Real-time Coastal Observing Systems for Marine Ecosystem Dynamics and Harmful Algal Blooms: Theory, Instrumentation and Modelling

Edited by Marcel Babin, Collin S. Roesler and John J. Cullen

20

Use of real-time observations in an operational ocean data assimilation system: the Mediterranean case

N. Pinardi, C. Fratianni and M. Adani

20.1 Introduction

Real-time observations are essential for operational forecasting that in turn can be used to predict changes of the state of the ocean and its associated biochemical fields. In addition, real-time observations are useful to detect changes in the past with the shortest delay, to standardize practices in data collection and to exchange data between remote regions of the ocean and seas. The drawback is that real-time observations could be less accurate than their delayed mode counterparts due to the time constraints for data dissemination. *In situ* real-time data are usually decimated to be transmitted in real time (loss of accuracy and resolution), whereas satellite data are corrected with approximate algorithms and less ancillary data. Delayed mode quality control analysis increases the value of the observational data set, flagging outliers and producing climatological estimates of the state of the system. Thus real-time data, together with a modelling system and the climatological estimates, can give also the appropriate information for scientific studies and applications.

The principles of operational science started to develop in the 1940s and 1950s, based on the combined use of real-time data and modelling systems that can extend the information from observations in space and time. Operational science is based on a sound knowledge of the dynamics and processes for the space/timescales of interest and operational meteorology and oceanography have started to implement these principles to weather and ocean forecasting.

In the past 20 years, operational meteorology has become a reality with a network of *in situ* and satellite observations that has made the weather forecast capable of extending the theoretical limit of predictability of the atmosphere (only one-two days theoretically, now forecasts are useful for more than five days on average). Today meteorological observations are mainly used in their assimilated form even if observations are still collected for specific process-oriented studies. Recently the meteorological re-analysis projects (Gibson et al., 1997; Kalnay et al., 1996) have released a wealth of data to be understood and analysed. These data sets are coherent and approximately continuous (daily), filling the observational gaps in space and time with a dynamical interpolation scheme. The model and the real-time observations are fused in one best estimate of the state of the system by data-assimilation techniques that have been developed to a great degree of sophistication in recent years (Lorenc, 2002). The re-analysis data are now forming the basic reference data set to understand climate variability in the atmosphere and upper oceans.

Dynamical interpolation/extrapolation of observational data for operational forecasting in the ocean began to be investigated at the beginning of the 1980s and the first successful forecasts were carried out in the open ocean (Robinson and Leslie, 1985). These exercises required real-time data that were initially collected with rapid ship surveys realizing adaptive sampling schemes and collecting a combination of traditional recoverable and expendable instruments (CTD, XBTs). At the same time but in a totally independent way, shelf scale and coastal real-time data from moored and drifting sensors such as meteorological buoys and sea-level stations started to be used for shelf scale storm surge operational forecasting (Prandle, 2002). Operational oceanography is now building on this experience and considers real-time measurements from opportunity platforms and satellites in a manner very similar to operational meteorology.

This chapter aims to show the use of real-time observations in a state-of-the-art ocean-predicting system realized in the Mediterranean. We discuss the pre-processing schemes required to properly assimilate the observations into an operational nowcasting/forecasting system, elucidate the role and impact of different observations in the assimilation system and show the use of real-time data to evaluate the quality of the modelling system.

We start with the description of the Mediterranean Forecasting System (MFS) real-time observing system and pre-processing quality control in Section 20.2, we then describe the modelling and assimilation system in relation to the impact of different real-time observations in Section 20.3. In Section 20.4 we evaluate the consistency, quality and accuracy of the forecasting system using model-data intercomparison and Section 20.5 offers conclusions.

20.2 Real-time observing system for operational forecasting in the Mediterranean

Based on the earlier open ocean forecasting experience, a Mediterranean Forecasting System (MFS) began to be implemented in the Mediterranean basin, the topography of which is reproduced in Figure 20.1. The average depth of the basin is 1,500 m and in several regions we reach over 3,000 m. Thus the initial observing system was chosen to be consistent with the main components of the Global Ocean Data Assimilation Experiment (GODAE) observing system[96] (Smith and Lefevre, 1997).

The major elements are: (a) satellite remote sensing for sea surface height (SSH) and sea surface temperature (SST); (b) voluntary observing ship (VOS) for temperature profiles (Rossby et al., 1995); (c) moored buoy systems such as the TAO array in the Pacific (McPhaden et al., 1995) but modified for the Mediterranean needs; (d) subsurface drifting and profiling floats such as ARGO. In the Mediterranean all these elements have been considered and implemented during the Pilot Project phase of MFS (Pinardi et al., 2003), with the exception of the ARGO floats that have been implemented starting from October 2004, in the second phase of the MFS programme (http://www.bo.ingv.it/mfstep).

The design of the real-time observing system is based on knowledge of the large-scale structure of the circulation. Figure 20.2 shows the simulated mean sea level from 1993 to 1999 with the operational Ocean General Circulation Model (OGCM) of MFS. The OGCM used is described in Pinardi et al. (2003), to which the interested

[96] http://www.bom.gov.au/bmrc/ocean/GODAE/

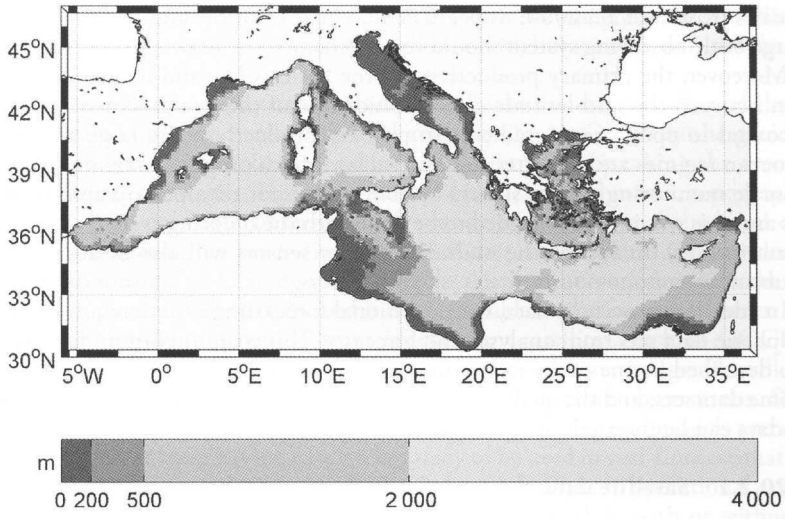

Figure 20.1
Mediterranean bottom topography (1/60) in metres.

reader should refer for details. The basin scale circulation is characterized by large-scale meridional SSH gradients that the observing system should sample. In addition, it is composed of large-scale sub-basin scale gyres (cyclonic in the North and anticyclonic in the south) that have intensified currents and open ocean jets at their borders. Most of these sub-basin scale gyres have large amplitude variations at seasonal and interannual

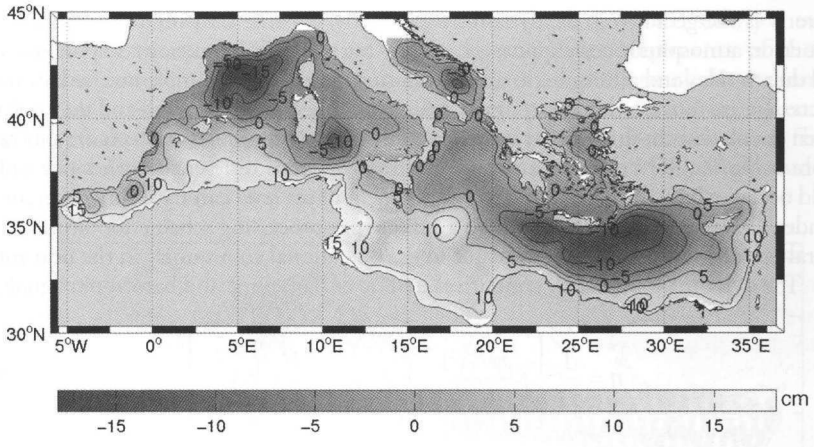

Figure 20.2
1993–99 average sea surface height (SSH) simulated by the MFS operational OGCM. The negative values correspond to SSH depressions due to waters heavier than their surroundings; the positive values correspond to warmer/fresher and lighter waters that 'expand the water column'.

timescales (Korres et al., 2000; Molcard et al., 2002; Demirov and Pinardi, 2002) and the large-scale observing system should resolve them.

Moreover, the primary productivity of the basin is low and inserts the Mediterranean between the mid-latitude oligotrophic areas of the world oceans. This means that coastal to open ocean gradients in primary productivity are large and that the open ocean regimes are characterized by a subsurface chlorophyll maximum. Thus the large-scale monitoring system should be able to monitor the open ocean biochemical fluxes and this was initially decided to be done with the mooring system and the satellite data (colour). In future, the multidisciplinary sensors will also be added to VOS and subsurface autonomous vehicles, such as gliders.

In addition to oceanic data, the operational forecasting system requires real-time atmospheric data sets from analyses and forecasts. This is an important data set and it is also described in one of the following sections. The information content of all these real-time data sets, and the quality control procedures and preprocessing needed before these data can be inserted into the assimilation system, are described below.

20.2.1 Satellite data

Satellite SSH and SST data compose nowadays the basis for the real-time monitoring of the open ocean and coastal areas. Even if the SSH data are not accurate enough near the lateral boundaries of the basin (data normally stop about 20–30 km from the coast), this data set is essential to initialize the internal flow field of local models and give the correct open boundary conditions wherever necessary. SST instead is at high resolution (1 km) and also high frequency (twice a day at least) and thus it is an important component of the coastal and open ocean observing system.

Following Le Traon (2002) the altimetric signal can be decomposed into four parts:

$$\mathrm{SSH} = N + \eta + \Sigma + \varphi, \tag{20.1}$$

where N is the geoid, η is the dynamic topography, Σ is the measurement errors (due to orbit error, atmospheric corrections, etc.) and φ are the high-frequency components of sea level due to tides and atmospheric surface pressure (sometimes a simple inverse barometer) effects. We are interested in the η signal, which is connected to the wind and thermohaline driven circulation. In this chapter we will assume that high-frequency effects are subtracted before using the SSH data for assimilation. This choice is different in other parts of the world ocean, where the high-frequency component of sea level can have very different amplitude. This makes it difficult to have a unique pre-processing scheme for altimetry, but several options should be made available to the operational community in the near future.

The dynamic topography contains the steric or baroclinic and barotropic signals, i.e.

$$\eta = \frac{1}{H}\left\{ f\frac{\psi}{g} + \frac{1}{\rho_0}\int_{-H}^{0} \rho\, z\,\mathrm{d}z + \frac{H}{\rho_0}\int_{-H}^{0} \rho\,\mathrm{d}z \right\} \tag{20.2}$$

where geostrophic balance has been assumed and the symbols are explained in the footnote[97] (Pinardi et al., 1995). The first term on the right of (20.2) is the barotropic

[97] Symbols: ρ is the water density and ρ_0 its constant value, H is the bottom depth supposed to be constant, f is the Coriolis parameter, g is the gravity acceleration, ψ is the barotropic streamfunction in Sverdrup ($10^6 \frac{m^3}{s}$) defined as $\bar{\vec{u}} = \frac{1}{H}\hat{k} \times \nabla\psi$ and $\bar{\vec{u}} = \frac{1}{H}\int_{-H}^{0} \vec{u}\,\mathrm{d}z$ and $\hat{k}$ is the vertical unit vector.

component while the last term is the bottom pressure. The integral in the middle term is a 'potential energy' term, as defined by Mellor (1996), but we refer to it simply as the baroclinic term.

The dynamic topography can be decomposed into a mean part $\overline{\eta}$, and the variability, indicated by η and called sea level anomaly (SLA). Generally the geoid N is not known with sufficient accuracy and it is then subtracted by taking the average of the sea level along the tracks. This eliminates N and the mean dynamic topography $\overline{\eta}$ from (20.1). Then only SLA observations are made available in real time and operationally. This is a major pre-processing of the SLA data, that requires the knowledge of the long-term mean of SSH along satellite tracks. Such mean contains both the geoid and the mean dynamic topography and it is not possible to distinguish between them. Only the addition of information from independent satellites, such as GOCE and GRACE (Le Traon, 2002) that measure independently the marine geoid, will allow the estimation of the mean dynamic topography from satellite observations. Several attempts have been made in the past, using mainly large-scale geoid models and *in situ* data but they are far from having enough accuracy to be used in real-time estimation of along track SSH for assimilation into dynamical models. Thus in operational systems, the real-time satellite altimetry data are given in terms of SLA, with or without the high-frequency component, φ, subtracted.

Two pre-operational satellites were working for the past 10 years, Topex/Poseidon and ERS-2, and they covered the Mediterranean quite extensively (Figure 20.3). For the Mediterranean, the geoid and mean dynamic topography was calculated from the average of along track SSH for the period 1993–99. SLA is then released weekly with an estimate of the orbit error since the value of Σ depends above all on the precision of the satellite orbit computation and this requires environmental ancillary data that are not available in real time (Le Traon and Ogor, 1998). It has been shown that the accuracy of real-time SLA and delayed mode data is now comparable (Buongiorno et al., 2003).

The SLA signal described by (20.2) contains the large-scale, slowly moving components of the sea level, also called the geostrophic components of the sea level. These components are mainly due to the seasonal thermohaline changes, in turn due to air/sea buoyancy fluxes and their penetration in the water column. The baroclinic signals due to the air/sea physics (also momentum fluxes due to wind stress) are strongly modulated by the mesoscale eddy field and the

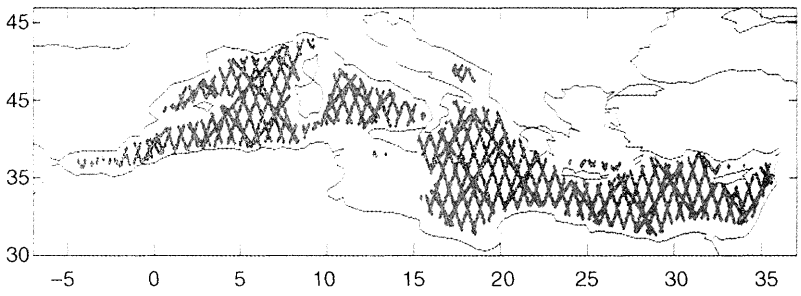

Figure 20.3
Topex/Poseidon and ERS-2 superimposed tracks with respectively 10- and 35-day repeating cycles, 2002.

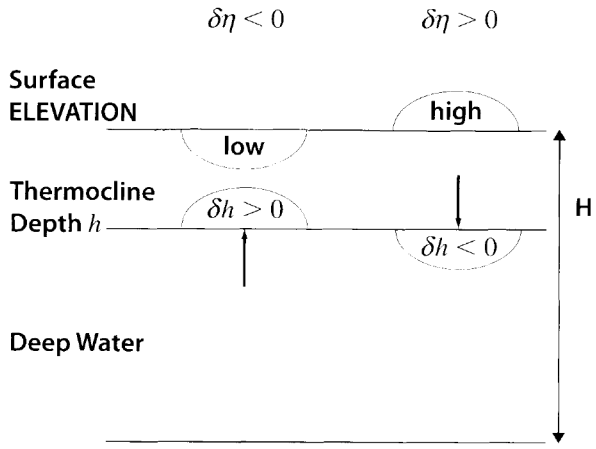

Figure 20.4
Schematic of the vertical displacement that correlates the surface elevation with the isopycnal displacements (δh).
Source: Redrawn from Haines (2002).

sub-basin scale gyres that compose the Mediterranean Sea circulation structures (Pinardi and Masetti, 2000). The barotropic signals are mainly wind driven and they couple with the baroclinic effect at the eddy field space scales. Assimilating SLA in a general circulation model, as described below, means that the model should be corrected for this slowly moving component of the sea level. The typical geostrophic relationship at these time and space scales is that SLA is high where subsurface temperature is also high, and vice versa. This means that the slope of isotherms and SLA have an opposite sign. Haines (2002) showed this concept with the diagram reproduced in Figure 20.4 where the surface elevation is correlated with isopycnal depths (δh).

Mellor and Ezer (1991) demonstrated for the first time that SLA described by (20.2) is strongly correlated to the thermocline depth in the Atlantic subtropical gyre and Masina et al. (2001) showed that this is true also for the tropical regions of the ocean. This means that this information should be used for insertion of the data into an assimilation system. As seen from (20.2) all components of SLA are integral quantities and assimilation should correlate the SLA signal with the model state variables, in particular temperature and salinity profiles, either through the observational operator H, described in the next sections, or with statistical correlations contained in the background error covariance matrix. The SLA for short-term forecasts is then the integrated effect on the geostrophic timescales (approx. two to three days for the Mediterranean) of thermocline displacements due to mesoscale or sub-basin scale gyre variability.

The other real-time satellite data set that is important for assimilation into forecasting models is the sea surface temperature (SST). This is the oldest real-time data set available but pre-processing algorithms for the space radiometer signal are under continuous development. The algorithms for SST retrieval from radiances use information from *in situ* temperatures to calibrate the parameters of the algorithm itself.

The most used sensor is AVHRR (Advanced Very High Resolution Radiometer), flown by NOAA satellites, and the best algorithm both for the Mediterranean and the world ocean is now the Pathfinder algorithm.[98] The algorithms can be applied rapidly to night time and daytime images but the latter are more of a problem as sun glitter effects and high humidity in the air layer adjacent to the sea surface, especially in the Mediterranean, can affect the SST signal received by the radiometer. A new project, the Global High Resolution Sea Surface Temperature Pilot Project, is being launched where the quality of AVHRR real-time data will be assessed for different users in the context of GODAE.[99]

SST contains information about two major processes occurring in the ocean: the first is the warming/cooling processes due to air/sea interaction physics and the second is the mesoscale/gyre structures that produce local changes in temperature due to geostrophic isotherm displacement. One example of the latter process is the detachment of cold and warm core rings across the Gulf Stream front that produces SST anomalies due to the dynamical instabilities of the oceanic jet that transfer water of different SST properties across the Gulf Stream front.

With regard to the air/sea interaction processes, another way to say that SST is affected by exchange of heat at the air/sea interface is that SST contains the information about the dynamics of the surface mixed layer. This layer can mask the geostrophic SST contribution, as it happens for cold core rings that frequently, after their birth, have their 'geostrophic SST anomaly' masked by intense air/sea interaction heat exchanges.

The most common way to use real-time SST or assimilate SST in operational models is to use SST to correct for inaccurate air/sea fluxes at the surface boundary of the models. Normally the correction is carried out via a restoring term, also called a nudging term, that adds to the heat flux term, such as

$$Q_{\text{corr}} = Q - \left.\frac{\partial Q}{\partial T}\right|_{T=T^*} (T - T^*) \tag{20.3}$$

where Q is the net heat flux at the air/sea interface, T^* is the observed SST, T is the SST produced by the model when only Q is used. The coefficient $\left.\frac{\partial Q}{\partial T}\right|_{T=T^*}$ in (20.3) is taken to be constant and different values have been chosen from the basin scales to the subregional Seas (Pinardi et al., 2003).

20.2.2 The *in situ* platforms and sensors

One of the main components of the *in situ* large-scale real-time monitoring system for the world ocean is based on the Voluntary Observing Ship (VOS) system that relys on commercial ship lines for the deployment of expendable temperature sensors, such as eXpendable BathyThermographs (XBT). The ship tracks implemented in the first phase of MFS are reproduced in Figure 20.5: data were transmitted in real time from the ships to the collecting centre (Manzella et al., 2003) and then to the modelling centre.

The temperature profiles on all tracks were taken outside the 200 m depth areas and were collected down to 700 m to resolve the subsurface temperature maximum associated with the Levantine Intermediate Waters in the Western Mediterranean. These data are normally decimated since the satellite telecommunication system used (ARGOS) has a low transmission speed. This is done for the world ocean and it was also tried

[98]Information available at: http://podaac.jpl.nasa.gov/sst
[99]http://www.ghrsst-pp.org/

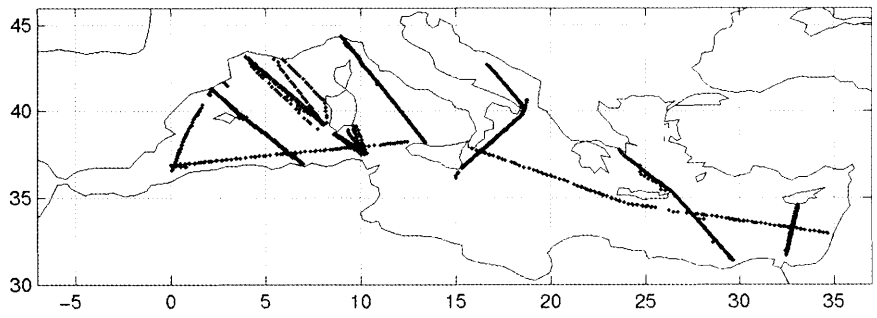

Figure 20.5
VOS-XBT system set up from September 1999 to December 2000 in the Mediterranean.

for the Mediterranean. However, it was found that the standard automatic decimation system for VOS did not cope for drastic changes in the vertical stratification, as they occur in the Mediterranean. The worldwide VOS-XBT programme considers the transmission only of 10–15 points along the temperature profile, computed at the profile's inflection points. When the vertical profiles become quasi-uniform with depth, many XBT samples were badly decimated by the standard on board ship-software. In the Mediterranean, winter mixing is very intense and stratification is rapidly lost for about three months each year. The XBT automatic decimation algorithm is poorly adaptive to these conditions and almost 30% of the data were lost due to the decimation software failures.

Smith et al. (1999) stated: 'all upper ocean thermal data are to be distributed as soon as is practical after measurements (preferably 12 hours). The strong preference is to keep intervention to a minimum, perhaps just automated processes. There should be a well-supported second stream, which allows for improved quality control and scientific evaluation of the data sets.' The MFS VOS system started to send full resolution profiles in real time and realized a new system of real-time data quality control that will store quality checked XBT data in oceanographic archives (Manzella et al., 2003).

The near-real-time quality control procedure contains seven steps that in synthesis are:

- position control;
- elimination of spikes;
- interpolation at 1 m intervals;
- gaussian smoothing;
- general malfunction control;
- comparison with climatology;
- visual check, confirming the validity of the profiles and providing an overall consistency.

The XBT data set is then inserted into the assimilation system. The vertical temperature profile is a basic state variable of the physical system, containing information about several processes and in particular the vertical density distribution. In addition, being the data collected in a section-like track, the temperature field gradients along track give approximately the geostrophic velocity field across the track itself. In the Mediterranean, as well as in other temperate seas, temperature should be adequately combined to salinity to describe the basin water masses and derive the geostrophic velocity field across the VOS ship track. Most of the data assimilation systems, as it will

be seen later, can update both temperature and salinity profiles starting from the single temperature profile, keeping the historical water mass relationship in consideration.

Raicich and Rampazzo (2003) simulated the impact of VOS-XBT data in the reduction of initial errors in the ocean general circulation model. It was found that repeated long tracks, such as the trans-Mediterranean track going from Gibraltar to Haifa, had a very positive impact on the error reduction due to the consistent improvement of the geostrophic velocity field associated with the coherent temperature section given by the track.

Another kind of observing system that nicely complements the VOS is the moored buoy system. Such a system was first developed in real time for the Pacific ocean (McPhaden et al., 1995) and the Bermuda testbed mooring (Dickey et al., 1998) and is now being developed for other areas. The data collected by such a measuring system are multidisciplinary and at high temporal frequency for the physical and biochemical components of the marine ecosystem. These point-like measurements should be mainly used as an independent data set to validate both model and data assimilation components. For the Mediterranean such a system was first developed in the Cretan Sea and it is described in Nittis et al. (2003). Here it is sufficient to say that the system allows the correlation between the physical and biochemical components of the marine ecosystem at high time frequency (normally a few minutes) to be explored for the first time. A network of buoys is being developed under the second phase of MFS.

Finally the ARGO profilers, collecting temperature and salinity profiles from 700 m depth to the surface, have been deployed in the Mediterranean.[100] This completes the 'open ocean' basic monitoring components for the region.

20.2.3 Atmospheric forcing data

Ocean forecasts are driven by atmospheric forecasts. Atmospheric forcing is also used, by means of 'analyses', during intermittent data assimilation steps. The atmospheric analyses are an optimal combination of observations and atmospheric general circulation models outputs, that is they are the best estimation of the past and present state of the atmosphere. They have substituted for many purposes the direct observations for the real-time assessment of the atmospheric state. These analyses are now produced twice a day by the major meteorological offices around the world.[101] Forecasts are also produced at least twice a day from the major meteorological centres and the accuracy of the analysis scheme is constantly increasing with time.

The atmospheric forcing for ocean models is derived from atmospheric surface variables using interactive bulk formulas that relate the model SST with the air temperature, relative humidity, cloud cover and winds at the sea surface (Castellari et al., 1998). The short-term ocean forecast is driven by the atmospheric forecast surface fields (see Figure 20.6). Any error in the input of atmospheric data will affect the quality of the ocean forecast. Some authors think that such error is large compared with other errors, that is initial conditions specification, that it should specifically considered in the assimilation procedure. This error is mainly due to the offline coupling between the atmosphere and the ocean, that is the atmospheric surface variables do not see the ocean forecast SST and currents since they are coupled after the atmospheric forecast has been performed. This is why the correction in (20.3) is usually applied to the computed air/sea fluxes. Only fully or synchronous atmosphere-ocean coupling

[100] http://www.moon-oceanforecasting.eu
[101] http://www.ecmwf.int

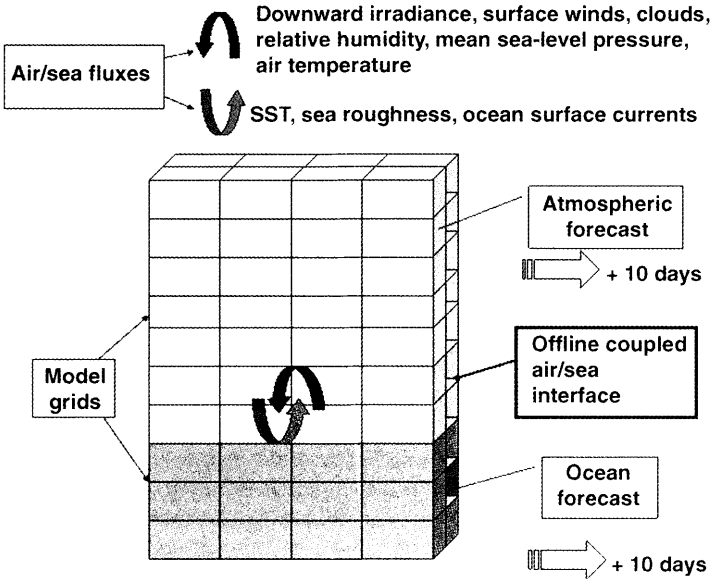

Figure 20.6
Schematic of the offline air/sea coupling for ocean forecasting.

will alleviate this error and seasonal forecasts are carried out nowadays with the fully coupled ocean-atmosphere system without heat flux corrections.[102]

The most inaccurate data set in meteorological models is the accumulated precipitation field that modifies the water flux into the ocean. However, this error drives very long-timescale ocean responses and short-term forecasting should not be greatly affected by such inaccuracy. More important is the wind stress or momentum flux error produced by the coarse atmospheric winds resolution and the bulk formula parameterizations used at the air/sea interface. Such error impacts the tropical Pacific predictability so that, in the past, a special wind stress data set was developed to account for higher space and time frequency of the forcing (FSU tropical Pacific winds, Legler et al., 1988). In the Mediterranean several empirical factors were found that normally increase the wind stress amplitude to reach better agreement between modelled and measured waves (Cavaleri et al., 1992). The increased resolution of numerical weather prediction models and the improvement in data assimilation schemes will make obsolete the use of such empirical factors. Recent developments have shown, however, that the high wave number content of scatterometer winds (Milliff et al., 1999) is far from being reproduced by the present atmospheric forecasting systems and new blending procedures are being developed (Milliff et al., 2001) that produce high wave number content wind fields to realistically force the ocean.

The timely provision of meteorological analyses and forecasts is at the basis of real-time delivery of ocean forecasts. Future developments will involve the coupling of high-resolution non-hydrostatic meteorological models with the ocean counterpart, but this crucial area is still in its infancy.

[102]http://www.ecmwf.int/products/forecasts/seasonal/

20.3 Modelling and assimilation system

20.3.1 Combination of different sources of information: observations and models

Gauss in 1809 wrote: '.. since our measurements and observations are nothing more than approximations to the truth, the same must be true of all the calculations resting upon them, and the highest aim of all computations made concerning concrete phenomena must be to approximate, as nearly as practicable, the truth.' With this statement, he solved the problem of the determination of the planet's motion by fitting via the least-squares method several parameters to the measurements. Astronomical theory gave insight on which parameters to fit and measurements determined their optimal values. Both theory and measurements gave rise to the 'best estimate' of the truth. In modern times, theory is substituted with numerical models based on the general equations of hydrodynamics. Gauss's statement can be taken as the basis of modern oceanic and atmospheric estimation theory. In what follows we review the work of Lorenc (2002) and Daley (1996) about data assimilation.

Data assimilation is defined by Lorenc (2002) as 'the process of finding the model representation which is most consistent with the observations'. This concept goes back to Gauss's statement in the sense that in order to obtain the best estimate of reality, two different approximations of truth are used, one from observations and the other from the dynamical model. Data assimilation in the ocean and atmosphere fuses these two approximations of reality in an optimal estimate of truth with a least-square method approach (errors should be unbiased, random and normally distributed).

In atmospheric and oceanic data assimilation systems, the 'truth' or 'true state' is generally assumed to be a state of the atmosphere or the ocean that has the fast motion filtered out (sound waves and fast barotropic gravity waves are not considered). Furthermore, the atmosphere and the ocean are considered to be close to horizontal non-divergence (geostrophic approximation) and the flow is assumed to be 'smooth', that is sharp changes are not allowed within few model grid points. This means that observations, taken at finite time and space resolution, can give information about the truth and the model needs to have appropriate parameterizations of sub-grid scale phenomena that will not drive the solution too far from the geostrophic balance.

Let us take the 'true' ocean state vector to be

$$X = \begin{bmatrix} T \\ S \\ U \\ V \\ W \\ \tilde{n} \\ \theta \\ \zeta \end{bmatrix} \tag{20.4}$$

where the symbols indicate (from top to bottom): three-dimensional fields of temperature, salinity, zonal, meridional, vertical velocity components, water density and

two-dimensional fields of barotropic streamfunction (defined in Note 97) and free surface elevation or SSH. Some of these state variables are prognostic, that is their time evolution can be described by an equation containing their time rate of change, some are simply diagnostic, such as density, vertical velocity and free surface elevation. Vertical velocity is diagnostic because the hydrostatic approximations has been adopted in the OGCM; SSH is considered to be diagnostic as we regard it as only the slowly moving component of sea level and fast barotropic external gravity waves are filtered out. SSH can be mathematically a prognostic variable but we consider in the data assimilation that the true state of the ocean contains only the diagnostic components of SSH as written in (20.2).

Suppose that we have two approximations of $\mathbf{X}$ which we now consider only to be a scalar value X, corresponding to one of the state variables contained in (20.4) at one single grid point. The first approximation is a numerical model solution, called X^{b}, with an error $E^{\mathrm{b}} = X^{\mathrm{b}} - X$ and the second is an observation, called Y^{o}, with an associated error $E^{\mathrm{o}} = Y^{\mathrm{o}} - X$, occurring at the same location of the model first guess. We assume that the error probability distribution is Gaussian, i.e.

$$p(E) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\left(\frac{E^2}{2\sigma^2}\right)}, \tag{20.5}$$

where $\sigma^2 = \langle E^2 \rangle$ and the brackets indicate the expectation operator (Daley, 1996). The joint probability distribution for the two approximations to truth is:

$$p(E^{\mathrm{b}})p(E^{\mathrm{o}}) = \frac{1}{\sigma_{\mathrm{b}}\sigma_{\mathrm{o}}\sqrt{2\pi}} \mathrm{e}^{-\left(\frac{E^{\mathrm{b}2}}{2\sigma_{\mathrm{b}}^2} + \frac{E^{\mathrm{o}2}}{2\sigma_{\mathrm{o}}^2}\right)} \tag{20.6}$$

If we want to make this probability maximum, we need to impose that the exponent in (20.6) has a minimum value. Calling the exponent I, written as

$$\mathrm{I} = \frac{1}{2}\left[\frac{(X^{\mathrm{b}} - X)^2}{\sigma_{\mathrm{b}}^2} + \frac{(Y^{\mathrm{o}} - X)^2}{\sigma_{\mathrm{o}}^2}\right], \tag{20.7}$$

the minimum is achieved at the value X^{a}, so-called analysis, that is:

$$X^{\mathrm{a}} = X^{\mathrm{b}} + \left(\frac{\sigma_{\mathrm{b}}^2}{\sigma_{\mathrm{b}}^2 + \sigma_{\mathrm{o}}^2}\right)\left(Y^{\mathrm{o}} - X^{\mathrm{b}}\right). \tag{20.8}$$

The second term between brackets is called the *misfit*, Y^{o}–X^{b}, that is the difference between the model solution and the observations. The analysis value, or best estimate X^{a}, is then the weighted average of the first guess, X^{b}, and the misfit between the observations and the model. The misfit clearly cannot be very large and this means that the model and the observations should be as close as possible. This means that model should be quite realistic, reproducing the 'bulk' of the physical processes contained in the data.

In general, the observations will be at different locations with respect to the model grid points and they could be related to state variables contained in $\mathbf{X}$ by a complex operator (indirect measurements of state variables). This means that in general we can write the misfit as:

$$\mathbf{d} = \mathbf{Y}^0 - H(\mathbf{X}^b) \tag{20.9}$$

where H is the observational operator and $\mathbf{d}$, $\mathbf{Y}^0$, $\mathbf{X}^b$ are now multivariate vectors in the four-dimensional space. Normally H is an interpolation from the model grid to the observational point but in case of diagnostic quantities, such as SSH, it can be a non-linear operator containing a combination of $\mathbf{X}^b$ state variables. The generalization of (20.7) and (20.8) considering (20.9) for the fully multivariate case are

$$\mathrm{I} = \frac{1}{2}\left[(\mathbf{X}^b - \mathbf{X})^T \mathbf{B}^{-1}(\mathbf{X}^b - \mathbf{X}) + (\mathbf{Y}^o - H(\mathbf{X}))^T \mathbf{R}^{-1}(\mathbf{Y}^o - H(\mathbf{X}))\right] \tag{20.10}$$

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{K}(\mathbf{Y}^o - H(\mathbf{X}^b)), \tag{20.11}$$

where

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \tag{20.12}$$

is the Kalman gain matrix. In (20.12) the observational operator appears as the linearized operator, $\mathbf{H}$ of H. The background error covariance matrix is then defined as:

$$\mathbf{B} = \left\langle (\mathbf{X}^b - \mathbf{X})(\mathbf{X}^b - \mathbf{X})^T \right\rangle, \tag{20.13}$$

which represents the error variance for all the model state variables and their cross-correlation. The observational error covariance matrix is represented by $\mathbf{R}$.

The similarity between the simple form of weights in (20.8) and (20.11) should be noted: in both cases the weight on the misfit is given by the error in the background field divided by the sum of the error in the observations and the error in the background field again. If the time variations of $\mathbf{B}$ are parameterized only by changes in the variance of the error fields, then the form (20.12) defines an optimal interpolation (OI) scheme. The essence of sequential data assimilation is to find a sound representation of the multivariate aspects and space-time variability of $\mathbf{B}$. Normally the variations and the multivariate character of $\mathbf{B}$ is chosen *a priori*, on the basis of the knowledge of the relevant processes included in the model and in the observations.

Evaluation of (20.11) can be done 'intermittently', collecting observations within a certain interval of time and inserting them at the end of such interval. Figure 20.7 illustrates a typical intermittent assimilation cycle that in the atmosphere is taken to be 12 or 6 hours, while for the ocean it is generally taken to be several days (Pinardi et al., 2003). In the atmospheric forecasting community, the continuous data assimilation or variational assimilation scheme (Daley, 1996) is now used but it is not described here.

Figure 20.7 illustrates the analysis and forecast cycle for a generic time interval Δt. Every interval cycle, (20.12) is evaluated and the dynamical model is initialized

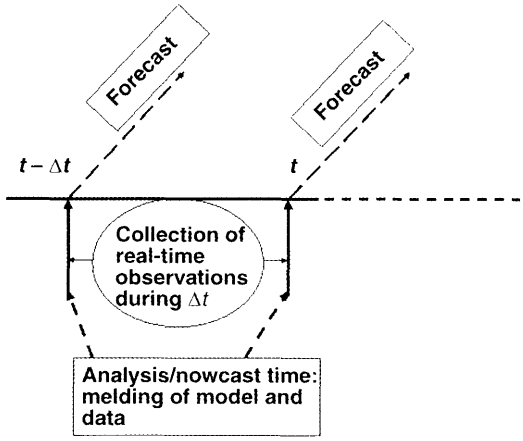

Figure 20.7
Intermittent analysis-forecast cycle.

with $\mathbf{X}^a$. The ocean forecast is started using forecasted surface atmospheric parameters and solving for several time steps, δt, a nonlinear equation:

$$\mathrm{X}^b(t+\delta t) = M\,\mathrm{X}^a(t), \tag{20.14}$$

where M is the state transition matrix corresponding to the model equations (belonging to the OGCM in this case). The background or forecast fields $\mathbf{X}^b$ predicted by (20.14) is used as a first guess, in the successive Δt assimilation cycle, to compute a new $\mathbf{X}^a$.

20.3.2 What does it mean to assimilate different data sets and correct model state variables?

The quality and efficiency of the estimation algorithm outlined above is clearly connected to the details of the background error covariance matrix, **B**. First of all, we do not know the truth of X contained in (20.13) and thus several approximations are normally taken. To estimate **B** we can consider

$$\mathbf{B} = \left\langle (\mathrm{X}^b - \mathrm{X}^a)(\mathrm{X}^b - \mathrm{X}^a)^T \right\rangle \tag{20.15}$$

or

$$\mathbf{B} = \left\langle (\mathrm{X}^b - \bar{\mathrm{X}}^b)(\mathrm{X}^b - \bar{\mathrm{X}}^b)^T \right\rangle \tag{20.16}$$

or

$$\mathbf{B} = \left\langle (\mathbf{Y}^o - \bar{\mathbf{Y}}^o)(\mathbf{Y}^o - \bar{\mathbf{Y}}^o)^T \right\rangle, \tag{20.17}$$

where the 'bar' above the state vectors indicates a suitable time mean. The brackets indicate the ensemble mean of the different realizations available to estimate **B**. Let us discuss in turn each of these approximations. The first in (20.15) indicates that

the error is estimated from the difference between the forecast and the analysis: this will give the most adequate approximation to (20.13). The second and the third approximations contained in (20.16) and (20.17), mean that the error is due to the variance containing part of the state variables with respect to a predefined mean. From a mathematical point of view, the worse approximation to **B** is given by (20.17) as we assume that the variance contained in the observations is the same as in the model. However, most practical data assimilation systems use (20.17) to estimate **B** as the model could have drifted from reality and the use the covariances deduced from the model itself could affect the efficiency of the assimilation system (20.11).

Hereafter, we consider that **B** is calculated from a form of either (20.16) or (20.17), which we now rewrite as

$$\mathbf{B} = \left\langle \tilde{\mathbf{X}}' \, \tilde{\mathbf{X}}'^{T} \right\rangle,$$

where the primes now indicate anomalies with respect to a suitable mean and $\tilde{\mathbf{X}}$ can be either model or observational fields. The important information contained in **B** is the cross-correlation between all the state variables of our system. In other words, **B** is multivariate and composed of block matrices:

$$\mathbf{B} = \begin{vmatrix} \langle T'T' \rangle & \langle T'S' \rangle & .... & .... & \langle T'\eta' \rangle \\ \langle S'T' \rangle & \langle S'S' \rangle & ... & & \langle S'\eta' \rangle \\ \langle U'T' \rangle & & \langle U'U' \rangle & & .... \\ ... & ... & & .... & .... \\ \langle \eta'T' \rangle & \langle \eta'S' \rangle & ... & ... & \langle \eta'\eta' \rangle \end{vmatrix}, \tag{20.18}$$

where the tides have been neglected for simplicity and several sub-matrices have not been explicitly written in (20.18). Each block is quite dense but banded, it contains the variance of the error associated with each field and it represents the cross-correlation between state variable errors in space. The brackets indicate a given ensemble average as before.

For the ocean, De Mey and Benkiran (2002) suggested a possible expression for **B**, that is

$$\mathbf{B} \cong \mathbf{F}\,\mathbf{Br}\,\mathbf{F}^{T}, \tag{20.19}$$

where **F** contains multivariate *vertical* empirical orthogonal functions (v-EOF) and **Br** is a two-dimensional correlation matrix, normally computed from analytical expressions. This means that horizontal and vertical components of the error covariance matrix have been separated. Such separation is allowed in the open ocean and it was found to work especially for quasigeostrophic assimilation of SSH (De Mey and Robinson, 1986).

The separation of **B** into vertical and horizontal modes allows control of the dominant structures of the background error field, as in the ocean the vertical part has a low vertical modal structure while the horizontal part can be quite complex, especially near the coastal margins. In horizontal, the open ocean auto-correlation scales of temperature, salinity and dynamic height fields (Nittis et al., 1993) is given by the 'exponential correlation function', that is

$$C(x,y) = A\left(1-\frac{r^2}{a^2}\right)e^{-\frac{r^2}{2b^2}} \tag{20.20}$$

where for the horizontal zero crossing distance and the correlation decay scale, respectively, $r^2 = x^2 + y^2$ is the square of the distance measured in a (x, y) coordinate plane and A is the variance of the field. This form is used in the **Br** matrix.

In the vertical, the error field is dominated by errors in the mixed layer and in the thermocline, where the highest temporal variance of the **X** fields is found. A low vertical error modal structure is to be expected in view also of the fact that temperature and salinity profiles, as well as dynamic height, are dominated by the first baroclinic modes (Masina and Pinardi, 1994) and background correlation errors may reflect the incorrect energy content of these modes.

Thus the separation (20.19) can be rewritten in mathematical form as

$$B(x,y,z) = \sum_j F_j^2(z)C_j(x,y), \tag{20.21}$$

where F_j are the vertical modes and C_j their respective horizontal structure. Equivalence (20.19) and (20.21) cannot be shown to be true in general but they are convenient approximations of the **B** matrix that show efficient assimilation of ocean data, as we show later. These ad hoc formulations should also be tried for ecosystems, in order to control the model structure of the corrections for the relevant state variables.

The **F** matrix represents then the vertical error variability in all the state variables and their vertical cross-correlation. If we substitute (20.19) into (20.12) we obtain

$$\mathbf{K} = \mathbf{F}\,\mathbf{Br}\,\mathbf{F}^T\mathbf{H}^T(\mathbf{H}\mathbf{F}\,\mathbf{Br}\,\mathbf{F}^T\mathbf{H}^T + \mathbf{R})^{-1}. \tag{20.22}$$

The three-dimensional estimation problem contained in (20.22) can now be decreased in order due to the fact that **F** can be described by EOF. In fact, the relevant vertical EOFs (v-EOF) will certainly be less than the number of vertical model levels used for the estimation of the correlation matrix **F**. The physical reason for this is that the temperature variance below the thermocline is low and, on subseasonal timescales, the variability between the surface and the intermediate and deep water are uncorrelated. In turn this means that the number of significant v-EOFs to represent the state variable variance in the vertical is less than the number of model levels. Normally if the model levels are m, the vertical EOF explaining most of the variance are only n, with $n \langle\langle m$, as shown by Sparnocchia et al. (2003). Thus we can use a 'reduced order space' for the v-EOF, that is the first n modes accounting for most of the variance. The estimation problem is then given by

$$\begin{aligned}\mathbf{K}^{\mathrm{ROOI}} &= \tilde{\mathbf{F}}\,\mathbf{Kr}\\ \mathbf{Kr} &= \mathbf{Br}\,\tilde{\mathbf{F}}^T\mathbf{H}^T(\mathbf{H}\tilde{\mathbf{F}}\,\mathbf{Br}\,\tilde{\mathbf{F}}^T\mathbf{H}^T + \mathbf{R})^{-1},\end{aligned} \tag{20.23}$$

where now the dimensions of $\tilde{\mathbf{F}}$ are much less than in **F**.

It is now easy to understand that, even if the misfit in (20.9) is given for few state variables, the correction in (20.11) will be carried out for all the state variables contained in $\mathbf{X}^b$, given that we consider all the cross variances between state variables. This might not be wise every time, and several cross-correlations could be eliminated on the basis of the specific data set used or physical assumptions. For example, in the case of temperature observations it is not advisable to use the cross-correlation of $\langle \boldsymbol{\psi}' \mathbf{T}' \rangle$,

that is to change the stream function on the basis of temperature data. This correlation in fact cannot be understood on the basis of sound physical processes within the limit of the data availability and the processes represented in a primitive equation model.[103] In general this correlation is small but this cannot be guaranteed *a priori* because it depends on the details of how we calculate **B** and in any case its presence in **B** will introduce noise in the analysed field. Thus for every data set to be assimilated, we could study the most important cross-correlations to be considered.

20.3.3 Example 1: assimilation of satellite altimetry only

Let us suppose that we have only SSH observations and that **K** contains **F** defined as follows:

$$\mathbf{A} = \begin{vmatrix} \langle \mathrm{T}'\mathrm{T}' \rangle & \langle \mathrm{T}'\mathrm{S}' \rangle & \langle \mathrm{T}'\boldsymbol{\Psi}' \rangle \\ \langle \mathrm{S}'\mathrm{T}' \rangle & \langle \mathrm{S}'\mathrm{S}' \rangle & \langle \mathrm{S}'\boldsymbol{\Psi}' \rangle \\ \langle \boldsymbol{\Psi}'\mathrm{T}' \rangle & \langle \boldsymbol{\Psi}'\mathrm{S}' \rangle & \langle \boldsymbol{\Psi}'\boldsymbol{\Psi}' \rangle \end{vmatrix} \tag{20.24a}$$

$$\mathbf{A} = \mathbf{F}\boldsymbol{\Lambda}\mathbf{F}^{\mathrm{T}},$$

where **A** contains only the vertical multivariate covariances and **F** are the v-EOF of **A**.

This means that the SSH misfit will be projected into temperature, salinity profiles and ψ amplitudes and that the corrections will be done on three model state variables, not the SSH itself. This is another way of saying that SSH in assimilation should be considered as a diagnostic variable. Another interesting correlation matrix for SSH assimilation is

$$\mathbf{A} = \begin{vmatrix} \langle \mathrm{T}'\mathrm{T}' \rangle & \langle \mathrm{T}'\mathrm{S}' \rangle & \langle \mathrm{T}'\Psi' \rangle & \langle \mathrm{T}'\boldsymbol{\eta}' \rangle \\ \langle \mathrm{S}''\mathrm{T}' \rangle & \langle \mathrm{S}'\mathrm{S}' \rangle & \langle \mathrm{S}'\Psi' \rangle & \langle \mathrm{S}'\boldsymbol{\eta}' \rangle \\ \cdots & \cdots & \cdots & \langle \Psi'\boldsymbol{\eta}' \rangle \\ \cdots & \cdots & \cdots & \langle \boldsymbol{\eta}'\boldsymbol{\eta}' \rangle \end{vmatrix} \tag{20.24b}$$

The vertical modes of the correlation matrix (20.25b) are forced to be covariant also with SSH anomalies. In MFS, we use the v-EOF constructed from the state vector

$\mathbf{X}^{a} = \begin{vmatrix} \mathrm{T} \\ \mathrm{S} \\ \boldsymbol{\psi} \\ \boldsymbol{\eta} \end{vmatrix}$ from both the form (20.24a) or (20.24b). The reason for using the covariance

[103] A primitive equation system is composed of the Navier-Stokes equations considering the Boussinesq and hydrostatic approximation. For the ocean, incompressibility is also assumed. These equations contain physical processes such as baroclinic instability that transfer energy between baroclinic (temperature-dominated) and barotropic modes (vertical integral of velocity field represented by the streamfunction). The baroclinic instability process is connected to horizontal gradients of temperature and vertical gradients of the velocity field but not simply to the temperature anomaly. Thus using the simple correlation between temperature and streamfunction field does not accurately represent the baroclinic process and this correlation should be disregarded if the only information is from temperature profiles.

between all these fields is connected to the analytical expression (20.2), where we see that SLA is a complicated function of all the T, S and ψ model variables.

Following the geostrophic relationship written in (20.2) the observational operator **H** contains explicitly that expression (Demirov et al., 2003). In this way we ensure that the correction is carried out on the model state variables T, S and ψ weighted by their respective scaling coefficients that produces the correct value for η or the SSH. The same procedure and understanding should pave the way for the assimilation of other observations that use the observational operator and the cross-correlations contained in **B** to transform the measurement information into the state variables contained in $\mathbf{X}^{\mathbf{b}}$.

As an example of application of this concept for the assimilation of SSH, Masina et al. (2001) computed cross-correlations between η' and T' only. The derived profiles, also called 'synthetic XBTs' in the paper, were assimilated in a global model with a two-dimensional variational algorithm. The success of their assimilation is based on the fact that SSH is strongly correlated to vertical profiles of temperature and salinity below the mixed layer and that it was possible to derive an error estimate for the synthetic XBT. This is another way to express the relationship of Figure 20.4 and reduce the order of **F**.

In conclusion, the degree with which one observed state variable controls the changes or updates on the other state variables is given by the degree of correlation between the errors in the state variables and the projection of **F** on **H**.

20.3.4 Example 2: assimilation of temperature profiles only

In the Mediterranean and the world's oceans, XBT profiles are collected in real time, as explained in previous sections, and thus assimilation of this data set is a priority for any forecasting system.

This time the vertical model error covariance matrix can be now taken to be

$$\mathbf{A} = \begin{vmatrix} \langle \mathrm{T}'\,\mathrm{T}' \rangle & \langle \mathrm{T}'\,\mathrm{S}' \rangle \\ \langle \mathrm{S}'\,\mathrm{T}' \rangle & \langle \mathrm{S}'\,\mathrm{S}' \rangle \end{vmatrix}, \tag{20.25}$$

which we also call the 'vertical water mass cross-correlation matrix'. Again, **F** are the v-EOF of **A**. Thus for assimilation of XBT in the Mediterranean, bivariate v-EOF were computed as EOF of (20.25) considering the several subregions of the basin, where different water masses can be identified. The state variable anomalies were computed using historical data (Sparnocchia et al., 2003).

The v-EOF for summer and for the Levantine basin region are shown in Figure 20.8. We note the presence of the high values of the errors in the 50 m surface layer corresponding to a well-known subsurface low salinity signal, normally referred to as Atlantic Water. The first salinity EOF changes of sign below 50 m and presents a subsurface maximum correspondent to the Levantine Intermediate Water signal. Thus, looking at the v-EOFs it is possible to recognize the major surface and intermediate water masses of the Mediterranean Sea where variance of the temperature and salinity signal is contained and thus, in our interpretation, also the model errors. In conclusion, a misfit in temperature will induce corrections in salinity that have a known water mass T-S relationship contained in **F**.

To conclude, let us discuss the assimilation of SST observations. If **A** were full, then the misfit in SST would induce changes in the whole water column for T, S, U, V, and others, and this is clearly not representative of a realistic ocean process, except

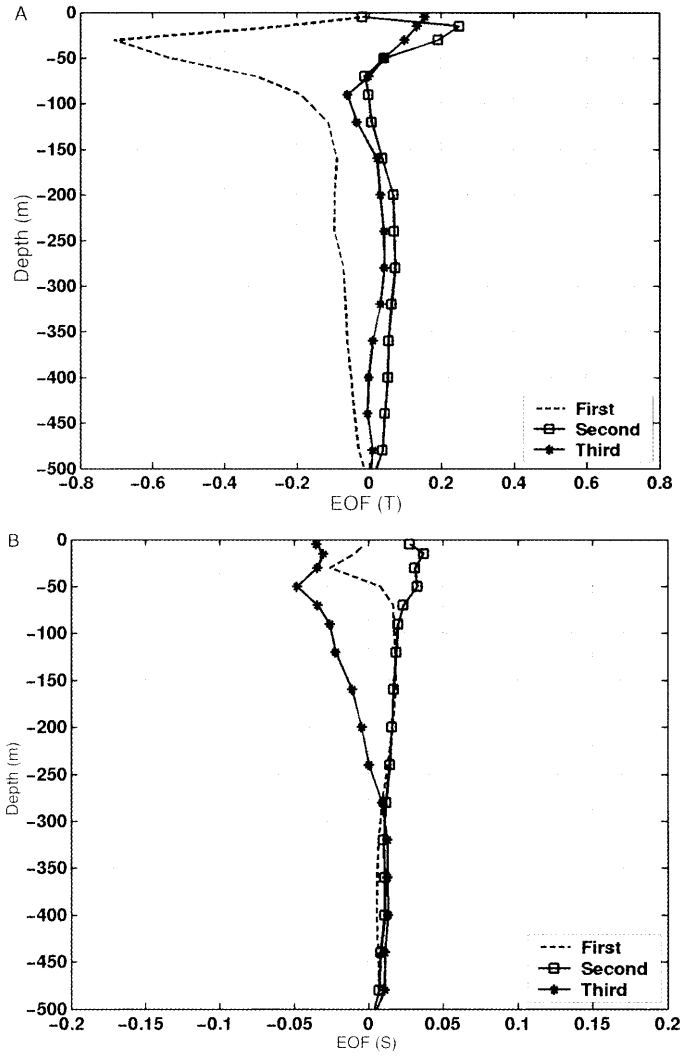

Figure 20.8
A, First three summer season temperature bivariate v-EOF for the Levantine basin region in the Mediterranean (east of 30°E and north of 30°N), calculated by Sparnocchia et al. (2003) from historical hydrographic data. The three modes account for more than 80% of the temperature-salinity variance in the water column.
B, First three summer season salinity bivariate v-EOF for the Levantine basin region in the Mediterranean (east of 30°E and north of 30°N) calculated by Sparnocchia et al. (2003) from historical hydrographic data. The three modes account for more than 80% of the temperature-salinity variance in the water column.

for few limited areas (deep convection areas) and severe winter periods. On the other hand, the error cross-correlation between state variables is the only information that allows a change in the prognostic state variables on the basis of a limited amount of observed state variables. In the case of SST, it is advisable to carefully study the assimilation impact on the complete state vector before using automatically a full state **F** vector in **B**.

20.3.5 Example 3: assimilation of combined altimetry and temperature profiles

Having decided the satellite altimetry and temperature profiles assimilation assumptions, MFS has implemented a combined assimilation of both measurements using a two-step approach, described below (Demirov et al., 2003). The method is based on the hypothesis that it is important to maintain the different expressions of **A** and thus **F** for the two kind of measurements, as they contain different physical information about the water column stratification. The multivariate v-EOFs for Ψ, T, S are used for assimilation of SLA and the bivariate T, S v-EOFs are used for assimilation of XBT profiles. This way each data set is assimilated with its optimal vertical error modes.

In order to combine the two assimilation procedures we envisage the following steps, illustrated in Figure 20.9. The time window of assimilation is two weeks and an analysis is calculated once a week (day 'J' in the diagram). Both past and future observations with respect to J are considered, that is the OI is used in the 'smoother mode'. One week we use only SLA observations to generate the present week analysis starting from the previous week analysis which instead assimilated only XBT observations. After the first cycle it will be unimportant which data set is assimilated first and each week the analysis will benefit from both SLA and XBT observations.

The working of the combined assimilation of satellite and temperature profiles is shown in Section 20.5 where the comparison between the observed, modelled and corrected profiles, by means of (20.11), is given.

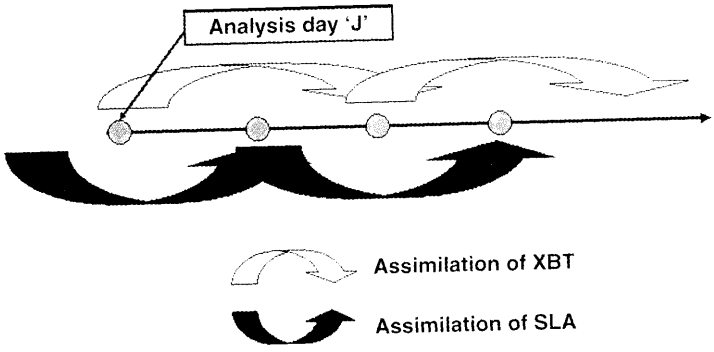

Figure 20.9
Combined assimilation of XBT and SLA profiles used in the MFS operational assimilation system.
Source: Demirov et al. (2003).

20.4 The coastal forecasting and assimilation problem

The MFS approach to coastal forecasting uses nesting and downscaling of the large-scale flow field up to the required resolution. Coastal environmental forecasting will make use of the analysed physical flow fields to advect passive or active tracers in the coastal areas with the maximum accuracy. Specifically MFS could provide a basis to examine the transport and fate of planktonic populations and dissolved constituents in the water column. Potential users would greatly benefit from an explicit statement of the space and timescales over which the MFS can provide realistic estimates of transport. The benefit from MFS is maximum when the variability will be forced heavily by the offshore forcing, which is the case for several coastal regimes in the Mediterranean and the world ocean.

In a recent review paper, Pinardi et al. (2006) have shown that more than half of the Mediterranean coastal areas are limited by narrow shelves that allow for strong control by the open ocean flow field. In addition, large portions of the coastal areas do not have any local runoff control mechanism and thus coastal productivity will be totally dominated by the discharges from coastal towns and offshore inputs of nutrients. In the latter case, a system such as MFS is crucial to understand the primary productivity of the coastal areas. Even in extended shelf areas, the remote effects of open ocean ecosystem structure cannot be neglected to understand coastal productivity. This is evident for the case of Adriatic mucilage phenomena where it is believed that the arrival from the open ocean areas of highly refractory dissolved material could influence aggregation rates of particles in the water column and thus influence the mucilage itself.

MFS has also developed the coastal downscaling and the necessary data assimilation for the shelf and coastal domains of interest. In particular, we refer here to the Adriatic coastal areas where a 5 km OGCM was nested within the MFS OGCM in order to arrive properly to the near coastal areas (Oddo et al., 2005). In addition, the assimilation system described above for temperature profiles was applied to coastal CTD profiles (measuring both salinity and temperature) collected in the very near coastal areas in order to control the near-shore coastal field. It was found that the separation of vertical and horizontal modes of the background error covariance matrix allowed a consistent assimilation even in complicated coastal dynamical regimes (Grezio and Pinardi, 2006). In future the coastal data assimilation system may be different from the open ocean one allowing for the higher space and time variability of the background error covariance matrix in these regions.

20.5 Forecast consistency, quality and accuracy

In this section we show how to carry out an assessment of the quality and accuracy of the assimilation system using three basic indicators:

- *Consistency indicator*: the qualitative (visual) correspondence of circulation structures in the analysis fields with features known from phenomenological studies or observations only.
- *Quality indicator*: comparison between observations and model before data insertion. This indicator can be expressed in terms of the root-mean-square of the misfit, defined by

$$\text{rms\_misfit}(\mathbf{d}) = \frac{1}{N}\sqrt{\Sigma\, \mathbf{d}^2}$$

$$\text{nms\_misfit}(\mathbf{d}) = \frac{1}{\sqrt{\Sigma \mathbf{Y}^{o2}}} \sqrt{\Sigma\, \mathbf{d}^2},$$

and **d** is defined by (20.9) and the sum is done over different realizations in space at a given time.

- *Accuracy indicator*: comparison between analysis and forecast, i.e. forecast skill score. Two can be defined:
  (a) the first is $\text{rms\_fcst}(\mathbf{X}^b - \mathbf{X}^a) = \frac{1}{N}\sqrt{\Sigma(\mathbf{X}^b - \mathbf{X}^a)^2}$, referred to simply as rms of forecast error;
  (b) the second is $\text{rms\_pers}\left[\mathbf{X}^b - \mathbf{X}^a(t=0)\right] = \frac{1}{N}\sqrt{\Sigma\left[\mathbf{X}^b - \mathbf{X}^a(t=0)\right]^2}$, rms of persistence error.

Figure 20.10 shows the comparison for April 2002 of SLA from the forecast system analyses with respect to objective analysis done only with the satellite data. In both

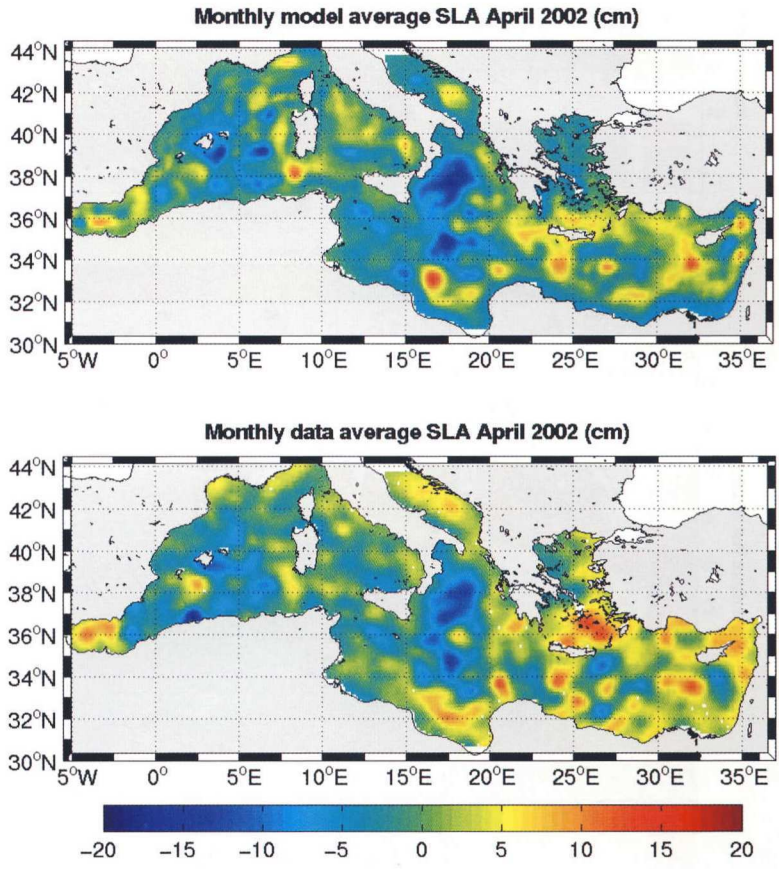

Figure 20.10
Correspondence between A, model; B, satellite objective analysis of SLA (cm) for the month of April 2002.

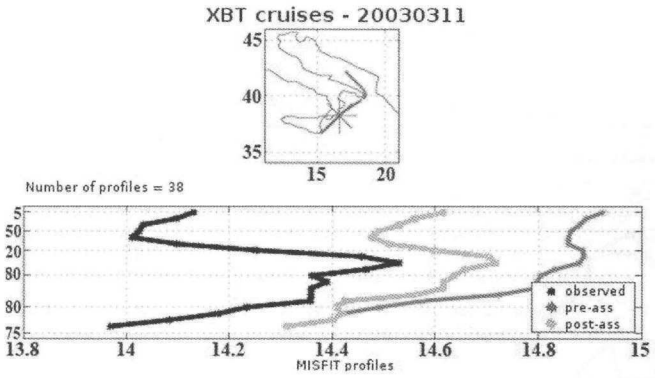

Figure 20.11
Comparison between model temperature before data insertion, temperature observation, model corrected solution for 11 March 2003. The star in panel A indicates the position of the observation along a VOS line.

estimates we recognize the cyclonic anomaly area in the Ionian, which is known from other data to be dominant in these years.

More important is the quality indicator and the study of the structure of the misfit. We show an example of misfit for the case of XBT profiles (Figures 20.11, 20.12) and SLA (Figure 20.13).

Figure 20.11 shows the observed profile, the model profile before data insertion and the corrected profile or analysis done with (20.11). The model profile is shifted towards the observed one and the misfit is of the order of -0.5°C. This is a good result for the data-assimilation scheme, confirming that the assumptions made for **F** and **B**

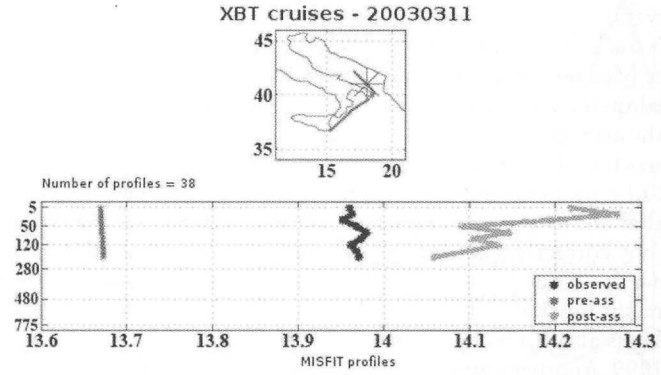

Figure 20.12
Comparison between model temperature before data insertion, temperature observation, model corrected solution and vertical misfit for 11 March 2003. The star in panel A indicates the position of the observation along a VOS line.

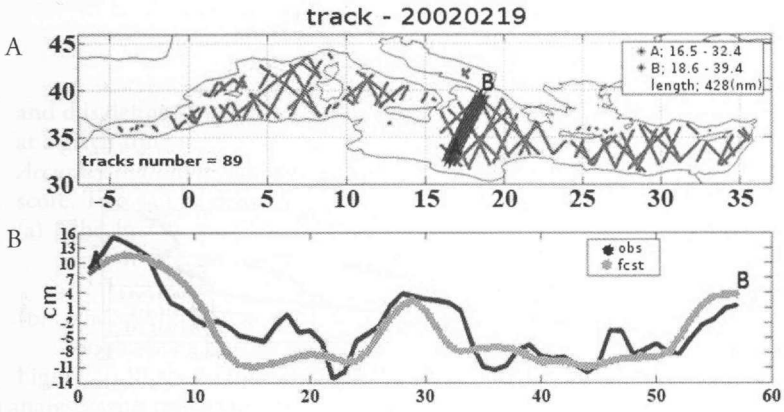

Figure 20.13
Model and observational SLA values along a Topex/Poseidon track on 19 February 2002. Panel B shows the model solution before data insertion with superimposed observations for each point of the track from point A to B.

are generally correct. However, Figure 20.12 shows that sometimes the assimilation fails to produce a sensible corrected profile. In fact the new profile is closer in absolute value to the observations but it overshoots the observed one and has a noisy vertical structure. We argue that this is due to the inadequate choice of v-EOF for this region and for this time of the year. This points out to the fact that **B** should be more space and time varying while in our case we have a unique v-EOF for the whole southern Adriatic area and for the three months of the winter season. Future improvements include v-EOF that are calculated point by point and with higher time frequency (monthly and interannually).

In Figure 20.13 we show the misfit and the comparison of the model solution versus observations before data insertion. We see that the model follows quite closely the altimeter track values but misfit is still of the order of 5–10 cm which is large with respect to the Mediterranean anomaly signal which varies between ± 30 cm.

A global quantitative estimation of the quality of the forecasting system may come from the average of the *rms_misfit* and *nms_misfit* for the whole Mediterranean region as shown in Figures 20.14 and 20.15, for SLA data and XBT respectively. In Figure 20.14 we note that the *rms_misfit* error varies from months to months indicating that the temporal scale of variability is not correctly captured by the model and not correctly inserted in the analysis by the assimilation scheme. We suppose that a shorter assimilation cycle (now one week) could improve this situation. The rms of misfit for the XBT (Figure 20.15) at 30 m indicates the reasonable value of 0.6°C. This is also decreasing rapidly with the continuous insertion of data after September 1999. Another important issue in assimilation is concerned with the length of time it will take the system to 'converge' towards the observations after the initial time when model and data are added together. This is a difficult question to analyse here but our estimate is that several months are needed to show the improvement, as we show in Figure 20.15. This naturally will depend also on the data scarcity and the measuring network.

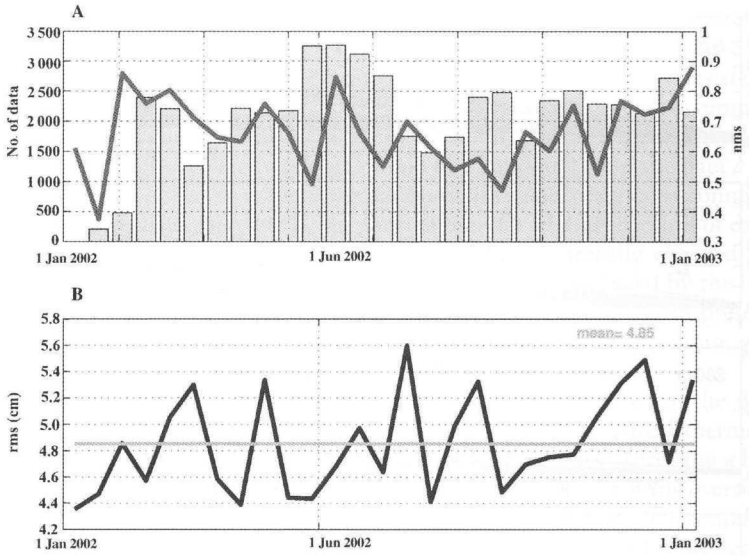

Figure 20.14
A, weekly number of SLA data points for Topex/Poseidon altimeter every two weeks with overlaid the nms_misfit.
B, rms_misfit *for* the period January 2002 to January 2003.

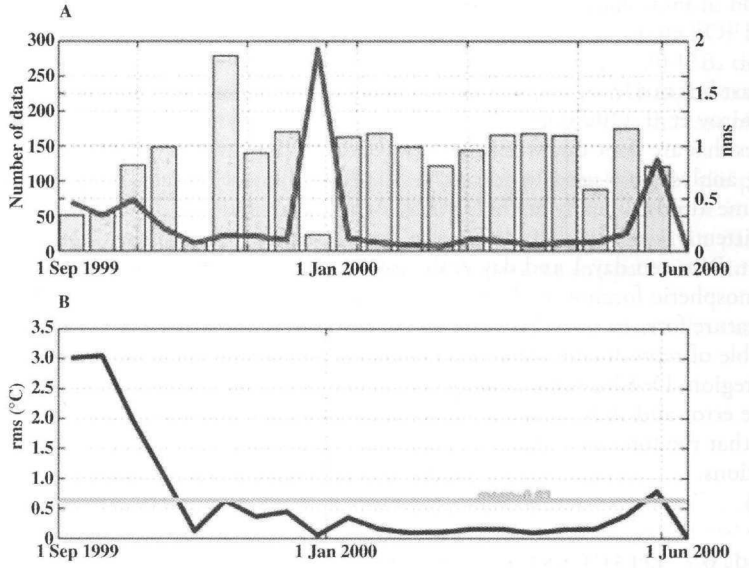

Figure 20.15
A, weekly number of XBT data used in the analysis with overlaid the nms_misfit for temperature at 30 m.
B, rms_misfit for the period September 1999 to June 2000.

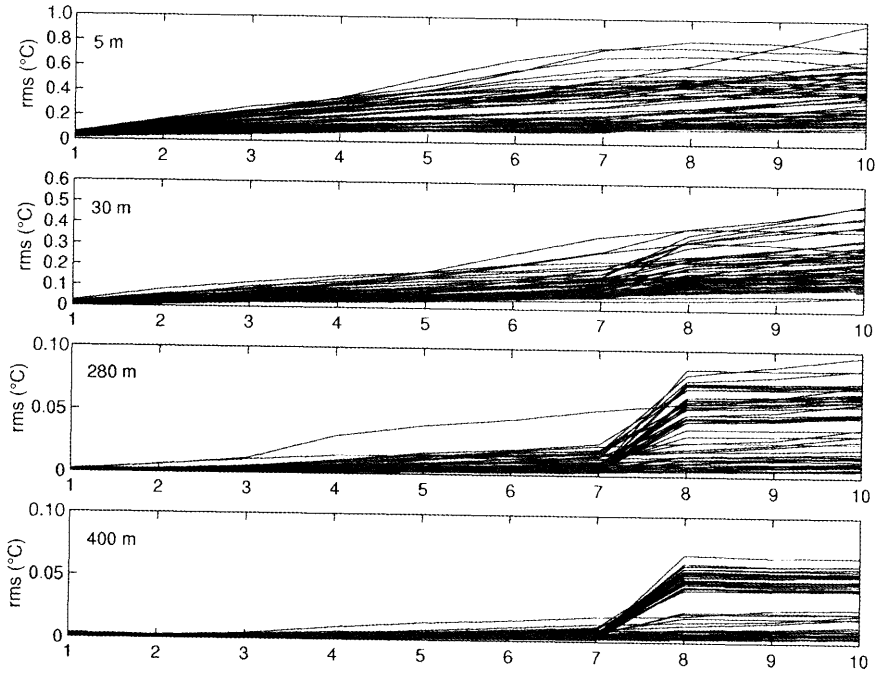

Figure 20.16
10 days root-mean-square (rms) forecast error growth for the weekly forecasts in 2000, at different depths.

Last but not least, the accuracy indicator is shown in Figure 20.16. As described in Demirov et al. (2003) in MFS we use to run a simulation experiment between the analyses that are done one week apart. Since this is done in delayed mode, atmospheric forcing analyses are used to force the 'hindcast' between analyses. The error grows with time almost linearly at the surface. At the levels below, the data insertion occur intermittently every week and thus the error suddenly grows at day 7 of each 10-day forecast. Between day 1 and day 7, the forecast error growth is due only to the different atmospheric forcing used in the forecast and the hindcast. The value of the rms temperature forecast error is reasonable showing that the atmospheric forecast forcing is capable of reproducing some of the essential features of the atmospheric variability of the region. Demirov et al. (2003) compare this rms forecast error with the rms persistence error and show that the latter is always higher that the former. This basically shows that the forecast is needed in order to reach reasonable accuracy in the 10-day predictions.

20.6 Discussion and conclusions

This chapter has described the ocean state estimation problem set up in the MFS operational scheme. Apart from the necessary regionalization issues, a few general statements about assimilation of real-time data have been made.

First of all, it is recognized that sophisticated preprocessing of satellite and *in situ* data is needed before observations can be inserted into the model. Above all, the quality control procedures should be as much as possible automatic and consistent with the physical assumptions made in the assimilation scheme. One of these assumptions is that the observations contain only the slow timescale variability (larger timescales than the day) and thus high-frequency signals should be eliminated from the data *a priori*.

Second, we have seen that the assimilation quality and practice is connected to the assumptions made to calculate **B**, the form chosen for **H** and the kind of data that we assimilate. The multivariate character of **B** should be carefully checked against physical processes that are contained in the cross-correlations induced by the inserted data. The time and space variability of **B** is generally underestimated in present systems, like the Mediterranean Sea, and inconsistencies may occur that will not produce the optimal merging of background or model data with observations.

Order reduction of **F** is strongly recommended especially if it can be done, as in the ocean physics, on the basis of process assumptions. The vertical thermohaline structure of the ocean is low mode, with few vertical EOF modes expressing a big part of the variance of the error field in the vertical. Thus separation of **B** into vertical and horizontal structures seems to be advisable. However, the v-EOF are horizontally non-homogeneous and that effect should be considered. As an example of more recent developments, we show in Figures 20.17A and 20.17B the percentage variance explained by the first bivariate v-EOF mode of **A** written as in (20.25) considering the definition (20.16) and different mean fields subtracted. The figures show that, depending on which $\tilde{\mathbf{X}}^b$ is used, the percentage variance explained by the first v-EOF changes as well as its horizontal distribution. **A** is constructed from a 35-day temporal time series of $\mathbf{X}^b$ subtracting $\tilde{\mathbf{X}}^b$ calculated by a mean over several years (a climatological $\tilde{\mathbf{X}}^b$) for Figure 20.17A and only a 35-day mean for Figure 20.17B. It is important to note that several regions have almost 100% variance explained by only one bivariate EOF but this value changes depending on which average is subtracted. In Figure 20.17B, the areas with a large proportion of variance explained by the first EOF have changed extension, and more modes are needed in general to explain the same variance in the case of Figure 20.17A.

Schemes for ecosystem data assimilation of the same complexity of the physical state estimation problem discussed here are being developed. However, it is recommended that the structure of **B** for these systems is studied in detail, in particular because biogeochemical observations are much more scarce in space and time than observations for the physical state variables.

Let us make an example of how matrix **B** could look in ecosystem data assimilation. First of all the state vector could be indicated by:

$$\mathbf{X} = \begin{bmatrix} \mathbf{T} \\ \mathbf{Chl} \\ \mathbf{N} \\ \mathbf{Z} \\ \mathbf{D} \end{bmatrix} \tag{20.26}$$

where this time the biochemical state variables, **Chl**, chlorophyll concentration, **N**, dissolved nutrients, **Z**, zooplankton biomass and **D**, detritus, have been introduced in

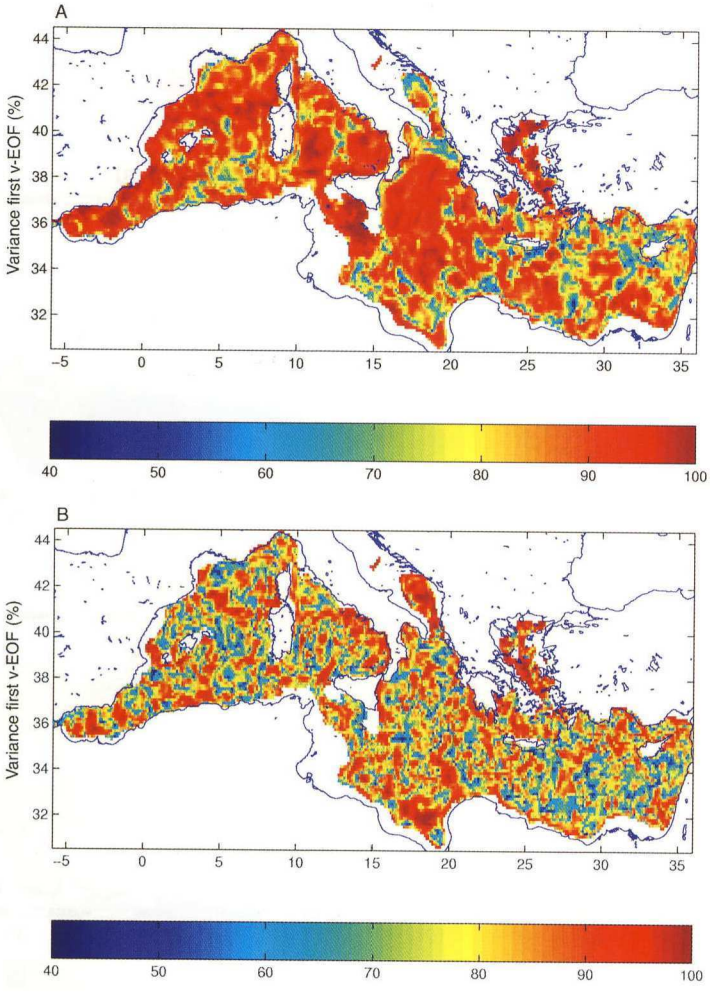

Figure 20.17
A, percentage of variance (colours of the palette) explained by bivariate v-EOF calculated from a 35-day time series of temperature and salinity profiles at each model grid point – matrix of (20.26) – and with a climatological monthly mean value of T,S subtracted at each level to compose **A**.
B, percentage of variance (colours of the palette) explained by bivariate v-EOF calculated from a 35-day time series of temperature and salinity profiles at each model grid point (matrix of (20.26)) and with the 35-day mean value of T,S subtracted at each level to compose **A**.

addition to temperature, **T**. Chlorophyll concentration here is a model state variable which corresponds to a phytoplankton group and its concentration may be thought to be proportional to the phytoplankton biomass. Decomposing the **B** matrix into horizontal and vertical modes, we now take a possible definition of **A** and **F** as follows:

$$\mathbf{A} = \begin{vmatrix} \langle \mathbf{T}'\mathbf{T}' \rangle & \ldots & \ldots & \ldots & \ldots \\ \langle \mathbf{Chl}'\mathbf{T}' \rangle & \langle \mathbf{Chl}'\mathbf{Chl}' \rangle & \ldots & \ldots & \ldots \\ \langle \mathbf{N}'\mathbf{T}' \rangle & \langle \mathbf{N}'\mathbf{Chl}' \rangle & \langle \mathbf{N}'\mathbf{N}' \rangle & \ldots & \ldots \\ \langle \mathbf{Z}'\mathbf{T}' \rangle & \langle \mathbf{Z}'\mathbf{Chl}' \rangle & \langle \mathbf{Z}'\mathbf{N}' \rangle & \langle \mathbf{Z}'\mathbf{Z}' \rangle & \ldots \\ \langle \mathbf{D}'\mathbf{T}' \rangle & \langle \mathbf{D}'\mathbf{Chl}' \rangle & \langle \mathbf{D}'\mathbf{N}' \rangle & \langle \mathbf{D}'\mathbf{Z}' \rangle & \langle \mathbf{D}'\mathbf{D}' \rangle \end{vmatrix} \tag{20.27}$$

and

$$\mathbf{A} = \mathbf{F} \Lambda \mathbf{F}^{\mathrm{T}} \tag{20.28}$$

F is then a vector containing the v-EOF of **A**.

We have explicitly written only half of the matrix because we know that is symmetric. Let us consider the satellite ocean colour transformed into surface chlorophyll as the input observational data set. If the analysis system (20.11) is used, then the chlorophyll misfit will produce corrections on all the system state variables listed in (20.26) and **F** will infer the vertical corrections to be introduced, as in the case of the physical system explained above.

The matrix (20.27) expresses the vertical covariance between **T**, **N**, **Z**, **D** and **Chl** state variables. The auto-correlation for **Chl** is a very important part of the matrix (20.27): it means that surface chlorophyll will produce a vertical profile of chlorophyll, sometimes containing a subsurface chlorophyll maximum if the statistics allows it. The other cross-correlations can be interpreted as follows:

- the covariance between **T** and **Chl** points to the functional relationship between photosynthetic activity corresponding to a representative phytoplankton group and water temperature. This correlation may be weak in subtropical regions;
- the covariance between **Chl** and **N** points to the functional relationship between the dissolved nutrients and chlorophyll concentration changes. This covariance in the vertical will have a subsurface maximum in open ocean areas while it will be surface intensified in coastal areas;
- the covariance between **Chl** and **D** points to the functional relationship between detritus variance (related to the mortality rate of the phytoplankton in simple models) and chlorophyll variance. This cross-correlation is complicated since there is a delayed response of detritus to the increase in chlorophyll error variance;
- the covariance between zooplankton variance and chlorophyll is also very complex, as again the covariance has a time delay.

Simplifying the structure of (20.27) could involve the deletion of the cross-correlations between **Chl** and all the other ecosystem state variables except **Chl** itself, **N** and **T**. This will produce corrections by (20.11) on only three of the system state variables given in (20.26) while the others will be changed by the model time stepping, as given by (20.14). In this way the corrections to the detritus and zooplankton, in response to the insertion of surface chlorophyll observations, will be made following the dynamical equations contained in the ecosystem model, after chlorophyll and nutrients have been updated by the observations in the whole water column. This scheme seems to be reasonable instead of changing the zooplankton biomass directly as a consequence of a change in surface chlorophyll: this operation in fact may not be justified within the limits of the assimilation cycle chosen and the dynamical response of the system.

Ecosystem data assimilation is at its infancy but the premises of a workable data assimilation system for the physical components of the marine environment makes it possible to think that in the next years primary production estimates in the ocean will benefit from the optimal merging of observations and predictive ecosystem models.

References

Babin, M., Roesler, C. S. and Cullen, J. J. (eds). 2006. *Real-time Coastal Observing Systems for Marine Ecosystem Dynamics and Harmful Algal Blooms: Theory, Instrumentation and Modelling*. Paris, Intergovernmental Oceanographic Commission of UNESCO. (Monographs on Oceanographic Methodology XX.)

Castellari, S., Pinardi, N. and Leaman, K. 1998. A model study of air-sea interactions in the Mediterranean Sea. *J. Mar. Syst.*, 18, pp. 89–114.

Daley, R. 1996. *Atmospheric Data Analysis*. Cambridge, UK, Cambridge University Press.

De Mey, P. and Benkiran, M. 2002. A multivariate reduced order Optimal Interpolation Method and its application to the Mediterranean basin-scale circulation. In: N. Pinardi and J. D. Woods (eds), *Ocean Forecasting: Conceptual Basis and Applications*. New York, Springer-Verlag.

Demirov, E. and Pinardi, N. 2002. The simulation of the Mediterranean Sea circulation from 1979 to 1993. Part I: The interannual variability. *J. Mar. Syst.*, 33–34, pp. 23–50.

Demirov, E., Pinardi, N., Fratianni, C., Tonani, M., Giacomelli, L. and De Mey, P. 2003. Assimilation scheme of the Mediterranean Forecasting System: operational implementation. *Ann. Geophys.*, 21, pp. 189–204.

Dickey, T., Frye, D., Jannasch, H., Boyle, E., Manov, D., Sigurdson, D., McNeil, J., Stramska, M., Michaels, A., Nelson, N., Siegel, D., Chang, G. C., Wu, J. and Knap, A. 1998. Initial results from the Bermuda Testbed Mooring Program. *Deep Sea Res.*, 45, pp. 771–94.

Gibson, J. K., Kallberg, P., Uppala, S., Hernandez, A., Nomura, A. and Serrano, E. 1997. ERA description. *ECMWF Re-Analysis Project Rep. Series*, 1, 71 pp.

Grezio, A. and Pinardi, N. 2006. Data assimilation of temperature and salinity profiles in the Adriatic Regional Model. *Acta Adriatica*, 47, pp. 149–68.

Kalnay, et al. 1996. The NCEP/NCAR 40-year re-analysis project. *Bull. Am. Meteorol. Soc.*, 77, pp. 437–71.

Korres, G., Pinardi, N. and Lascaratos, A. 2000. The ocean response to low frequency interannual atmospheric variability in the Mediterranean Sea. Part I: Sensitivity experiments and energy analysis. *J. Climate*, 13(4), pp. 705–31.

Legler, D. M. and O'Brien, J. J. 1988. *Tropical Pacific Wind Stress Analysis for TOGA, IOC Time Series of Ocean Measurements*, Vol. 4. Paris, IOC–UNESCO. (Intergovernmental Oceanographic CommissionTechnical Series 33.)

Le Traon, P. Y. 2002. Satellite oceanography for ocean forecasting. In: Pinardi and Woods (eds), op. cit., pp. XX–XX.

Lorenc, A. C. 2002. Atmopsheric data assimilation and quality control. Pinardi and Woods (eds), op. cit.

Manzella, G. M. R., Scoccimarro, E., Pinardi, N. and Tonani, M. 2003. Improved near real time data management procedures for the Mediterranean Ocean Forecasting System-Voluntary Observing Ship Program. *Ann. Geophys.*, 21, pp. 49–62.

Masina, S. and Pinardi, N. 2004. Mesoscale data assimilation studies in the middle Adriatic Sea. *Cont. Shelf Res.*, 14(12), pp. 1293–310.

Masina, S., Pinardi, N. and Navarra, A. 2001. A global ocean temperature and altimeter data assimilation system for studies of climate variability. *Climate Dynamics*, 17, pp. 687–700.

McPhaden, M. J. 1995. The tropical atmosphere ocean array is completed. *Bull. Am. Meteorol. Soc.*, 75(5), pp. 739–41.

Mellor, G. L. 1996. *Introductory Dynamical Oceanography.* Oxford, UK, Pergamon Press.

Mellor, G. L. and Ezer, T. 1991. A Gulf Stream model and an altimetry assimilation scheme. *J. Geophys. Res.*, 96(5), pp. 8779–95.

Milliff, R. F., Feilich, M. H., Liu, W. T., Atlas, R., Large, W. G. 2001. Global ocean surface vector wind observations from space. In: C. J. Koblinsky and N. R. Smith (eds), *Observing the Oceans in the 21st Century*, Melbourne, Bureau of Meteorology, GODAE Project Office, pp. 102–19.

Milliff, R. F., Large, W. G., Morzel, J. Danabasoglu, G. and Chin, T. M. 1999. Ocean general circulation model sensitivity to forcing from scatterometer winds. *J. Geophys. Res.*, 104(5), pp. 11337–58.

Molcard, A., Pinardi, N., Iskandarani, M. and Haidvogel, D. B. 2002. Wind driven general circulation of the Mediterranean Sea simulated with a Spectral Element Ocean Model. *Dynam. Atmos. Oceans*, 35, pp. 97–130.

Nittis, K., Pinardi, N. and Lascaratos, A. 1993. Characteristics of the summer 1987 flow field in the Ionian Sea. *J. Geophys. Res.*, 98(6), pp. 10171–84.

Oddo, P., Pinardi, N. and Zavatarelli, M. 2005. A numerical study of the interannual variability of the Adriatic Sea (2000–2002). *Sci. Total Environ.*

Oddo, P., Pinardi, N., Zavatarew, M. and Coluccelli, A., 2006. The Adriatic Basin Forecasting System. *Acta Adriatica*, 47, pp. 169–184.

Pinardi, N., Allen, I., Demirov, E., De Mey, P., Korres, G., Lascaratos, A., Le Traon, P. Y., Maillard, C., Manzella, G. and Tziavos, C. 2003. The Mediterranean Ocean Forecasting System: first phase of implementation (1998–2001). *Ann. Geophys.*, 21, pp. 3–20.

Pinardi, N., Arneri, E., Crise, A., Ravaioli, M. and Zavatarelli, M. 2006. The physical, sedimentary and ecological structure and variability of shelf areas in the Mediterranean Sea. *The Sea*, 14. Boston, Harvard University Press.

Pinardi, N. and Masetti, E. 2000. Variability of the large scale general circulation of the Mediterranean Sea from observations and modelling: a review. *Palaeogeogr., Palaeoclimatol., Palaeoecol.*, 158, pp. 153–73.

Pinardi, N. and Woods, J. D. (eds). 2002. *Ocean Forecasting: Conceptual Basis and Applications.* New York, Springer-Verlag.

Prandle, D. 2002. Predictions in the North Sea. In: Pinardi and Woods (eds), op. cit.

Raicich, F. and Rampazzo, A. 2003. Observing system simulation experiments for the assessment of temperature sampling strategies in the Mediterranean Sea. *Ann. Geophys.*, 21, pp. 151–66.

Robinson, A. R. and Leslie, W. G. 1985. Estimation and prediction of oceanic fields. In: J. Crease, W. J. Gould and P. M. Saunders (eds), *Progress in Oceanog*raphy, Vol. 14. Oxford, UK, Pergamon Press, pp. 485–510. (Essay on Oceanography.)

Rossby, T., Siedler, G. and Zenk, W. 1995. The volunteer observing ship and future ocean monitoring. *Bull. Am. Meteorol. Soc.*, 76(1), p. 5.

Smith, N. and Lefebvre, M. 1997. The Global Data Assimilation Experiment (GODAE). *Proc. Biarritz International Symposium*, October 1997.

Sparnocchia, S., Pinardi, N. and Demirov, E. 2003. Multivariate Empirical Orthogonal Function analysis of the upper thermocline structure of the Mediterranean Sea from observations and model simulations. *Ann. Geophys.*, 21, pp. 167–87.