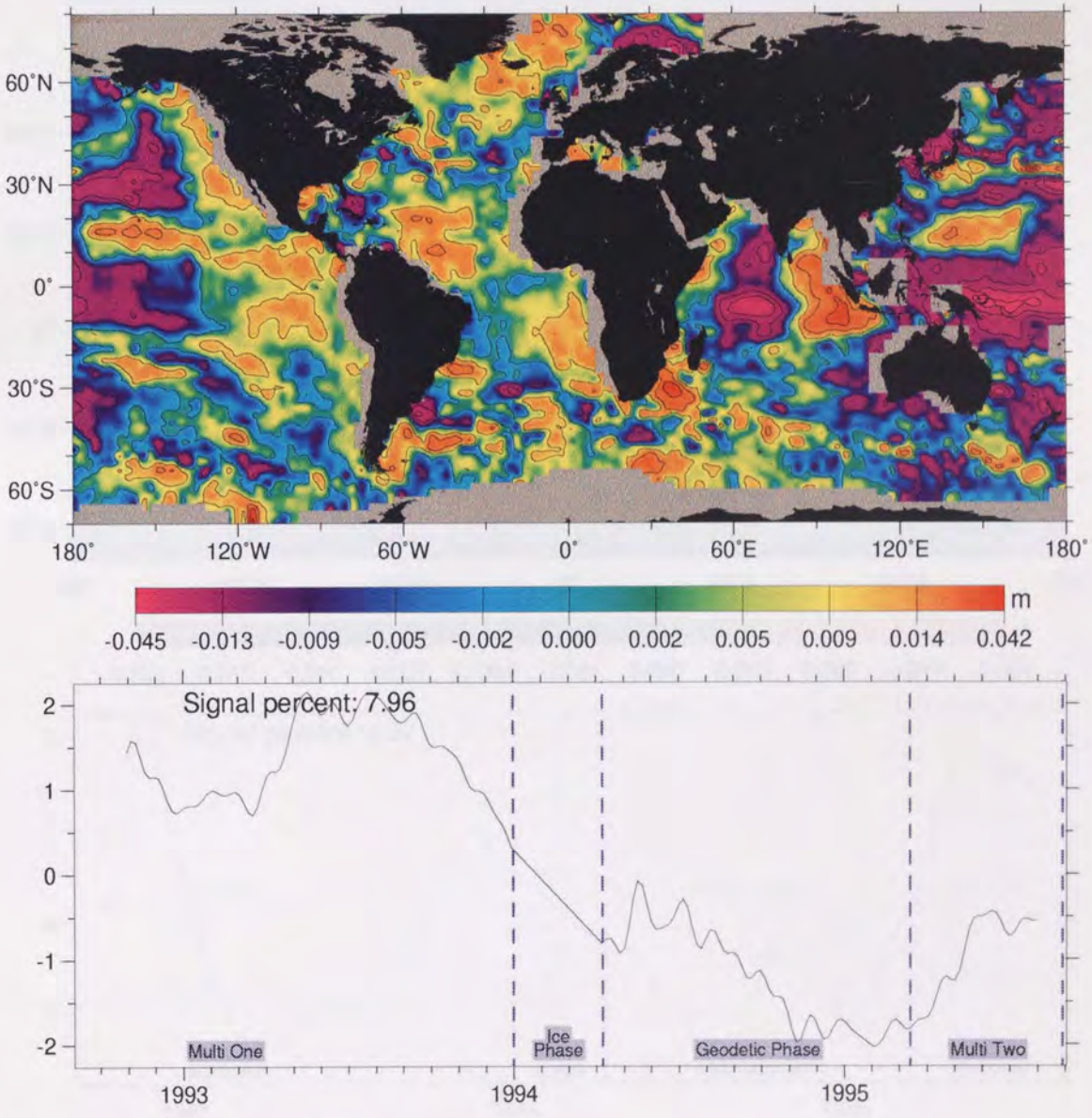

Figure 23. EOF 3 from the Combined data set.

EOF 4.

The spatial mode for this EOF (Figure 24) is similar in appearance to that of the previous EOF, suggesting this is also part of the El Niño cycle. The amplitude time series, however, shows an approximately 18 month cycle of some kind, though in the absence of the peak which may have occurred during the ERS-1 ice phase (phase D) it is difficult to specify its period. The spatial mode shows strong signals across the equators of both the Pacific and the Indian oceans. The amplitude time series also shows some oscillation with a period of about 35 days throughout, which is a result of residual ERS-1 orbit errors. All further EOFs contain some element of this orbit error signal.

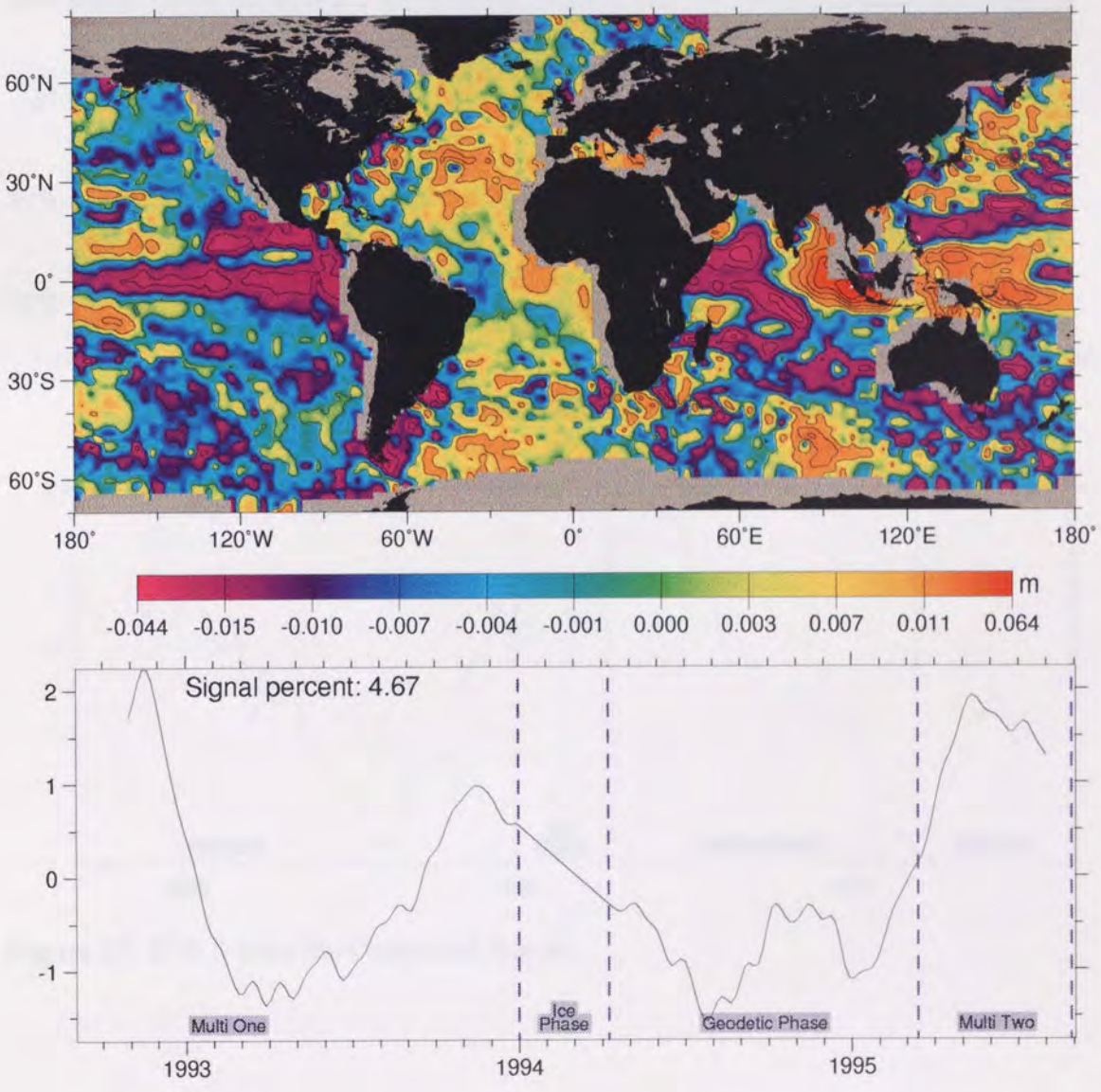

Figure 24. EOF 4 from the Combined data set.

EOF 5 (Figure 25) seems to be a complementary signal to that observed in EOF 4. The amplitude time series reveals similarly an approximately 18 month cycle (though with phase around 6 months behind that of EOF 4) and the spatial mode bears many similarities to that in EOF 4. The main difference between the spatial modes, however, is altered phase in the eastern and southern Indian Ocean and the Atlantic, with other regions showing the same phase as the EOF 4 spatial mode.

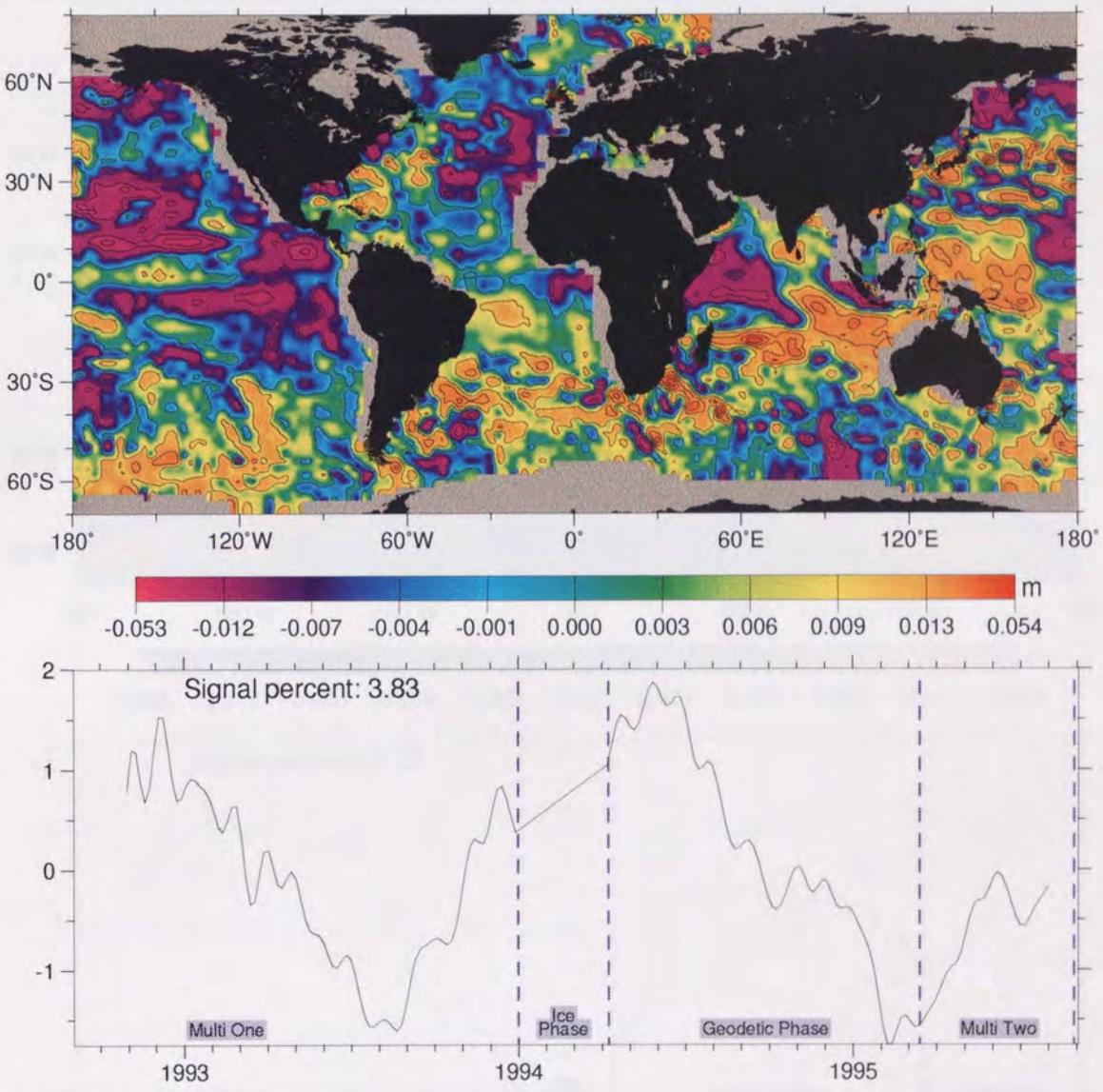

Figure 25. EOF 5 from the Combined data set.

EOF 6.

The amplitude time series (Figure 26) shows what may be a long period signal (with approximately two-year wavelength) though again the gap during the ERS-1 ice phase makes it difficult to decide upon the true nature of this signal. This EOF may represent a second component of the El-Niño signal, due to its long period. The spatial mode itself is not so clear, with most values remaining less than ±1cm, and highs only occurring in the Indian ocean and at the eastern and western boundaries of the equatorial Pacific (locations characteristic of El-Niño event maxima).

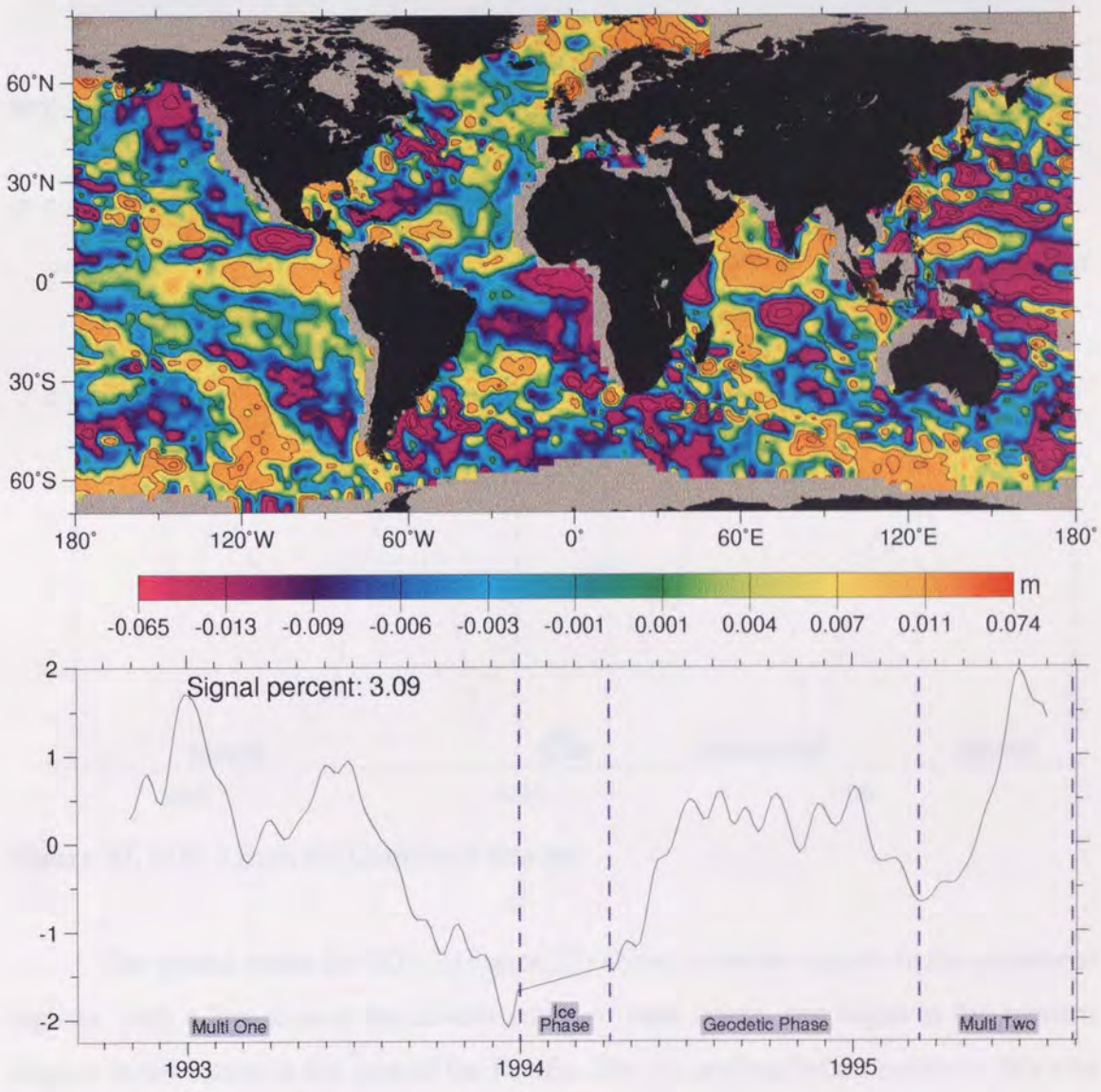

Figure 26. EOF 6 from the Combined data set.

EOFs 7 and 8.

These EOFs are both components of a semi-annual signal. The phase of the two signals differs by approximately one or two months.

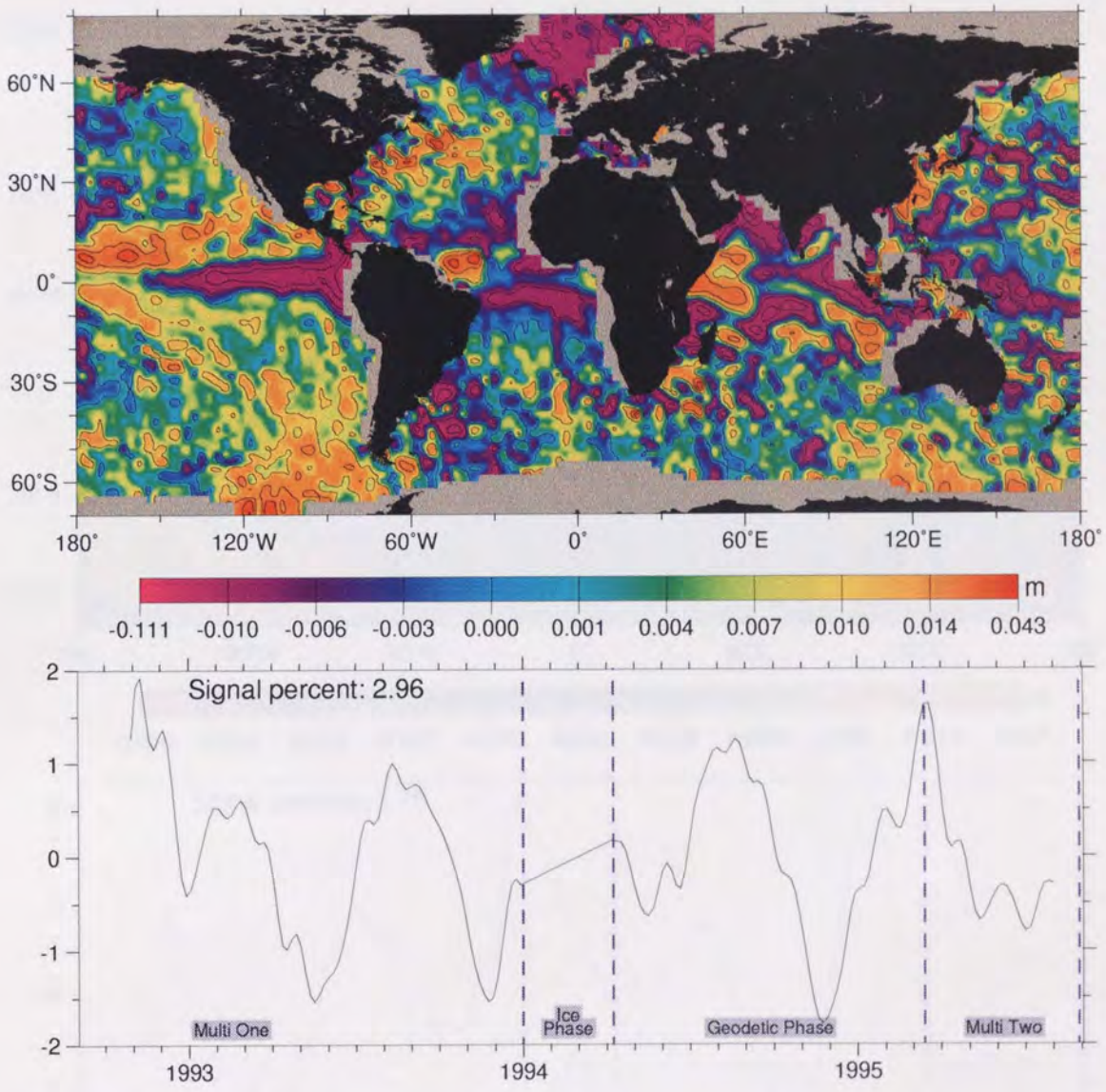

Figure 27. EOF 7 from the Combined data set.

The spatial mode for EOF 7 (Figure 27) shows coherent signals in the equatorial regions, with a low area at the eastern edges of each ocean, and highs at the western edge or in the centre in the case of the Pacific. The sea surface height peaks on this plot appear to follow the sub-tropical gyres in each ocean, i.e. they curve away from the equator as they travel eastwards. This holds for the signals present in EOF 8 (Figure 28) in the Pacific also, but the equatorial Pacific shows much finer patterns in this EOF.

Other oceans show different types of patterns in the two EOFs. EOF 7 appears to have high sea-surface height values in regions of the globe which are known for high variability, suggesting that much of this variability is a result of these 6-month cycles. Note it is possible that the semi-annual signal contains some component of the long periodic tidal modelling errors discussed in Section 2.2.4.

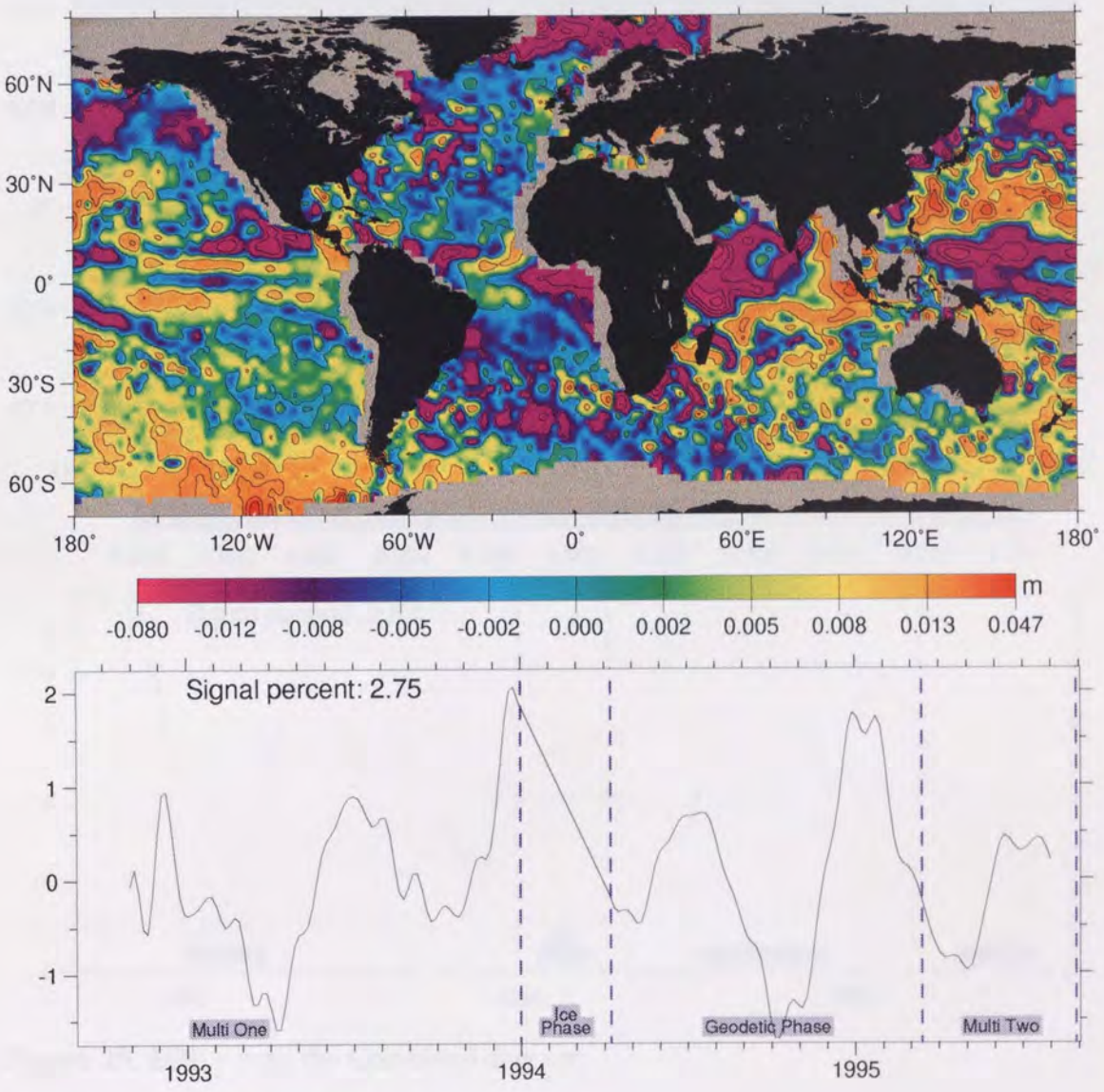

Figure 28. EOF 8 from the Combined data set.

EOF 9 (Figure 29) appears to have some coherent signals in the spatial mode in the Pacific, but the amplitude time series shows no consistent signal, and as such this EOF is most likely to be a result of some orbit error signal.

EOFs 10 onwards contain little coherence in either the spatial mode or the amplitude time series. These EOFs are difficult to analyse as a result, and much of the signal present is probably a result of residual ERS-1 orbit errors. The previously observed 35-day oscillation is present (strongly in some cases), which supports this hypothesis.

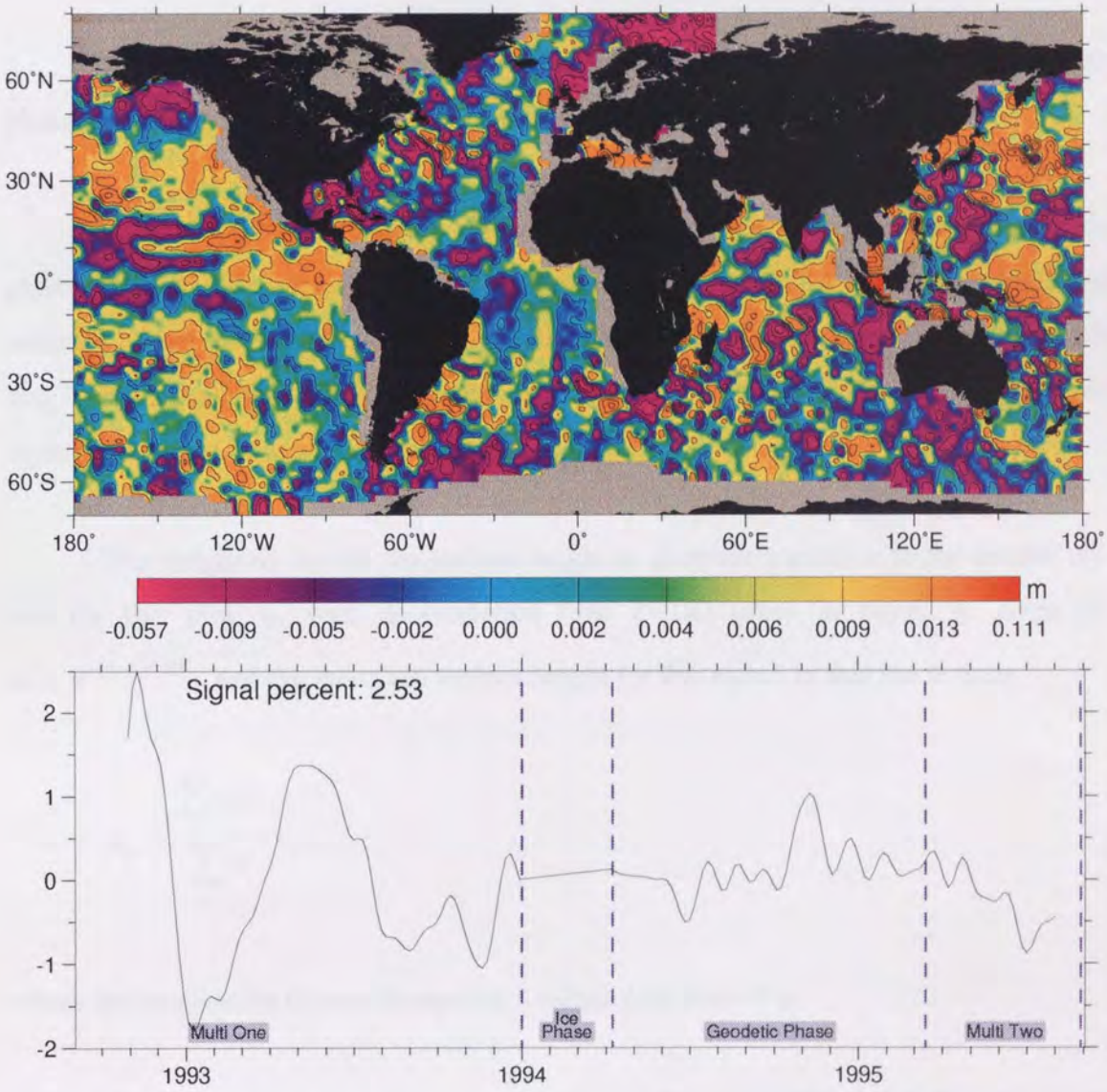

Figure 29. EOF 9 from the Combined data set.

9.5 Analysis Of The Optimally Interpolated Data.

This section contains observations of sea level change from global sea level anomalies derived from the combined corrected ERS-1 and T/P data set after temporal

smoothing by optimal interpolation. These observations provide support for those based upon the EOFs in the previous section.

9.5.1 The Combined ERS-1–T/P Optimally Interpolated Data.

The combined data set includes 1.3 million observations, which represents 200 global sea surface anomaly maps (one every five days for nearly three years).

In an effort to present this enormous data set in a systematic way, quarterly global plots were produced. These were generated by producing a Gaussian-weighted mean of all sea-surface height observations within 45 days of the central epoch in each bin. The decorrelation time for weighting purposes was set to half the time span between epochs, i.e. approximately 45 days.

The weight w_i for the sea surface height h_i at epoch t_i relative to the central epoch for this plot, t_0, with decorrelation time D (all times in days), is given by $w_i = e^{-(t_i - t_0)^2 / D^2}$, and the mean sea surface height for this epoch in this bin is then

$$h_{t_0} = \frac{\sum_i w_i h_i}{\sum_i w_i}$$

where the summation is over the epochs t_i within $D/2$ days of t_0.

This procedure reduced the number of plots from over 200 to 12. The drawback of this reduction is that short term sea surface variability is ignored, but this was considered a reasonable compromise due to the amount of data involved.

This reduced data set is dominated by the annual oscillation. The plots in Figure 30 are sea surface anomalies approximately six months apart (MJD 49045 and MJD 49225), and they clearly show the north/south oscillation on a global basis. In order to make it easier to see the underlying signals, this annual variation was removed. Given

that an Empirical Orthogonal Function decomposition has been carried out and the first two EOFs usually represent the dominant annual signal, it is possible to subtract these from the data and thereby remove this signal. However, as has been seen in Section 9.4 above, these EOFs may also contain other signals and other EOFs may contain parts of the annual signal (the EOF procedure does not isolate temporal modes but rather spatial modes) and hence this procedure is not ideal. The preferred method is to remove an empirically solved sinusoidal annual signal from each bin via a least squares procedure.

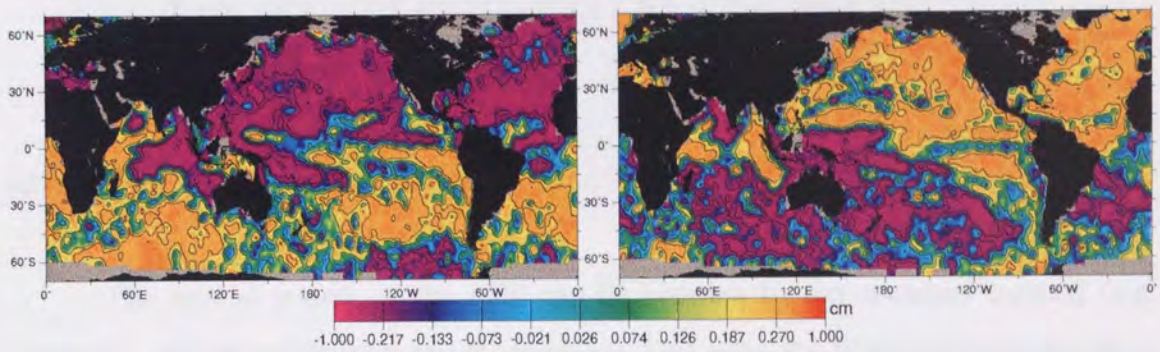

Figure 30. Examples of sea surface anomaly plots before annual filtering. The left plot is as at MJD 49045 (27th Feb 1993) and the right at MJD 49225 (26th August 1993).

9.5.2 Removal Of Annual Signal By Least Squares.

An observed sea-surface height time series y_i at time x_i (in years) is assumed to include a sinusoidal function

$$y_i = \alpha \sin(x_i + \beta). \qquad 28.$$

In Equation 28 α represents the amplitude of the sinusoidal and β the phase. This can then be rearranged to the form in Equation 29.

$$y_i = A \sin(x_i) + B \cos(x_i) . \qquad 29.$$

Therefore, to find a least squares fit to this data,

$$Z = \sum_i \left(A\cos(x_i) + B\sin(x_i) - y_i \right)^2$$

must be minimised.

Upon taking derivatives with respect to A and B, setting equal to zero and writing in matrix representation, Equation 30 is derived.

$$\begin{bmatrix} \sum \cos^2(x_i) & \sum \cos(x_i)\sin(x_i) \\ \sum \sin(x_i)\cos(x_i) & \sum \sin^2(x_i) \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \sum y_i \cos(x_i) \\ \sum y_i \sin(x_i) \end{bmatrix} \qquad 30.$$

(all summations are over the i observations).

The normal matrix on the LHS can be inverted by any standard method (e.g. Choleski), and the values of A and B calculated. This gives the fit for each current set of x and y values. Once values for A and B are known, the sea surface height values for the current bin are adjusted to remove this annual signal and the data are written to a new file.

9.5.3 Discussion of the Annual Signal.

The nature of the annual signal can be studied more closely as a by-product of this procedure. The solved A and B values for each bin (from Equation 29) can be converted to the form given in Equation 28 to reveal an amplitude (α) and phase (β) in each bin. These values can be plotted globally to show more information about this particular signal. The plots in Figure 31 and Figure 32 show respectively the amplitude and phase of the annual signal found in the optimally interpolated data set.

Figure 32 clearly shows the differing phase between the Northern and Southern hemispheres. As expected due to thermal expansion effects from solar heating, the Southern hemisphere shows a March or April peak and the Northern Hemisphere shows

a peak in September or October (*Pattulo et al.*, 1955). More complex patterns are observed on and around the equator across the globe.

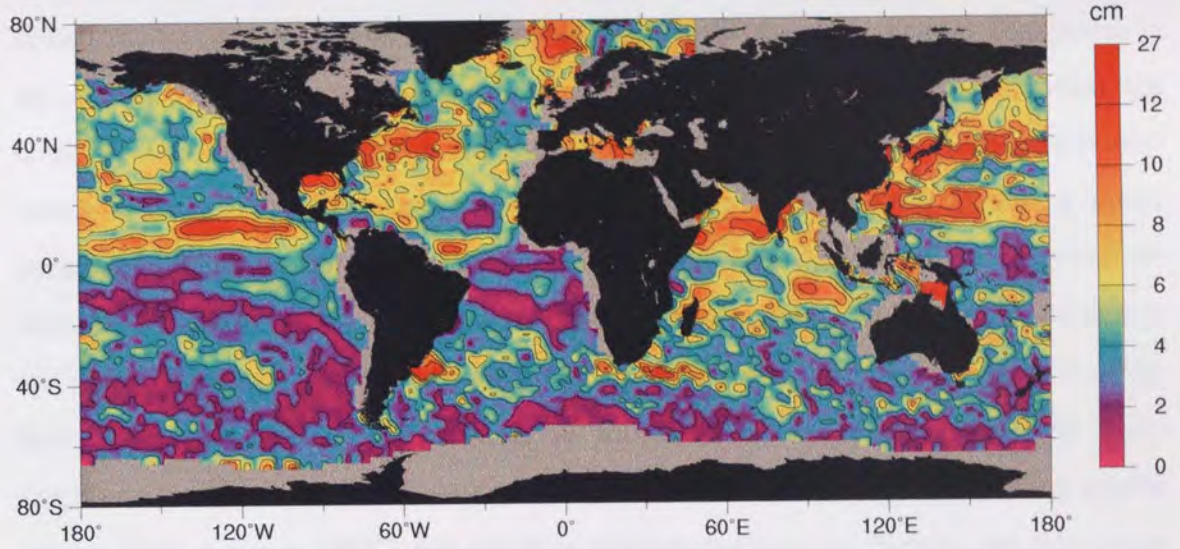

Figure 31. Amplitude of annual signal in cm.

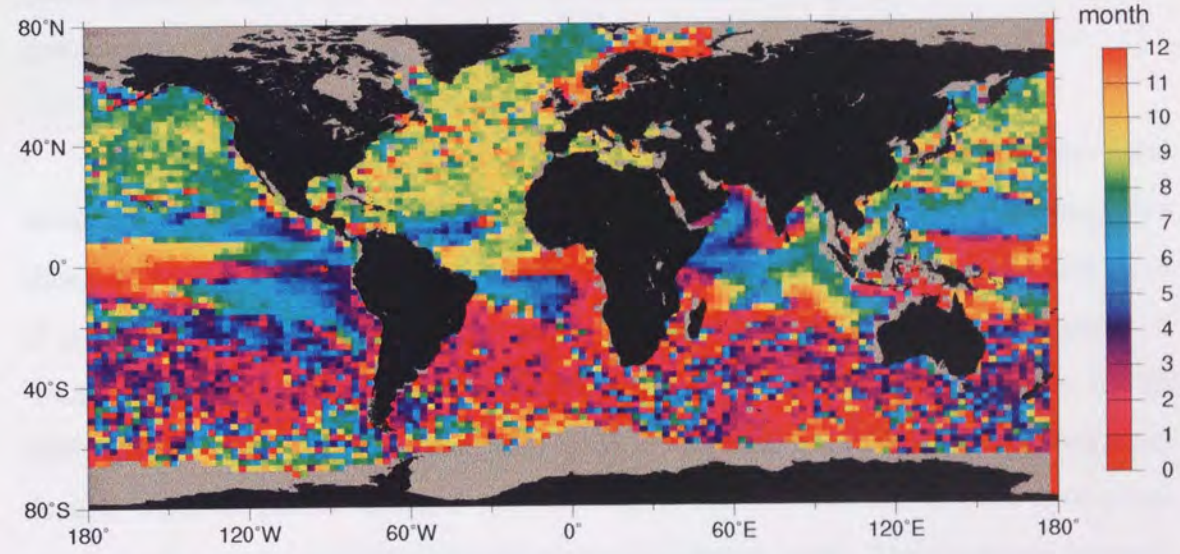

Figure 32. Peak phase of annual signal in months. Scale starts on January 1st.

The equatorial Pacific shows an early year peak (around January or February) in the western and central areas, moving towards a mid-year peak on the eastern boundary. The rest of the Pacific, however, follows closely the expected thermal expansion patterns. Amplitudes in the Pacific Ocean (Figure 31) show peaks along a line at about 10°N across the entire ocean, coinciding with the latitude bands most strongly affected by the trade winds and their annual cycle.

The Atlantic Ocean shows similar phase changes to the Pacific above and below the equator, but the characteristic patterns occur 5-10° South due to Africa's land mass intrusion into the ocean causing the current paths to be diverted. Annual amplitudes in the Atlantic peak at the equator (again likely to be a result of the trade wind cycle), the north west (in the region of the Gulf stream) and the extreme north. This area in particular is likely to experience extreme seasonal changes since ice coverage during winter periods produces corrupted altimetric measurements. Note that the values observed are nonetheless unusually high and are thought to be an artefact of the DXO correction technique (Chapter 4): since there are no DXOs in this region, the cubic spline is effectively unconstrained. This hypothesis is supported by the presence of this high-amplitude feature in all DXO-corrected ERS-1 data sets (including those using orbital heights from Aston or DUT), but not in an uncorrected data set. Neither the tide gauge analysis in Chapter 4 nor the DXO analysis in Chapter 5 are likely to pick up this particular problem due to the low-latitude concentration of tide gauges and the latitudinal spread of the DXOs.

The Indian Ocean has a strong seasonal cycle due to the dominance of the monsoons, as is clear from the high values observed on the amplitude plot. The phase plot shows differing phases across the western Indian Ocean, with peaks near the east coast of Africa in May or June and on the western coast of India in December or January. A similar phase shift can be observed in the eastern Indian Ocean. These effects are all likely to be results of the monsoon wind cycle. The southern area of this ocean has a noticeably different pattern which is probably due to the southern sub-tropical gyre transporting colder water north toward the warmer equatorial water. The boundary between the two is visible running from northern Madagascar to the south western tip of Australia. Similar effects are observed in both other ocean basins with similar causes.

High annual signal amplitudes are observed in several locations, generally concentrated in the Northern hemisphere. This is expected: the large land masses in the northern hemisphere exacerbate the annual thermal cycle within the atmosphere, which feeds variability in the oceans. High amplitudes can also be seen around 10°N across the Pacific; in much of the Indian Ocean; in the north western Pacific and north western

Atlantic (both are intense western boundary currents so have similar variability patterns); extreme north Atlantic (greater than 70° N, though see note above) and off the eastern coast of South America at approximately 40°S. All these are typically regions with high variability.

After the annual signal filtering described above, variability within the data set is much reduced in certain areas. The full global variability pattern (after optimal interpolation) is plotted in Figure 15 and as a comparison the same image after removal of the annual signal is plotted in Figure 33. The mean RMS variability globally is 5.78cm, and after removal of the annual signal this drops to 4.92cm, a change of approximately 3cm RMS or around 27%. The EOF decomposition performed on the same data set shows that the first two EOFs (the two spatial modes dominated by the annual signal) were responsible for 29.66% of the overall signal. This suggests that these two EOFs do not therefore represent *only* the annual signal, and that (as hypothesised above) there are other signals which may also be represented, but further analysis would be required to confirm this. This discrepancy may also be due in part to the effect of long-periodic tidal modelling errors (§2.2.4) or sea state bias modelling problems (§2.2.2).

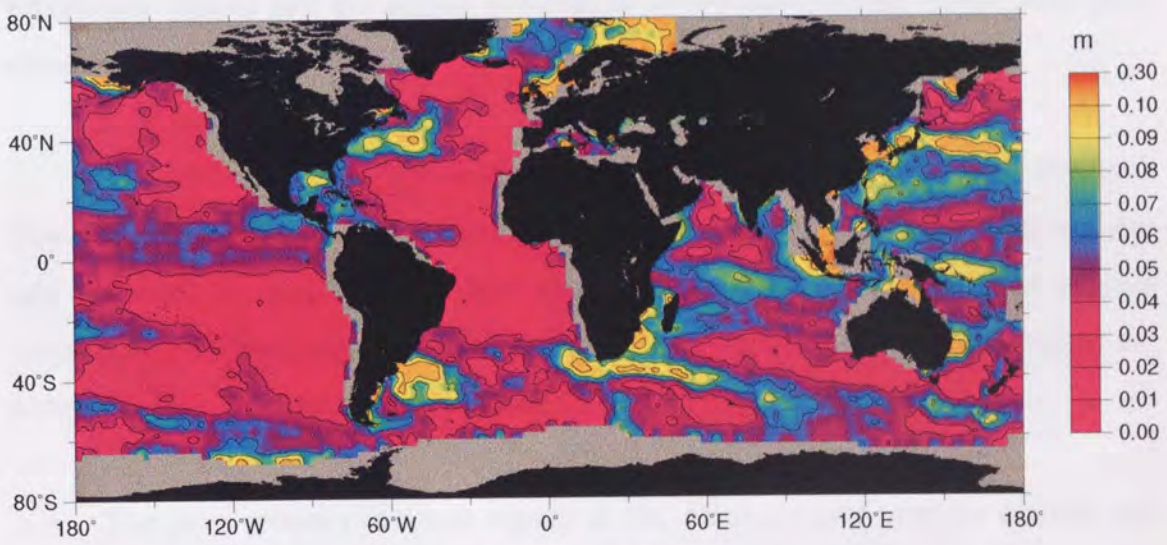

Figure 33. Global RMS variability after annual filtering.

The most significant effect visible upon comparing these plots is the difference in signal strength in the equatorial regions. The removal of the annual signal has re-

moved most of the variability from the equatorial Pacific and Atlantic and has significantly reduced variability in the northern Indian Ocean.

9.5.4 Analysis of Sea Level Anomalies.

Global plots of quarterly mean sea level anomalies after removal of the annual signal are presented in Figure 34, Figure 35 and Figure 36. These 12 individual plots represent a time period of approximately three years, from 1992.73 through to 1995.71. Each plot is produced using the same colour scale and contours are marked at 0.1m.

Noticeable in these plots is the continual change in the far south, due to the presence of the ACC. This is difficult to interpret further because of the high variability, which seems to be particularly strong in the far southern areas between 30° and 120°E.

A further obvious signal in these plots is on the equatorial Pacific. This area exhibits continual complex east-west oscillations, but shows no consistent long-term repeat pattern. This effect is a combination of the previously observed El Niño long-wavelength signals and the shorter wavelength semi-annual signals which also occur along the equator.

The northern Pacific (and similarly the northern Atlantic) shows little coherence. These areas have low variability in the central-to-eastern regions, and as such any signals are small in amplitude and short lived. The western boundary currents of each ocean, however, have high variabilities. As a result, signals in both of these areas are difficult to interpret in plots such as these.

The most spatially coherent signals in the Atlantic occur near the equator and extend as far as 20° south. These signals are also likely to be El Niño related, since the effects which signify the El Niño are found globally. The southern Atlantic is affected by the ACC to a large extent, creating high variability which is incoherent on plots such as these due to the comparatively long time span between images.

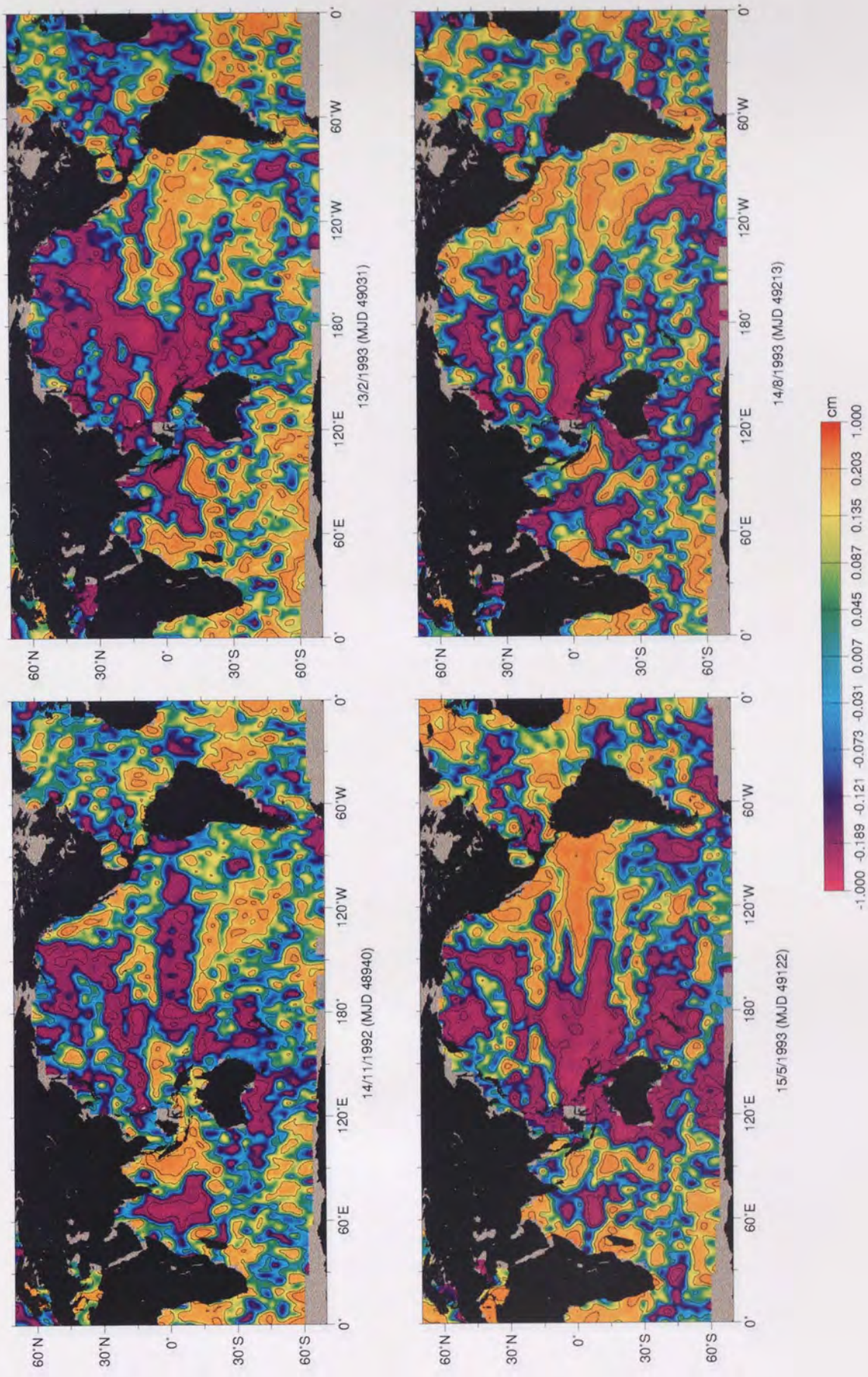

Figure 34. Sea surface anomaly plots 1 to 4 (November 1992 to August 1993).

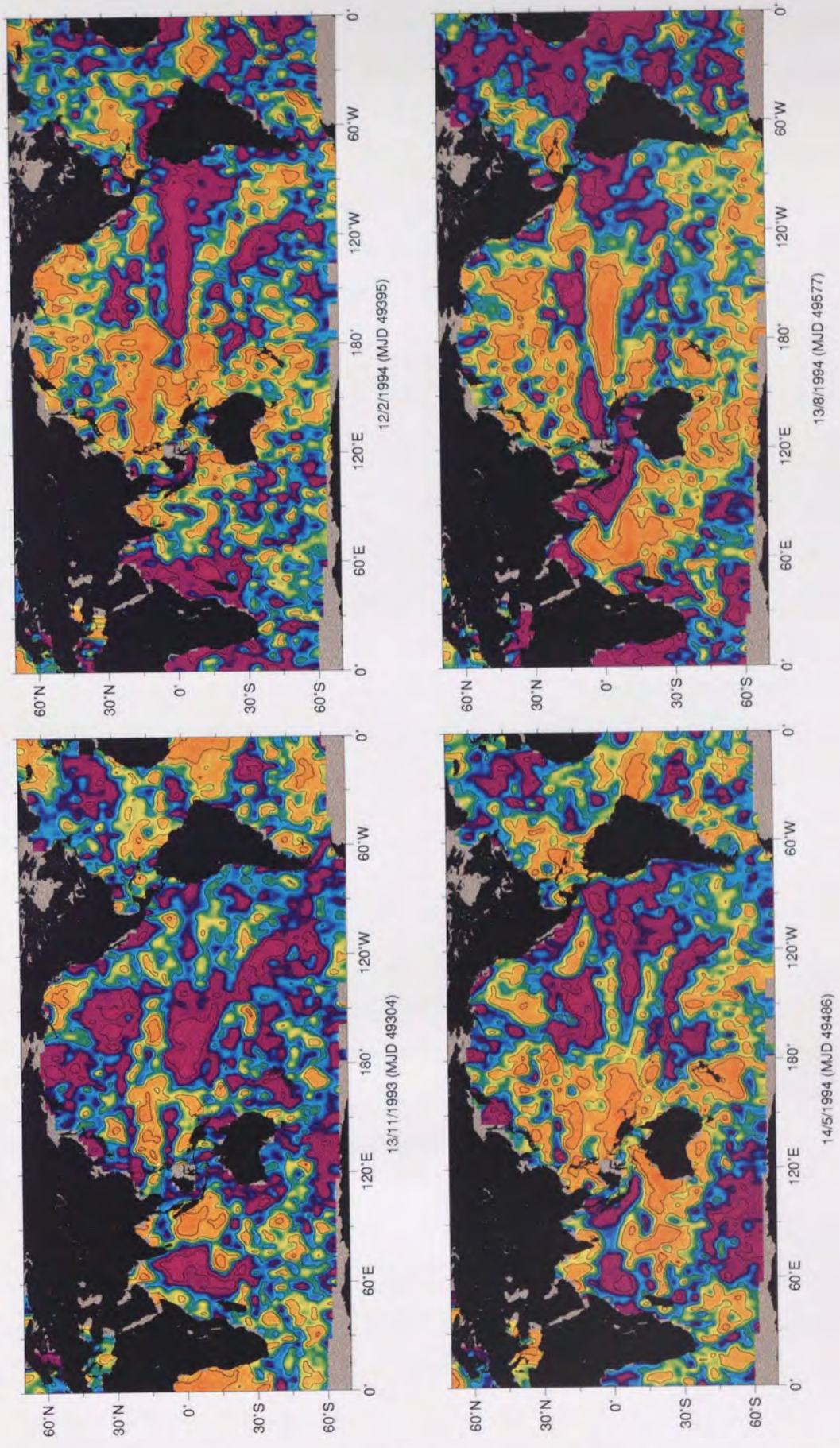

Figure 35. Sea surface anomaly plots 5 to 8 (November 1993 to August 1994).

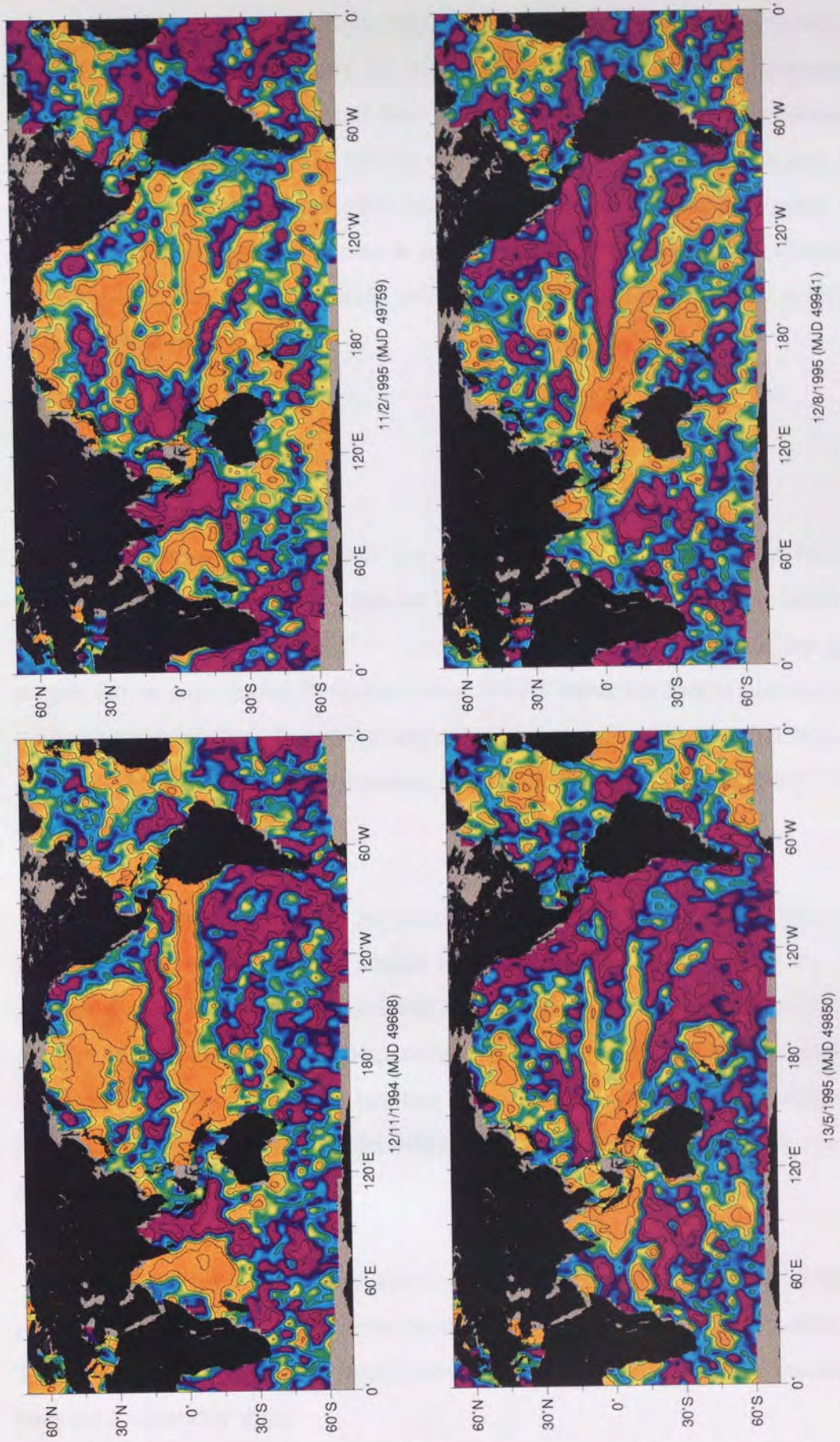

Figure 36. Sea surface anomaly plots 9-12 (November 1994 to August 1995).

The Indian Ocean displays complex patterns during this time. It appears to alternate between a straightforward east-west tilt and a more complex east-central-west variation. The north-south axes of these tilts are also mobile and vary in position by up to 30°, from 60°E (e.g. at MJD 49395) to 90°E (e.g. at MJD 49668). The area is highly dynamic but appears to contain quite coherent signals. As discussed previously, this area is dominated by the annual monsoon season, but a strong 6-month oscillation is also prevalent, and this in combination with the long term El-Niño effects results in the changes found in these plots.

9.6 Conclusions.

It has been shown that the sea surface anomaly data set derived from ERS-1 crossovers is a complementary data set to that from T/P. Both data sets have inherent strengths and weaknesses. The T/P repeat pass data are highly accurate, but problems remain due to gaps during Poseidon cycles. ERS-1 crossover data is also accurate and used alone would allow the production of a finer resolution data set, but still contains remnants of 35-day orbit related errors, along with the difficulties in spatial resolution during phase D.

The crossover technique has proven capable of providing information from the ERS-1 mission where previous studies have been unable to, i.e. during the Geodetic mission. Furthermore, the data produced are not poor relations to that from highly accurate T/P repeat pass data. The comparative studies in section 9.2 have shown that there is only a 1.5cm RMS difference between the two sources of sea-surface anomaly information, and that the majority of the differences occur at times when one of the two satellites is providing no data.

If ERS-1 orbits were improved to the extent where ERS-1 data are of truly comparable accuracy to T/P data, without the need for some kind of DXO correction *against* T/P, then this SXO technique would allow the use of the whole ERS-1 mission as the basis for a variability study.

The combination of the ERS-1 crossover data and T/P repeat pass data have proven, through comparison with known oceanographic truths (by both EOF analysis and observation of raw sea surface anomalies), that they are reliable and flexible. When ERS-1 and T/P data are combined, the sheer quantity of data is enough to control errors effectively within either data set, and provide accurate and reliable information.

Of concern, however, is the presence (in Section 9.2.1 and 9.2.2) of an apparent first-order spherical harmonic difference between the ERS-1 data and the T/P data. This suggests a first-order gravity modelling problem (as was also observed in chapter 5), which could be resolved by the use of a newer model such as the DUT DGM-4 or the model produced at Aston (*Boomkamp and Moore,* 1997). Further possible causes of this type of difference could be the use of different reference frames for (e.g.) the SLR tracking data or a differing geophysical correction (as listed in Chapter 2) for each satellite. These would require further analysis and are worthy of future study.

10. Global Rossby Wave Observation

10.1 What Is A Rossby Wave?

Rossby waves have been observed by many altimeter missions. They are observable as a long wavelength, fairly slow moving wave pattern which propagates in a westward direction almost exactly along a line of latitude - there is little latitudinal drift.

The name "Rossby" was coined after the discoverer of the phenomenon and the name is accepted for general use but a more accurate title would be geostrophic (or planetary) eddies. Rossby waves can occur in both water or air movements and are responsible for transferring large quantities of energy around the Earth's eco-system.

Further information on Rossby waves (in addition to the summary presented in Sections 10.2 and 10.3) can be found in many texts such as *Chelton and Schlax* (1996), *Hughes* (1995) or *Polito and Cornillon* (1997).

10.2 Vorticity

Rossby waves form as a result of the conservation of potential vorticity. Vorticity is the 'tendency to form vortices' - in some ways analogous to angular momentum. Any current shear (a change in velocity at right angles to the direction of flow of a current) has a component of vorticity and therefore has a tendency to rotate. By convention, clockwise rotation is considered to be negative vorticity, and anti-clockwise rotation is positive. Note that vorticity is a separate concept from actual rotation since vorticity can be caused by different effects, and two opposing effects could be causing separate *vorticities*, whilst cancelling each other out and there being no overall *rotation*.

Planetary vorticity is that provided by the rotation of the Earth. Any point on the Earth's surface has an angular velocity proportional to the rate of rotation of the Earth

and the latitude at which the phenomenon is observed. At the poles there is a maximum of angular velocity equal to the angular speed of rotation of the Earth about its axis, Ω ($7.29 * 10^{-5}$ s^{-1}), whereas at the equator there is none. In general the angular velocity, ω, at latitude ϕ is given by

$$\omega = \Omega \sin \phi$$

and the vorticity of a parcel of fluid is defined to be twice this angular velocity. Planetary vorticity is given the symbol f.

Relative vorticity (ζ) is the vorticity of a parcel of fluid relative to the Earth's surface below it. The total or absolute vorticity is the vorticity relative to free space which is the sum of the planetary and relative vorticities, and this is the quantity which is preserved, i.e.

$$\text{Absolute Vorticity} = (f + \zeta) . \qquad 31.$$

Since the absolute vorticity must remain constant, Equation 31 shows that any increase in planetary vorticity brings about a decrease in relative vorticity and vice versa e.g. if a parcel of water with zero relative vorticity moves away from the equator and gains planetary vorticity, it must also gain increasingly negative relative vorticity and it will start to rotate.

In the oceans, a further complication applies in that the conserved vorticity is the sum of the vorticities of all the particles which are rotating. If the case of a column of water is considered and the column increases its depth, the volume remains unchanged, and the column thus becomes narrower. The average radius about the rotational axis will decrease, and in order to retain angular momentum, the rotational velocity must increase (c.f. an ice skater in a spin). Because of the changes in length of the column of water, the actual quantity which is conserved is not the absolute vorticity but the potential vorticity, or $(f + \zeta) / D$ where D is the depth of column involved.

In fact, in the open ocean away from boundary currents and other areas of strong current shear, $f \gg D$ and thus any change in column depth will force a change in f and hence latitude much more than affecting the local relative vorticity.

10.3 Vorticity and Rossby Waves

If a parcel of water at latitude ϕ in the northern hemisphere initially has no relative vorticity and no current shear, and is displaced northwards, it enters a region of higher positive planetary vorticity. Its potential vorticity must remain constant so its relative vorticity will become more negative to compensate and the parcel will start to rotate clockwise. The opposite occurs if it moves southwards, and overshoots its original latitude, i.e. the parcel will rotate anti-clockwise. This parcel of water can swing back and forth about its original latitude, consecutively gaining negative and positive relative vorticities, therefore rotating clockwise and anti-clockwise as it progresses. A row of such parcels in a current or airstream is what is commonly observed as the Rossby wave phenomenon, oscillating north and south of the original latitude. In the ocean, such oscillations will only be to a scale of hundreds of kilometres, whereas atmospheric Rossby waves can oscillate by up to 20,000 km. It is the continual changes in rotation which cause these features to propagate westward at rates of between about 1.27 ms^{-1} at 5° from the equator, 0.08 ms^{-1} at 20° N or S and around 0.02 ms^{-1} at 40° N or S. Therefore a Rossby wave at 5° N will take approximately six months to cross the Pacific, but over 20 years at 40° N.

10.4 Rossby Wave Extraction From The Optimally Interpolated Data Set.

The most straightforward method for location and analysis of Rossby waves in a sea level anomaly data set (such as the optimally interpolated data set) is to construct plots showing the anomalies along a line of latitude with time as the Y-axis. On such a time–longitude plot, any westward moving anomaly will show up as a clear diagonal signal from which the wave speed can be calculated.

The full global data set was filtered to remove any data pertaining to bins that were outside the required latitude band. At this stage, much filtering had already been applied to this data set, so it was unnecessary to add any further criteria. A single latitude band of bins was used as the basis for this type of plot to avoid visual confusion over pixel sizes. The plot was then created using longitude as the X axis and time as the Y axis. In plots such as this, many signals were visible at once, not the least of which was the dominant annual signal. In order to clarify the plots, it was useful to remove the annual signal from the data, which was carried out via the least squares fitting of a sinusoidal signal to the data (as in section 9.5.2).

The reduced signals in the data then provide clearer indications about the location and propagation rates of Rossby waves. Plots for each basin were produced along all available latitudes between 66° S and 66° N using these techniques and searched visually for signs of Rossby wave patterns. Any such patterns found were analysed to ascertain the location and approximate propagation rate of the observed waves.

10.5 Results

The full plots are shown for several sets of Rossby wave observations. It is not feasible to show all such plots, so a selection of examples are shown in Figure 37, Figure 38, Figure 39 and Figure 40. A global plot of the wave speeds found at each geographical location is given in Figure 41, where the speed is indicated by the colouring of each 5° by 2° pixel. This information is smoothed and resampled onto a finer grid before display (see *Wessel and Smith*, 1991). Contour lines are displayed at 0.05m intervals.

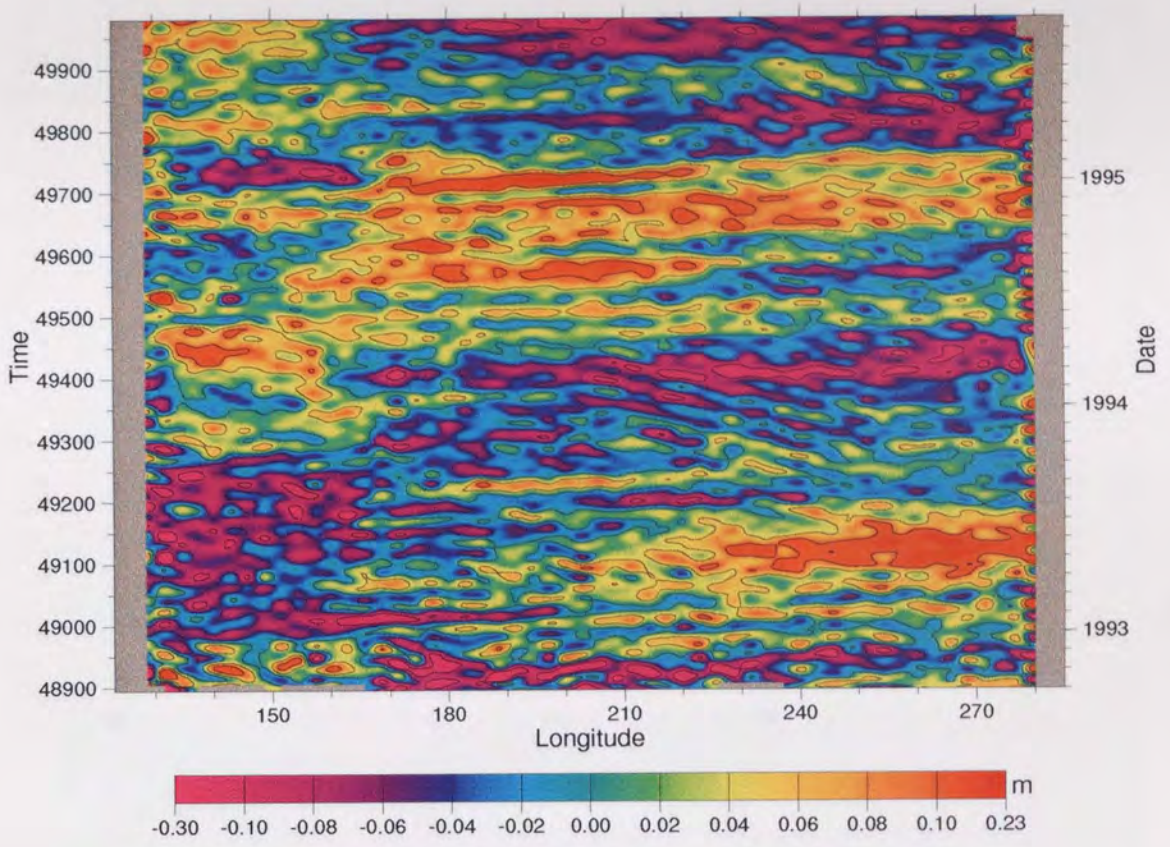

Figure 37. Time/Longitude plot along the equator in the Pacific.

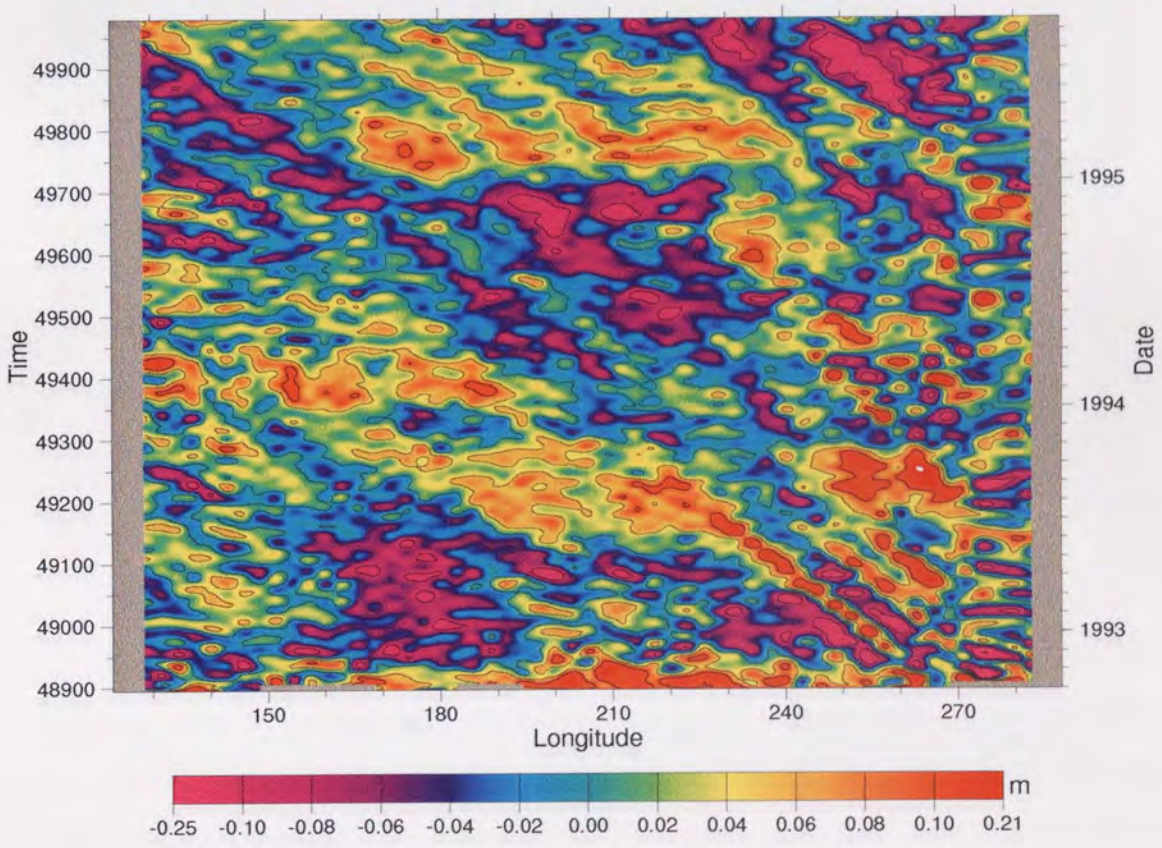

Figure 38. Time/Longitude plot at 10° north in the Pacific.

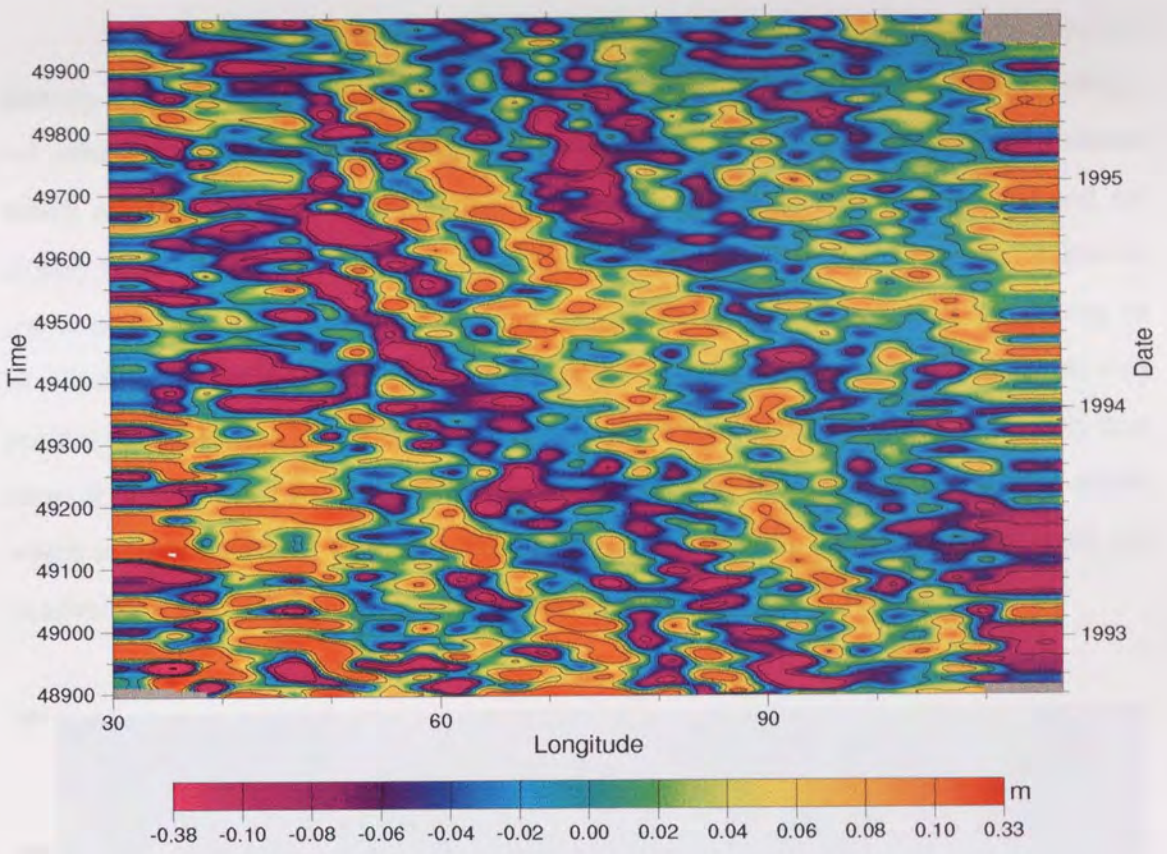

Figure 39. Time/Longitude plot 25° south in the Indian Ocean.

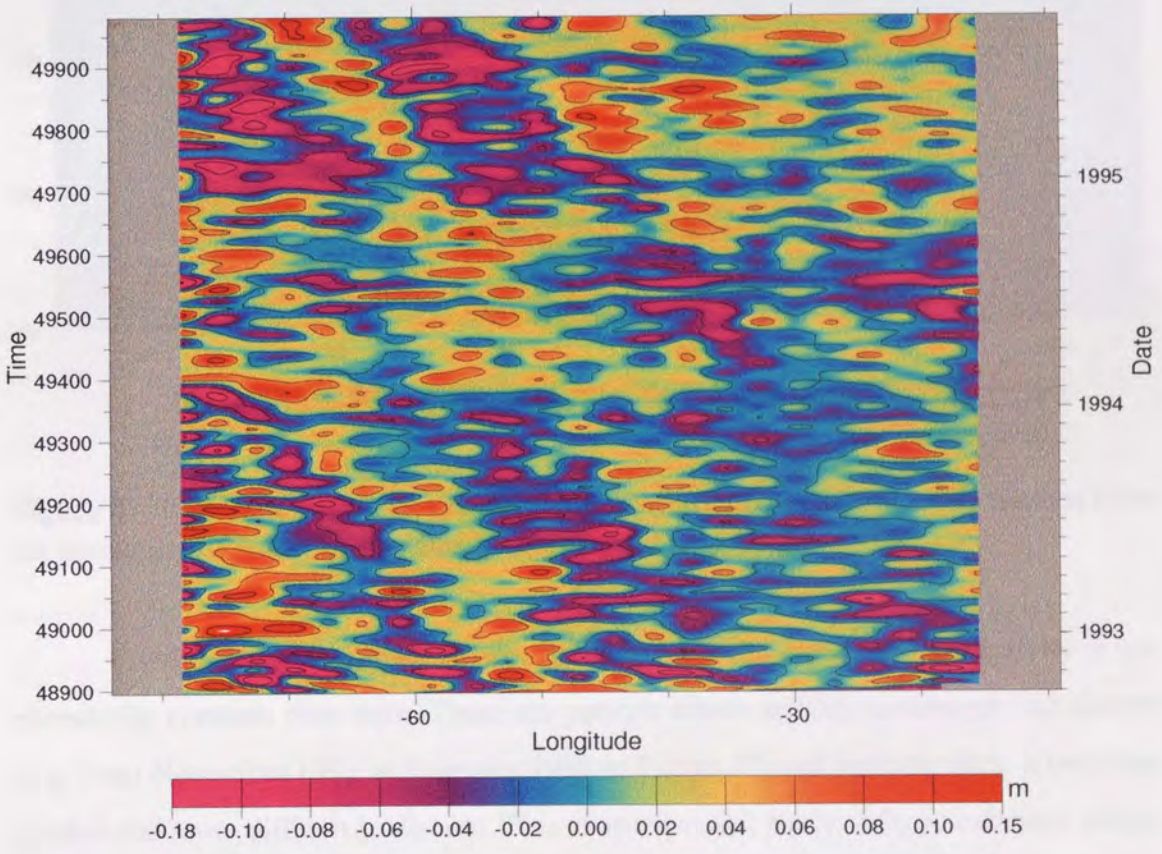

Figure 40. Time/Longitude plot 30° north in the Atlantic ocean.

A further feature of the figures for latitudes near the equator is a rapid eastward moving anomaly representing the progress of so called 'Kelvin Waves' (e.g. the diagonal effects at about 5-10° from the horizontal visible in Figure 37). These are features which are found mostly in equatorial regions (in the equatorial wave guide) and are closely related to Rossby waves in that one feeds off the other and vice versa. A Kelvin wave, usually produced by unusual wind stresses along the equator, possibly during an El Niño event, moves eastwards rapidly to the eastern boundary of its basin. At this point it may split into two poleward moving coastal Kelvin waves which eventually turn away from the coast as the thermocline shallows, and become diffuse Rossby waves which radiate westward from this Eastern boundary. Kelvin waves are accelerated due to advection from the equatorial undercurrent and travel at up 2.5 m/sec.

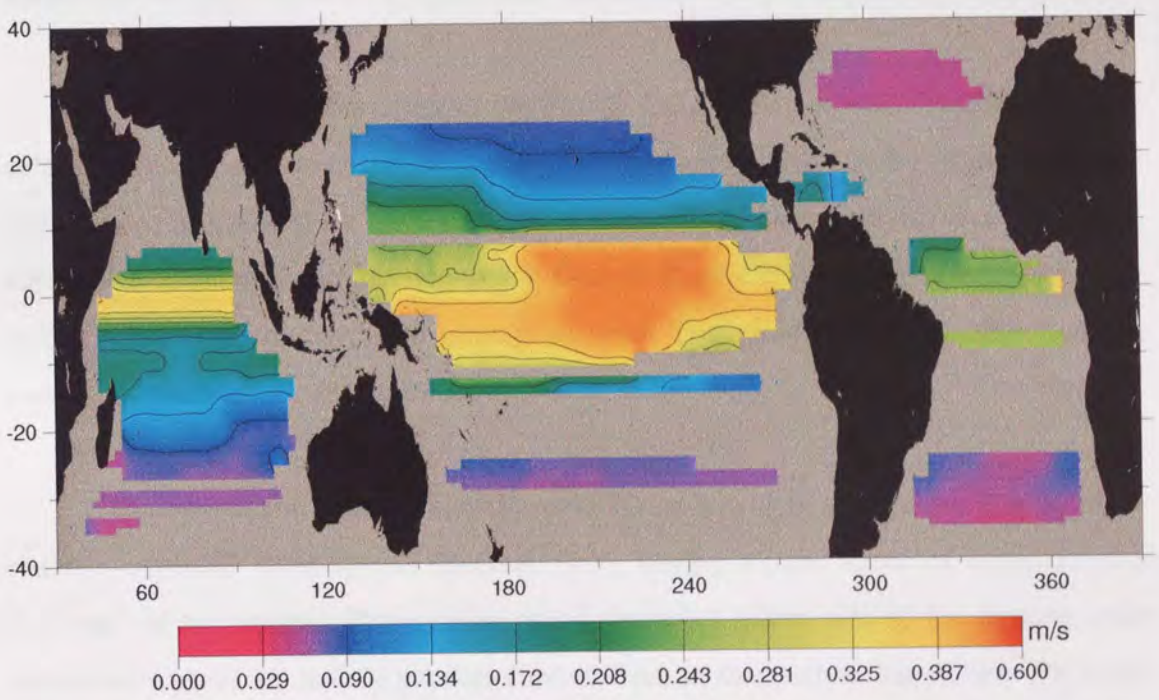

Figure 41. Locations and wave speeds of observed Rossby-Wave type phenomena from the combined ERS-1 & T/P optimally interpolated data set.

From the time-longitude figures it is clear that any Rossby Wave activity is not necessarily constant over time. There are periods where activity is stronger and clearer (e.g. from November 1993 to February 1994 in Figure 37) and periods when it becomes weaker and more difficult to discern. This is as expected, particularly in locations where a strong annual signal is present (such as in the Indian Ocean), which will affect both Rossby wave production and propagation patterns.

10.6 Discussion.

As can be seen from Figure 41, the region showing most rapid propagation of Rossby waves is near the equator, as is expected from the comments in section 10.3, with much slower transport occurring at latitudinal extremes. This is due also to the existence of the equatorial wave-guide, an effect whereby waves tend to travel faster in a narrow latitude band along the equator than one would expect by extrapolating from the observed rates at other latitudes.

The maximum velocity observed for Rossby wave transport is around 0.4 ms^{-1}, which is in contrast to the quoted value of 1.27 ms^{-1} in section 10.3. Currently no explanation can be found for this inconsistency, and further work is required to resolve this.

Further effects are evident in the Pacific Ocean (PO) as well. Rossby waves accelerate towards the centre of the Ocean, but then slow down in the Western Pacific. This pattern is repeated for most latitudes in the central Pacific and the deceleration occurs at approximately the location of the Central Pacific Basin, suggesting some interaction between water depth and propagation rate. This phenomenon has also been observed and analysed by *Killworth et al.* (1997).

The movement in the Indian Ocean (IO) appears to fit more closely to the theoretically expected wave speeds (Section 10.3), having a peak speed of approximately 0.35 ms^{-1} at the equator. Propagation speed decreases either side of the equator quite consistently (note the mainly parallel contour lines in this part of the globe). The wave speed drops to around 0.10 ms^{-1} by 20° from the equator and falls to 0.03 ms^{-1} at the southern extreme. The IO receives an input of warm water from the complex shallow water current systems around the Philippines and Papua New Guinea which also affects Rossby wave production in this area.

The Atlantic Ocean (AO) shows comparatively little evidence of Rossby waves, but some are observed in the equatorial region, as well as between 28° and 36° South and between 24° and 34° North. Since the longitude spans of this ocean are comparatively short, it is difficult to see any fine structure in the wave speeds observed, other

than in the north-south direction. In most cases however, a slight increase in wave speed is observed as the waves approach the Western boundary.

The Antarctic Circumpolar Current may also be producing eddies (formed by the meandering of the current which creates the dynamic equivalent of an oxbow lake) which are moving west in a Rossby-wave-like manner. Some authors also suggest that the ACC acts as a wave guide for Rossby waves (*Hughes*, 1996). Rossby waves are evident in all three basins in the southern areas. These are very slow moving phenomena with wave speeds of less than 0.03 ms^{-1} in places, which once again fits the scenarios outlined in section 10.3.

10.7 Conclusions.

Westward moving sea-surface-anomaly features have been found in many regions of the globe. Propagation speeds observed suggest they are Rossby waves. These speeds do not exactly fit the standard linear theory of Rossby waves and further research is being carried out into this phenomenon. It appears that (as hypothesised above) ocean depth should also be considered when modelling Rossby Wave propagation speeds, which is supported by recent work such as *Killworth et al.* (1997).

11. Mean Sea Level Change

The mean sea surface (or Quasi-stationary Sea Surface Topography) is the topography which includes effects due to the dynamic movement of water i.e. the general mean circulation. This is distinct from the geoid, which can be thought of as the topography existing in the absence of any forces other than gravitational. The mean sea surface is thus calculable from any source of sea surface height information if it is possible to remove (or average) the effects of short and long term ocean variability.

The mean sea *level* is the globally temporally and spatially averaged sea surface height. This can be calculated from the mean sea surface derived above by averaging across every geographical location.

11.1 Mean Sea Level Change

In recent years there has been much speculation about the possibility of some systematic rising of global mean sea levels due to man's impact on the environment. It is suggested that this has been brought about through the burning of fossil fuels in the industrialised countries of the world (the so-called 'greenhouse effect'), and the proliferation of Chloro-Fluoro-Carbon (CFC) molecules in the upper atmosphere destroying the Earth's protective ozone layer so that more heat energy is allowed into the atmosphere. The extra heat retained by the Earth's ecological system as a result is thought to have the potential to cause the melting of the polar ice caps and flood millions of tonnes of water into the remainder of the world's oceans, thus raising the mean sea level. Evidence of the break up of polar ice is clear from recent expeditions to the area and remote sensed evidence. Evidence from other sources has been scarce, however. There is some evidence indicating that the mean sea surface temperature is rising, but no mean-sea-level rise has been proven.

The satellite altimeter is an invaluable tool in studies of mean sea level, due to the global coverage and continuous data flow. Traditional methods involve taking

measurements *in-situ*, using either a research ship equipped with accurate satellite tracking electronics, or some other technique such as tide gauges or ocean buoys. All these techniques have limitations in that they only take a single measurement at a time, and cannot reveal information about anything other than the local sea level. A network of buoys and tide gauges large enough to give global coverage equivalent to that provided by a radar altimeter on a satellite would be an enormous undertaking and unlikely to be worth the investment of time and money. Furthermore, with missions such as the ERS-1 Geodetic Mission (GM), and the earlier but less accurate Geosat GM, the spatial density of satellite observations is very high (8km ground track spacing at the equator for the ERS-1 GM). Extremely highly detailed mean sea surfaces can therefore be derived.

One disadvantage of a satellite altimeter is that the data it produces do not represent simultaneous readings at all points on the Earth's surface, which a network of buoys would represent. Therefore the information can become corrupted by sea surface variability. In orbital configurations such as the TOPEX/Poseidon 9.9 day repeat period, sea surface variability is easily removed, and over such a short period of time the sea surface is essentially static.

The mean sea surface is defined as the sea surface height at locations over the oceans averaged over a suitably long period. This period must be long enough to average out the effects of any long term sea surface variability, and in most cases 2 years is sufficient since the majority of the variability is in an annual or seasonal cycle (as seen in Chapter 9), and is certainly concentrated in temporal wavelengths of less than 2 years. In order to resolve some form of mean sea level variation, it is necessary to solve for an 'instantaneous' mean sea level over some short period of time, so that any long term trend can be plotted. Obviously, short term solutions will be corrupted by mid- and long-term variability, but over a long enough period of time, a trend can be solved for which will remove the effects of this variability.

11.2 Mean Sea Level From Optimally Interpolated Time Series

In order to investigate any possible sea-level rise it is necessary to obtain a long-term time series of global mean sea level. This series should cover as long a time span as possible. The optimally interpolated data set provides an ideal source of such information. It also has the advantage that, by this stage in the data processing, the main sources of error have been accounted for by a combination of the orbit error correction procedures and the optimal interpolation procedure itself.

The mean sea level for a given epoch could be found by simply summing the sea surface height values in each bin and taking the arithmetic mean. This method fails to take account of the differing surface area of bins at the equator compared with those nearer to the poles, and will therefore place excessive weight on observations nearer to the poles. The data are therefore subjected to a weighting scheme.

Since the area of a bin is proportional to the cosine of the latitude of that bin, a cosine function is used to provide a weight for each bin. i.e. the weight (w) at latitude ϕ is given by the formula $w = \cos \phi$. The mean sea level $\overline{h}(t)$ at a given epoch is then

$$\overline{h}(t) = \frac{\sum_i w_i h_i(t)}{\sum_i w_i},$$

where $h_i(t)$ is the sea level in the ith bin at the time t, and w_i is the weight assigned to the ith of n bins.

This weighted averaging procedure was carried out on the combined ERS-1–T/P optimally interpolated data set, which produced 207 data points at five day intervals. Although a line of best fit through this data (by least-squares) was found to have a slope of 0.7 ± 0.6 mm/year (tabulated in Table 12), this trend contains a contribution from the annual signal due to the time span being a fractional number of years and cannot be relied upon. It is therefore necessary to use data from a whole number of years - either to

use a 2-year time span or switch to the uncorrected ERS-1 data set, which runs for just over 3 years. (The orbit correction procedure rejects data at the start of the ERS-1 mission due to the fact that T/P was not launched at the time and hence no correction is possible for several months.)

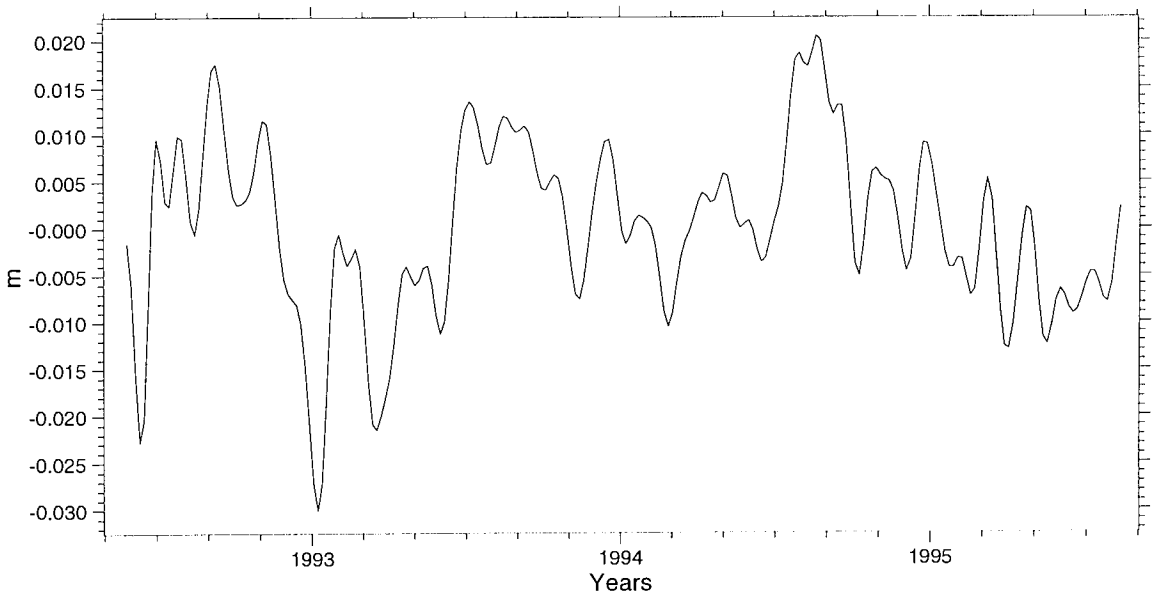

Figure 42. Mean Sea Level trend from DUT JGM3 uncorrected orbits after optimal interpolation.

The uncorrected ERS-1 data (from DUT JGM3 orbits) has a time span of 3 years and 2 months (Figure 42) and a line of best fit for the central 3 years has a slope of -0.17 $\pm$ 0.71 mm/year. The observed slope depends on which segment of the data is used, and has values for different two year periods ranging from -5.4 to +8 mm/year. This effect and the trends visible in Figure 42 suggest the presence of a longer term signal in the data. Notice that there is a pronounced drop at the start of 1993 which is not repeated at the start of 1994 or 1995, the two later year boundaries both being significantly higher than the 92/93 boundary. By visual inspection it is possible to estimate the presence of a 4 or 5 year signal within this data, showing a trough at the start of 1993 and a peak near the end of 1995.

On removing the annual signal from the DUT JGM3 uncorrected data set (as detailed in Section 9.5.2), the time series (Figure 43) is dominated by both the long term effect noted above and further oscillations: one with a wavelength of approximately 30

days and another with slightly larger amplitude of around 60 days wavelength. The estimated slopes in each of these cases are given in Table 12.

Data set	Time Span (years)	Overall slope (mm/year)
Combined ERS-1 - T/P	2.8	0.7 ± 0.6
Uncorrected ERS-1	3.0	-0.17 ± 0.71
Annual-filtered uncorrected ERS-1	3.0	1.29 ± 0.56

Table 12. Estimated sea level rise rates from various data sets.

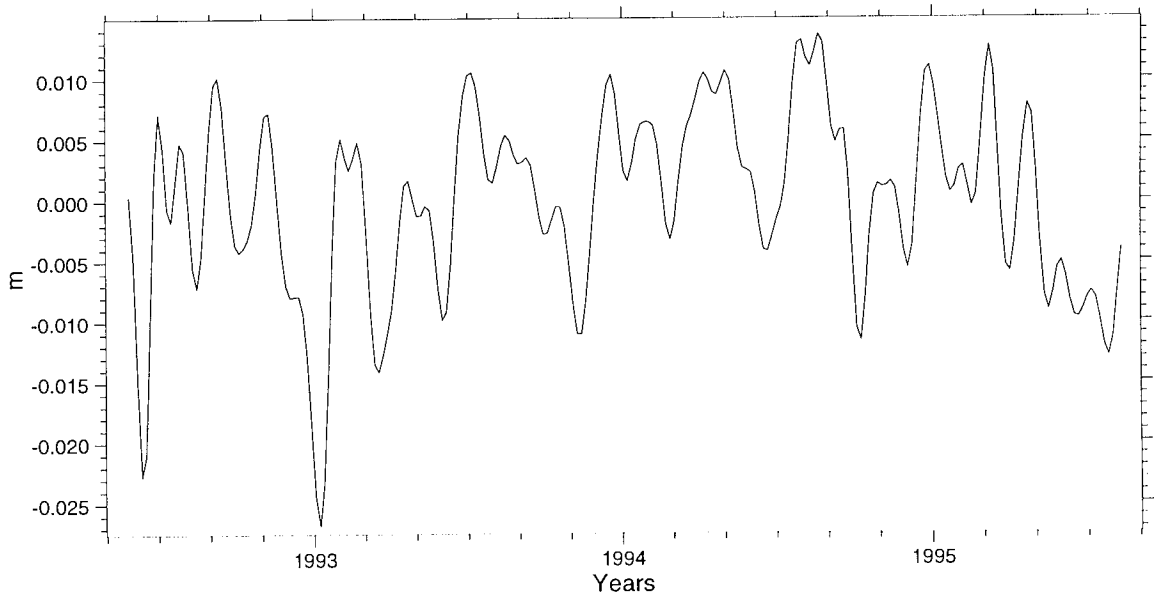

Figure 43. Mean Sea Level trend from ERS-1 uncorrected orbits after optimal interpolation and removal of annual cycle.

11.3 Variation Of Mean Sea Level From EOF Data.

As a source of further information on sea level rise, it is possible to use the global EOF data set to produce a time series of mean sea level. This data set has the advantage for this particular application that further insight can be gained by looking at the effect on mean sea level of each Empirical Orthogonal Function separately.

Since there is no direct sea-level anomaly information in an EOF, this information must be calculated from the EOF and its accompanying amplitude time series. The formula for the instantaneous mean sea level for the kth EOF therefore becomes

$$\bar{h}_k(t) = \frac{\sum_i w_i a_k(t) F_k(\mathbf{x}_i)}{\sum_i w_i},$$

where $F_k(\mathbf{x}_i)$ is the sea surface height level of the kth of K EOFs in the ith of n bins (at location $\mathbf{x}_i$), $a_k(t)$ is the amplitude coefficient for the kth EOF at the current epoch (t) and w_i is the assigned weighting for the ith bin, as before. The overall mean sea level at epoch t is therefore

$$\bar{h}(t) = \frac{\sum_k \bar{h}_k(t)}{K}.$$

From the uncorrected ERS-1 data set, 93 EOFs were produced. Of these, only the first 12 represent any more than 2% of the overall sea level variability each and the first ten together represent over 50% of the total.

The results gave the overall calculated mean sea level rise from the EOF data set as 0.10 ± 0.76 mm/year (c.f. the figure of -0.17 mm/year reported from the optimally interpolated data set above which is within 1 σ of this value). The rise rates due to each separate EOF are mostly small, with a single exception. EOF 3 is responsible for a significant rise in sea level and a plot of this particular EOF shows a large slope on the amplitude time series with a jump occurring during the ice phase. The DXO corrected EOFs also reveal a similarly strong change in this period (as shown in Figure 23), and with other evidence presented in this chapter, this would seem to confirm the presence of a long wavelength signal. As noted in Chapter 9, this EOF represents the El Niño phenomenon, and as such is part of a four (or more) year cycle. This has been observed by other researchers working with TOPEX or ERS-1 data such as *Hendricks et al.* (1996), and *Nerem et al.* (1997b). Both studies concluded (from respectively two and three years of data) that the El Niño long wavelength signal prevents accurate sea level rise rate assessment. Both showed that (as above) a single EOF is responsible for the

majority of the reported sea level rise, and that this EOF is caused by El-Niño-related effects.

11.4 Conclusions

The optimally interpolated data set can reveal nothing more than the existence of a long term, as yet unresolvable, signal with a wavelength in excess of three years. Removal of the annual signal reveals more strongly the presence of this long term signal. A line of best fit to three years of ERS-1 data yields a mean sea level *fall* of 0.17 ± 0.71 mm/year. This signal cannot however be assumed to be accurate since it is still impossible to resolve the aforementioned long wavelength signal. The EOF data set shows that a single EOF is responsible for a large proportion of the estimated rise. This EOF has previously been shown to represent the El Niño phenomenon. Therefore the conclusions from this study can only be that given the currently available data set, it is not possible to obtain an estimate for long term sea level rise.

Future efforts must involve data collected over longer periods than could be included in this study in order to completely resolve the El Niño cycle and any other long term cycles. The author suggests a minimum time period in excess of 6 years in order to be able to separate the El Niño effect from any longer term cycles. The extended T/P mission should provide a good basis for this type of work in future, since at the time of writing it would provide over 5.5 years of data and this mission is still on-going. Also, in conjunction with both ERS-1, -2, the ENVISAT radar altimeter, the Geosat Follow On mission and the Jason satellites there will exist many better data sets in the future. The crossover technique described in chapter 7 will enable the linking of multiple consecutive satellite missions into a single, long term data set which will hence allow the resolving of very long term signals.

12. Conclusions.

12.1 Review Of The Aim Of This Project.

The intention of this project was to produce a method which allowed extraction of sea-surface-variability information from the entire ERS-1 data set, which has previously proved difficult due to the differing orbital phases and the ERS-1 orbit error.

The orbit error problem was tackled by using a DXO orbit correction procedure with T/P data. A fit of less than 6cm RMS was achieved to the DXOs by using a cubic-spline based correction. The ERS-1 phase boundary problems were overcome through the use of a SXO-based technique, with the exception of the ERS-1 ice phase which remains difficult to integrate due to extremely low spatial resolution. Sea surface variability data extracted from the ERS-1 data set were then successfully validated by comparison with T/P data from the same time period under similar processing conditions.

12.2 Progress.

Through the use of the SXO processing technique (Chapter 7), it has been shown that the ERS-1 mission can provide sea-surface variability information of similar accuracy to that provided by T/P. Previous work on variability has concentrated on the T/P satellite due to its straightforward applicability to this particular use. Either ERS-1 has proved too difficult to work with due to the varying mission phases and accompanying changes in orbital type, or the Geodetic Mission (GM) and ice phases have been ignored in order to use the 35-day repeat pattern orbits in phases C and G. Though the SXO technique allows the inclusion of the GM alongside the two multi-disciplinary phases, the ice phases still prove highly problematic, and this is unlikely to be overcome due to the extremely low spatial resolution afforded.

The statistical difference between the ERS-1 and the T/P data sets was small and was observed to be due mainly to the specific methodology in use. In particular the lack

of Poseidon data in the repeat pass time series left gaps which produced spurious effects after optimal interpolation, and the poor spread of DXOs used in the orbit correction procedure caused erroneous signals at latitudinal extremes and poor resolution of certain very long-wavelength orbit error signals.

There is still the problem that ERS-1 orbits are not resolved to the same accuracy as T/P orbits, even with the addition of the DXO cubic spline orbit correction method detailed in Chapter 4. This technique's main weakness is in the use of DXO data: the DXOs are necessarily concentrated at higher latitudes, thus leaving the equatorial region poorly constrained; and DXOs do not extend beyond ±66° latitude, due to the T/P latitude limits. The full method proposed by *Le Traon et al.* (1995) included the use of ERS-1 SXOs in the cubic spline determination process, and this may have reduced the problems observed in Section 9.5.3 at the latitudinal extremes of ERS-1 arcs. This is unlikely, however, to have improved the overall situation since the original problem of poor data spread will still prevail.

Further analysis of the ERS-1 orbits through the use of ERS-1–T/P DXOs revealed that the DXO correction technique can significantly improve the ERS-1 orbits (RMS DXO residuals with T/P reduced from 10-13cm to less than 6cm - see Chapter 4), and is in some ways superior to the improved gravity model DUT DGM-04 (within the framework of the specific analysis carried out) as used in Chapter 4. The DGM-04 model could be improved further with the addition of extra data types to its derivation process with little difficulty. Refinement of the gravity model prior to use in the orbit determination process is still a useful tool, however, allowing the more general improvement of all ERS orbits, not just the section for which DXO data are available. The uncorrected JGM3 orbits (and to a lesser extent the DGM-04 orbits) were shown to contain significant (in excess of 5 cm RMS) first-order spherical harmonic errors associated with first-order gravity model terms or a shifting of the reference geocentre. Similar effects were noted in work comparing the combined ERS-1–T/P optimally interpolated data set and the corrected ERS-1 data set, where an EOF decomposition revealed several EOFs which resembled first-order spherical harmonics. This suggested that the DXO correction was not removing these effects entirely, and this is again thought to be a result of the poor spread of DXOs used in the process.

The oceanographic implications produced from these data have shown the presence of a long term signal (in addition to the normal annual and bi-annual signals) which is thought to be a result of El Niño sea surface signatures. This has been supported by both EOF decomposition analysis and observation of sea surface anomalies. Such inferences should be made with care, however, due to the possibility of remaining tidal- and SSB- related errors (Sections 2.2.2 and 2.2.4).

A derivation of mean sea level change from both sea surface anomalies and an EOF decomposition also provided support for the presence of a longer term signal than was fully resolvable with data collected over only 3 years. This revealed that any estimate of sea-level rise rate is subject to an error margin which precludes exact knowledge of even the *direction* of the change. This is further hindered by the presence of the aforementioned long term signal. It is therefore not possible to state, with the given data set, whether the mean sea level is in fact rising or not, and further work must be carried out to clarify this issue.

12.3 Suggestions For Future Work.

The launch of ERS-2 and the continued operation of T/P provides the opportunity to extend the work carried out in this project to up to six years in time span. This would allow the resolution of all but the longest cyclic signals in the ocean.

Preliminary studies carried out at Aston have shown the potential to resolve long term signals with a great deal of accuracy from the extended T/P data set alone. The addition of ERS-1 and 2 data would further improve the resolution of the system.

A unification of the altimetry data from all three satellites would require improvements to be made in ERS gravity and geophysical modelling in order to bring the orbits into line with the consistently high quality offered by the T/P satellite. Reduced inconsistencies in altimetry processing for different satellite missions prior to user level is also a priority, to enable more straightforward combination of satellite data sets.

References

AVISO (1995). AVISO CD ROM user manual: Altimeter products for space oceanography. AVI-NT-02-400-CN, Edition 1.

Andersen O. B. (1995). Global ocean tides from ERS-1 and TOPEX/POSEIDON altimetry. *JGR*, 100, C12, pp 25,249-22,259

Aksnes K. (1976). Short-period and long-period perturbations of a spherical satellite due to Direct Solar Radiation. *Celestial Mechanics.* 13, pp 89-104.

Blanc F., P. Y. Le Traon and S. Houry (1995). Reducing orbit error for a better estimate of oceanic variability from satellite altimetry. *J. Atmos. Oceanic Tech.*, 12, No 1, pp. 150-160.

Boomkamp H. J. and P. Moore (1997). A gravity field solution based on unified ERS-1 and TOPEX/Poseidon altimetry. *Third ERS Symposium on Space at the service of our environment, Vols II and III,* 414, Ch 219, pp. 1567-1572.

Brouwer D. and G. M. Clemence (1961). Methods of Celestial Mechanics. *Academic Press,* London, UK, pp. 249-251.

Carnochan S. (1996). Altimeter and orbit corrections for the ERS satellites through the analysis of single and dual satellite crossovers. PhD thesis. Aston University, Birmingham, UK.

Cazenave A., P. Schaeffer, M. Berge, C. Brossier, K. Dominh and M. C. Gennero (1996). High-resolution mean sea surface computed with altimeter data of ERS-1 (geodetic mission) and TOPEX/Poseidon. *Geophysical Journal International*, 125, No 3, pp 696-704.

CERSAT (1994). Altimeter products user manual, C1-EX-MUT-A21-01-CN.

Chelton D. B. and M. G. Schlax (1996). Global observations of oceanic Rossby waves. *Science*, 272, pp 234-238.

Cullen R. A. (1998). Mean sea surface modelling from multi satellite altimetry utilising frequency domain techniques. PhD Thesis. Aston University, Birmingham, UK.

Davis R. E. (1976). Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, 6, pp 249-266.

Eanes R. J. and S. V. Bettapur (1995). The CSR 3.0 global ocean tide model. *CSR-TM-95-06*, Center for Space Research, University of Texas at Austin.

Ehlers S. (1993). Various Techniques and Procedures for Refining ERS-1 Orbits. PhD Thesis. University of Aston in Birmingham, U.K.

ESA (No date). Earthnet online [online]. Available: *http://earth1.esrin.esa.it/*, search for 'USO' or 'SPTR'. [1999, May 20th].

Fu L. L. and D. B. Chelton (1985). Observing Large-Scale Temporal Variability of Ocean Currents by Satellite Altimetry: With Application to the Antarctic Circumpolar Current. *Journal Geophysical Research (JGR)*, 90, pp 4721-4739

Gaspar P., F. Ogor, P. Y. Le Traon and O. Z. Zanife (1994). Estimating the sea state bias of the TOPEX and Poseidon altimeters from crossover differences. *JGR*, 99, pp 24,981-24,994.

Glazman R. E., A. Greysukh and V. Zlotnicki (1994). Evaluating models of sea-state bias in satellite altimetry. *JGR*, 99, C6, pp 12,581-12,591.

Haines B. J. and Y. E. Bar-Sever (1998). Monitoring the TOPEX microwave radiometer with GPS: Stability of columnar water vapor measurements. *Geophysical Research Letters*, in press.

Hayes J. G. (1974). Numerical methods for curve and surface fitting. *J. Inst. Maths Applics.*, 10, pp. 144-152.

Hendricks J. R., R. R. Leben, G. H. Born and C. J. Koblinsky (1996). Empirical orthogonal function analysis of global TOPEX/POSEIDON altimeter data and implications for detection of global sea level rise. *JGR*, 101, C6, pp. 14131-14145.

Hughes C. W. (1995). Rossby Waves in the Southern Ocean: A comparison of TOPEX/POSEIDON altimetry with model predictions. *JGR*, 100, C8, pp. 15933-15950.

Hughes, C. W. (1996). The Antarctic Circumpolar Current as a waveguide for Rossby waves. *J. Physical Oceanography*, 26, No. 7, 1375-1387.

Jolly G. W. and P. Moore (1995). Accuracy assessment of TOPEX/Poseidon Dual Crossover correction of ERS-1 altimetry. *Advances Space Research*, 16, No 12, pp 115-118.

JPL (1993). TOPEX/Poseidon Research News, Issue 1, *JPL*, September 1993.

JPL (No date). TOPEX/Poseidon main screen [online]. Available: *http://topex-www.jpl.nasa.gov/*. [1999, May 20th].

Kaula W. M. (1966). Theory of satellite geodesy: Applications of Satellites to Geodesy. Blaisdell Publishing Company, Waltham, Massachusetts, U.S.A.

Killworth P. D., D. B. Chelton and R. deSzoeke (1997). The speed of observed and theoretical long extra-tropical planetary waves. *J. Phys. Oceanogr.*, 27, 9, pp.1946-1966.

Kubitschek D. G. et al (1997). Calibration methods for the TOPEX and Poseidon altimetry systems (abstract). *Annales Geophys.*, 15, Suppl 1, p. C196

Lambeck K. (1980). The Earth's Variable Rotation. Cambridge University Press, U.K.

Lerch F. J., R. S. Nerem, B. H. Putney, T. L. Felstenberger, B. V. Sanchez, J. A. Marshall, S. M. Klosko, G. B. Patel, R. G. Williamson, D. S. Chinn, J. C. Chan, K. E. Rachlin, N. L. Chandler, J. J. McCarthy, S.B. Lutcke, N. K. Pavlis, D. E. Pavlis, S. J. Robbins, S. Kapoor and E. C. Pavlis (1994). A geopotential model from satellite tracking, altimeter, and surface gravity data: GEM-T3. *JGR*, 99, B2, pp 2815-2839.

Llewellyn S. K. and R. B. Bent (1973). Documentation and description of the Bent ionospheric model. *AFCRL-TR-73-0657.*

Le Traon P. Y., P. Gaspar, F. Bouyssel and H. Makhmara (1995). Using TOPEX-Poseidon data to enhance ERS-1 orbit. *J. Atmos. Oceanic Tech.*, 12, 161-170.

Le Traon P. Y. and F. Ogor (1996). ERS-1/2 orbit improvement using TOPEX-Poseidon: The 2cm challenge. *JGR*, 103, No C4, pp 8045-8057.

Levitus S. (1984). Annual cycle of temperature and heat storage in the world ocean. *J. Phys. Oceanogr.*, 14, pp 727-746.

Marshall J. A., N. P. Zelensky, S. M. Klosko, D. S. Chinn, S. B. Luthcke, K. E. Rachlin and R. G. Williamson (1995). The temporal and spatial characteristics of TOPEX/Poseidon radial orbit error. *JGR*, 100, C12, pp 25,331-25,352.

Moore P., S. Ehlers, C. M. Murphy and M. D. Reynolds (In Prep.). Investigation of the Stability of the ERS-1 Range Bias Through Tide Gauge Augmented Altimetry. Submitted to *JGR.*

Moore P., S. Carnochan, M. D. Reynolds and R. J. Walmsley (1999a). Inferences from dual crossover analysis between concurrent satellite missions. presented at COSPAR 1998, Nagoya, Japan. To appear *Adv. Space Res.*, 23/4, pp 753-756.

Moore P., H. Boomkamp, S. Carnochan and R. J. Walmsley (1999b). Multi-Satellite orbital dynamics. To appear *Adv. Sp. Res.*, 23/4, pp 785-795.

Nerem R. S., B. J. Haines, J. Hendricks, J. F. Minster, G. T. Mitchum and W. B. White (1997a). Improved determination of global mean sea level variations using TOPEX/Poseidon altimeter data. *Geophys. Res. Lett.*, 24, No 11, pp 1331-1334.

Nerem R. S., K. E. Rachlin and B. D. Beckley (1997b). Characterisation of global mean sea level variations observed by TOPEX/POSEIDON using empirical orthogonal functions. *Surv. In Geophys.*, 18, No 2-3, pp. 293-302.

Pattulo J., W. Monk, R. Revelle and E. Strong (1955). The seasonal oscillation in sea level. *J. Marine Research*, 99, 14, pp. 88-155.

Pickard G. L. and W. J. Emery (1990). Descriptive Physical Oceanography An Introduction, 5th Edition. Pergammon Press Ltd, Oxford, UK.

Polito P. S. and P. Cornillon (1997). Long baroclinic Rossby Waves detected by TOPEX/Poseidon. *JGR*, 102, C2, pp. 3215-3235.

Pond S. and G. L. Pickard (1983). Introductory Dynamical Oceanography, 2nd edition. Pergammon Press Ltd, Oxford, UK.

Rosborough G. W. and B. D. Tapley (1985). Geographically Correlated Orbit Error and its effect on satellite altimetry missions. *JGR*, 90, C6, p1817 et seq.

Rothwell D. A. (1989). A Description of Altimeter and Crossover Processing Software. Internal Report, ESRU, Aston University, Birmingham, UK.

Scharroo R. (1999). ERS Precise Orbit Determination: Orbits [online]. Available: *http://www.deos.tudelft.nl/ers/precorbs/orbits/*, [1999, May 20th].

Scharroo R. and P. N. A. M. Visser (1998). Precise orbit determination and gravity field improvement for the ERS satellites. *JGR*, 103, C4, 8113-8127.

Schum C. K., P. L. Woodworth, O. B. Andersen, G. D. Egbert, O. Francis, C. King, S. M. Klosko, C. Le Provost, X. Li, J.-M. Molines, M. E. Parke, R. D. Ray, M. G. Schlax, D. Stammer, C. C. Tierney, P. Vincent and C. I. Wunsch (1997). Accuracy assessment of recent ocean tide models. *JGR*, 102, C11, pp 25,173-25,194.

Sinclair A.T. and G. M. Appleby (1986). SATAN - programs for the determination and analysis of satellite orbits from SLR data. Royal Greenwich Observatory, SLR technical note number 9.

Summerhayes C. P. and S. A. Thorpe (1996). Oceanography An Illustrated Guide. pp. 352, Manson Publishing Ltd, London, UK.

Tapley B. D., M. M. Watkins, J. C. Ries, G. W. Davies, R. J. Eanes, S. R. Poole, H. J. Rim, B. E. Schutz, C. K. Shum, F. J. Lerch, J. A. Marshall, S. M. Klosko, N. K. Pavlis and R. G. Williamson (1996). The Joint Gravity Model 3. *JGR*, 101, 28029-28049.

Tomczak M. and J. S. Godfrey (1994). Regional Oceanography: An Introduction. Elsevier Science Ltd, Oxford,UK.

Vonder Haar T. H. and V. E. Suomi (1971). Measurements of the Earth's radiation budget from satellites during a five year period. *Journal of Atmospheric Science*, 28, 3, pp 305-314.

Wagner C. A. and J. Klocknik (1994). Accuracy of the GEM-T2 geopotential from Geosat and ERS-1 crossover altimetry. *JGR*, 99, 9179-9201.

Wessel P. and W. H. F. Smith (1991). Free software helps map and display data. *EOS Trans. Amer. Geophys. U.*, 72, pp. 441, 445-446.

Appendix - The software used in this project.

A.1 Introduction.

The software written for the processing of the crossover data used in this project was split into several sections. In that this software is likely to be used in the future, this appendix will give the user all the information needed to successfully use this software. All software can be found in a directory tree on the UNIX computer systems under /data/src/matt_code.

1. Time Series Program.

This program reads the crossover residuals, grids according to the chosen grid size, corrects the orbits using a specified B-Spline representation of the error and/or applies DUT orbit heights to the data. Then it solves for a time series in each bin, outputting the time series to a set of data files a latitude band at a time.

2. Optimal Interpolation Program.

This software reads the gridded ERS-1 data produced by the time series program, and/or the gridded T/P data produced from repeat pass analysis, carries out the optimal interpolation (OI) procedure on all the data that has been read and outputs a temporally smoothed time series. Where T/P data has been included in the procedure, the time series must be filtered to remove the effects of tidal modelling errors which are present in T/P data using the 'filt' program. The OI data set can be analysed on a temporal basis and split into various wavelength signals using the 'spect' program.

3. Empirical Orthogonal Function Program.

This reads the OI data set and filters it according to the spread of the arcs which were read by the OI software. It then creates the matrix **C** as defined in section 8.2.1, and uses a NAG routine to solve for the eigenvalues and eigenvectors which are the Empirical Orthogonal Functions. These are output to a data file.

Each separate piece of software uses the same text input file format, though each program requires different parts of the information contained within this file. To unify the reading procedure across all programs a single parsing routine was written which allows a program to load values according to the contents of a second text (or statement) file. Thus the routine can be altered to accept different formats very quickly and easily, and can be used totally generally - it is not specific to this or any other piece of software. This routine is explained below in section A.2, along with details of the requirements of each of the SSV programs in section A.2.1.

A.2 The input file parsing routine.

In order to overcome the usual Fortran inflexibility in reading from text files, the subroutine `readinp.f` was written which allows almost free formatting of the input. This section of the appendix makes up the documentation for this routine.

The routine searches the input file for keywords which are listed in a statement file and reads the rest of a line which contains a keyword into a specific place in a specific type of array (e.g. character, real*8, integer or logical). The statement file instructs the routine which slots in which array each keyword is associated with. The routine has the advantage that it can quickly and easily be adapted to work with any piece of software since the keywords that it recognises are totally re-definable.

Once a call statement is added to the program, the routine is used by creating four arrays each with at least one element: real*8 (or double precision), integer, character and logical. String length within the character array is sensed by the subroutine. It is necessary to define long enough strings for the longest possible value (e.g. a full filename). These arrays store and return the values that are read from the input file and parts are allocated to each keyword. The routine uses a statement file (the full filename of which is passed to the routine) to control the keywords it will search for and the types of data and parts of arrays which each represents.

In constructing the statement file, slots in the storage arrays are allocated to each keyword by specifying a keyword, the type (or types) of data which it will require, and the slots this data will take up in the storage arrays.

For example: if ten filenames (character strings) and two integers are to be read, arrays would be created of those sizes, and the first slot in the character array would be allocated to be filename one, the second slot would be filename two and so on. The statement file would then possibly contain the following:

```
filename one:c 1
filename two:c 2
```

Anything before the colon ':' is assumed to be the keyword and whatever is immediately after is the type of variable to be read. These are shown in Table 13. A maximum of 1 space is allowed between the colon and the type of variable and a minimum of 1 space must be between the type of variable and the slot number since the routine will fail otherwise.

Letter	Variable Type
r	Double Precision (real*8)
i	Integer (*4)
c	Character (*(*))
l	Logical(*4)

Table 13. The variable type each letter is interpreted as by the readinp routine.

Entries can be made in any order (though it is sensible to keep them in order to simplify the debugging process in the event of a problem), and the system is not case sensitive or column sensitive so the following will also work:

```
  FileName THREE     :c 3
filename four        :c  004
```

In cases where there are to be multiple values for a certain keyword (e.g. a keyword which represents a list of filenames called 'bspline' below) a format such as this in the statement file:

```
Bspline              :c 51- 55
```

will allocate a block of entries within the character storage array starting with slot 51 and ending with slot 55, i.e. 5 separate values. An error is signalled if the storage arrays are exceeded during the allocation procedure (e.g. if in the above example the character array has less than 55 elements). In such a case, the user is notified of which definition is at fault. The routine also checks that each array slot is used only once, through the use of a dynamically allocated workspace array. A warning is signalled if the routine is unable to allocate sufficient memory for this, but this is not a fatal error - checking is simply not carried out.

Within either the statement file or the input file itself a blank line or a line with a '#' as the first item is ignored and treated as a comment (with a few specific exceptions mentioned below), similarly to shell script conventions. Comment lines can therefore either be marked as such or simply not contain a recognisable keyword.

The routine has the ability to read several items of data from a single line for items such as:

```
#  Tide gauge - Gauge ID, Name of gauge, Latitude, Longitude
Tide gauge    : 022       Easter       27.15  109.45
```
(Example 1).

This type of keyword can be specified in the statement file by a simple extension of the previously explained format. For example the keyword above would be listed in the statement file as:

```
Tide Gauge  :icrr  1
```

which means that an integer, a character and two reals (in that order) should be read from each line with the keyword 'tide gauge'. Note that since there are *two* reals to be read for this keyword, the first will take array slot r1, and the second will take array slot r2. Care should therefore be taken when including entries such as this to avoid confusion over which array slots are taken for which keywords, particularly in the event of multiple instances of such a keyword being required. For example if the following is defined in the statement file:

```
Tide Gauge  :icrr 1-20
```

then the 20 integers and the 20 character strings will be in slots 1-20, whilst the *first* real on each line will be in slots 1-20, and the *second* real on each line will be stored in slots 21-40. This system extends to a total of 15 data items per line.

A further option is to set a 'default' data type which is to be read in the absence of a recognisable keyword on a line in the input file. A default item is set up in the statement file in this manner:

```
default Tide Gauge :icrr 1-20
```

This tells the routine that any line with the correct number of fields (in this case 4) which does *not* contain a colon ':', and is not commented out should be read in the same way as in the previous example. Care should be taken when using a default data type to comment out any lines which may be accidentally read because of their field count. The name 'Tide Gauge' is stored as with any other keyword, and is thus recognised as both the above Example 1 and:

```
022    Easter       27.15  109.45
```

Values will be read into the allocated slots from the input file in the order in which they are found. If the number of slots allocated in the statement file to a particular keyword is not enough for the current input file, an error will be signalled, and the problem will be explained to the user.

After parsing the statement file and storing all the keywords and types found, the routine reads a line from the input file, and looks for a colon in that line. If one is found, it will check the string at the start of the line against all the keywords in the statement file. If a match is found, it will read the appropriate type of variable from the rest of the line. For example, if the routine finds the line

```
Filename TWO          :/Path/To/File
```

it will read '/Path/To/File' into the character array in slot two, exactly as specified before in the statement file. Then the calling program must use the value at slot 2 in the character storage array as the location of the file. Spaces before or after the keyword are ignored, but the keyword must be exactly as it was written in the parameter file, otherwise it will not be recognised. Keywords are not case sensitive, but character variables are stored exactly as written in the input file (important for filenames and the like). Also, if more than one variable is to be read from a line and one of them is a character string, the routine reads only the next *word* as that string. For example if the following appears in the input file (with the appropriate statement file entries):

```
Gauge Name:  Easter Island
Tide Gauge: 022 Easter Island 27.15 109.45
```

in both cases the string 'Easter' will be stored, and the routine will fail on the second line since it will attempt to read 'Island' as a real. A string can be quoted using double quotes, however, so the following would succeed:

```
Tide Gauge: 022 "Easter Island" 27.15 109.45
```

Logical values are read similarly to character strings, and the value returned is set to .true. if there is either a 'y' (e.g. ' yes '), a '1' or a 't' in the string. Therefore any of the following will result in the variable 'reject ice' being set to .true. (even for the final example)

```
Reject Ice     :  Yes
Reject Ice     :  y
Reject Ice     :  1
Reject Ice     :  .true.
Reject Ice     :  noWay
```

whilst all the following will result in a .false. value

```
Reject Ice     :  No
Reject Ice     :  n
Reject Ice     :  0
Reject Ice     :  f
```

A line with unclear meaning e.g. 'NoYesNo' will always result in a .true. value being returned, since the routine simply finds the 'y' and ignores the rest.

A variable can be made 'required' in one of two ways. Setting the value of the first appropriate slot in the storage array to the value shown in Table 14 or placing an '!' at the end of the line for that keyword in the statement file will both result in the routine signalling an error if that particular keyword is not found at least once. Note there must be spaces between the last digit and the exclamation mark in the statement file if that method is used.

Variable Type	Value for 'Required'
Real	1.0d0
Integer	1
Character	'1'
Logical	.true.

Table 14. Value to be used to make a variable 'required' in the readinp routine.

If the variable 'ierr' has the value -1 or -3 when the routine is called, a report of the parsing process will be written to the file READINP-[program name].REP, including a display indicating whether there are free slots (and therefore wasted memory) in the storage arrays. If this file already exists, it will be overwritten.

Further special tokens are recognised in input and statement files to control the parsing procedure. In the statement file, if it is necessary to comment out a block of text without placing a '#' at the start of every line, the special token '#STOP STATEMENT' at the start of the line will force the routine to not parse any further lines until it finds the token '#START STATEMENT' (again, at the start of the line), at which point parsing will continue. This can be repeated as many times as is necessary in the statement file. Similarly in the input file itself, the tokens '#STOP READ' and '#START READ' can be used with identical effects. These commands are not case sensitive, but must be placed at the start of the line. These options make the routine flexible enough to be able to read the same file as both a statement file and an input file, enabling for example a commonly read data file to contain not only the data, but also the instructions to the routine on how to read that data.

Another possibility is to include the statement file in the Fortran code of the calling program itself. This is achieved by observing the standard Fortran comment syntax for the segment of the file which contains the statement file, and when the readinp routine is called, setting the variable ierr to -2 or -3. (See subroutine header, for more details on this.) The routine will also automatically switch to Fortran mode if a line with the token '#FORTRAN-MODE' after the comment character (usually a 'c') is found. This should be the first line in the file, since Fortran-style comments will not be recognised otherwise. Upon switching to Fortran mode in this manner, the system automatically goes to no-parsing mode to prevent the accidental reading of the Fortran header code. Before and after the statement list, a pair of start/stop statement lines must be used, since the routine will attempt to read the code otherwise.

The subroutine header gives details of the variables that must be passed.

A.2.1 The input file as required by the SSV software.

Once the readinp routine was set up, a suitable statement file for the sea surface variability software was created. Included below is an example input file which was used to produce some results, with notes on what each section is relevant to.

```
#   Input file for sea surface variability programs.
#
Resid file          :/vol1/matt/Xovers/resids/mul1/resid1.ers1
Resid file          :/vol1/matt/Xovers/resids/geod/residE.ers1
/ ..etc.. /
Resid file          :/vol1/matt/Xovers/resids/mul2/resid2.ers1
Resid file          :/vol1/matt/Xovers/resids/mul2/resid3.ers1
```

Residual Files:

List of full path names to residual files.

```
Data Directory      :/vol1/matt/data/delft/
```

Data Directory:

The path to the directory which will be used to store the time series output files. There may be up to 180 of them (one per latitude band), so they are placed away from the current working directory.

```
# Set the rejection level, shallow water, orb error correction
Reject Level(m)     :  -1.0
Absurd Reject(m)    :   1.0
```

Rejection Level:

This is the threshold in metres which is applied to the residuals before any further processing is done. An iterative 3σ routine is used if the value specified here is '-1' and no rejection criteria is used at all if '0' is specified. The absurd rejection level is also a threshold level in metres which is used to reject any bins which have a calculated RMS variability in excess of the value specified here. No 3σ option is available in this case.

```
Reject shallow?     :    No
```

Reject Shallow:

Set this to true if it is necessary to reject any residuals which are flagged as shallow water values.

```
Delft orbits?       :    Yes
Delft Dir           :/data/delft_orbits/JGM3
```

Delft Orbits:

Set to true if the delft orbit heights are to be applied to the crossovers. The Delft Dir variable contains the full path name required for the DUT orbit routine. This is the location of the root of the directory tree which contains the DUT orbit files. This can therefore be used to specify whether the delft JGM3, DGM3 or DGM4 orbits are used.

```
#   List the BSpline files.
Bspline File        :/data/ORB_CORR/B_SPL/Amul1_B_SPL
Bspline File        :/data/ORB_CORR/B_SPL/Bice_B_SPL
Bspline File        :/data/ORB_CORR/B_SPL/Cgeod_B_SPL
Bspline File        :/data/ORB_CORR/B_SPL/Dmul2_B_SPL
Correct orbits?     :    Yes
```

B-Spline Correction Files:

List of the full path names to the B-Spline correction files. Also there is a logical variable which toggles the use of orbit corrections.

```
#Specify the width and height of the bins (In degrees) and the limits
# for the current run.  If width is greater than 10, we specify the
# number of bins.
Height(deg)          :   2.0
Width(deg)           : 127.0
```

Grid Dimensions:

Dimensions of grid (see comments below)

```
Lat Min              : -90.0
Lat Max              :  90.0
Long Min             :-180.0
Long Max             : 180.0
```

Grid Limits:

Latitude and longitude maximums and mimimums. This allows a run concentrating on a particular area of the globe, or alternatively it allows the splitting of a full global run into latitude bands to reduce run time and memory usage. Note the grid limits are checked to ensure that it is possible to get an integer number of bins of the specified size into the specified grid. This caused problems with the grid spacing of T/P giving rise to rounding errors which caused this check to fail, even though it was approximately correct, so an alternative method of specifying bin width was introduced. If the width is specified to be greater than 10, it is taken to be the *number* of bins to fit across the specified longitude range, so when T/P data is used, a value of 127 can be put as the 'width', and the software will interpret this to mean there are to be 127 bins across the -180° to 180° longitude range, and will calculate the actual spacing internally.

```
#   Topex Boxpass file
Boxpass File         :/vol1/pmoore/TOPEX/boxpass.out
```

Topex Boxpass File:

This is the full path name to the file which contains the gridded repeat pass T/P data.

```
Start Opt Int        : 48925.0
End Opt Int          : 49925.0

Optint File          :/vol1/delft/jgm3/OI.delft-jgm3
```

Start / End Opt Int:

Sets the start and end dates which the optint program attempts to interpolate between. Optint File is the full path to the file which will contain the optimally interpolated time series data.

```
Max Time Diff       :  80.0
Decorrelation Time  :  20.0
Max Variability(cm):  30.0
Altimeter noise(cm):   5.0
```

Max Time Diff:

Sets the maximum time between raw observations that is allowed in a bin when the optimal interpolation program runs.

Decorrelation Time/Max Variability/Altimeter Noise:

Various parameters controlling the way the optimal interpolation system works. See Chapter 8 for an explanation of these terms.

```
Start EOF           :  48960
End EOF             :  49890
```

Start EOF / End EOF:

Start and end dates within which to carry out the Empirical Orthogonal Function decomposition. (N.B. Different to Start / End Optint, since optint procedure will work to the limits of the given data, but EOF decomp. will reject any bin which does not contain data from Start EOF date to End EOF date inclusive.

```
Max Average Tdiff   :   10.0
Compute EigenVec?   :    Yes
Write COMPH?        :    Yes
Reject Ice Phase?   :    Yes
```

Max Average Tdiff/Compute Eigenvec/Write COMPH/Reject Ice Phase:

Max average tdiff is the maximum average time allowed between raw observations in a particular bin. If this is exceeded, the bin is rejected. The other three are logicals which control the functionality of the EOF program (see below for more information).

The readinp subroutine and all its required function routines are available pre-compiled in a library file on the computer system.

A.3 Time Series software.

Due to the nature of crossovers, this program (called 'ts') has to store the entire residual data set in memory in order to sort it chronologically. This ensures minimum computer run time when calculating the delft orbit heights, and is also necessary when applying a B-Spline orbit correction.

In the data set used in this project, there are over 9 million crossover residuals. This requires over 216 Megabytes of RAM to store in its entirety, so it is necessary to use this program in shorter runs. The arrays are set up to store 5 million residuals, and this allows a simple North/South split. Due to the land imbalance between the hemispheres, there are more SXOs in the southern hemisphere and the split should be made 20° south of the Equator.

If it is necessary to alter the array sizes, they are stored in the program directory the file 'arrays.com'. The program is compiled by typing 'make' in the program directory. If there are any .o files pre-compiled it will be necessary to delete them before re-compiling when array sizes have been changed. This can be done by typing 'make clean'.

Certain variables are required by the time series program. The residual file list and the three logicals which control applying of delft orbits, orbit corrections (and obviously the B-Spline file locations if corrections are being applied, and the Delft Grav directory location if the Delft orbits are to be used) and whether or not to reject shallow water are necessary for the first part of the program which reads the residuals. In a case where the delft orbits are applied or the orbits are corrected using B-Splines, the raw residuals are altered before any rejection criteria are used. Also needed are the grid dimensions, the data directory location, and the absurd RMS rejection threshold.

The program writes a verbose report file containing information about all aspects of its run parameters to a file called 'TSER.REPORT'. If this file already exists, it will not be overwritten. Instead, the program will create a new file 'TSER.REPORT01' and so on until there are 100 such files in the current working directory at which point it will exit signalling an error. The program functions similarly for the output summary file 'TSERIES.OUT' (more information below).

The entire residual data set is read, rejecting many of the points due to geographical location or extreme residual values. The points are then sorted chronologically, and the appropriate corrections are applied. The rejection criteria, either a three sigma scheme or a fixed threshold is then used to filter the residuals.

Note that the program uses pixel registration, rather than grid line registration. For example if the grid runs from -180° to 180° longitude and -90° to 90° latitude and is a 3° by 2° grid, the first bin is centred on (-178.5°,-89.0°), all the crossovers in it will have longitude between -180.0 and -177.0 and latitude between -90.0 and -88.0, and it will be referred to by the co-ordinates of its centre.

The program then runs through the latitude bands, opening a separate output file for each band. This file is named 'TSER_###.#.DAT' where ###.# is the latitude at the centre of the current latitude band. This file contains a list of epochs and associated sea levels for each bin in the latitude band in order west to east, along with the RMS variability for each. A summary of the output data set is written to the file 'TSERIES.OUT', which contains the locations, number of residuals, number of arcs and the RMS variability in each bin.

For each bin, a matrix is created and then inverted using an LU decomposition. The arc heights are then calculated and written to the output file for the latitude band. A failure to invert the matrix, an RMS above the absurd rejection level, or a significantly non-zero mean in the output data all result in a bin being rejected. It was found that a non zero mean only occurs very near to coastlines, and therefore is most likely the result of tidal modelling problems giving rise to inconsistencies in the residual data set. A value of 1 metre is most commonly used as a catch-all absurd RMS rejection level.

The program then writes some statistics about the maximum values of various parameters which are useful when setting array sizes for subsequent programs and exits.

If the program exits with an error, the following programs in the suite will refuse to run (useful when the software is being run as part of a script or batch job). When a program finds an error, it creates a file called 'BIN.ERR' which contains notification of the error. Each program checks for the existence of this file, and refuses to run if it exists. The Time Series program will run even if the error file exists, and will delete it if it completes successfully. To force a program to run when the previous program has flagged an error, rename or delete the error file.

The output data set will use a substantial amount of disk space. Actual usage for a particular run depends on the number of bins which contain data, and the number of arcs per bin. A typical run using a 3° by 2° grid globally with 9 million crossovers over a thousand days for the ERS-1 satellite will need approximately 200 - 250 Megabytes of disk space. If the program crashes due to an error, since it outputs a latitude band at a time, it is possible to restart at the latitude band where it stopped.

If more than one run is required to complete the global data set, it is necessary to collate a complete TSERIES.OUT file from the 2 or more smaller files which have been written. This is done by tagging the subsequent files onto the end of the first. This is best done in the correct latitude order (the program runs from the southern extreme to the northern extreme and from west to east), and any overlaps where a bin has been written twice (e.g. in a case where a run was stopped part way through a latitude band) must be deleted.

The per-latitude band output files, and the TSERIES.OUT file are required to use the optimal interpolation software. The complete TSERIES.OUTxx file must be renamed to 'TSERIES.OUT' before continuing.

A.4 The Optimal Interpolation software.

This program ('optint') reads the data set output by the time series software and performs optimal interpolation temporal smoothing on the time series in each bin and outputs evenly spaced data to a single large file (the 'optint file').

This software reads a bin at a time from the previously produced data set, so there is no need to store large amounts of information in the computer's memory. The largest array required is a N by N array where N is the number of observations in the bin. This must be set to a value greater than the 'most points in a bin' value reported by the Time Series program.

Similarly to the Time Series program, it is possible to split the usage into shorter latitude-grouped runs. This enables use of multiple CPUs and accelerates the run time considerably.

Setting up for the run requires setting several variables in the input file. The data directory location is required, otherwise the data files will not be found. The 'Optint File' file name and location is required. If this file exists, a new file will be created with a numerical extension to indicate the version number.

If T/P data is to be included, the full path to the filename of the Boxpass file must be added. If this keyword is not found, no T/P data will be read.

Start Opt and End Opt are the dates to attempt to interpolate between. These dates must be an integer multiple of five days apart, and should be set to dates beyond the ends of the observations available, since the procedure will not interpolate to a point beyond the last observation. If the dates are within the possible range that could have been achieved, the whole procedure must be rerun to extend the data set - it is better to get as much information as possible in one run.

Normal values for the other parameters are as follows:

- Max Time Diff

  80 days. Lower than this causes problems with the ice phase where some bins have no data for several months.

- Decorrelation time

  20 days was chosen as a good compromise between losing variability information and having a noisy data set.

- Max Variability

  30cm.

- Altimeter Noise

  5cm is used. This is the RMS of the noise in the altimeter's measurements. This is mainly due to noise inherent to the electronic hardware on board the satellite.

The Optint program will write a verbose report file containing all the information about the current run to the local file 'OPTINT.REPORT'. A summary of the interpolation procedure in each bin will be written to a file 'OPTINT.OUT'. This lists the location of the bin, the number of arcs, average number of days between arcs, number of arcs rejected and the RMS variability of the output arcs. Again, if either of these files exist already, a numerical extension will be added to the filename.

The file 'OPTINT.OUT' is required for all further programs. Therefore where multiple runs have been performed in order to complete the optimal interpolation data set, it is necessary to concatenate the multiple 'OPTINT.OUTxx' files into the single file 'OPTINT.OUT', ensuring that no bins are repeated in this file (e.g. due to an overlap in latitude bands or a crash mid way through a latitude band).

Where T/P data has been included in the processing, whether directly from a boxpass file or indirectly by using DXO corrected ERS orbits, it is necessary to filter the data to remove artefacts of T/P tidal modelling problems (§2.2.4). Problems are known

to occur at various discrete wavelengths, so the program 'filt' solves for these wavelengths in each bin and removes them before writing the time series out to a new file which is named identically to the original with the extension '.filt'. This file is identical in format to the optint file, and it should either be renamed to the name in the input file for the optint file or the input file should be altered to point to the new file.

It is not possible to accidentally run the filt program twice, since it checks for the presence of a tag in the input optint file which indicates that it has already been filtered. If the tag is not found, the filtering process goes ahead, and the first line written to the output file is the tag line.

The sea surface anomaly time series present in the optimally interpolated data set can be analysed using the 'spect' program which solves for discrete 1, 2, 4, 6 and 8 cycles-per-year signals in each bin. This is output to the file 'SSV_SPECT.OUT' in the current working directory. Further information on the annual signal is output to two other files: 'ANN_AMP.OUT' and 'ANN_PH.OUT' contain the amplitude and phase of the annual signal in each bin respectively.

A.5 Empirical Orthogonal Function software.

This is the final part of the software package that was used in this project, and it reads the entire optimally interpolated data set and performs an Empirical Orthogonal Function decomposition on the data, outputting a series of EOFs with the accompanying Time Series for each, and various other statistics.

This program requires the following information from the input file:

- Optint File.

  Full path to the file which contains the optimally interpolated data set.

- Start / End EOF.

Start and End dates for performing the EOF procedure within. In order to perform an EOF decomposition, it is necessary to have a complete set of data within the time specified. Every bin must have the same number of points and there must be no gaps in the data. These dates are therefore generally set in from the ends of the optimal interpolation dates and experimentation will reveal how long a period is viable for a particular optimally interpolated data set.

- Max Average Tdiff.

This controls which bins are accepted when attempting an EOF decomposition on a data set. Included in the OPTINT.OUT file is record of the frequency of raw observations in each bin and the bin is rejected if the average number of days between observations exceeds the value given to this parameter. 10 days is generally used in order to reject particularly bad bins (mostly in problem areas), but to still ensure sufficient bins are retained to achieve good global coverage.

- Compute Eigenvec.

Specifies whether or not the software computes the EOFs or simply reads them from a previous run and performs various statistics on them.

- Write COMPH.

Specifies whether or not the software writes out a file which contains a list of sea heights at times and locations equivalent to those in the optimally interpolated data set, with heights calculated from the EOF representation of the data. This can be used as a tool to check that the EOF procedure is working correctly by comparing this file against the original optint file.

- Reject Ice Phase.

Specifies whether the ice phase is to be included in the calculations or not. Problems occur when including the ice phase due to the lack of observations for long periods of time in certain bins. This has been seen to add extra, spurious modes of variability to the data set.

The program will read the 'OPTINT.OUT' file from the current working directory, and will signal an error if this file does not exist. The start and end dates must be spaced an integer number of 5 day periods apart, as for the optimal interpolation software.

The following files are output at the end of the run. If these files already exist, they are overwritten.

'EOF.evectors'.

This file contains the empirical orthogonal functions themselves. They are listed in descending order (most significant first) as a height per bin per EOF. The first column contains the EOF number, the next two are the latitude and longitude and the final column is the height above mean sea level in metres.

'EOF.coeff'.

This contains the time series associated with each EOF listed in the same order as the vector file. The first column is the EOF number as in the vector file. The next are the date as MJD (Modified Julian Date - the number of days since November 17th 1858), the amplitude value for this epoch, and finally the date in years to three decimal places (e.g. 1994.279).

The above two files can be separated into multiple files - one per EOF using the shell script 'spliteofs'. By default this will output files for every EOF found in the vector and coeff files, but if an numeric argument N is given, it will only output the first N EOFs. They will be written to files eof.X and coeff.X where X is the EOF number in a sub-directory called 'eofs'. If this directory does not exist it will be created. By default these files will be gzipped. If the required input files exist as gzipped files instead of plain text, the gzipped files will be used. The files produced differ slightly in format to the standard coeff and vector files in that no EOF number column is present.

'EOF.coefftab'.

This file contains the coefficient time series for each of the first 10 EOFs tabulated in columns.

'EOF.scale-factors'.

This contains the maximum absolute values of each coefficient time series. This is used in the chart plotting routines as a scaling factor so the time series can be plotted on a scale from -1 to 1.

'EOF.evalues'.

This is a list of the eigenvalues for each EOF. These are an indication of the proportion of the total signal that each EOF represents.

'EOF.globalav'.

This gives the global average sea level which is calculated from all bins with a weighting applied according to the area of each bin (i.e. by the square of the cosine of the latitude).

'EOF.percent'.

This lists the percentage of total variability signal that each EOF represents, along with the cumulative total.

A.6 Visualisation tools.

There are various programs which can be used to produce postscript plots of this data. Most are shell scripts which use the GMT plotting tools.

'anphch' and 'anampch' produce graphics representing the phase and amplitude of the annual signal from the data output by the spect program.

'tsch' will output a chart representing the RMS sea surface variability overlayed on a map of the world from the TSERIES.OUT file which is output by the time series program.

'eofch' will produce a chart of the EOF number given as the first command line argument. This requires that the EOF vectors and coefficient time series files have previously been split up using the 'spliteofs' command. This will automatically use the gzipped versions of these files if the non-gzipped ones do not exist in the 'eofs' subdirectory.

All these scripts require the presence of a file '.chart_val' in the current working directory which contains the dimensions of the current grid written 'xdim/ydim'. E.g. where a 3° by 2° grid is used, the file would contain just the line '3/2'.

Using this package of software, it is possible to process the crossover residuals and produce graphical representations of all aspects of the resulting data set.