SECTION

7

THE FUTURE

Chapter 7.1
TOWARDS A WOCE SYNTHESIS
Lynne D. Talley, Detlef Stammer and Ichiro Fukumori
525

Chapter 7.2
NUMERICAL OCEAN CIRCULATION MODELLING: PRESENT STATUS AND FUTURE DIRECTIONS
J. Willebrand and D. B. Haidvogel
547

Chapter 7.3
THE WORLD DURING WOCE
Bob Dickson, Jim Hurrell, Nathan Bindoff, Annie Wong, Brian Arbic, Breck Owens, Shiro Imawaki and Igor Yashayaev
557

Chapter 7.4
OCEAN AND CLIMATE PREDICTION – THE WOCE LEGACY
Neville Smith
585

7.1

Towards a WOCE Synthesis

Lynne D. Talley, Detlef Stammer and Ichiro Fukumori

7.1.1 Exploiting the WOCE data set

The primary goal of WOCE is to obtain an understanding of the large-scale circulation of the world ocean, its time variability, and its impact on climate. This knowledge is necessary in its own right and will also help to improve coupled models for climate diagnosis and prediction. A global WOCE survey of the circulation has been completed. Analysis of these data sets is now proceeding, as are improvements in ocean general circulation models and developments in constraining models by the observations. The thrust of these efforts, current progress and future plans are described briefly in this chapter. The WOCE bibliography, maintained by the WOCE International Project Office, already numbers more than a thousand articles, many of which contribute significantly to understanding regional oceanography.

Description and understanding of the oceans at basin and global scales is being addressed during the WOCE Analysis, Interpretation, Modelling and Synthesis (AIMS) phase, as discussed extensively in various chapters of this book. Central issues include the ventilation and the renewal rates of water masses, the determination of oceanic transport of heat, fresh water and tracers, their divergences, and a description of ocean variability during the WOCE period. These activities build on the existence of and widespread access to the comprehensive and pre-assembled WOCE data set, as well as continued development of ocean general circulation models with ever-more sophisticated algorithms and enhanced resolution, and development of data assimilation methods. A number of regional process experiments were undertaken as part of WOCE to address specific scientific questions such as the detailed mechanisms of ventilation through subduction and convection, thermocline and abyssal mixing rates, and abyssal upwelling quantification and processes.

From global WOCE (including altimetry) and pre-WOCE data sets (e.g. existing long oceanographic time series) and through improved modelling capabilities, it is evident that the ocean has vigorous variability on all time and space scales. Increasing emphasis is being put therefore on the changing ocean state and its relation to changing atmospheric forcing conditions at time scales of intraseasonal to decadal. In order to proceed towards understanding this variability, a requirement of the WOCE synthesis phase is to provide the best possible description of the time-evolving ocean state during the WOCE period. WOCE syntheses therefore merge information from various WOCE data sets to enhance our understanding of the ocean circulation in a way that is more complete than can be done from one data set or a model alone. Three associated and overlapping activities are central to synthesis activities: (1) analysis, which includes the assembly and synthesis of data with ocean models, which can be as simple as geostrophy; (2) model testing and improvements, which include comparison of model results with the WOCE data set and the estimate of formal model error descriptions as input for, e.g. ocean state estimation; and (3) ocean state estimation (assimilation), by which ocean circulation models are combined with WOCE and other data sets.

OCEAN CIRCULATION AND CLIMATE
ISBN 0-12-641351-7

The two goals of this chapter are to review the current status of data-based analyses including production of climatologies and atlases (Section 7.1.2), and to give an in-depth description (Section 7.1.4) of current progress in WOCE state estimation (data assimilation) related to regional and global estimates of the ocean state. Progress with ocean general circulation models is reviewed briefly in Section 7.1.3, as modelling is a major component of synthesis; Böning and Semtner (Chapter 2.2) provide a more complete review. We give a tentative prediction of where WOCE synthesis activities will lead and especially what progress can be expected from ocean state estimation over the next few years (Section 7.1.5).

7.1.2 Data-based analyses

Most observation-oriented research is being carried out initially using observations in combination with simple analytical or inverse models. The process of data assembly and availability status are described by Lindstrom and Legler (Chapter 3.5), as well as in the WOCE AIMS plan (WOCE Synthesis and Modelling Working Group, 1997). Many of the data collection and assembly efforts will continue past the end of the WOCE AIMS phase and will provide a significant heritage for future organized programmes such as CLIVAR (Climate Variability and Predictability) or GOOS (Global Ocean Observing System). Links to the WOCE data sets and a large amount of information about WOCE, as well as the WOCE bibliography, can be found at http://oceanic.cms.udel.edu/woce. The 'online' data sources listed through the rest of this text can be accessed at time of writing through this central address.

Primary global WOCE data sources are hydrographic and velocity observations from research ships including Conductivity-Temperature-Depth (CTD) data, chemistry and Acoustic Doppler Current Profiler (ADCP) measurements, eXpendable Bathy Thermographs (XBTs) and XCTDs deployed from merchant ships, subsurface floats, surface drifters, and current meter arrays in restricted locations. With respect to ocean variability unprecedented information is being obtained from the precise TOPEX/POSEIDON altimeter data set (Fu, Chapter 3.3). The global data sets are being combined to answer questions about large-scale processes and to provide a basic description of the ocean during the 1990s. A number of focused regional process studies combined various observational techniques to tackle specific problems.

Where are we now with analysis of WOCE data sets? How healthy is the synthesis phase given the natural tendency for investigators to move on to the next large programme? A more general question relates to the overall 'success' of the WOCE observational programme – how appropriate and well designed were the observational programmes? Is there anything that we as a community would do differently if the planning were to begin again, at the same stage of technical development and computer power as was available in the mid-1980s? The subsections here address the large-scale circulation, process experiments and variability studies. Most of these topics are dealt with in much greater detail in other chapters of this book. Inverse model results are emphasized here because they are not included separately elsewhere in the book.

7.1.2.1 Large-scale circulation and water mass descriptions

Circulation and direct transport estimates relevant to the long-time average ocean circulation are derived from WOCE hydrographic observations and average velocity fields from surface drifters and subsurface floats. Results from many of the WOCE hydrographic sections have been published, including water mass descriptions, transports and regional studies. Some of these results are summarized in Section 4 of this book. This work will continue over the next several years, and forms one basis for the ongoing basin-oriented and global work. Tracer analysis is approached initially on a more basin-wide scale than hydrography because the value tracers add to traditional hydrographic analysis is an inventory of water mass ages and ventilation paths, which are larger-scale questions than are easily answered with individual sections. Current progress in using tracers is described by Schlosser *et al.* (Chapter 5.8). These are areas of continuing major activity for WOCE investigators and are well represented in the large WOCE bibliography.

Work on synthesizing the WOCE circulation observations into basin-wide descriptions has also begun. These take the form of property mapping and merging of all available information to produce the best estimate of circulation at all vertical levels, including both geostrophic calculations using hydrographic data and direct velocity observations.

Lagrangian measurements

Surface drifters have revealed large-scale convergences and divergences as well as the average 'Ekman' layer flow. Using surface drifters drogued at 15 m, Ralph and Niiler (1999) produced the first large-scale map of the ageostrophic component of near-surface flow and a complete Ekman spiral, providing a basis in observations over a very large region for one of the fundamental tenets of large-scale oceanography, that of Ekman-like response to the winds. They were able to show that the ageostrophic component is proportional to wind speed rather than wind stress, and inversely proportional to the square root of the Coriolis parameter. Their result has been confirmed with a much larger data set and the results are shown in Fig. 4.1.6 (Niiler, Chapter 4.1).

Subsurface floats were deployed on a global scale for the first time during WOCE and are a major element of WOCE circulation mapping. Floats have already revealed or validated major circulation patterns in many regions (Davis and Zenk, Chapter 3.2). Results from this new and growing global data set are being produced at an increasing rate. The following are a few examples. In the tropical Pacific, Davis (1998) has shown vigorous zonal circulation with a significant seasonal component throughout the tropical region in strong contrast with the gyre circulations of the mid-latitude South Pacific. He found general agreement with the mid-latitude circulation patterns produced by Reid (1997) based on geostrophy and careful selection of reference velocities based on tracers.

In the northern North Atlantic, where subsurface float density exceeds that in any other region, Lavender *et al.* (2000) have shown that the cyclonic boundary circulation of the Labrador Sea is accompanied by a weak but permanent anticyclonic circulation offshore of the boundary circulation (Fig. 3.2.8), with potential major ramifications for the location of mid-depth convection producing Labrador Sea Water. The stability of the North Atlantic Current and its meanders along the western boundary and of its eastward turn as it separates have been demonstrated with subsurface floats (Rossby, 1996). In the South Atlantic, Davis *et al.* (1996) used floats to examine the validity of general circulation model results and showed the importance of eddies in the confluence of the Brazil and Falkland Currents. Boebel *et al.* (1997) used floats and other direct velocity data sets to show the westward flow of intermediate waters towards the western boundary and bifurcation of the flow at the coast (**Fig. 3.2.7**), confirming that the source of intermediate waters to the tropics comes from within the subtropical gyre rather than from the western boundary.

Geostrophic circulation analyses

Full assimilation of hydrographic and velocity data with models is rapidly developing (see Section 7.1.4 below). However, methods employing hydrographic observations short of full assimilation are and have been used fruitfully, including geostrophic velocity reference level choices based on geostrophic shear and water mass properties (as in McCartney, 1992; Reid, 1994), box inverse models (as in Wunsch, 1978), beta-spiral methods applied over a large area (as in Olbers *et al.*, 1985), and grid-based adjoint inverses (as in Schlitzer, 1995). These methods can also incorporate direct velocity measurements as constraints or as initial velocity choices.

The first global box inverse models have been produced recently, using zonal, pre-WOCE, hydrographic sections (Macdonald and Wunsch, 1996; Macdonald, 1998) and with WOCE data including nutrients (Ganachaud and Wunsch, 2000). Global mass transports from these two efforts are shown in Figs 7.1.1 and 7.1.2 (see Plates XX and XX). The deep water formation rate for the North Atlantic in both works is comparable to earlier estimates. Macdonald tested the sensitivity of the global inverse to variations in Pacific–Indian transport through the Indonesian archipelago, showing that the region of influence of the throughflow is limited to the Indian and South Pacific Ocean, without a significant effect on the global overturning cell magnitude. Ganachaud and Wunsch's new result of a bottom water formation rate of 22 Sv (1 Sverdrup $= 10^6\,m^3\,s^{-1}$) for the Southern Ocean is made possible by the new circumpolar hydrographic observations, and indicates that the higher range suggested by Schmitz (1995) is likely valid.

On basin scales, analyses including WOCE data are well underway. Perusal of the abstracts from any of the WOCE basin workshops indicates the large amount of synthesis work, including a number of box inverse models. Reid (personal communication) has nearly completed an Indian Ocean circulation analysis similar to his previous Pacific and Atlantic works (e.g. Reid, 1994), and incorporating all WOCE Hydrographic Programme (WHP) data along with historical data. Ganachaud *et al.* (2000) focus on the Indian Ocean portion of their global inverse, showing a somewhat higher Indonesian Throughflow and lower inflowing deep water than in previous estimates, thus somewhat reducing the strength of the overturning circulation compared with previous results. They also use the inverse model results to calculate surface heat and fresh water fluxes, providing a useful constraint on climatologies. Sloyan (1997) used hydrographic sections, including several WHP sections, in a box inverse of the Southern Ocean, including air–sea and dianeutral fluxes to constrain the water mass transformations. She diagnosed the formation rate of SubAntarctic Mode Water (SAMW) north of the Antarctic Circumpolar Current (ACC), and showed that the sources of SAMW are thermocline water north of and surface water south of the ACC. The southward meridional heat transport across 30°S was dominated by the Indian Ocean.

A consortium in the US and Australia (Robbins, Wijffels, Toole, Johnson and Talley) is currently making a box inversion of the Pacific WHP data using initial reference velocity fields based on examination of all of the sections, and incorporating constraints from direct velocity measurements. Initial results for heat transports and convergences are encouraging. Robust results can be obtained despite the use of intersecting, non-synoptic sections, showing that zonal heat redistribution is larger than meridional in the South Pacific. Heat flux convergences for the many boxes formed by the intersecting sections have been compared with many different air–sea flux climatologies, showing that model-data-based climatologies are better than data-only climatologies in the data-poor southern South Pacific, and providing bias estimates for the net heat fluxes computed from the various climatologies.

For the South Atlantic, Holfort and Siedler's (2000) inversion of WOCE sections provides results for mass, heat, fresh water and nutrient transports and the conclusion that the warm water that feeds the North Atlantic Deep Water overturn is both upper ocean and Antarctic Intermediate Water, with a higher fraction of the upper ocean waters. B.A. King (personal communication) is using a box inverse to study the southwestern Atlantic circulation; results thus far suggest that excess intermediate water enters at Drake Passage and excess surface water leaves the box, supporting Rintoul's (1991) result of net upwelling in the southwestern Atlantic.

Cunningham (2000) used a Bernoulli inverse to construct streamlines for the North Atlantic, with emphasis on the North Atlantic Current and subpolar region. He found a northward influx of 14 Sv into the eastern subpolar region and that the Irminger Basin circulation is equally fed from the North Atlantic Current west of the Reykjanes Ridge and westward flow south of Iceland. Constraints on the outflow from the Mediterranean are being produced using an inverse model in a box around the outflow, with initial results showing almost all flow exiting to the north and west despite the prevalence of Meddies to the southwest (D. Slater and H.L. Bryden, personal communication).

Work on global inversions will continue with the complete WOCE data set, providing improved estimates of air–sea fluxes, diapycnal velocities, and complete circulation schemes. These inversions are of increasing complexity, which in its limit is equivalent to state estimation. For instance, a group at the Alfred-Wegener-Institut in Germany is using a number of complementary inversion and data assimilation techniques to study the mean and time-varying circulation of the Atlantic. Continuing section and basin-oriented work will contribute to further improvement of the global-scale work. Results may also eventually be used to test basic hypotheses about ocean dynamics.

Basin-scale property mapping as an aid to circulation analyses is included in some of the abovementioned works, particularly that of Reid. The basin atlases, described below, will include property maps. The potential vorticity field is important for planetary wave propagation (e.g. deSzoeke and Chelton, 1999) and also provides information like that of other tracers for circulation. O'Dwyer and Williams (1997) mapped the global potential vorticity field using historical data, showing the zonality of potential vorticity at mid-depth and large regions of potential vorticity that deviate from zonal at abyssal depths, including near homogeneity in the deep North Pacific. WOCE hydrographic data allows local features to be more closely defined, as in the Indian Ocean potential vorticity fields mapped using WHP and historical data (McCarthy and Talley, 1999). The Indian Ocean potential vorticity field is zonal at all depths in the tropics; at mid-latitudes the zonality fills most of the mid-depth layer while at greater depths contours of potential vorticity follow the bathymetry; homogenization does not occur at any depth or in any region.

Transport calculations

For heat and freshwater transport calculations, the continuing battle to improve surface fluxes has not yet been won, although significant progress in producing higher-quality flux data sets has occurred in WOCE (Josey *et al*., 1999; WGASF, 2000). The remaining large error bars on surface fluxes mean that direct transport estimates based on *in-situ* measurements remain the best. Heat and freshwater transport estimates from WHP sections and from surface fluxes are summarized in Section 6. These section-based works are preliminary, as almost all are based on single section analyses without an attempt at a unified circulation picture for at least an entire basin. Basin-wide and global transport estimates based on consistent analyses of the complete data set will likely improve the current estimates, and will also make direct ocean transport estimates available for many additional sections.

7.1.2.2 Process studies

Dedicated local observational studies clarify issues important for ocean circulation and hence ocean- and climate-related modelling. WOCE process studies have been carried out with multiple observational tools, many with integrated modelling studies. Completed are the subduction experiment in the eastern North Atlantic (Price, Chapter 5.3), a mixing experiment in the same region (Toole and McDougall Chapter 5.2), and an experiment on abyssal circulation, upwelling and mixing in the geographically restricted Brazil Basin (Hogg, Chapter 4.6). Other regions that have been studied intensively with respect to a specific process are the Labrador Sea for convection (Lazier *et al*., Chapter 5.5), the Weddell Sea (Rintoul *et al*., Chapter 4.7), and flow through the Indonesian passages (Gordon, Chapter 4.8). Many of these latter experiments were carried out with both WOCE and non-WOCE resources; all contribute importantly to understanding ocean circulation processes and hence the goals of WOCE.

A regional study of North Atlantic Deep Water (NADW) formation in the subpolar North Atlantic is in the synthesis phase. Field work was started late in WOCE (in 1996) and is continuing, now as part of CLIVAR and national research programmes. Components of this work include transformation of upper ocean waters in the subpolar gyre, inflow and outflow over the Greenland–Scotland ridge, flow into and out of the Labrador Sea, and ongoing study of Labrador Sea Water formation. Learning how the thick surface layer changes properties, especially temperature and salinity, is central to the overall NADW problem. It is apparent from the new observations that this subpolar layer does not change properties in the continuous and large-scale fashion suggested by McCartney and Talley (1982) using a coarser and non-synoptic data set; rather the North Atlantic Current (subarctic front) and underlying circulation following the topography provide a complex circulation for the surface layer, with preferred regions of convection near topographic features (Talley, 1999).

Exchange between the Nordic Seas and the North Atlantic is a central practical problem for basin- and global-scale ocean models, which do not resolve straits sufficiently to model the physics correctly there. This problem is described by Saunders (Chapter 5.6). One thrust of the ongoing NADW work is to describe as completely as possible the overflow and to model it theoretically and with high-resolution numerical models. Recent results presented at the WOCE North Atlantic workshop show in particular the presence and strength of eddies in the Denmark Strait, which might influence interpretation of moored results within the strait (Krauss and Käse, 1998; Käse, 2000; Käse and Oschlies, 2000).

7.1.2.3 Variability studies

WOCE goals include observing, where possible, the ocean's variability at time scales relevant for climate. There are many open questions, ranging from simply characterizing the variability and its spatial structure, to understanding its source and relationship to forcing. Thus in addition to the long-term average circulation and water properties, studies include the adjustment of the ocean to changing forcing through planetary wave propagation and the eddy field, the seasonal adjustment of the circulation, and interannual (several years) and decadal phenomena using both historical and WOCE data. Relevant WOCE data sets include especially the XBT, altimetric and Lagrangian data sets, as well as extensions of hydrographic time series and current meter statistics.

A question that has been posed since the start of WOCE is to determine the representativeness of the WOCE data set. One way to phrase this is to ask how useful the large-scale fields determined in WOCE are for understanding the long-term ‘mean’ circulation. One obvious measure is where the WOCE period fits relative to various known interannual and decadal climate phenomena and then to determine the size of the WOCE ‘mean’ signal compared with the extrema of these climate phenomena. The latter requires detailed examination in every region, although one rule of thumb is that outside the tropics, temperature variations due to climate variations are on the order of 1°C in the upper ocean. These regional studies are being carried out, most notably in the North Atlantic and some regions of the southern hemisphere, as attention has shifted from describing the mean fields to unravelling the climate variability signal. WOCE provides an extremely good sampling for one period for all of these studies.

Some of the interannual and decadal climate phenomena that affect ocean properties and that have received major attention are El Niño, the Atlantic tropical dipole, the North Atlantic Oscillation, the Pacific North American mode (decadal Pacific oscillation), the Arctic Oscillation, southern hemisphere annular modes and the Antarctic Circumpolar Wave. Indian Ocean tropical and subtropical dipoles have also been described recently (Saji *et al.*, 1999; Behera and Yamagata, 2000). Knowledge of these phenomena is expanding rapidly, including their interconnections. It is likely that other patterns will also be described and that understanding and hence quantification of already-known climate phenomena will be greatly advanced in the near future. This is particularly true for the southern hemisphere, which has not yet received the full scrutiny given to the tropics and the northen hemisphere. These interannual and decadal indices are shown in Fig. 7.1.3. Dickson (Chapter 7.3) provide a complete discussion of the WOCE period in the context of several of these well-described climate cycles.

The baroclinic and barotropic planetary wave fields at seasonal and shorter time scales are evident in any basin-scale general circulation model (e.g. Semtner and Chervin, 1992; Böning *et al.*, 1996) and are observed with satellite altimetry (e.g. Stammer and Wunsch, 1999). The latter found that seasonal variations in surface slope are maximum in the tropics and northeastern Atlantic and northeastern Pacific, and are correlated with changes in wind only in the latter two regions, contradicting earlier ideas that most of the energy outside the western boundary currents is due to local wind effects. There is much geographic variation in surface height and eddy energy, both in amplitude (e.g. historical studies by Wyrtki, 1975, and more modern equivalents) and in wave propagation as demonstrated with altimetry (Chelton and Schlax, 1996). The latter showed that phase speeds depend on more than Coriolis parameter variation, and hence must be a function of the ocean’s stratification, circulation and possibly topography (Killworth *et al.*, 1997; deSzoeke and Chelton, 1999).

Gyre-scale seasonal variations in the thermocline have been shown from the WOCE high-resolution XBT programme in the Pacific. Although the phasing of the wind field with these variations has been studied, the dynamical relationship between them has not yet been unequivocally established and remains an open question. Morris *et al.* (1996) reported an annual cycle in the subtropical gyre in the western South Pacific, with isotherm slopes steepest and the geostrophic circulation most intense in late spring, 2–3 months after the maximum Ekman pumping. McCarthy *et al.* (2000) showed a puzzling phase relation between subsurface isotherm depth, Sverdrup transport and sea surface height on interannual time scales across both the South and North Pacific subtropical gyres, with surface height increasing as expected with increased Sverdrup transport, but the subsurface isotherms shoaling at the same time.

Variations in meridional heat transport at midlatitudes are also being studied using high-resolution XBT transects. Roemmich *et al.* (2000) used a large number of transects at 24°N in the Pacific to obtain 0.76 PW of northward heat transport, with a robust error estimate of 0.12 PW; they are the first to estimate the interannual variations, at 0.3 PW, a substantial fraction of the mean.

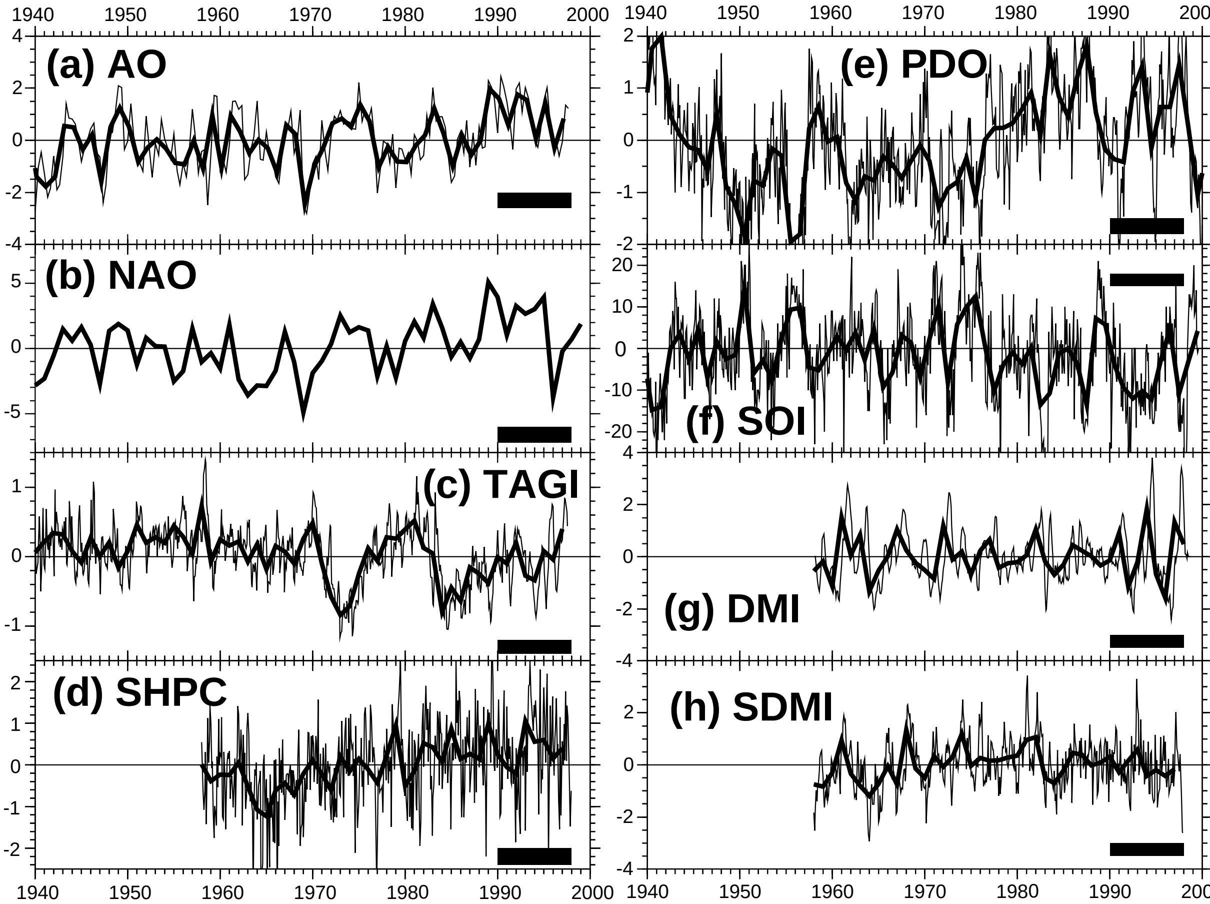

Fig. 7.1.3 Various climate indices since 1940 and the WOCE observational period (horizontal bar). (a) Arctic Oscillation index (AO) (Thompson and Wallace, 1998); (b) North Atlantic Oscillation (NAO) (http://www.cgd.ucar.edu/cas/climind/nao_winter.html;, Hurrell, 1995); (c) Tropical Atlantic Gradient Index (TAGI) (Chang *et al*., 1997); (d) Southern Hemisphere principal component (SHPC) (Thompson *et al*., 2000); (e) Pacific Decadal Oscillation (PDO) (http://tao.atmos.washington.edu/pdo; Mantua *et al*., 1997); (f) Southern Oscillation Index (SOI): Sea Level Pressure (SLP) at Tahiti minus SLP at Darwin, from http://tao.atmos.washington.edu/data_sets, based on Ropelewski and Jones (1987); (g) Indian Ocean dipole (DMI) (Saji *et al*., 1999); (h) subtropical Indian Ocean dipole (SDMI) (Behera and Yamagata, 2000). Heavy line, annual average; light line, monthly values (for b–h) and January–March values (AO only). See Dickson *et al*. (Chapter 7.3) for more analysis of the climate states during WOCE.

At seasonal to interannual time scales, variation in the upper ocean structure of the tropical Pacific was a central focus for the Tropical Ocean Global Atmosphere (TOGA) programme. In other tropical and mid-latitude regions, WOCE data sets as well as time series of hydrographic, XBT and surface data dating back several decades have been used profitably to learn about these time scales. In the North Pacific, the oceanic signature of the decadal time scale has been clarified using the XBT data set collected since the 1970s and continued in the 1990s as part of WOCE (Deser *et al*., 1996 and Schneider *et al*., 1999, showing the subsurface propagation of temperature anomalies around the subtropical gyre; Tourre *et al*., 1999, describing the decadal cycle of temperature anomalies throughout the mid-latitudes).

The magnitude of the Indonesian Throughflow is important qualitatively for the global thermohaline circulation and quantitatively for global heat budgets. Several major field projects are placing bounds on the transport, which has a large interannual range but which appears to average around 9 Sv through Makassar Strait with a range of 2.5 Sv based on current meter measurements (Gordon and Susanto, 1999; Gordon *et al*., 1999). Transports were minimum during an El Niño and maximum during a La Niña during the limited observational period. The Throughflow is estimated at 10–15 Sv for all passages based on pressure gauge pairs

(Chong *et al*., 2000). Variations in the geostrophic transport through the Indonesian Throughflow based on XBTs show a mean of 5 Sv and a large seasonal and interannual component, with the transport reaching as high as 12 Sv (Meyers *et al*., 1995), the latter in agreement with current meter measurements in Makassar Strait (Gordon *et al*., 1999; see also Gordon, Chapter 4.7).

Monitoring the transport of the Kuroshio south of Japan has been carried out with *in-situ* current and hydrographic observations (WOCE PCM5 array) and TOPEX/POSEIDON altimetry data (Imawaki *et al*., 1997, 2000, Fig. 7.3.11). The 7-year long record (1992–99) shows a mean transport of 57 Sv with large variability. Of this, 42 Sv flows on through, excluding recirculations, and has only a small seasonal signal compared with the transport of the western boundary current inferred with the theory for a wind-driven, flat-bottom ocean. Thus most of the variability of the 57 Sv is due to variations in the recirculation.

Variations in production rates and properties of various components of the NADW cell in the North Atlantic have been studied using WOCE and historical data. Some studies have employed current meter arrays in strategic locations such as the Denmark Strait Overflow and most include repeat hydrography. These studies are continuing as part of CLIVAR, and should be a part of any climate-monitoring system.

Interannual to decadal variation in water properties have been studied in most oceans where there are repeat hydrographic data, including considering the WOCE survey as a 1990s 'data point', for comparison with previous decades (Dickson *et al*., Chapter 7.3).

Because the mid-latitude North Atlantic is the seat of a major decadal climate phenomenon, the North Atlantic Oscillation (NAO), considerable effort has been given to characterizing upper ocean anomalies of temperature and salinity in the North Atlantic, their relation to circulation, and relation to the NAO. Hansen and Bezdek (1996) and Sutton and Allen (1997) demonstrated the propagation of surface temperature anomalies originating near the western boundary eastward with the North Atlantic Current.

Labrador Sea Water (LSW) is a major component of the NADW that departs southward from the North Atlantic into the southern hemisphere. Ongoing hydrographic work along a section in the Labrador Sea during WOCE has continued the time series started as Ocean Weather Station Bravo, but with coast-to-coast coverage including the boundary currents. Major shifts in LSW properties (temperature, salinity and thickness) were apparent from earlier data, but the shifts towards fresher and colder LSW in the 1990s were especially large (Lazier, 1995). The response of LSW production and properties to the NAO both within the Labrador Sea and downstream several years later in the subtropical gyre has been clearly demonstrated (Curry *et al*., 1998), with LSW thinning and warming when the NAO index decreases, and vice versa. Sy *et al*. (1997) showed that the entire subpolar North Atlantic at mid-depth responded within 5 years to changes in convection properties in the Labrador Sea. Koltermann *et al*. (1999) compare mass, heat and freshwater transports in the North Atlantic for the 1950s, 1980s and 1990s, concluding that there was highest northward transport of warm water and southward transport of deep water in the 1980s, coinciding with a decrease in production and transport of Labrador Sea Water.

Time series at point locations were continued during WOCE at several locations, including at Bermuda and at Ocean Weather Stations Mike (Norwegian Sea) and Papa (northeast Pacific). A new time series was established at Hawaii. Ongoing hydrographic work in the western North Pacific continues long-time-series hydrographic sections there. Repeat sections have also been established in the Southern Ocean at key locations crossing the Antarctic Circumpolar Current. Many of the results obtained from incorporation of the WOCE time series, and from considering the WOCE one-time survey as a decadal data point, are described by Dickson *et al*. (Chapter 7.3).

7.1.2.4 Climatologies and atlases

The WOCE observational programmes produce both basic data sets and averaged fields (climatologies). The latter will be substantial improvements to currently available climatologies and will serve as a basic description of the mean ocean state, helping to improve and initialize circulation models. Averaged fields include means over different time periods and regions, and various sets of statistics. These statistics are necessary for the state estimation efforts described below, which will use basic data for the most part rather than climatological data.

Ultimately we hope that the best climatologies will become available from model–data combinations using the WOCE and historical data sets. Most of the WOCE data sets and expected climatologies are described here: hydrographic, upper ocean thermal, velocity and surface fluxes.

Hydrographic data

Hydrographic climatologies are being produced by a number of different groups; the principal efforts are described here. The most widely used climatologies are from the National Oceanographic and Atmosphere Administration (NOAA) (Levitus, 1982; Levitus *et al*., 1994), utilizing the global pre-WOCE hydrographic data set with quality control; the 1994 version includes monthly and annual climatologies of temperature and salinity on a 1° latitude by 1° longitude grid at 33 standard depths. It is now understood that averaging along isopycnals/isoneutral surfaces is superior to averaging at standard depths, as the latter results in spurious climatological features in the neighbourhood of the strong isopycnal slopes that occur at major fronts (e.g. Lozier *et al*., 1995). New climatologies employing evolving methods and the ever-enlarging hydrographic data set will continue to be produced at NOAA.

Climatologies of the South Atlantic (without WOCE data) and South Pacific (with WOCE and historical data) have been produced by the WOCE Hydrographic Programme Special Analysis Center (WHP SAC) in Hamburg (Gouretski and Jancke, 1995, 1996), with a preliminary version for the world ocean. The input data were subjected to major quality control, and many data were adjusted using regional average temperature–salinity relations. Along-isopycnal mapping was used, using objective analysis, prior to interpolation to standard depths for the output climatology. The mapped fields are provided on a 1° grid at up to 43 standard depths in the vertical. The quality-controlled and adjusted data and the averaged fields are available online.

The Lozier *et al*. (1995) climatology for the North Atlantic is the basis for the first interactive (electronic) climatology/atlas, which is being developed further as Hydrobase by R. G. Curry (personal communication, 1999). The goal in this work is to provide the most flexible access to the data sets, allowing the investigator to construct whatever average is most useful, along with error estimates. This then allows comparison of fields from different time periods more easily than does a single climatology (e.g. Curry *et al*., 1998). Rigorous quality control of the input data sets is performed; this is the step that requires the most time prior to making the atlas available. The quality-controlled data sets are available through the atlas. Averaging is along isopycnals, although output can be at standard depths as well as on isopycnals. Bin averaging is used in the horizontal, but is being migrated to optimal interpolation. The interactive atlas is available for the Atlantic Ocean, and is continually updated with newly collected data sets. A North Pacific version has been produced (Macdonald *et al*., 2000). The remainder of the world ocean is under construction and will be available in late 2000 (R. G. Curry, personal communication, 1999).

WOCE has produced the most comprehensive hydrographic data set ever collected, both in terms of global coverage and in types of observations. To present both a quick, accessible view of the data sets in cross-section and isopycnal maps of properties, atlases of hydrographic data are being constructed for each ocean basin (Pacific, Atlantic, Indian and Southern), in both printed and interactive electronic formats. Atlases are being produced by different groups for each ocean basin working in close collaboration to produce uniform products. The WOCE one-time survey station sampling strategy was optimized to resolve features along sections with a minimum of spatial aliasing, including the all-important density field for geostrophic calculations. The survey was intentionally not optimized for horizontal (along-isopycnal) coverage, given the existence of a historical data that provide reasonable coverage of most regions for the basic properties (temperature, salinity, oxygen and nutrients). Therefore the isopycnal maps will incorporate high-quality historical data as well as WOCE data. The electronic atlases for the WOCE hydrographic data set will include the information in the printed atlases, links to the original data sets, and interactive plotting capability.

Upper ocean thermal fields

Upper ocean thermal fields are constructed primarily from XBT observations collected by merchant ships. This data set is much larger than the research ship hydrographic data set. A new source of upper ocean thermal data are profiles from profiling subsurface floats, which collect temperature data when they cycle to the surface. WOCE's contribution to upper ocean thermal data sets has been three-fold: collection of XBT data; development of the profiling float; and provision of the framework for quality control and archiving of the datasets. In these efforts, only the high-resolution XBT programme (closely spaced profiles along selected seasonally repeated merchant ship transects) was entirely a WOCE programme; all other XBT work has been joint with other climate programmes such as Tropical Ocean Global Atmosphere (TOGA), which recognize the importance of regular thermal sampling in the upper ocean for climate studies. Float work and especially technical development was also supported by other programmes in addition to WOCE, although the primary impetus and much of the funding came from WOCE.

The thermal data are used as inputs to developing coupled prediction models. The Global Temperature–Salinity Pilot Project (GTSPP), run by the National Oceanographic Data Center (NODC) in the US, is the central resource for this programme at this time. After collection of XBT data at NODC, data are quality-controlled by three data assembly centres: the US's Scripps Institution of Oceanography (SIO) (Pacific Ocean), Australia's Commonwealth Scientific and Industrial Research Organization (Indian Ocean), and the US's NOAA Atlantic Oceanographic and Meteorological Laboratory (Atlantic Ocean). Data are available online from the NODC. Profiling float data, for which collection began around 1995, are still in an experimental stage, particularly with regard to archiving; at this time only the temperature profiles are sent to NODC.

Climatologies, including monthly averages of temperature, heat storage and mixed-layer depth fields and anomalies, have been produced by the Joint Environmental Data Analysis Center (JEDAC) at SIO (White, 1995) and are easily accessible online at SIO and through the World Meteorological Organisation/Intergovernmental Oceanographic Commission site at Lamont-Doherty Earth Observatory. Analyses for the Indian Ocean are available online from CSIRO in Australia. The fields are used for basic research and comparison with models (e.g. Auad *et al*., 1998).

Velocity data sets

Velocities are measured directly in three ways in WOCE: current meters at fixed sites, Lagrangian drifters and floats, and acoustic Doppler current profiling. Current meter records, including means and statistics, constitute a type of climatology, albeit based only on the time period of each current meter record. Pre-WOCE (pre-1989) current meter statistics were compiled by Dickson (1990), with an update from the WOCE period compiled by Dickson, Medler and Wooltorton; both are available online through the WOCE Current Meter Data Assembly Centre (CMDAC). The CMDAC also maintains statistics for WOCE current meters.

WOCE provided a major expansion of subsurface float coverage compared with the limited deployments prior to WOCE and critical breakthroughs in float technology with the aim of global coverage. Many of the subsurface floats at mid-depth deployed in WOCE are still functioning and data are still being recorded. The next few years will see continuation and expansion of the WOCE profiling float coverage (the Argo programme), largely for the temperature (and salinity) profiling capabilities of the floats, as part of international climate programmes. Velocity (displacement) means and statistics will be available for much of the world at the end of the deployment period; results for the tropical Pacific (Davis, 1998), southwestern Atlantic (Davis *et al*., 1996) and Labrador Sea (Lavender *et al*., 2000) are already published, as described in the previous subsection. Maps will likely be included in the hydrographic atlases along with isopycnal property maps. These groups also construct dynamic topography maps from the averaged float displacements, for better comparison and integration with hydrographic data sets.

TOGA and WOCE, as well as many other funding agencies in many countries, have provided the most comprehensive surface drifter coverage yet available. These observations include water temperature as well as velocity, and many drifters also measure meteorological parameters. Surface drifter data and products are available online at this time through the surface drifter data assembly centre.

Surface fluxes

Surface flux products, including winds and components of heat and water fluxes, are computed and compiled by a number of different organizations, described by Lindstrom and Legler, Chapter 3.5, Bryden and Imawaki, Chapter 6.1, and Wijffels, Chapter 6.2. Climatologies are available from these many different sources. The surface flux Special Analysis Center has its own data products for the North Atlantic, tropical Pacific and tropical Indian Oceans, including monthly means and anomalies for many years as well as overall means. Other primary sources are the NCEP and ECMWF re-analysis projects, climatologies produced by the Southampton Oceanography Centre (Josey *et al*., 1999), and a large number of products accessible through the online LDEO Data System.

Discriminating between these many different products, which are relatively easy to access through the WOCE Special Analysis Center website, has been a large challenge for WOCE. A major and much-desired recent activity of the air–sea flux community has been the comparison of the many available flux products and parameterizations. An international working group designated by SCOR (Scientific Committee on Ocean Research) and the WCRP has recently completed this comparison. This Working Group on Air–Sea Fluxes (WGASF) considered the requirements for surface flux data sets, determined what data sets are currently available, and analysed their strengths and weaknesses. It is not possible to summarize their results briefly here; the reader is referred to their report (WGASF, 2000).

7.1.3 Model evaluation and development

WOCE synthesis involves the development and testing of models for a range of purposes including: (1) revealing and studying the physics of processes relevant to the general circulation; (2) simulation of the ocean circulation with good representation of eddy statistics important for defining the semi-permanent features of the circulation; (3) good simulation of the ocean circulation at resolutions that are useful for coupled climate modelling; and (4) use in ocean state estimation as a way of providing the most consistent representation of the observed fields. The first three topics are covered well in other chapters in this book.

Models used for ocean state estimation must be evaluated with respect to their realism in simulating observed processes. The models must be reasonably consistent with ocean data prior to using them for state estimation. Such comparisons are a usual step in using the models to understand and simulate processes that have not been observed as a result of undersampling. By incorporating data in the models through state estimation, another stage of model evaluation is reached, ultimately leading also to model improvements where needed (e.g. see Fu and Smith, 1996; Stammer *et al*., 1996). The two examples discussed here that indicate that models, at least in some of their state variables, show an intriguing agreement with what has actually been observed in the ocean.

A comparison of sea surface height time series simulated by the POCM 4C model (R. Tokmakian, personal communication; see Chapter 2.2) with data from five tide gauge stations on the west coast of the Americas is given in Fig. 2.2.3 of Böning and Semtner (Chapter 2.2). The model noticeably simulates most of the observed climate events, at least qualitatively – the correlation coefficients vary between 0.57 and 0.88. A similarly encouraging result can be found from a comparison of velocity fields from the MIT model (Marshall *et al*., 1997a,b) with TOGA-TAO velocity data at 0°N, 140°W (**Fig. 7.1.4**; T. Lee, personal communication, 1999). The comparison indicates reasonable skill in simulating the time mean and variability of the upper equatorial Pacific circulation, an important prerequisite for using the model to study seasonal to interannual variability. The mean flow field shows some biases, however, in the core of the undercurrent.

Uncertainties in prescribed mixing coefficients and surface forcing fields are believed to cause these observed model–data differences. Model development efforts aim at reducing such differences in order to provide a best possible model for process-oriented studies and data assimilation. See Böning and Semtner (Chapter 2.2) and Willebrand and Haidvogel (Chapter 7.2) for detailed discussions of ongoing and anticipated model improvement activities.

It should be noted that ocean state estimation often includes estimation of surface flux fields that bring the model into consistency with ocean observations. Resulting estimates need to be consistent with meteorological centre-provided fields and their uncertainties. Inconsistencies in turn will shed light on errors in the atmospheric and ocean models.

7.1.4 Ocean state estimation

While data-based studies make use primarily of simple models, such as geostrophy and mass

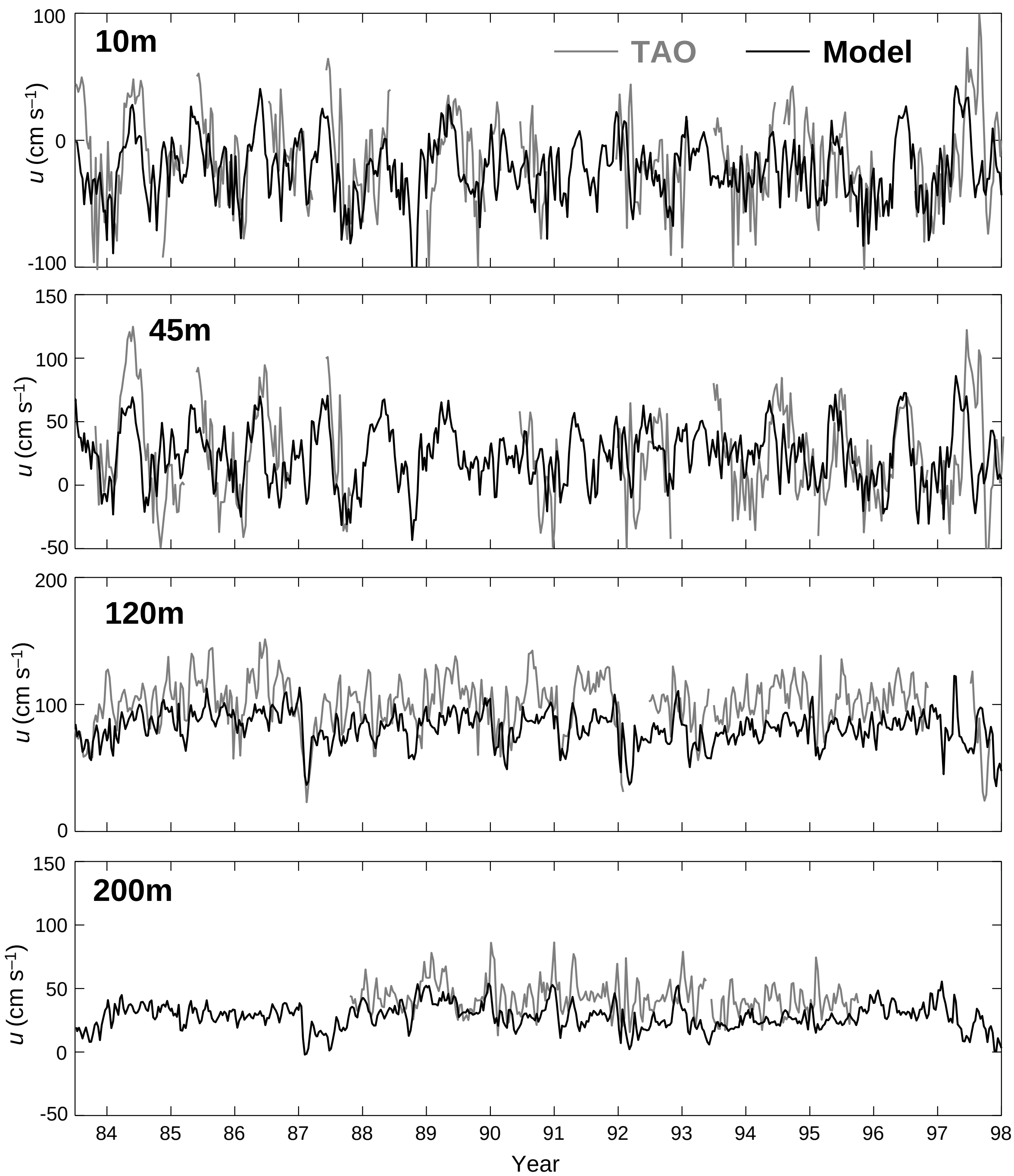

Fig. 7.1.4 Comparison of zonal velocity at four depths in the MIT model (Marshall *et al*., 1997a,b) with current meter measurements obtained from a TOGA-TAO mooring at 0°N, 140°W. The comparison indicates reasonable skill in simulating the time mean and variability of the upper equatorial Pacific circulation, an important prerequisite for using the model to study seasonal to interannual variability. Figure courtesy of T. Lee, personal communication (1999).

conservation, numerical ocean circulation models provide a comprehensive means of analysing observations based on a more general physical principle. Moreover, a truly optimal analysis of the ocean will be achieved by utilizing all available measurements, including altimetry, hydrography, float and drifter velocities, XBT temperature profiles, and ultimately tracer and nutrient data. Ocean state estimation, often referred to as data assimilation, involves combining these diverse observations with ocean models. We expect that such an approach will lead to estimates of the time-evolving ocean state and its uncertainties, as well as insight into missing model physics and uncertain model parameters, such as mixing. Considerable experience is needed in assessing the realism of ocean models and in evaluating the adequacy and design of climate-observing systems. Both areas benefit greatly from data analyses and ocean state estimation during the WOCE synthesis phase, as will be described below.

Although the field of ocean state estimation is still in its early stage of development, the substantial progress achieved over the last years has given rise to great expectations of what can be attained in the near future. Several attempts are now underway to estimate the time-evolving ocean state for up to a decade from basin to global scales. Examples include synthesis activities in the Indian, Pacific and Atlantic Oceans, with many accompanied by efforts on the global scale.

7.1.4.1 The estimation problem

Thorough treatments of ocean state estimation can be found in the text books by Bennett (1992) and Wunsch (1996) and recent applications are summarized in Malanotte-Rizzoli (1996). Marotzke and Willebrand (1996) discuss details of a variety of ocean inverse models. Procedurally, ocean state estimation interpolates and extrapolates data information in space, time, and among different variables exploiting the model equations. In the process, the information is combined with other data, which further improves the description of the oceanic state. In essence, state estimation is a dynamic extrapolation as well as a synthesis and averaging process.

Figures 7.1.5 and 7.1.6 (see Plates XX and XX) illustrate the extrapolation and averaging effects of data assimilation. Figure 7.1.5 shows sea-level anomalies as a function of longitude and time in the tropical Pacific Ocean (Fukumori, 1995). The figure illustrates how assimilation (b) corrects the model (c) to be consistent with TOPEX/POSEIDON observations (a), including the magnitude of seasonal changes and the westward-propagating waves. At the same time, the assimilation eliminates noise from the data and interpolates over data dropouts (grey area in the left panel). Figure 7.1.6 (see Plate XX) illustrates how assimilation extrapolates data information (sea level) into improvements of different properties of the ocean, such as subsurface temperature (a) and subsurface zonal velocity (b). Sea-level information (TOPEX/POSEIDON) corrects the model (red) into an estimate (blue) that is in closer agreement with independent *in-situ* measurements (black), consistent with formal error estimates (bars) (Fukumori *et al*., 1999).

As opposed to numerical simulations, ocean state estimation is mathematically an inverse problem. That is, instead of computing properties of the ocean (e.g. model equivalent of data) given the state (numerical simulation), assimilation estimates the state based on observations. The difference between state estimation and simpler inversions described in Section 7.1.2.1 is in the underlying model dynamics. Instead of simple balances and conservation statements applied over large regions or with a steady-state assumption, data assimilation employs the full, time-dependent, non-linear physics embodied in numerical ocean circulation models at every model gridpoint. As such, the unknowns ($\mathbf{x}$) are independent variables of the model, such as temperature, salinity and velocity, over the entire model domain, and various model parameters, as well as uncertainties in external forcing and boundary conditions.

From a practical standpoint, the distinguishing property of data assimilation in relation to other inverse problems is its enormous dimensionality. Typical ocean models contain on the order of several million independent variables at any particular instant. The size of the problem precludes any direct method of solution and, as a result, devising methods of assimilation has been one of the central issues in ocean state estimation.

Many of the so-called ‘advanced’ assimilation methods originate in estimation and control theories (e.g. Bryson and Ho, 1975), which in turn are based on ‘classic’ inverse methods. These include the adjoint, representer, Kalman filter and related smoothers, and Green’s function methods. These techniques are characterized by their explicit assumptions under which the inverse problem is solved consistently. The assumptions include, for example, the data and model weights used in the inversion and specific statistical criteria in choosing particular ‘optimal’ solutions, such as least-squares, minimum error variance and maximum likelihood. As with ‘classic’ inverse methods, these assimilation schemes are equivalent to each other and result in the same solution so long as the assumptions are the same. Using specific weights allows for explicitly accounting for uncertainties in models and data, as well as evaluation of *a posteriori* errors. However, because of significant algorithmic and computational requirements in implementing these optimal methods, many studies have explored alternate simpler approaches.

In practice, most inversions, and hence assimilation, amount to solving a least-squares problem. A solution is sought that minimizes a sum of differences between the knowns and unknowns, which, in state estimation, is typically written as a sum over time (t) of model–data misfits and deviations from model equations:

$$\mathfrak{J} = \sum_t (\mathbf{y}(t) - \mathbf{H}\mathbf{x}(t))^{\mathrm{T}} \mathbf{R}(t)^{-1} (\mathbf{y}(t) - \mathbf{H}\mathbf{x}(t)) + \sum_t (\mathbf{x}(t+1) - \mathfrak{F}[\mathbf{x}(t)])^{\mathrm{T}} \mathbf{Q}(t)^{-1} (\mathbf{x}(t+1) - \mathfrak{F}[\mathbf{x}(t)]) \quad (7.1.1)$$

Matrices $\mathbf{R}$ and $\mathbf{Q}$ are weights that, statistically, correspond to error covariances of data ($\mathbf{y}(t) \simeq \mathbf{H}\mathbf{x}(t)$) and model evolution ($\mathbf{x}(t+1) \simeq \mathfrak{F}[\mathbf{x}(t)]$) constraints, respectively. Such an objective (cost) function mathematically defines the state estimation problem.

Recent innovations in estimation theory, combined with improvements in computational capabilities, have enabled applications of optimal estimation methods feasible for many data-assimilation problems. Various approximations have been put forth to reduce the computational requirements of statistically optimal methods such as the adjoint method (e.g. Courtier *et al*., 1994) and Kalman filtering and smoothing (e.g. Fukumori and Malanotte-Rizzoli, 1995). Adjoint model compilers have been advanced (e.g. Giering and Kaminski, 1998) reducing programming efforts in model adjoint code generation. Examples of global data assimilation using general circulation models and employing these advancements can be found, for example, in Stammer *et al*. (1997) and Fukumori *et al*. (1999). Results from an extension of the former study are further described in the following section.

The simpler approaches include optimal interpolation, '3Dvar', 'direct insertion', 'feature models', and 'nudging'. Many of these techniques originated in practical atmospheric weather forecasting, sequentially modifying model fields with observations. The methods are characterized by various *ad hoc* assumptions to effect the simplification, but their relationship to the underlying mathematical problem of equation (7.1.1) is at times obscured. Many of the simpler approaches do not account for model and data uncertainties. Observations that lie formally in the future are generally not used in the estimate.

An important related question is how to use ocean data containing mesoscale eddy signals in the assimilation context. Many existing examples, including the one discussed below, are coarse-resolution approaches and do not permit eddies to develop in the models. In such situations eddies are considered part of data error (more precisely, model representation error) and the respective information is filtered out in the cost function. On the one hand, such removal of information can be achieved through prior data smoothing. On the other hand, prior smoothing might not be practical in general circumstances in which measurements are sparse. A proper choice of prior error covariance matrices $\mathbf{R}$ and $\mathbf{Q}$ can account for uncertainties as a function of geographic location and spatial scales, down-weighting signals associated with eddies. In comparison, there is no prior smoothing or down-weighting necessary for assimilating eddy signal containing data into eddy-resolving models so long as the two are consistent with each other. Further discussion of the importance of prior error prescription is given in Section 7.1.4.3.

7.1.4.2 Work in progress

Most of the present focus in ongoing ocean state estimation is on the time-evolving large-scale circulation using primarily altimetric observations (see Fukumori, 2000, for a review). Precise and accurate TOPEX/POSEIDON (T/P) sea surface height observations are now available on a routine basis over most of the globe from September 1992 through the present. At present information from WOCE *in-situ* data, such as XBTs, floats and the WOCE hydrography, are mostly being used as independent information to test the altimetric assimilation.

Here we mention only a few examples that seem relevant for WOCE synthesis and that can illustrate the full potential of ocean state estimation for the synthesis problem. One example is the global ocean state estimation as it results from an ongoing effort at the Massachusetts Institute of Technology, NASA's Jet Propulsion Laboratory and Scripps Institution of Oceanography. The state estimation system is based on a general circulation model and a dual assimilation approach utilizing the model's adjoint and an approximate Kalman filter and smoother. The forward component is the general circulation model described in Marshall *et al.* (1997a,b). The adjoint component is obtained from the forward model by using the Tangent-linear and Adjoint Compiler of Giering and Kaminski, (1998). This compiler has proven to be extremely flexible since it allows one to easily regenerate the adjoint code whenever a change in the forward code is necessary (Marotzke *et al.*, 1999). The approximate filter is an extension of the reduced-state filter described by Fukumori and Malanotte-Rizzoli (1995).

Ongoing computations include a 6-year estimate of the time-evolving ocean circulation (1992 through 1997) with up to 1° spatial resolution, a complete mixed-layer model (Large *et al.*, 1994) and an eddy parameterization (Gent and McWilliams, 1990). Data include the absolute and time-varying T/P data from October 1992 through December 1997, SSH anomalies from the ERS-1 and ERS-2 satellites, monthly mean SST data (Reynolds and Smith, 1994), time-varying NCEP re-analysis fluxes of momentum, heat and fresh water, and NSCAT estimates of wind stress errors. Monthly means of the model state are required to remain within assigned bounds of the monthly mean Levitus *et al.* (1994) climatology. To bring the model into consistency with the observations, the initial potential temperature (θ) and salinity (S) fields are modified, as well as the surface forcing fields. Changes in those fields (often referred to as 'control' terms) are determined as a best-fit in a least-squares sense of the model state to the observations and their uncertainties over the full data period. In the current configuration, there are 10^8–10^9 elements in the control vector.

A schematic of this experiment is shown in Fig. 7.1.7 listing the data used to constrain the model, the control variables (initial conditions and time-varying surface forcing) that bring the model into agreement with the data, and the WOCE data set that at this point is being used for testing the results. Once convergence is obtained, the next step is to include more and more of those data into the optimization procedure, starting with XBT and P-ALACE (Profiling-Autonomous Langrangian Circulation Explorer) profiles of temperature and salinity, and continuing all the way to Lagrangian velocity measurements. Because this optimization is not fully converged yet, we will show here only a few preliminary results from the ongoing work. More details are given in Stammer *et al.* (1997, 2000).

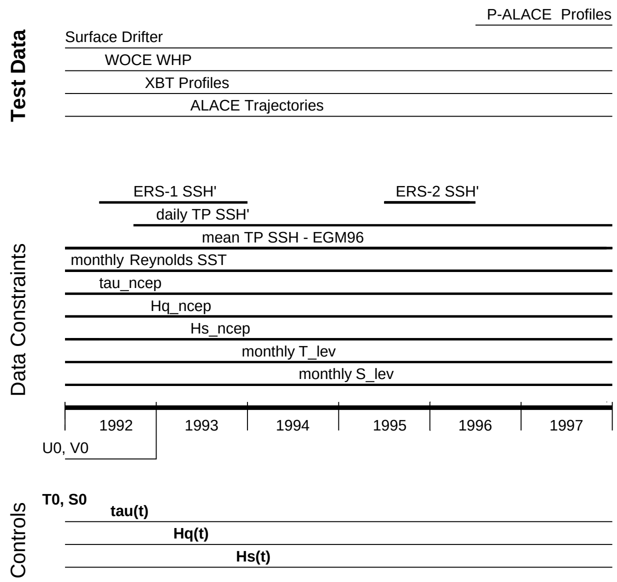

Fig. 7.1.7 Schematic of the ongoing MIT/SIO optimization. The middle part of the figure shows the data constraints and their distribution in time. The lower part shows the 'control' parameters, which are the initial T and S fields, and the time-varying surface forcing (wind stress, heat and freshwater fluxes). Those parameters are adjusted to bring the model into consistency with the data. The top part shows the WOCE data set, which at this point is being used for testing purposes of the results. From Stammer *et al.* (2000).

In Fig. 7.1.8 (see Plate XX) we display the time-mean of the estimated (time-varying) changes of the National Center for Environmental Prediction (NCEP) fields of the zonal windstress τ_x and surface heat flux H_q that are required to minimize model-data differences over the 6-year long assimilation period. The figure is taken from 43 iterations and is still subject to changes, as for all following figures. Changes in mean net heat flux are $\pm 20\,\mathrm{W\,m^{-2}}$ over the bulk of the ocean, but can be as high as $80\,\mathrm{W\,m^{-2}}$ near boundary currents. Overall those changes are fully consistent with current understanding of uncertainties in the meteorological fields and agree especially in the tropical East Pacific with known deficiencies of stratus cover in the atmospheric model. Changes in wind stress show a strengthening of the easterlies over the tropical Pacific where the NCEP re-analysis is known to be too weak. It should be noted that the small-scale features in the wind stress changes arise to some extent from the lack of a full wind stress error covariance function in the current optimization. The displayed changes and similar changes in meridional wind stress and fresh water flux will be used to understand and improve uncertainties in meteorological forcing fields after a fully optimized solution has been obtained.

Estimated mean fields of sea surface height and the near-surface flow field and the flow and temperature fields at 610 and 2450 m depth are shown in Fig. 7.1.9 (see Plate XX). All major ocean current structures are well represented and the largest deficiency – the lack of small-scale structures – can be primarily attributed to the still coarse resolution of the optimization. (The figure actually stems from iterations being performed on a 2° spatial resolution.)

The model–data misfit of absolute sea surface height is displayed in Fig. 7.1.10 (see Plate XX). Differences are as large as 40 cm, especially in the Southern Ocean. There is a close correlation of the residuals with independent information about the error of the EGM96 geoid model over many parts of the global ocean. The associated difference between the first-guess mean flow field at 575 m depth and that obtained from the constrained model is shown in the lower panel of the figure. In many events, changes in the large-scale current structure are apparent with maximum amplitudes about 1.2 cm s^{-1}, or a full 30% of the mean.

An important climate-related issue is the question of the amount of heat the ocean carries and how variable it is in time. To illustrate what can be expected from the estimated fields for a WOCE synthesis, Fig. 7.1.11 shows the simulated convergence of zonally integrated heat transports as they result in the North Atlantic between 36 and 30°N (a) and between 25 and 10°N (b). There is a significant amount of high-frequency variation present in the regional heat flux convergence superimposed on seasonal and interannual variability. This needs to be analysed in terms of its relation to (local and remote) air–sea interactions, heat storage and, ultimately, its climate implications.

7.1.4.3 *Prescribing* a priori *errors*

The weights **R** and **Q** in equation (7.1.1) define the mathematical problem of ocean state estimation. Suitable specification of these weights is essential to obtaining sensible solutions, and is a fundamental issue in ocean state estimation. In fact, while further advancements in computational capabilities will directly solve many of the technical issues of estimation, the improvements will not resolve the weight identification. Using different weights amounts to solving different problems, which thereby leads to different solutions. Poor specification of the weights will result in inaccuracies or even outright 'failures' of the estimation.

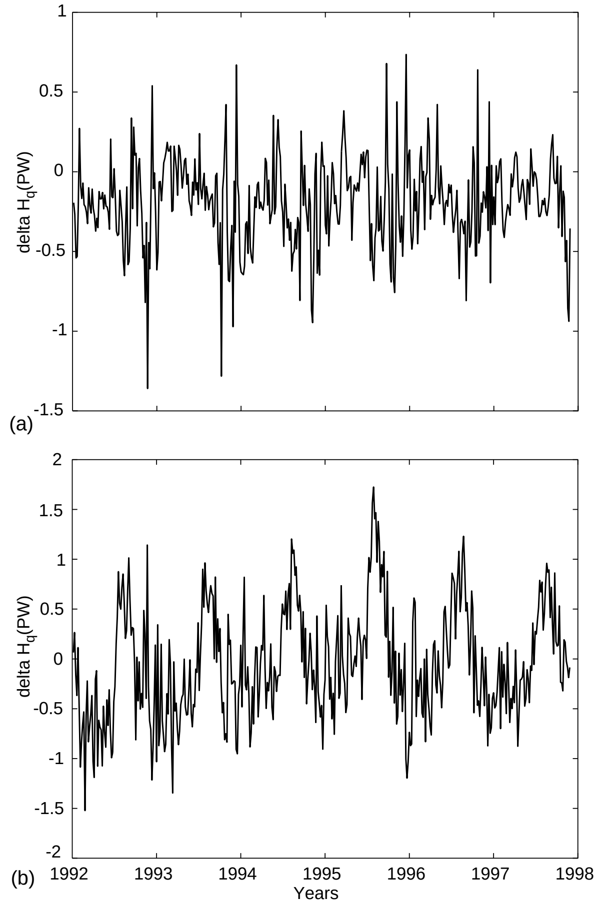

Fig. 7.1.11 The convergence of zonally integrated heat transports in the North Atlantic in the constrained MIT model. Panel (a) shows the convergence between 36° and 30°N; panel (b) shows the same field, but between 25° and 10°N. There is a significant amount of high-frequency variation superimposed on seasonal and interannual variability. This needs to be analysed in terms of its relation to (local and remote) air–sea interactions, heat storage and, ultimately, its climate implications.

Statistical considerations suggest a suitable choice of the weights being *a priori* error covariances of the respective terms in equation (7.1.1). Specifically, weights **R** and **Q** should be regarded as errors in data and model constraints rather than merely data and model errors. In particular, so-called representation error (e.g. Lorenc, 1986), which corresponds to real processes that are measured but are not represented or resolvable by the models, should be considered part of the uncertainties of the data constraint (**R**) instead of those of the model constraint (**Q**). Consequently, data constraint error **R**, like model constraint error **Q**, is model dependent.

Representation errors are processes beyond the model phase space. Such errors are not resolvable by models and thus inconsistent with them. Such errors include, for instance, eddies for non-eddy-resolving models, and tides and diabatic processes for quasi-geostrophic models. Forcing models to fit representation errors necessarily increases the model constraint error, **Q**. Consequently, such corrections can result in so-called 'overfitting' of models to data that lead to model degradation rather than improvement by the assimilation, in particular, for aspects of the model that are not directly measured. Therefore, representation errors should generally be considered part of data constraint error, i.e. part of the model error that cannot be corrected by data assimilation. It should be noted, however, that possible effects of representation error on processes within the modelled phase space can be corrected, because such effects, as opposed to their sources, are within the model space. For example, eddy effects on large-scale circulation as simulated by a non-eddy-permitting model can be corrected, even when the eddies themselves cannot be resolved by assimilation.

A better understanding of what models and observing systems do and do not represent is arguably one of the most urgent and important issues in estimation. While some components of representation errors are distinguishable (e.g. eddies for non-eddy-resolving models), many are not as easily identifiable because it is not necessarily clear what the model phase space really is. Even when distinguishable, a quantitative description of representation errors is often lacking. Data representation is also an issue. Many *in-situ* measurements are sparse and information on what the observations accurately resolve is often unavailable. For instance, some island tide gauge stations may not reflect conditions of the open ocean nearby. Variabilities associated with geographically or temporally offset hydrographic stations or mooring measurements are not accurately known but can provide information of the measurement's representativeness.

Advancing model and data representations can only be achieved by improving model physics, forcing, boundary conditions, etc., or making more comprehensive observations that contribute more information. Therein lie the relative merits of modelling, observations and assimilation. Assimilation is not a panacea that can correct all model errors and compensate all deficiencies of observations. Data assimilation does not add information independent from observations or model physics, but merely allows compensation of one by the other within their common phase space essentially through averaging the two. Although assimilation provides a new dimension to ocean state estimation, results are ultimately limited to what models and observations resolve and especially our understanding of their nature.

A major difficulty in assessing model and data uncertainties has been a lack of global oceanic data sets of sufficient quality and duration. The advent of TOPEX/POSEIDON altimetry and the other systematic large-scale ocean observing systems, including those of WOCE, have made significant contributions to this end. Fu *et al.* (1993) devised a simple means of evaluating model and data errors by comparing observations and model simulations. Menemenlis and Chechelnitsky (2000) and Tokmakian and Challenor (2000) extended the approach by utilizing time-lagged model–data differences and multiple data sets, respectively. Adaptive Kalman filter methods have also been explored to estimate simultaneously from observations prior errors in addition to the oceanic state (Blanchet *et al*., 1997).

However, error estimates resulting from these empirical methods are limited due to duration and extent of the data sets, and the methods only provide partial guidance for error specification. The validity of these and other error specifications must be carefully assessed in terms of their consistencies (e.g. Fukumori *et al*., 1999). In the absence of a wide user community, such as exists for meteorological analyses, accurate ocean state estimation will require a concerted effort within the general oceanographic community, i.e. modellers, observationalists and assimilators alike, to quantify and assess the various assumptions that form the basis of data assimilation. Namely, what do available observations and numerical models represent, how accurately do they do so, and how do we improve upon them?

7.1.4.4 Observing system design and sensitivity studies

In order to achieve maximum return from long-term climate observations, a well-designed observing system is required. Ocean models and data assimilation systems are beginning now to support the design of such an observing system and to test its usefulness to measure climate-relevant quantities in a quantitative and yet cost-efficient way.

A posteriori errors with and without particular data sets provide one measure of the impact of certain observing systems. For instance, by comparing separate ocean estimations using different data sets, Carton *et al*. (1996) found TOPEX/POSEIDON altimeter data to have a larger impact in resolving intraseasonal variability of the tropical Pacific Ocean than data from a mooring array or a network of XBTs.

Forward simulations also allow the testing of sensitivity of model solutions to varying forcings, initial conditions, and/or model parameters, by deducing effects of particular aspects of the circulation on other elements of the model. In comparison, the model's adjoint can be utilized to examine the reverse sensitivity, namely the dependence (rather than effect) of properties on other elements of the model. An example of such use of the adjoint is shown in Fig. 7.1.12 (see Plate XX; from Marotzke *et al*., 1999). The figure shows the sensitivity of the MIT model's annual mean heat transport across 29°N in the North Atlantic with respect to changes in initial temperature and salinity fields near the surface and at depth, respectively. The figure clearly identifies the influence of hydrography along the boundaries on the estimated heat flux and identifies measurements along the western and eastern boundaries as the most important measurement for improving heat transport estimates at this latitude. Likewise one can study the impact of local and remote forcing fields on heat transports, or, for example, on the strength of the meridional overturning circulation or any other scientifically important quantity. Such information will lead to a better understanding of physical adjustment processes in the ocean and allow us to define key measurements required to understand the response of the ocean to atmospheric changes and to accurately measure climate fluctuations in the atmosphere–ocean system.

7.1.5 Summary and outlook

WOCE is contributing enormously to our ability to describe and understand the ocean, and to creating the technical capability needed through development and large-scale use of new observing methods, ocean models and state estimation. These tools and understanding are being carried forward as a basis for study of climate variations. Our treatment of the WOCE synthesis above was separated into a review of the more purely data-based analyses and a progress report on ocean state estimation, combining data with models and including new results. Our conclusions are also separated into these two areas.

7.1.5.1 Observation-based synthesis

WOCE was originally conceived of and designed to produce a description of the large-scale ocean circulation and its transports of heat, mass and fresh water, to understand the climatic state of the ocean and to understand the relevant processes that set that state. The practical design was for global *in-situ* coverage at the space and time scales relevant for climate, with temporal as well as complete geographic coverage, and with complementary global satellite observations. It was recognized at the outset that description of a 'mean' state is a moving target in the sense that any specific period of observations will produce a somewhat different answer, with the differences being complex and regional, depending on the state of the many different, largely uncoupled, climate modes. Several major categories of WOCE observations help to define the climate context, particularly the upper ocean thermal and satellite observations. However it is also clear, and remains clear after WOCE, that a 'mean' state is a relevant concept. The WOCE data sets and modelling are providing a major step in learning about these largest-scale processes, which one can think of as also including the statistics of the shorter time-scale variability that depends on the existence of this mean state. Interannual to decadal climate variations provide in most cases only about a 10% change to the mean state, with the exception of the strong ENSO conditions that affected the upper ocean in the tropical Pacific during WOCE.

The comprehensiveness of the WOCE data set both in types of measurements and geographical coverage are truly exceptional. Never before has a global view been possible, from ongoing satellite coverage of the surface to ongoing large-scale coverage of the upper ocean, to the global description of properties, tracers and circulation at depth. Because of the massive scale of WOCE, it has generally been much more difficult to isolate specific urgent questions, leading to concerns about its being an unfocused experiment, unlike for instance the tropical Pacific experiments with a unified goal – to describe and predict El Niño. The WOCE data sets are being used to describe a plethora of questions of the same scope as El Niño; some of these problems that are first being addressed with WOCE are blossoming into larger ongoing projects. We have made an attempt in this chapter, and the attempt is made in more depth throughout this book, to cover the many different thrusts taken with WOCE data. The results are clearly accretionary. Taken as whole, even at this time, and clearly further along in the analysis, WOCE will provide many significant advances for large-scale oceanography.

A major strength of WOCE has been the collection of many different types of complementary data and development of models and investigator expertise for every basin. WOCE 'synthesis' means approaching individual ocean descriptions or processes using as many of the data types and modelling tools as are available and useful to the specific problem. The existence of concurrent upper ocean and satellite time series along with all of the background information at the largest scale for the WOCE hydrographic programme, as well as basin-scale and global models, has allowed a description of the time-varying response of the ocean that was not possible before. The coexistence of a complete mid-depth mean circulation from floats, along with hydrographic and process-oriented measurements and model analysis of the same regions, has meant that new and potentially important circulation features have been exposed. The coexistence of full tracer data sets, temperature/salinity and direct velocity measurements at depth is allowing significant advances in describing the mean circulation and ventilation time scales. Well-designed local current arrays with hydrographic and basin-scale float measurements have enabled major circulation elements to be quantified. One of the major advances of WOCE, quantification of the diapycnal diffusion rate (Ledwell *et al.*, 1993) and its geographic variations (e.g. Ledwell *et al.*, 2000), has been considerably strengthened by the existence of concurrent larger-scale data sets in addition to those observations specifically designed for measuring diffusivity.

Was it useful to have undertaken the whole exercise of observing the global ocean? Models are becoming global, climate-relevant processes are global, and the existence of global data sets to test hypotheses and provide context for regional process descriptions is critical. The latter has not perhaps received proper credit. Every analysis of every regional process (using WOCE data) requires a larger-scale context (provided also by the larger WOCE data sets). Many basin-scale analyses are underway, all drawing to some extent on the context/boundary conditions provided by the larger data sets. The importance, persistence, and basic physics of various ocean processes are much more clearly revealed if they can be found and studied in more than one setting, meaning more than one basin. The particular scientific questions that will be answered by the global data sets *per se* are of the same ilk as those for each region – describing the global thermohaline circulation, global heat and property transports, global variability. All require global coverage. But how important was synopticity, even in the loose WOCE sense of coverage over about 10 years? It was important for two reasons – the most efficient way to obtain enough information to make the global estimates and the regional descriptions with global context was to do it as quickly as possible, using the same technology throughout with the same sampling philosophy. Some data sets, such as transient tracers used for ventilation studies, require a reasonable amount of synopticity.

What would have been better done differently, if anything? Better coordination could have reduced the amount of time required to cover regions. Satellite programmes that were an integral part of the description did not begin until after WOCE had started, resulting in unfortunately incomplete coverage overlap. Throughout the design and execution of WOCE there was pressure to have modellers and observers coordinate their efforts more seamlessly. However, at a practical level, this can only be accomplished if investigators have an immediate common goal that requires the integration of different approaches that each provide enough valid information to allow synergy. That this did not happen was perhaps a consequence of the state of ocean modelling, in particular computational capabilities, at the beginning of WOCE. Remarkable advances in regional and global-scale modelling during WOCE are closing the gaps between observers and modellers. Analysis of many WOCE data sets now takes place with reference to modelling, and most importantly, the maturation of models and the collection of global data sets is allowing ocean state estimation to develop rapidly.

Should global observations continue to be taken? The answer is a resounding yes, but clearly in a different way than during WOCE. The technologies that were developed and put into widespread use during WOCE should be included, so that the next set of large-scale questions may be addressed. The community has chosen to continue working with problems that are more clearly the variability of the ocean and atmosphere, given the large data set that now exists to describe the 'mean' state. The appropriate technologies have been selected, mostly involving autonomous measurements from satellites, floats, drifters and large moored arrays with telemetry. Processes important to climate that were not explored fully in WOCE, particularly in more remote regions such as the southern hemisphere, should be pursued. Should there be a future WOCE-type hydrography survey? A complete repetition of WOCE hydrography in the near future is unlikely and unnecessary given the long time scales that these measurements address. Monitoring of global carbon inventories and studying the physical and biological processes that affect them through use of transient and other tracers are best accomplished from large-scale research ship coverage at present and likely will remain so for at least a decade, thus requiring at least partial coverage through the intervening years until a new global survey is undertaken.

7.1.5.2 State estimation

Ocean data assimilation has long been a focus of theoreticians and modellers. The tremendous efforts of the community have brought us well along the way to success. There are now several ongoing WOCE-related assimilation efforts in the US, Europe and Japan. Maturation of this ongoing international work will result in fully successful WOCE syntheses. Ongoing efforts in section and box inversion and in ocean state estimation are still largely preliminary. Model resolution, spatial extent, and employed physics must further be expanded and the entire WOCE data set must eventually be incorporated. The results so far illustrate the potential for model–data combination. Among many possible applications, one expects the production of much better estimates of the magnitudes and mechanisms of property flux divergences, and the study of oceanic biogeochemical cycles that are dependent upon the circulation.

What progress is there in ocean state estimation? At the beginning of WOCE there was little confidence in the prospect of bringing models and data together in a full ocean state estimation. It has now been shown that such a synthesis is technically feasible. Moreover, the field is rapidly migrating towards operational applications of ocean state estimation, allowing us to estimate changes in the ocean circulation on seasonal to longer climate-relevant time scales, similar to re-analysis projects in the atmospheric community.

What are the gaps in knowledge needed for state estimation? The attempt to accurately prescribe *a priori* model and data errors is as much a state estimation problem as it is a fundamental issue in data analysis and numerical modelling. Such complete statistics are likely not possible for every data type, and estimating them requires a close collaboration between observers and modellers. For climate applications, our knowledge about details of the connection of the changing atmospheric forcing with changes happening in the deep ocean over hundreds of years is only rudimentary. However, if we are to estimate the deep ocean state in the presence of large climate fluctuations, such as ENSO events, model drift has to be separated from actual changes in the deep ocean due to climate changes over, say, the last 500 years. There are also many technical issues to be addressed, ranging from questions about efficiencies of methodologies, to the representation of the ocean's topography and its impact on connection between basins, e.g. in the eastern North Atlantic where the deep water is supplied through a narrow fracture zone. The representation of mixing in general is a complex question and there is hope that ocean state estimation will be able to contribute substantially to estimates of the three-dimensional and possibly time-varying mixing tensors.

What are the strategies for the future for state estimation? A number of groups around the world are developing the capability to assimilate oceanographic data. These efforts include those aimed at estimating the state of the general circulation as part of WOCE as well as those focusing on operational applications such as mesoscale forecasting and predictions of seasonal-to-interannual climate variability. Several of the major efforts focus on operational assimilation and/or mesoscale forecasting. Various remote sensing and *in-situ* upper ocean data sets are used, such as satellite altimetry, sea surface temperature, upper ocean temperature profiles from XBTs and profiling floats, and some velocity information. Many of these data sets were collected as part of WOCE and are continuing as part of a global ocean observation system and/or CLIVAR. Recent development and expansion in assimilation activities along these lines is remarkable, and is continuing with substantial individual and agency investments of time and funding. For instance, the Global Ocean Data Assimilation Experiment (GODAE) is an international initiative centred on data assimilation and aims to conduct regular, near-real-time global data assimilation to demonstrate the practical utility of ocean observations.

Only a few ocean state estimation efforts include one of the basic WOCE data sets, hydrography, and are focused on estimating a general circulation rather than on prediction. An example of one such effort in the US, in which the coauthors (DS and IF) are engaged, is a consortium funded by NSF, NASA and ONR – 'Estimation of the Circulation and Climate of the Ocean' (ECCO), which is building on the already existing MIT/SIO ocean state estimation efforts and parallel ones at JPL. The central technical goal of ECCO is a complete global-scale ocean state estimation over the 15-year period 1985–2000 at the highest possible resolution (given available computer resources) along with a complete error description. But to develop the technical understanding required to incorporate all available data types, develop associated error covariances, and to continue to improve models in support of the global estimate, ECCO is beginning with coarser-resolution global and basin-wide models and developing near-real-time capabilities in support of the climate community needs.

What is required for the future for successful state estimation? The importance of accurately prescribing *a priori* model and data errors has been discussed above in detail. Other issues include suitable access to data and computational resources. Ready availability of observations is an obvious requirement for the estimation process. Appropriate data reduction and quality control must be applied to the observations. In light of the extent of the WOCE observational data set, such processing, including data assemblage amenable for model integration, is a non-trivial task.

A major limitation for ongoing state estimation efforts in some countries is the available computational resources. In spite of advances made in practical estimation, data assimilation is computationally intensive. Even moderate-sized general circulation models strain extant supercomputers when used in rigorous state estimation. The paradigm of shared resources at supercomputing facilities has so far failed to meet the requirements of ocean state estimation in the US. An alternate means of providing the necessary computational resources is urgently needed.

Acknowledgements

Grant support from the National Science Foundation OCE-9712209 (LT), OCE-9730071 (DS) and from NASA NAG5-7857 (DS) is acknowledged as well as support from the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (IF).