ENSO AND SEASONAL FORECAST SYSTEMS

D.L.T. Anderson(1), T.N. Stockdale(1), M.K. Davey(2), M. Fischer(3), M. Ji(4),
A. Rosati(5), N. Smith(6) and S.E. Zebiak(7)

(1) ECMWF, Shinfield Park, Reading, UK; (2) UK Met Office, Bracknell, UK; (3) IMGA-CNR, Bologna, Italy; (4) NCEP, Washington D.C., US; (5) GFDL, Princeton, US; (6) BMRC, Melbourne, Australia; (7) IRI, Lamont, US.

1 INTRODUCTION

The best known application of ocean data to seasonal forecasting is in the attempts to predict El Nino events, or, more generally, ENSO variability. The title of this paper is a reminder that one should not concentrate on ENSO to the exclusion of other phenomena, however. We view our task as the general assessment of data requirements for seasonal or longer term prediction, and hence consider the requirements of global forecast systems. In reality the equatorial Pacific contains the strongest source of interannual variability in the earth's climate, and has the best developed ocean observing system, and so this region acts as a good guide to the data uses and requirements of a global system. One can go further and say that the strategy needed for seasonal forecasting can act, with some limitations, as a guide to even longer-term prediction: the multi-analysis/multi-model infrastructure being pursued at ECMWF, for example, could easily be extended to the decadal time-scale as well.

In the tropics there are good reasons for believing that the ocean state is strongly controlled by the forcing. There are exceptions: one example is the tropical instability waves, although even these might be partly controlled by the forcing, another might be smaller-scale features (such as squirts and jets). Nonetheless the dominant picture is of an ocean state, which can be determined to some considerable extent by forcing, at least on larger spatial scales and seasonal time scales. Variability on decadal time scales may not be so 'deterministic' and so to capture some decadal variability of ENSO, for example, might put demands on the observing system additional to those required for seasonal prediction. In addition to the forcing, it is necessary to have a good ocean model. We will return to this point later. If one does not have either a good history of the ocean forcing fields or a good ocean model, then the ocean state has to be determined by ocean observations. To a large degree this was the case for many users during the last El Nino: the TAO observing system was very important for providing an easily accessible picture of the ocean to those with no access to a model analysis.

For those who do have an ocean model and fast access to the forcing fields, what use data? There are three main applications: the primary one is to create initial conditions from which one may launch forecasts using coupled models. A secondary one is to produce an ocean analysis in its own right. The experience from meteorology is that such analyses can be useful for understanding processes at work in the fluid since in some sense they give the most comprehensive integrated picture of the fluid evolution. A third reason for data is for model validation and improvement. For this paper, a good place to start is to look at how data are used currently in real-time applications of seasonal forecasts and what possible extensions are being considered.

2. CURRENT "QUASI-OPERATIONAL" SYSTEMS: THE FORCING REQUIREMENTS

At ECMWF a global analysis of the ocean is performed daily and a two hundred-day coupled model forecast performed from each analysis. The direct use of ocean data in producing the ocean analysis will be considered in the next section; here we consider the role played by the forcing data. The ocean is forced with the fluxes from the atmospheric analysis system. These are probably the best that are available in a timely (few hours) fashion and contain 'synoptic-type' variability such as intraseasonal variations and westerly wind bursts. The atmospheric analysis system produces instantaneous analyses every 6 hours. From every second analysis (i.e. at 12 hour intervals) a 10-day medium range atmosphere forecast is launched. There are various options available for obtaining the ocean forcing fields from such a system. We can derive stresses and fluxes from the 6 hourly instantaneous analyses (and this derivation could be made in various ways), or we can use the actual stresses and fluxes from the first 24 hours of the medium range forecasts, and again here there are various methods of extracting daily averaged values. The fields we ultimately need are the fluxes of momentum, heat and moisture.

If these forcing fields were perfect, the ocean model were perfect and the role of internal instabilities were weak, then we might have little need of in situ ocean data for model initialisation: the past history of the forcing used to drive an ocean model would be sufficient to produce the ocean state. Of course none of these conditions can be met. In fact the forcing fields are less well known than we would like, even with the TAO array producing wind measurements several times per day and with one scatterometer providing continuous data coverage for several years since early 1992.

It is not just improved measurements that are needed, but sustained measurements. Variations in the quality and quantity of past measurements in the meteorological system lead to low frequency variations in the quality of the forcing field, making it difficult for those using the data to calibrate their seasonal or ENSO forecasting system. Of course it is not just the observing system which has not been steady, it is the atmospheric analysis/forecast system too (i.e. the system used for weather forecasting, as opposed to seasonal forecasts). Typically either the forecast model, or the analysis procedure itself changes more than once per year, again adding low frequency variability into the ENSO forecasting system: even if the seasonal forecast system does not change, it relies on output from an atmospheric analysis/forecast system which does. To overcome this non-stationarity, meteorological centres such as ECMWF and NCEP have carried out extensive reanalysis of past data using a consistent processing/model/assimilation system. However, there is low frequency variability in the wind products which is probably spurious and in the case of ECMWF there is a jump at the end of 1993 when one changes from wind stresses from the reanalysis to those from the then operational system from Jan 1994. In addition to the stresses from the model analysis system, which are derived from 24 hour averages, we can deduce a stress from the analysed near-surface wind fields. Investigation suggests that this latter product is more stable and less prone to spin-up problems. Spurious jumps or trends in wind scan be particularly troublesome to those making seasonal forecasts especially if they were to occur at the same time as a major ENSO is developing.

Fig. 1a shows the large differences between the reanalysis/operational stress and the FSU products in the EQ2 region. The differences are larger than one might expect, and show some low frequency differences. In the central Pacific the interannual anomaly of stress is of the order of 0.02 Nm^{-2}, with uncertainties (differences between wind products) which are sometimes of similar size. How important are these differences? The easiest way to gauge that is to run an ocean model forced by the various wind products. The results for the depth of the 20C isotherm are shown in Fig. 1b for the region to the east, EQ1. The differences are often of the order of 20 metres again comparable to the size of the interannual anomaly. We have used other wind products also with similar results. The purpose of Fig. 1 is not to show any one product as better than another but to give a sense of the size of the uncertainty.

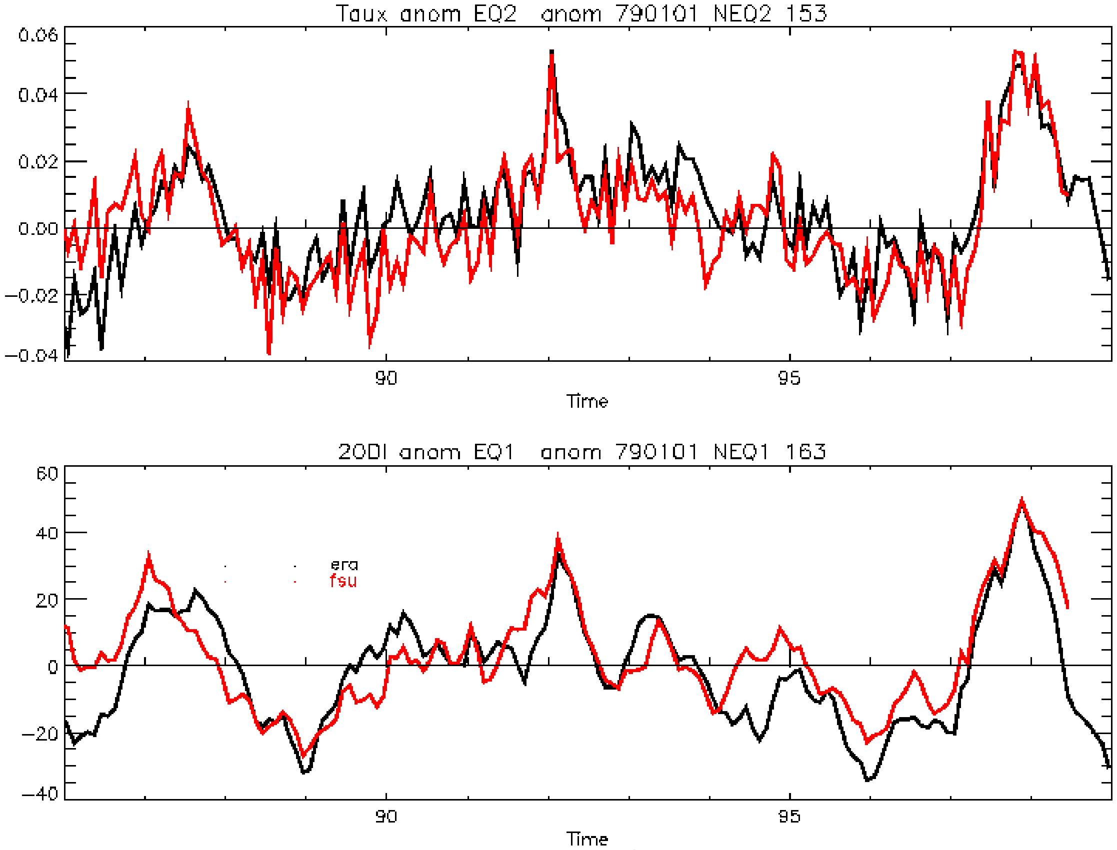

Fig.1. Top panel (1a): Time series, 1986-98, of monthly mean zonal wind stress anomalies in the central equatorial Pacific (EQ2: 5N-5S, 130W-170W) as given by the ECMWF analyses (black) and by FSU wind analyses (red). Bottom panel (1b): Anomalies in depth of the 20 deg C isotherm in the eastern Pacific (EQ1: 5N-5S, 90W-130W) when an ocean model is forced by these two wind stresses.

Further anecdotal evidence can be provided as to the perceived quality of the forcing fields. At NCEP, the ENSO forecasting group did not use either the wind stress or the wind from the NCEP analysis during their coupled model development process, but instead a hybrid product consisting of the mean seasonal cycle from Hellerman and Rosenstein and the anomalous winds from FSU. The story for heat and fresh water fluxes is not much better. At NCEP they use various observed climatologies, while at ECMWF the ERA and operational products are used, but with little confidence that they are correct, given their sensitivity to the atmospheric model physics and analysis procedure changes referred to above.

Although this meeting is about ocean measurements, we must put surface flux measurements, in particular surface stress, as a high priority since there is little prospect of sufficient ocean data to allow ocean analyses without involving the forcing fields. This does not necessarily mean the fluxes must be directly measured. In principle they could be diagnosed using atmospheric models fed with other measurements. In practice it is hard to see how one avoids direct measurements of at least some of the surface variables over at least part of the globe. To some degree steps are in hand to improve the situation: a 40-year reanalysis is scheduled to begin in the next few months using the almost latest technology for assimilation (3D-var). While it is likely that this will be a big improvement over what is currently available and will reduce some of the uncertainties highlighted in Fig. 1, it is naive to expect that this will produce a completely acceptable product. Oceanographers have been stressing the need for better fluxes for twenty years. Has the accuracy not improved? One would like to think it has, but as the output from atmospheric models is sensitive to many aspects of the physics, we can not expect a monotonic improvement. It is also well to remember that the demands put on the flux fields by equatorial ocean modellers is very severe. One should also note that we can not be sure that our ocean models have the correct sensitivity to surface forcing fields.

It is against this backdrop that one should now view ocean requirements. Data are needed both to improve the ocean model and to correct for model and forcing error, but it is frequently unclear whether errors are a result of forcing error or ocean model formulation/parameterisation. In particular, the largest impact of the data is often correcting the mean state of the ocean, which otherwise rapidly changes to a different state due to model/forcing inadequacies.

3. PRESENT-DAY "OPERATIONAL" OCEAN ANALYSIS SYSTEMS FOR SEASONAL FORECASTING

Here we describe some ocean GCM based analysis systems in use at operational centres. This is to give a specific context for the general discussion of ocean data which follows. There are a number of other ocean assimilation/initialization schemes used for real-time ENSO forecasting which are not discussed here, in particular those which are based on intermediate or simpler ocean models. The data needs of such models are generally either similar to or less than those of the most complex systems; the interaction between model system and data needs is discussed later.

3.1 The ECMWF system

The Optimum Interpolation (OI) scheme used is derived from that of Smith et al (1991), which in turn is a simplified version of the OI scheme formerly used at ECMWF for atmospheric analyses. Currently only thermal data are analysed. The data are bunched into a data window of 10 days, stretching from 5 days before to 5 days after the central analysis time. An OI is performed for the central analysis time and the increments derived are then fed in slowly to the model over the next 10 days until the next assimilation time. This procedure has some deficiencies, as there is no imposed multivariate relationship. In the extra-tropics, geostrophy or related balance could be used to deduce velocity corrections, though this would be difficult in the equatorial region. Because of the interest in the tropics, by putting the increments in gradually the model should be able to develop an appropriate velocity field consistent with the density field. However, spreading the increment over a few days blurs some of the information in the thermal data. For example, a Kelvin wave can travel more than 2000 km in the 10 days currently used to spread the increment. A shorter time window of 3 days in the TOGA TAO region has been tested but not yet implemented.

The data are quality controlled by comparing with analyses performed without a given datum. One departure from normal OI procedure is that data are interpolated to model level rather than model interpolated to datum level. This allows an OI to be performed model level by model level. No vertical coherence is imposed in the assimilation process. An important part of the OI system is an appropriate specification of the correlation scales, which define the range and strength of the influence of observations in the analysis system. Attempts have been made to relate the choice of these scales to the observed correlation scales, but the limitations of the data available mean that only approximate bounds can be diagnosed.

A very important part of the ECMWF ocean initialization is the use of a strong constraint on SST within the model. This is relaxed towards Reynolds' 1 degree analysis with a time scale of only two or three days, which ensures that the surface mixed layer of the ocean always has a temperature close to that observed. This in turn means that the coupled model forecast is able to start with a good representation of the atmospheric convection, and plays a large role in the skill of the forecasts in the early months. Small scale erroneous “blotches” over the oceans in ECMWF seasonal forecasts have been traced to inconsistencies in the use of SST analyses for initializing the ocean model. Even SST changes of 0.1-0.2 deg. C in the tropics can have a systematic impact on the subsequent evolution of coupled forecasts. A consistent and accurate SST analysis will be essential for a high quality and reliable seasonal forecasting system.

No salinity analysis is performed. This is clearly an undesirable feature as there are reasons to believe that salinity plays a nontrivial role in shaping the density field (Cooper 1988). It is imposed upon us as there are essentially no real-time salinity data with which to perform an analysis. Some steps are taken to control salinity. In particular, the surface salinity is forced with the best estimate of the fresh water flux that is available in a timely fashion. So, the precipitation and evaporation are taken from the ECMWF analysis system. Although it is one of the few such products available in a timely fashion, it is not without its problems (see comments on spin-up above). So in addition to the freshwater forcing, an additional relaxation to Levitus surface salinity is applied. The relaxation time scale is 30 days. This is a compromise between keeping the model from drifting in response to inadequate freshwater forcing, and not damping too much, any genuine interannual anomalies. See section 4.3 for a discussion of salinity.

3.2 The NCEP ocean analysis scheme

At NCEP, data are assimilated using a 3D variational scheme. Although this is formally equivalent to OI in an ideal world, there are several practical decisions taken at ECMWF and NCEP, which mean that the two schemes are not equivalent. The NCEP ocean model is only of the Pacific, and although in principle analysis is done for the extratropics, interaction with the atmosphere in the forecast is only in the restricted equatorial region of 15S to 15N. As in the scheme used at ECMWF, "appropriate" weight is given to the data and to the model when blending the two in the assimilation process. The scale of the first guess error covariance is now allowed to vary with latitude, being most anisotropic on the equator where the horizontal scale is ~18 degrees (Behringer et al 1998). It also varies with depth. It differs from ECMWF in that some vertical correlation is applied, whereas the scheme used at ECMWF is performed on a level by level basis, without vertical correlation. Because of the shortage of data, quite long windows are used. At NCEP a 4-week window is used, but with a linear taper applied to give data zero weight if more than two weeks distant from the analysis time. (Cf. the ECMWF procedure, which is to give full weight to data within +/-5 days of analysis time and zero weight outside this time). The minimisation of the cost function, equivalent in some sense to an analysis, is done quite frequently (every 3 timesteps out of 8), and increments are applied every timestep. As at ECMWF, the model is allowed to find its own velocity field in response to changes to the thermal field.

A further major difference relates to the handling of the salinity field. Originally, no forcing of S was made at NCEP and no analysis of S was performed, in which case the salinity could be badly wrong. More recently, (Behringer et al. 1998), a surface relaxation has been included and surface salinity is now relaxed to the seasonal cycle of Levitus with a time scale of 50 days. No fresh water flux is applied to the current NCEP ocean analysis, although this is available from the NCEP operational meteorological analysis and ready to be implemented.

3.3 BMRC coupled GCM ENSO forecast system

For a number of years the Bureau of Meteorology Research Centre in Melbourne, Australia has run an El Nino forecast system using an intermediate model and an ocean assimilation system. Recently, they have also developed a coupled GCM forecasting system. This is similar to the ECMWF system discussed above and uses the same data assimilation method. Both models are global. In BMRC's model, the increments are inserted every 10 days rather than introduced gradually. SST information is inserted with a short relaxation time which means that the SST boundary condition is effectively prescribed, obviating the need for good surface fluxes (short-wave penetration is added as a separate source term). At present, neither altimeter or salinity data are assimilated (a nudging to Levitus' salinity climatology is used subsurface). As with the intermediate model, there is a clear positive impact of subsurface data over the hindcast period 1980-1995. Major issues are the quality of the wind forcing (both models use a hybrid of FSU, reanalysis and real-time monthly-mean winds), initialization beyond the tropical Pacific, the Indian Ocean data stream and the assimilation of altimeter data.

In summary, thermal data from the TAO moorings, from XBT and PALACE floats are used at BMRC, NCEP and ECMWF, although only those data which are delivered in a timely fashion to the GTS. Dealing with salinity variability is still a challenge.

4. SPECIFIC DATA TYPES

4.1 SST

SST has already been mentioned as vital for seasonal forecasting. Present day operational analyses are typified by Reynolds' 1 degree global analysis, available as a weekly product in very near real time. The uncertainty in this field varies spatially, depending on the density/quality of in situ data and the existence of clear skies to allow satellite data to be gathered. In the 1990's the error in much of the equatorial Pacific is probably only 0.1 to 0.2 deg C on the monthly time scale. Errors are significantly bigger in the pre-satellite era, and also in regions with reduced in-situ data. A new analysis method (2DVAR instead of OI) and new quality control is about to be introduced at NCEP. The new analyses can differ from the old by up to 0.5 deg C in the equatorial Pacific, although typically only in small regions. Note that errors of 0.1 to 0.2 deg C are typically 10-20% of the interannual "signal" of SST variability. If the SST anomalies control the changes in wind stress that drive the coupled system, one might then suppose that the SST uncertainty produces errors in the "tendency" of the coupled system, and hence errors in the forecast evolution of SST, at a similar magnitude. At the moment other errors probably dominate, but it should be recognized that improved SST analyses are highly desirable, and may become more important as model and other errors are reduced.

The specifics of obtaining SST are discussed elsewhere in this volume (Reynolds et al., this Conference). We simply note that it is of the highest importance that adequate satellite and in situ measurements are made to allow accurate SST products to be derived; and that at the moment a limiting factor in the quality of available SST analyses is the availability of resources for making best use of existing data.

4.2 Sub-surface temperature data

This is the key data component of present day operational ENSO assimilation/forecast systems at the major NWP centres. Cost is an important issue in choosing the most appropriate methods for observing in situ temperature. Other things being equal, we would prefer cheaper methods that allow a greater volume of data to be returned with the resources available. Timeliness, accuracy and position are also important, however. The equatorial band is particularly important for model initialization, with its delicate balances and fast response times, and of course its importance for ENSO variability. Frequent, high quality data are important. Present day ocean GCM assimilation systems suffer from substantial systematic errors, the correction of which has the potential to be aliased with variability. Repeated data at fixed locations, such as is provided by fixed moorings, is useful in minimizing this problem.

One important issue that has not been much studied is the impact of data dropouts. TAO moorings have a noticeable failure rate, either of individual instruments or sometimes the whole mooring. Replacement time is typically many months, and during this time the observing network might be significantly less effective than it was designed to be. In assimilation systems where data are fighting model bias, a temporary lack of data can result in spurious anomalies in the analysis. An appropriate level of redundancy/resilience in the observing system design is important.

The highest priority must be given to the equatorial oceans, where the role of ocean data in contributing to seasonal forecast skill is already established. The value for seasonal forecasts of extratropical ocean data remains to be established, although the comprehensive modellers would certainly like data in these regions. A lower frequency of data return is likely to be adequate at these latitudes for operational purposes, due to the slower time scale of the signals we would hope to sample.

4.3 Salinity

As mentioned above, salinity data are not used at either NCEP or ECMWF. In assimilation, it is assumed that if you have no information on a variable it should not be updated, but in fact leaving salinity untouched when T is updated may not be a good idea: the salinity field drifts and fictitious water masses are created as a result. Yet the dearth of salinity measurements makes it difficult to envisage a full salinity analysis in the next few years. We can probably do better than is done currently by some use of a climatological T (S) relationship, or even inferring an S value for every T observation by statistical or water mass-conserving relationships. This may be less acceptable in the near surface layer, but some form of salinity analysis at the surface might be possible. In the near future such an analysis would be based on a combination of surface ship measurements and data from fixed moorings and profiling floats; see also altimetry below. In the longer term, salinity measurements from space might be possible, at least for the scale and amplitude of variability seen in the tropics.

How worried should we be about a lack of a salinity analysis in current ‘operational’ systems? At ECMWF and NCEP, the present analysis systems suffer from a drift in salinity. This was not identified as a major problem at ECMWF until they started testing the assimilation of altimeter data. NCEP also ignored the salinity issue until they began assimilating TOPEX/POSEIDON altimetry data in 1996; since then they have identified a large salinity signal believed to have lead to different ENSO prediction results (Ji et al., 1999). There is no doubt that salinity is a problem for ocean analyses, but the sensitivity of the coupled system to errors in the salinity field is largely unknown, and it remains unclear the degree to which the forecasts are degraded by salinity errors.

4.4 Altimeter data

In using altimeter data, it is necessary to project the data in the vertical or to use some trajectory assimilation procedure such as 4D-var. If we seek to project in the vertical, there are basically two procedures: one is to use a statistically-derived relation between sea-level and subsurface temperature and salinity (Vossepoel and Behringer 1999), or as is sometimes the case just temperature, totally ignoring any salinity influence on sea level; a second is to project the data, preserving local water mass properties as far as possible (Cooper and Haines 1996). Statistically-derived schemes have some advantage in that they build-in a climatology of sorts, but suffer from the fact that they can not adapt to changing water mass properties as may sometimes occur. Water mass schemes have the advantage that they can adapt to changing conditions, but have no memory of the climatology: so, if the model drifts then there is no means of correction. Some compromise in which a relaxation to climatology on a relatively long time scale may be necessary with such schemes. The purpose of this section is not so much to advocate any particular method, but rather to highlight the difficulties in solving the salinity problem. Further discussion can be found in Vossepoel and Behringer 1999, Alves et al 1999 and Segschneider et al 1999.

Data from Topex-Poseidon and ERS-2 are received in near real time (with a delay of 7 days). These data are compared with the model surface height field in near real-time, but are not yet part of the real-time assimilation system at ECMWF. However, the HH data (Historical Homogeneous data set from CLS) have been assimilated into the ECMWF ocean analysis system off-line using a water mass-preserving scheme. Early attempts to use the data showed up problems with the salinity field in the model. Further analysis shows that it is possible to use altimeter data constructively, provided a satisfactory geoid is used and the salinity drift is tackled. The drift is reduced by including a relaxation of S(z) towards Levitus climatology with a relaxation time which was initially set at one month. Not only did the analyses using altimeter data alone compare favourably with the OI analyses obtained by assimilating all thermal data, but also subsequent forecasts were promising. However, the quality of the geoid used is likely to have been dependent on data from the TAO array, given the way in which the geoid was constructed. In other areas where the geoid is less well known the benefit of the altimeter is likely to be reduced.

For seasonal forecasting, the bottom line is the quality of the forecasts. One may use subjective assessment of analyses to get an idea of the impact of a given observing system on the quality of the analyses, but there is no obvious way of converting that into an assessment of the forecast skill short of doing the experiments. To do the experiments requires a lot of computing time. The assessment of the impact of altimeter data on forecasts was based on 100 integrations, each of 200 days. The 100 integrations consist of a 5-member ensemble, over the 5 years of Topex for 4 seasons per year. One may argue that for some purposes this set is too small. The experiments in question are comparisons of forecasts from analyses using just TAO and SST, just altimeter and SST and a control using just SST. ECMWF is currently assessing the combined use of TAO, altimetry and SST. This may require a larger set of hindcasts, if altimetry and TAO show a significant redundancy in the tropical Pacific. Experiments at GFDL have found that with regard to forecast skill, little value is added by assimilating altimetry data over TAO/XBT.

At NCEP Topex/Poseidon altimetry data has been used in real time analysis and forecasting since 1996. To avoid needing precise knowledge of the geoid, sea surface height deviations relative to a three year mean of 1993-1995 are compared to the same deviations from the model. The difference in sea surface height field is used as an additional constraint when computing the increment to the model temperature field (Behringer et al. 1998). One effect of assimilating altimetry data has been correcting model sea level errors due to lack of knowledge in salinity. However, because the correction is made only to the temperature field, this was sometimes achieved at the expense of greater error in the temperature field.

Progress has been made during the past two years toward correcting the model salinity field using T(z), altimetry data and S(z) if available (Vossepoel et al, 1999; Vossepoel and Behringer, 1999). The new scheme is an extension of the existing 3D variational system to include capability of assimilating observed S(z) and derived pseudo S(z) profiles. The pseudo S(z) profiles are derived from SST, T(z) from TAO and altimetry data based on EOFs of temperature and salinity (Maes and Behringer, 1999). The first guess error covariance for assimilation S(z) in the vertical is derived from statistical T-S relationship and EOFs of salinity variability based on historical

CTD data. Preliminary results show significant impact on the salinity field of the analyses.

The UK Met. Office (UKMO) has been assimilating past and near-real-time altimeter data into a Pacific OGCM with the aim of improving ocean analyses and forecasts. The effectiveness of the altimeter data has been assessed by making forecasts with a CGCM with and without the altimeter data in the ocean initial conditions. The evidence (from a small number of cases to date) is that the altimeter data is beneficial.

The assimilation scheme used by the UKMO is the analysis/correction method (Lorenc et al. 1991), which is a sub-optimal OI-type multi-variate method. With this scheme, increments are assimilated repeatedly within a time window of several days, and spread spatially with an iterated filter to approximate a Gaussian footprint. The ocean scheme has been used to assimilate T and/or altimeter data. For altimeter data, upper ocean T and S increments are calculated by first determining sea level increments and then raising/lowering vertical profiles accordingly (Cooper and Haines 1996). The model and observations are given roughly 2/3 and 1/3 weight respectively, so the analysed sea level anomalies can depart substantially from the observed values in some regions. Surface salinity is weakly relaxed toward observed climatological values. SST is also relaxed toward observed values. No sub-surface relaxation terms are currently used, but such terms are being implemented.

4.5 Ocean velocity data

The ocean is strongly controlled by the earth's rotation. A measure of this is given by the radius of deformation which typically varies from ~200 km in the equatorial region to 20 km at higher latitudes. On scales longer than the radius of deformation, potential energy is more important than kinetic for many processes. Long Rossby waves for example store most of their energy as potential energy, meaning that it is more important to observe the density field than the velocity. Kelvin waves in an inviscid case are equipartitioned, although dissipation might shift the balance in favour of potential energy. Thus more emphasis has gone on using thermal data than velocity. This is a happy circumstance as the thermal field is better measured than the circulation field. Nonetheless, velocity data in the form of time series have been useful for model verification. Examples are the five moorings along the equator. Surface velocity data from drifters are not assimilated into any ENSO forecast system, although they could provide some spatial coverage. There are various reasons why the surface drifter data (drogued at 10m) are not assimilated: the data isn't quite dense enough - the decorrelation lengths for velocity are likely to be shorter than for the mass field; velocity is probably less effective in defining the ocean analysis, as discussed above; the data are at a single level, which meteorological experience suggests makes them less easy to use; and finally the data are located in the ocean boundary layer, which may not represent the deeper flow. The velocity data have been used for model validation, however, and these have led to changes to the NCEP analysis system as a result.

4.6 Timeliness of data

One may think that timeliness of data is relatively unimportant for seasonal forecasting, but this is not necessarily the case. Systems which make longer range forecasts of the approximate phase and amplitude of ENSO are relatively insensitive to data delays of several weeks, but for seasonal forecast systems which are attempting more detailed calculations of climate anomalies on timescales of one to several months, the timeliness of forecasts being issued is important, and lead time can make a difference. A one month delay causes a noticeable drop in expected skill for Nino-3 SST, for example. Present practice at eg ECMWF does not yet produce forecasts with the shortest possible lead-times consistent with the availability of data on the GTS, but more rapid use of the data is expected in the future. If coupled forecasts are being made within one week or less of real time, for example, then data availability might become an issue. Ignoring non-available data may not have a large impact on the forecasts themselves, particularly if the data is in more slowly varying parts of the ocean, but it would obviously be preferable to have the data available in near real time if cost or quality are not much affected. The present TAO and VOS distribution systems (using GTS) generally work well in making reasonable quality data available in a reasonable time, and it is hoped that ARGO data delivery will be similarly effective.

5. NEW METHODS OF ASSIMILATING DATA

The current method of assimilating altimeter data is to adjust the vertical profiles of T and S and to blend these increments with an OI or 3D-var system. There is no temporal or spatial coherence implied. But new methods of assimilating data are being developed. One promising approach is to use 4D-var, which we discuss here as an example of possible new methods. Other advanced schemes based on the Kalman filter and its variants are also being considered. In the 4D-var approach a trajectory through the data is required, which best fits the data. Practical applications of this require the use of the model, its tangent linear and its adjoint. A cost function representing the difference between the model-predicted fields and the observations is then minimised. A typical period over which the cost function is minimised is 1 month. An important component of such a system is the specification of the forecast error covariance, usually called the background and denoted B. B acts as a smoothing operator. Information from one assimilation cycle is carried forward to the next by using the outcome at final time of one assimilation as the starting point of the next.

What should be the free variables in the cost function? The initial conditions of the model, including salinity are obvious candidates. But in so far as we do not know the wind field sufficiently well, something about the wind field should be used and probably other components of the surface forcing fields. There are far too few observations in the ocean to allow us to take a full set of the wind at every time step at every grid point as being free but some subset of wind uncertainty could be used. In addition some parameters used in subgrid scale processes could in principle also be taken as variables to be determined by data. In a research environment these issues are all being investigated. For both the ocean models used at ECMWF, adjoints exist and assimilation experiments are in progress. It is likely to be a few years before these are used in real-time applications but the technology is developing fast.

Currently, fields for which there are few data, e.g. salinity, will be only loosely constrained. In fact if there are no data, they will only be constrained in the use of the first guess for the trajectory analysis and possibly through the background term. In current versions being tested at ECMWF, there is no effective control. So S can vary quite a bit, new water masses can be created. On the other hand as far as forecasts are concerned, we do not know if this is a problem. Letting S find its own values may not lead to bad forecasts: the sensitivity of forecasts to various errors in the initial conditions just is not known.

6. OTHER ISSUES

6.1 Shorter range forecasting

Although there is interest in pushing seasonal forecasting out to longer times, decadal and beyond, the infrastructure needed to do that is essentially that being set up at places like UKMO or ECMWF. But there is also an interest in looking at the intermediate range between a week and a month or so. For such a range, clearly SST is important and some analysis of the upper ocean is likely to be needed but to what precision is unknown. Timeliness of data transfer is likely to be even more acute for this. In as much as there is some predictability of the Intra Seasonal Oscillation (ISO), and this influences mid latitudes, there is some scope for monthly forecasts even in mid latitudes. At this stage, however, it is not clear what is possible or how useful such forecasts would be. Data requirements might be less wide-ranging than for seasonal forecasting, but more intensive. There might be a greater reliance on high quality forcing fields and surface data.

6.2 Longer range

There are two aspects here. The first is to capture the low frequency present in the structure of the ocean as it affects shorter-term processes such as El Nino. The requirements of data for this issue are basically that low-frequency changes in the depth and structure of the thermocline are captured and present in the initial conditions. Slow, large-scale changes in water mass properties may often have smaller amplitude than interannual variability, so there is an implication that the accuracy of data, particularly the bias, has more stringent requirements than those for sampling ENSO.

The second aspect is the possibility of extending forecasts from the seasonal to the decadal time scale. In this case we require knowing the ocean state to a depth appropriate to the processes being predicted. Much of the decadal signal may be contained within the upper 1000m of the ocean, considerably deeper than for the seasonal-prediction time scale, but not requiring full-depth observations. The degree to which this upper ocean can be shaped by the forcing fields is unknown, as is the impact of forcing field errors on those ocean initial conditions.

6.3 What are the data for?

In the above we have largely discussed the data in terms of providing initial conditions for forecasts. Initial conditions are certainly necessary. However, data are needed for more than that. They are needed for model development and validation. This is not an easy task as the model output is a function not just of the model physics, but also of the forcing fields of stress heat flux and fresh water flux. It is important that any observing system addresses the issue of forcing field error as well as the need for surface and subsurface ocean data: improvements in the forcing fields and model configuration might well lead to different assessments of what is needed from an ocean observing system. It is possible that the impact of data will also be a function of the scheme used to assimilate data. So 4D-var might give different results to OI. Each coupled model may have its own peculiarities and accept initial data in a different way. The best analysis and best initialized state may not be the same thing, although as models improve these converge to the same. Therefore, the observing system can not be fully evaluated on the basis of the analyses, but only on the basis of forecasts. Unfortunately, the forecast system is a function of the atmospheric model as well as the ocean model and the assimilation system, and even worse it is hard to generate statistically reliable results from the limited number of El Nino cases we have for testing, particularly in the “data rich” era in which our forecast systems now operate.

The state of ocean modelling has been hampered over the years, partly because of a shortage of data with which to compare the models, but more by an inaccurate specification of the forcing fields. These are still not adequate; probably even the winds are inadequate over much of the ocean. Certainly the combination of winds and ocean model reveals errors: either the model sensitivity to winds is wrong or the winds are wrong. In principle we have a means to address this in part by including models of the wave field or the barotropic response in the assessment process. These models do not depend on the heat and fresh water fluxes, thus isolating the assessment of wind, or more importantly, the transfer of momentum from the atmosphere to the ocean through the wave field. Both wave models and the barotropic response models will have sensitivities to wind which are different to those of the baroclinic response of the ocean used for seasonal forecasting. A proper test system requires much more effort than just running a single integration: wide ranging and detailed model/data comparisons are needed, and the availability of appropriate data is crucial.

6.4 Observing System Simulation Experiments (OSSEs)

There is a long history of performing OSSEs in meteorology, and more recently, Observing System Experiments (OSEs), the former ostensibly to be used to design an observing system, the latter to confirm its use once in place. There are serious limitations with these experiments. First some redundancy between parts of the observing system is highly desirable: without it, it is difficult to cross-validate the various components of the system. Secondly even though there might be considerable redundancy in the sense that parts of the system can be taken out without greatly degrading the system, this is not, per se, a valid argument for dropping part of the observing system. If the meteorological experience is anything to go by, it may well be that it is only on isolated occasions that parts of the observing system come into their own and prove to be important.

As an example of the difficulties of drawing conclusions from studies of this sort, we consider the paper of Carton et al (1996). On the face of it, this suggests that the TAO array is largely redundant in the presence of altimetry and the XBT network, but such a conclusion might be misleading. The study covered only a short period and was idealized in a number of ways, not least of which was the neglect of salinity. Thus altimetry was projected exclusively onto the thermal field. The TAO array data were spread over a 40-day window, defocusing much of the variability: no such defocusing was applied to the altimeter data, however. Further, at least from the perspective of seasonal forecasting, the wrong yardstick was used. If one was going to do an OSE then the final test is what impact there is on the forecasts. That might involve several hundred years of coupled integrations. It is not sufficient to evaluate an observing system purely on the basis of analyses, though this is undoubtedly a useful step on the way. The high cost of doing observing system experiments is not just a matter of computer resources, but also of manpower. There may well be some benefit from appropriate experimentation, but significant resources will be required, and the costs of specific studies may be difficult to justify if the interpretation of the final results is likely to remain ambiguous.

There are different systems for ENSO forecasting. Some models are structured so as to be dominated by only a few modes of the coupled system, and such systems are likely to have only modest data requirements. More complex forecast systems might be expected to need much more detailed data for initialization, since they purport to reproduce a much wider range of real-life variability. Some might argue that the extra complexity of the more realistic models is largely related to scales and types of variability that add little to predictability. There may perhaps be force in this argument with regard to the longer range of ENSO forecasts, perhaps in the 6-12 month range, although the spectre of decadal change and regime transitions must hang over any simplified forecast model. However, the real world is undoubtedly a lot more complex than the simpler models, and there is some evidence from the last El Nino that the more physically complete models tended to do better in the real-time forecasts. Regardless of one's beliefs in the merits of different modelling approaches, it is certainly true that the optimum balance of cost versus benefit of the degree of complexity for both model and initializing data has not been established. It would seem a counsel of despair, however, to decide in advance of trying that we had no hope of improving the forecasts by making them more realistic. We are still very much in the learning phase, and the continued availability of data is vital to allow this learning to take place.

For all of these reasons, we foresee only a limited role for OSSE-type experiments. We just cannot do the necessary experimentation to say whether this or that mooring in an array, or this or that XBT line is important, or whether this or that deployment should have salinity measurements. OSSEs are far too blunt instruments to give a reliable answer: common sense is a more practical way forward, as was used to design the TOGA observing system. OSEs based on impact on the analysis system might be useful, but great care is needed in interpreting the results of such experiments. We need experiments to learn about the way models and data work together, and we can hope to learn much about the general interactions and sensitivities. With our present capabilities, however, highly specific questions are hard to answer.

7. SUMMARY AND CONCLUSIONS

Our aim in this paper has been to develop a sense of the ocean data requirements for seasonal forecasting. It would be satisfying if we were able to produce hard and well-defended conclusions as to exactly which data were important, and where priorities should be given. As we have tried to show, this is not possible. Our knowledge is limited, as is our experience with initializing and using forecast models for El Nino. The work needed to give an adequate understanding of the interactions of possible data scenarios with differing forecast systems is not yet feasible. Quite apart from the limits on today's scientific knowledge and experience, there are three further important facts to be faced: (a) The data needed to make seasonal forecasts in ten years time will include the data we take today – if we underestimate what is required, it will hamper efforts to calibrate and validate systems for many, many years into the future; (b) our capabilities, especially with regard to ocean data assimilation, are likely to change relatively quickly, so we are aiming for a moving target; and (c) reliable experimentation on data requirements for real systems is always likely to be difficult and expensive. Together, these limitations mean that we need to take a physically based and common sense view of data requirements.

Given the caveats above, an attempt can still be made to outline what seem to be the most important ocean data requirements for seasonal forecasting in the near future.

1. SST. This is the single most important field for seasonal forecasting in its broadest sense. As a minimum the present quality of analysis needs to be maintained, but very probably an enhanced product is needed if seasonal forecasting is to fulfil its potential over the next one to two decades. An absolute accuracy of 0.1 deg C for weekly analyses is roughly what is needed. This level of accuracy will require careful definition of 'SST' – a nominal 2 metre depth bulk temperature is perhaps most appropriate.

2. Forcing fields, particularly wind stress. Wind stress is vital for initializing forecast models. Present day values, whether from NWP analysis systems or directly estimated from data, are less accurate than is needed to give a good ocean analysis. Theoretically there is a trade off between the quality of the ocean forcing and direct information on the ocean state, but in practice neither data stream is adequate on its own. The relationship between measured data, numerical models and techniques, and a final forcing field product is particularly complex, and so it is hard to make judgements on the data required. In situ data from TAO and other moorings are valuable both for present day wind analyses and for providing ground truth for future developments of assimilation systems. Scatterometry has good potential, and needs further work to develop its use in operational systems.

3. In-situ temperature and salinity. We need detailed information in the equatorial strip of the oceans, and a broad scale sampling of the upper ocean elsewhere. Salinity is needed as well as temperature, although the use of locally observed T/S relationships might allow less dense sampling of S in well-defined water masses. The equatorial requirements can probably only be met by fixed moorings, and the maintenance of TAO in the Pacific, plus some expansion of moorings in the equatorial Atlantic (PIRATA) and Indian Oceans are a high priority. The ARGO programme seems to be the most cost-effective means of providing global broad scale upper ocean sampling. Such sampling is vital both for accurately initializing the near-surface ocean and for providing the science base for understanding short term fluctuations in climate. Although we strongly support ARGO, we recognize the value of continuing XBT data, especially during the transition to ARGO and on well-established high density lines.

4. Altimeter. The big advantage of altimetry is its global coverage and high spatial and temporal resolution. It should be possible to combine this with broad scale in situ data in a way that adds real value to both data streams, by de-aliasing small scale signals on the one hand, and correctly interpreting the vertical structure on the other. A better knowledge of the geoid, and better techniques for assimilating the data into operational models, are needed to allow full benefit of the data.

5. Other data for model development and validation. Ocean data are needed not only for initialization of real-time forecasts, but also to help develop the models. Currents are particularly helpful, as well as field work aimed at better understanding mixing processes etc. More accurate surface forcing data, including

heat and moisture fluxes and turbulent inputs, are also needed for such studies. There is plenty of scope for model improvement, and the potential for substantial payback in terms of improved analyses and forecasts – our ability to use data is still relatively poor. Data needs for model development must not be neglected.

A final crucial point is that science needs to remain the key driver for seasonal forecasting for quite some years to come. Our knowledge is still poor, our techniques for using data still need considerable development. One day we might be able to design an optimally efficient long term operational observing system that will meet all of the needs of a mature seasonal forecasting system. It is likely to take us several decades, lots of data and much hard work to reach that point.

Acknowledgement

We are grateful to Magdalena Balmaseda and Frederic Vitart for providing Figure 1.

REFERENCES

Alves, J.O.S., K. Haines and D.L.T. Anderson, 1999: Sea level assimilation experiments in the tropical Pacific, J. Phys. Oceanogr., submitted.

Behringer, D.W., M. Ji and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: the ocean data assimilation system, Mon. Wea. Rev., 126, 1013-1021.

Carton, J.A., B.S. Giese, X. Cao and L. Miller, 1996: Impact of altimeter, thermistor, and expendable bathythermograph data on retrospective analyses of the tropical Pacific Ocean, J. Geophys. Res., 101, 14147-14159.

Cooper, N.S., 1988: The effect of salinity on tropical ocean models, J. Phys. Oceanogr., 18, 697-707.

Cooper, M. and K. Haines, 1996: Altimetric assimilation with water property conservation. J. Geophys. Res., 101, C1, 1059-1077.

Ji, M., R.W. Reynolds and D.W. Behringer, 1999: Use of TOPEX/POSEIDON sea level data for ocean analyses and ENSO prediction: some early results. J. Climate, submitted.

Lorenc, A.C., R.S. Bell and B. Macpherson, 1991: The Meteorological Office analysis correction data assimilation scheme. Q. J. Roy. Met. Soc., 117, 59-89.

Maes, C. and Behringer, D., 1999: Using satellite-derived sea level and temperature profiles for determining the salinity variability: a new approach, Geophys. Res. Letters, submitted.

Segschneider, J., D.L.T. Anderson and T. Stockdale, 1999: Towards the use of altimetry for operational seasonal forecasting, J. Climate, submitted.

Smith, N.R., J.E. Blomley and G. Meyers, 1991: A univariate statistical interpolation scheme for subsurface thermal analyses in the tropical oceans. Prog. Oceanogr., 28, 219-256.

Vossepoel, F.C, R.W. Reynolds and L. Miller, 1999: Use of sea level observations to estimate salinity variability in the tropical Pacific, J. Atm. Ocean. Techn., accepted.

Vossepoel, F.C. and D.W. Behringer, 1999: Impact of sea level asimilation on salinity variability in the western equatorial Pacific, J. Phys. Oceanogr., accepted.