Chapter 13

THE ECMWF PERSPECTIVE

David Anderson, Magdalena Balmaseda, and Arthur Vidard

ECMWF, Reading, UK

Abstract: ECMWF is involved in forecasting on time-scales from the medium range (days) to seasonal (6 months) ahead. This requires the preparation of initial conditions from which to start forecasts. For the medium range, a huge effort has been devoted to developing the most advanced assimilation strategies for analyzing the atmospheric state. Considerable effort has also been devoted to retrospectively analyzing the atmospheric state (ERA-40) using a more advanced assimilation system than was available at the time of the operational analysis. For forecasts beyond the medium range, coupled atmosphere-ocean models are used, requiring analyses of the ocean state. These in turn depend heavily on atmospheric analyses and reanalyses. Aspects of the atmospheric and oceanic assimilation systems used operationally are discussed, together with some limitations of current systems.

Keywords: ECMWF, forecasting, atmospheric analyses, reanalyses.

1. Introduction

In considering the merits of a data assimilation system, it is necessary to have the application in mind. That is after all what sets the measure by which one decides if the scheme works well or not. It is one thing to formulate an assimilation strategy theoretically but altogether a different matter to develop a practical implementation. No 'operational' assimilation system conforms to its theoretical optimum configuration.

At ECMWF (European Centre for Medium-Range Weather Forecasts), data assimilation is used in a number of applications. The key forecast applications are: deterministic forecasts for the medium range, ensemble prediction system (EPS) for the medium range, an ensemble of monthly forecasts and an ensemble of seasonal forecasts. The time ranges are 10 days for medium range, 31 days for monthly forecasts, 190 days for seasonal forecasts. As might be expected, the resolution of the model is not the same in all the applications. The model resolution decreases as one increases the

E. P. Chassignet and J. Verron (eds.), Ocean Weather Forecasting, 361-379.

forecast time or the ensemble size. Thus, for example, the deterministic forecast is made using a horizontal resolution of T511 (~40km), the EPS uses T255, the monthly uses T159, and the seasonal uses T95.

A further complication is that to make forecasts beyond a week or two requires information on the state of the ocean as well as that of the atmosphere. Forecasts of the monthly and seasonal range are made with coupled atmosphere-ocean models. The longer the range, the less important are the atmospheric initial conditions, but the more important the ocean initial conditions become (generally speaking). There might be a case for having an active ocean even for the short to medium range forecasts, as sea surface temperature changes associated with tropical cyclones and storm tracks might be important. This latter point requires confirmation, but regardless of whether the ocean impacts the atmosphere on the medium range, it is necessary to perform real-time ocean analyses for the monthly and seasonal forecast systems. At these longer forecast ranges, model error is sufficiently large that it cannot be ignored. One way to account for model error is to run the model for many realizations in the past to estimate the model climate and climate drift. This in turn brings in a need for extended analyses into the past as well as the present. ECMWF is thus heavily involved in reanalysis, both atmospheric and oceanic. It will be argued later that in fact reanalysis is an integral part of a forecast system, (at least for monthly and seasonal timescales and possibly also for medium range).

2. Atmospheric analyses

The atmospheric analysis is done using 4d-var (as discussed by F. Rabier in this volume). The 4 indicates the use of time as the 4th dimension. So the analysis is performed over a space-time box. It is performed over the highest resolution possible (this will be discussed later), producing an analysis which has the resolution of the first guess. The ECMWF model is formulated in spectral space rather than physical space i.e. in terms of spherical harmonics rather than grid points. In spectral space the resolution is T511 (about 20 km). The effective resolution is not this high, however. The T511 analysis is then used to provide initial conditions for the EPS, monthly and seasonal forecasts by truncating the analysis in spectral space to the appropriate resolution.

In addition to analyses made in near-real-time for the purpose of generating forecasts, analyses are made of past events as well, called reanalyses. Extended atmospheric reanalyses are major undertakings requiring years of dedicated work. So they are not undertaken lightly: in fact, at ECMWF, there have been only two to date. The first, denoted ERA-15 spanned the 15 years 1979-1993, and the second, denoted ERA-40, spanned the 40+ years Sept 1957- Aug 2002 (Uppala et al 2005). The analysis system used in these reanalyses should be the same throughout. For ERA-15, the scheme was OI, while for ERA-40 it was 3dvar. The expense in performing the reanalyses means that it is not possible to use the latest scheme, nor the highest resolution. So for ERA-40, the resolution was T159.

3dvar does not try to fit the model trajectory throughout the data time window. It shares with Optimal Interpolation (OI) the approach of performing the analysis at discrete times. Account is taken of the time however, in the sense that the observations are compared with the model first guess at the time of the observation, rather than bunching all the observations in a given data window to the central time. This procedure is called FGAT (First Guess at Appropriate Time). Formally, if all assumptions are the same, OI and 3d-var are equivalent. In practice this is never so.

The analyses from ERA-40 have been used as initial conditions for medium-range weather hindcasts (hindcasts are made as if they were forecasts but are made for past events). Results from these hindcasts have been discussed by F. Rabier (this volume). The improvement of the forecasts as the observing system changes can be clearly seen. The improvements in forecast skill since the end of ERA-40 are also shown. They can result partly from improved data assimilation procedures (4d-var cf 3d-var) and partly from improvements in the quantity and quality of data.

An interesting finding is that the skill of forecasts for the southern hemisphere now rivals or exceeds that for the northern hemisphere, whereas in the first decades of ERA-40 reanalyses, the southern hemisphere skill was relatively poor compared to that of the northern hemisphere. Since the in situ measurements for the southern hemisphere are sparse, the implication is that the improvement results from increased satellite coverage. Fig 1, from Kelly et al 2004, shows results of various Observing System Experiments (OSE's) which clearly indicate the importance of satellite data for the southern hemisphere: when satellite data are withheld, the forecast skill is much reduced. By contrast, there is a much smaller degradation in the northern hemisphere. Other components of the observing system have a smaller impact.

Although the use of reanalyses to provide initial conditions for hindcasts is informative, it is not the sole or even primary purpose of performing a reanalysis. The analyses are used for diagnostic work to understand physical processes in the atmosphere. (The physical assumptions/simplifications that went into the analysis are important here). The application of most use to climate modelers is that they provide an estimate of the surface wind, heat and fresh water fluxes. As we shall later see, these are crucial in developing a monthly or seasonal forecast system, as they provide a means of dealing with model error. For monthly forecasting, the atmospheric initial conditions from the reanalysis are more important than the fluxes used to generate the ocean analysis. As the forecast range increases this balance shifts and for seasonal forecasts the fluxes are more important than the atmospheric initial conditions, as they play a major role in determining the ocean initial conditions.

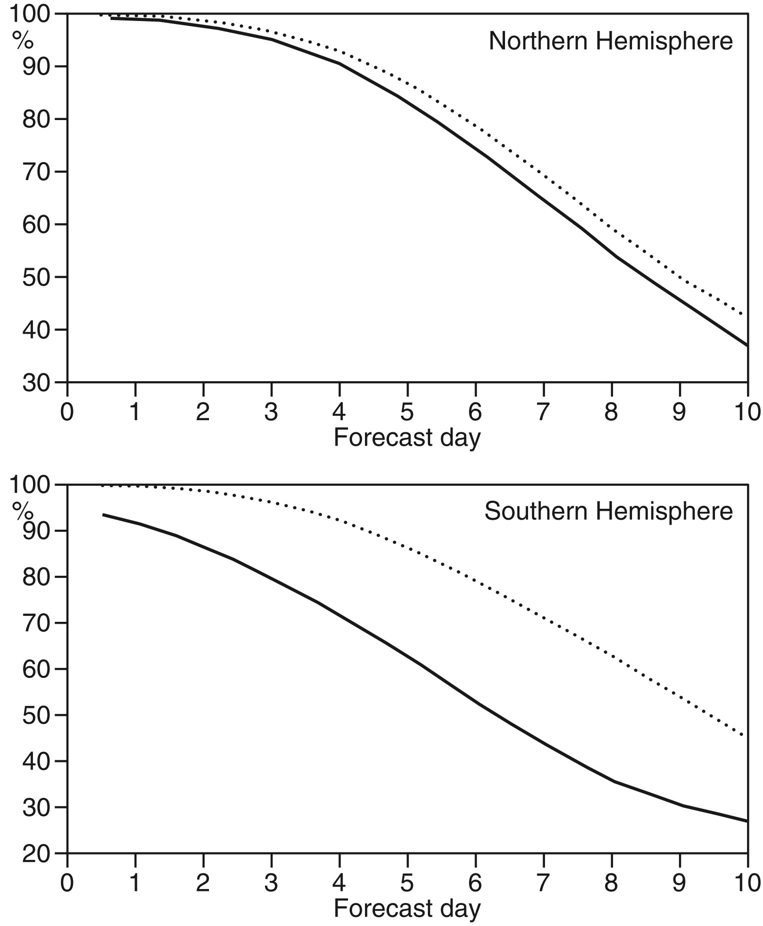

Figure 1. This plot shows the anomaly correlation between the predicted and analysed 500 hPa height field as a function of lead time out to 10 days for both the northern and southern hemispheres. The dotted line shows the skill when the full observing system is used: the solid line shows the skill when satellite data are withheld. Comparing upper and lower dotted curves shows that the skill of the SH is now commensurate with that of the NH. If one withdraws the satellite data, then the skill of the SH drops markedly whereas that for the NH drops only slightly. This OSE confirms the importance of satellite data, especially for the SH. From Kelly et al 2004.

The 4d-var assimilation system, seeks to minimise a cost function measuring the departure of the model trajectory from the data, subject to certain side-constraints, for example, that the departure from the first guess shouldn't be too large. The variables in the cost function are typically those governing the initial conditions. A data window (i.e. a time period over which the fit to data is minimized) is chosen. This is typically 6 to 12 hours for large-scale atmospheric models (Courtier et al 1994, Fisher 2005) but 30 days in the case of the ocean (Weaver et al 2003).

4d-var is cutting-edge technology and only a few operational weather forecasts centres have been able to adopt it. First was ECMWF, then Meteo France. The UK Met Office has recently introduced it. It is expensive, as one has to integrate the model forwards and its adjoint backwards through the data window (12 hours) of order 100 times. Simplifications are needed to make it feasible. First, an incremental approach is used (Courtier et al 1994). In this approach one assumes that only relatively small departures from the first guess will be made. The forward model is simplified, sometimes adjusting the physics to be more linear, so helping the derivation of the tangent linear (TL) and its adjoint. Although the TL can be a simplification of the forward model, the adjoint must normally be the exact adjoint of the tangent linear. The cost function, based on the tangent linear and its adjoint, is quadratic and convergence is faster than for a more general function. Further acceleration is achieved by using a lower resolution for the TL and its adjoint. At ECMWF, the resolution drops from T511 to T95 for the first 70 iterations of the inner loop. It is necessary to keep track of the full nonlinear model: so the trajectory is recalculated using the full nonlinear model at full resolution. This is followed by another 30 iterations of the inner loop at the slightly higher resolution of T159.

Because the ECMWF model is formulated in spectral space, rather than in gridpoint space, it is harder to impose covariances which vary geographically, e.g. to have different scales for the tropics to those for the extratropics and to have different degrees of isotropy. It can be done partially by using a variable such as vorticity, combined with linear balance constraints (Derber and Boutier 1998). Later, the implementation in an ocean context will be discussed and contrasted with the atmospheric case.

Although the same assimilation approach is used throughout the reanalysis, one can still get spurious low-frequency variability. This is because the observing system is not stationary. Although the number of observations generally increases as one approaches the present time, this increase is not monotonic. Some observing systems decline while other new systems come on stream.

An important component of any assimilation system is quality control: deciding which observations to keep and which to reject. In Rabier (op.cit), there is discussion of how to handle correlated observation error and how to thin potentially too-high-density observations such as those which can come from high-resolution satellites. Without getting into detail, as this is a difficult area (See Andersson and Jarvinen 1999), one counterintuitive example will be shown.

The late summer of 2004 has seen several intense hurricanes, several of which have caused substantial damage. When a hurricane is seen to be developing or tracking towards the Caribbean, additional meteorological observations are taken by flying aircraft into the storm and releasing dropsondes. These measure temperature, humidity and wind on their descent. In principle they should provide useful information on the location of the eye, of the central pressure and of the structure both of the near and far fields. In addition scatterometers can sense the near-surface winds. Fig 2 shows the location of the extra observations along the tracks of hurricanes Frances and Ivan. Grey indicates the measurements that were accepted, black those which were rejected. One can see that almost all the measurements close to the centre of the hurricanes were rejected. As indicated by central pressure differences, there were big discrepancies between the model estimates and the measurements which was why the data were rejected.

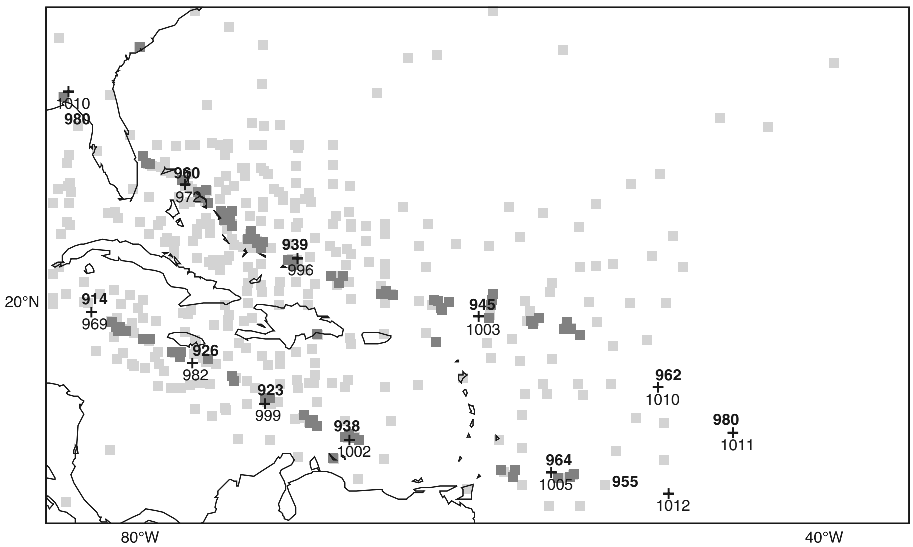

Figure 2. Dropsonde coverage for two tropical storms, Hurricanes Frances and Ivan in August September 2004. Bold numbers indicate the observed central pressure, normal numbers the analysed central pressure. Black squares indicate wind observations rejected by the analysis system as departing too far from the first guess. Light squares indicate observations which were accepted. Rejected data are mainly those close to the centres of the two storms. From Federico Grazzini, ECMWF.

Of course it is not clear that we are comparing like with like. The observation is a spot measurement, representative of only a small area of the storm. The model on the other hand represents an area average that is much larger. At best the grid is 20 km square but this is not the real resolution of the analysis system. First, the inner loops are run at T95/T159 (~180-120km). Secondly, the effective resolution of the analysis is considerably less than that as the structure functions are broader. In principle one could have adaptive structure functions with smaller scales around a storm but this is not done.

Often the analysis actually weakens the first guess. The model, when run forward, has a resolution of T511 and can develop i.e. intensify a tropical storm over the next 12 hours. When the analysis kicks in, with its lower effective resolution, the storm can get weakened. So although the observations say deepen this storm significantly, the analyses says fill it in. This conflict in analysing small-scale, but important, features is generic.

3. The ocean analysis

3.1 Introduction

Ocean analyses are needed to provide initial conditions for both monthly and seasonal forecasts. For the monthly and seasonal forecast systems we use essentially the same ocean analysis systems. They are not exactly the same, as the monthly system needs a faster analysis than the seasonal and also has different cut-off times for the receipt of data. The resolution of the ocean model is 1 degree except near the equator where it is 1/3 degree in the meridional direction to better resolve the equatorial waves which are important in processes such as El Nino. In future we will explore the need for even higher resolution models coupled to the medium range weather forecast model, with potential impact on tropical cyclone prediction, but that will not be discussed further here.

3.2 The observing system

Ocean observations are mainly of thermal data and mainly in the upper 500m. Some salinity measurements are now available from ARGO but only in the last few years. Other measurements of salinity e.g. from CTD are sparse and have not generally been available in real-time. There are almost no measurements of velocity, except in the surface layer. Fig 3 shows the thermal data coverage in a typical 10-day window. Several features are worthy of note. First, there is quite a substantial coverage in real-time. In fact most data are now received within a day or so at operational centres such as ECMWF through a network called the GTS (Global Telecommunications System). When data are received, each individual observation is checked. It is compared against the model first guess and also with an analysis performed without the datum being checked. The actual quality control is quite complex and will not be gone into in detail. See Smith et al, 1991 for further discussion. Data profiles on fig 3 which are rejected are in grey. These are mainly located near the coast since coastal points are not well represented in the ocean model. Some profiles are partially accepted; at some depth the model and data differ by too great an amount and the datum is rejected but data above and below this level might be accepted.

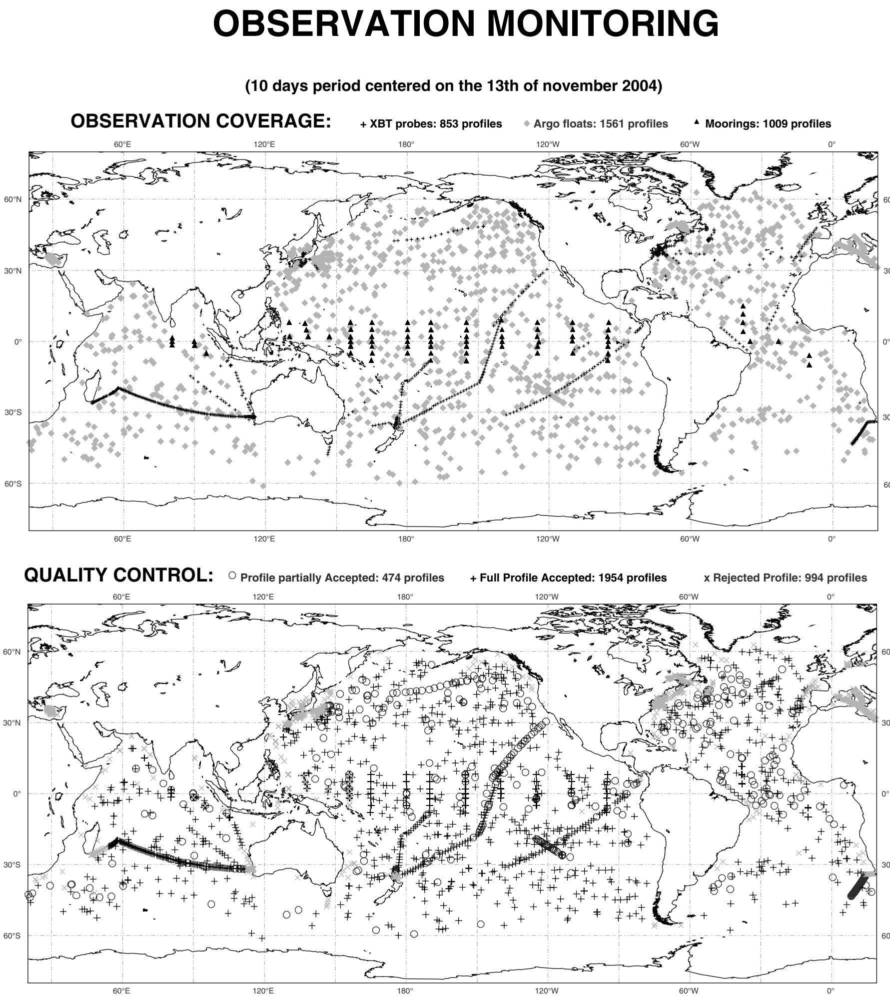

Figure 3. a) Data coverage for a 10 day period 9th-18th Nov 2004 inc. showing the TRITON/TAO/PIRATA mooring array in the Pacific/Atlantic oceans (Black triangles), the XBT network (black crosses) and the ARGO floats (grey diamonds). The moorings report daily in the Pacific to the east of the dateline and in the Atlantic. Hourly reports are received from moorings west of the dateline. b) This figure shows the profile data which have been fully accepted (black crosses), partly accepted (black circles) and completely rejected (grey crosses).

The array of triangles in the tropical Pacific and Atlantic oceans, seen most clearly in fig 3 is the TRITON/TAO/PIRATA. These report daily-mean values, so in a 10-day period one would expect ~10 measurements. For the TRITON array in the west Pacific the reporting is hourly. The straight or slightly curved lines of data are from merchant ships (called VOS or Voluntary Observing Ships) making XBTs measurements. XBTs are instruments which measure temperature to a depth of ~500m. A few measure salinity. The diamonds are ARGO float measurements from buoys which drift at ~1000m, but every 10 days pop up to the surface, measuring temperature and often salinity, and relay the information via satellite to a ground station where it is put on the GTS. The mooring and XBT data are also available in real-time. If there are many data points in close proximity in space and time, they can be 'super-obbed', a process by which they are combined into one 'super-observation' and given an increased weight. See Smith et al 1991 for discussion on super-obbing and quality control issues.

3.3 The ocean analysis system

The scheme currently in operational use is OI (Optimal Interpolation) using a time window of 10 days. FGAT has not been implemented in this scheme: rather all observations in the 10 day window are taken to apply at the centre of the window. It differs from a standard OI, however, in that the correction to the FG following an OI is not applied instantaneously; rather the increment is divided by the number of timesteps in a 10 day window and then that fraction of the increment is applied every timestep. The idea behind this was to allow the model to generate its own circulation field following an OI in which only thermal data were assimilated. Although this worked to some degree, a better circulation field can be produced. In principle, one can update the velocity field and salinity field even though only temperature data are being assimilated. This can be done through having multi-variate covariances. However, at the time of implementation we did not have these well tuned. Rather, we sought to improve the univariate assimilation of T by building in some corrections, done in physical space rather than through multi-variate covariances. So a geostrophic velocity field is calculated following the T-assimilation and then the velocity increments are also spread over 10 days. See Burgers et al 2002 and Balmaseda 2003 for more details.

Salinity is also corrected by applying an S correction such that the T-S relationship is preserved. In the ocean when T changes, so does S in such a way as to preserve the T-S relationship. There are regions in which this approximation doesn't work- for example in the surface layers where heat and fresh-water fluxes will change T-S. The true T-S relationship is not known at any time, so we use the model T-S. Unfortunately the model T-S can drift from truth; for example, if the model has too much or too little mixing or if the precipitation is wrong and the surface layer starts to interact with the layers beneath. Nonetheless, although not a perfect solution, it does give beneficial impact, especially in the equatorial Atlantic and Pacific. See Troccoli et al 2002 for further details on the strategy and its impacts.

A novel feature of the ocean analysis system is that not just a single analysis but multiple analyses are performed. The purpose of the analysis is to provide initial conditions for monthly and seasonal forecasts. Such forecasts must be probabilistic. This implies that an ensemble of forecasts must be made. In the case of the monthly forecast system, the ensemble size is 51, while for the seasonal forecast system, it is 40. The ensemble is there to sample uncertainty arising from the chaotic nature of the atmosphere. However, it should also take into account uncertainty in the ocean initial conditions. One method of representing this uncertainty is through running an ensemble of ocean analyses. In our case the ensemble size is 5. This ensemble is not to be confused with the EnKF in which the size might be ~100. Experiments are underway to assess the EnKF strategy as part of the EU project ENACT but results will not be presented here.

In the case of the atmosphere, almost all the information on which an analysis is based comes from observations of the atmosphere. In the case of the ocean, a substantial amount of information on the ocean state can be obtained not through ocean observations but through atmospheric observations - in fact all the observations that are involved directly or indirectly in defining the surface wind, heat and fresh water fluxes. For seasonal forecasting the most important of the surface forcings is the wind. The wind field that is used to force the ocean has uncertainty. We estimate that uncertainty and then force 5 ocean analysis streams with wind fields that are perturbed commensurate with the estimated uncertainty in the wind. In addition, the SST field is not known sufficiently accurately either. So perturbations to it are also applied. This is discussed more fully in Vialard et al 2005, who show the spread generated by different ensemble generation strategies.

Although the data coverage might look reasonable in fig 3 where the symbols are quite large, it is probably barely adequate even for today. Ten or twenty years ago the coverage was much worse. It is thanks in large measure to a major international programme called TOGA (Tropical Ocean Global Atmosphere) that the real-time coverage is as good as it is. Starting in 1985, this programme steered oceanography towards a free exchange of data in near real-time.

3.4 The value of data assimilation

There are surprisingly few clean sets of experiments to show that data assimilation improves the skill of seasonal forecasts. Any results probably apply only to the system being tested, as improvements in either the ocean model or the forcing fields through improved atmospheric analyses or reanalyses, could change the results. One clean set of experiments was performed by Alves et al 2004. Four sets of analyses were performed and four ensembles of forecasts were run from these analyses. Two different wind products were used and for each, experiments with and without data assimilation were performed. Fig 4 shows the growth of error (upper panel) and the anomaly correlation (lower panel) from these forecasts in the tropical Pacific, a key region for seasonal forecasts. The forecasts based on analyses with data assimilation are clearly better than those without: the rms error is smaller and the anomaly correlation is higher. The lower curves on fig 4a give a measure of the spread of the ensemble.

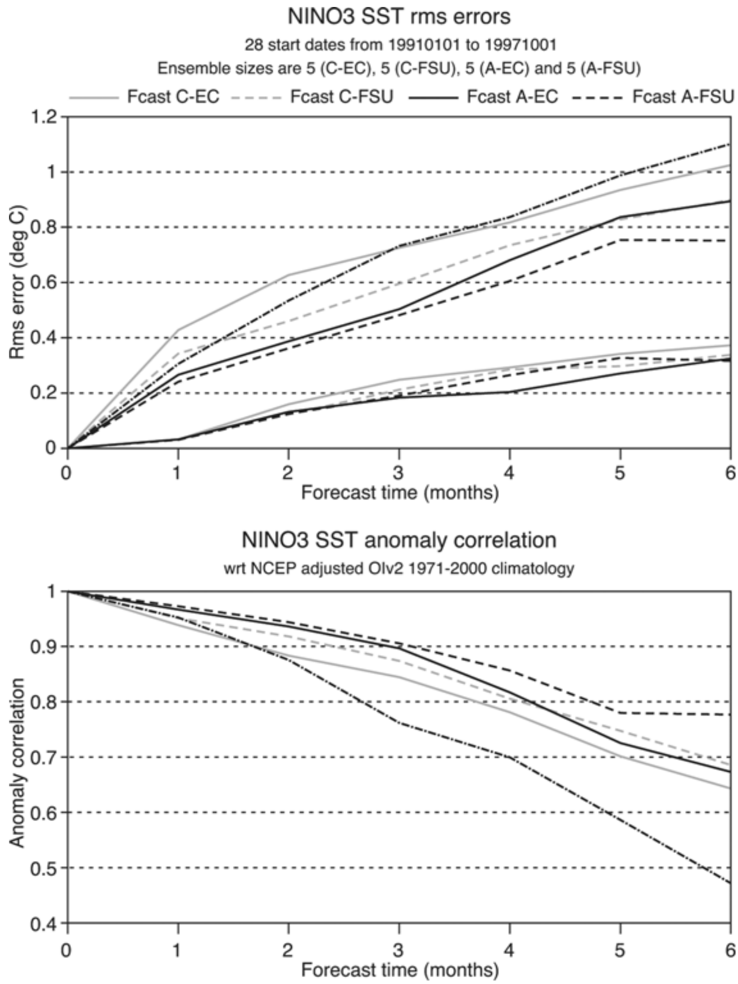

Figure 4. Upper panel: RMS error as a function of forecast lead time out to 6 months for the region NINO3 in the central east equatorial Pacific Ocean (upper set of curves). The lower set of curves on this panel shows the spread in the ensemble of forecasts. Lower panel: The anomaly correlation. Results for four experiments are shown. Two use ocean conditions with data assimilation (denoted A-); two do not assimilate sub-surface ocean data (denoted C-). Two different wind fields (FSU and ECMWF) are used to force the ocean during the assimilation process. This figure shows that the two runs with data assimilation have higher skill (smaller rms error and higher correlation) than those without, and that differences arising from differences in the wind field are reduced in the case of data assimilation. The dash dotted curve indicates the skill of persistence. From Alves et al 2004.

Vialard et al 2005 also show the growth of error in the forecasts. As for fig 4 which used earlier versions of both the ocean and atmospheric analysis systems and the coupled forecast model than that used by Vialard et al, the spread is considerably smaller than the rms error, indicating a problem in the analysis/forecast system. At this stage we do not know if this is due to the analysis or the forecast model. One way of interpreting the difference in the growth of error versus the growth of spread in the ensemble is that the forecasts are too confident, indicating that all the uncertainty in the forecasts is not being accounted for. There is another more optimistic interpretation: that the spread is the theoretical measure of predictability. If the model error were small and the initial conditions were correct, then this is how the forecast error should grow. This estimate can be model dependent so it is not a hard argument. In practice, it is likely that by improving the models and the initial conditions, the error growth can be reduced and by improving the ensemble generation the spread can be increased. At ECMWF we are developing a multi-model forecasting facility in collaboration with the UK and French Met Offices and in this multi-model system the separation between error growth and spread is reduced. See also Palmer et al 2004.

3.5 Problems with the winds

In section 1, we considered the importance of atmospheric reanalyses. There we showed that the medium-range forecasts from say 20 years ago are now much better than they were then, partly because of improvements in the current analysis system compared with what was done then. That is encouraging but is not particularly useful in its own right, unless one wants to use these past forecasts for calibration. Calibration on past events has not really taken hold in the medium range community, though some moves in this direction are afoot (Lalaurette 2003). But calibration on past events is a major feature of seasonal and monthly forecast systems. This is described in Stockdale et al 1998. When making forecasts to a few days ahead, there are plenty of cases on which to test the model forecasts. When making forecasts to 6 months ahead, there are very few cases of events such as El Nino that can be tested in real-time. To evaluate such a system, one has to go back in time and to make hindcasts from as far back as one can reasonably go. As mentioned earlier, there are insufficient ocean observations to make ocean analyses directly. But by using the forcing fields from the atmospheric reanalyses one can produce ocean initial conditions back say 15 or 20 years. There is a further reason for using these past hindcasts. All models have errors. For forecasts out to a few days, these have largely been ignored though realization that the model would benefit from calibration on past events is growing. For monthly and seasonal forecasts, the error is sufficiently large that it can not be ignored. To first order it is estimated from the past hindcasts and this information is used to calibrate the real-time forecasts. The atmospheric reanalyses are therefore very important in enabling ocean reanalyses to be performed. It seems that the ERA40 reanalyses are considerably better than the earlier ERA-15 reanalyses in that the ocean reanalyses using ERA-40 match the independent data set of sea-level from satellite altimeters such as TOPEX and Poseidon more closely and lead to better forecasts.

In fig 5, we show the zonal wind stress anomalies in the equatorial Atlantic averaged between 5S to 5N from both ERA15 and ERA40 as a function of time from 1987 to 2002 (ERA15 ended in 1993 so from then on we use the operational equivalent). Potential improvements in the assimilation system will therefore be present in the post 1992 era for ERA15. In contrast, ERA40 used the same 3d-var system throughout. One can see that the winds are substantially different between the two products and even post 2000 the differences are large even though both are using good (though not the same) assimilation methods and models. It is also clear that there were major differences in 1996 for reasons unknown, but possibly the ERA15/OPS product is better than ERA40 in this case.

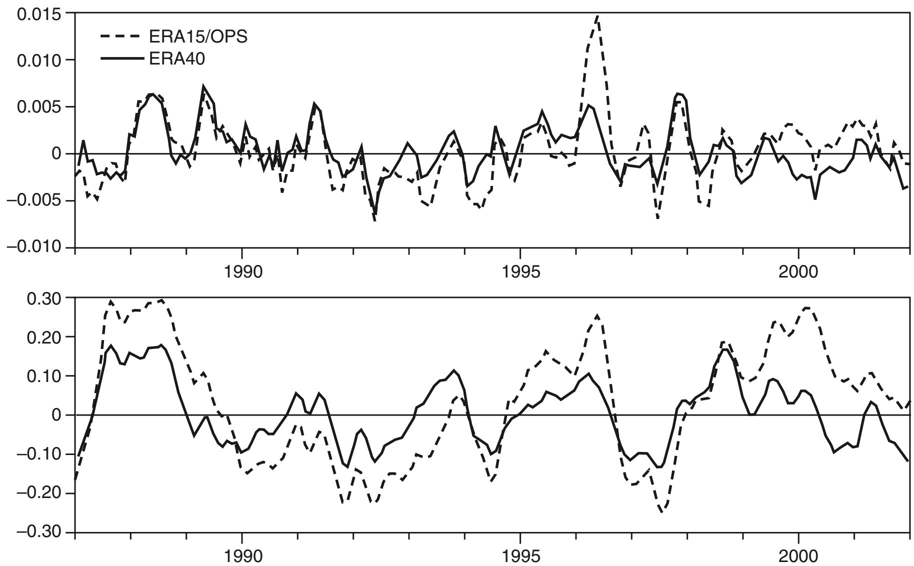

Figure 5. Upper panel: Plot of the wind stress anomalies in the equatorial Atlantic averaged from 5S to 5N. Two different reanalysis products are shown: ERA-40, solid line and ERA-15 dashed line. These wind fields differ considerably throughout the period but the differences are especially large in 1996 for reasons that are unknown. Lower panel shows the average temperature anomalies in the upper 300m for the same region. The ocean acts as a filter and integrator of noise. Thus the signal is redder than that of the wind. The differences are large throughout the integration, and not just in 1996.

What effect do these different winds have on the ocean? The lower panel shows the effect on the temperature averaged vertically over the top 300m and between 5S and 5N in the equatorial Atlantic. One can see that the effect is substantial; in fact the difference in the curves is nearly as large as the interannual signal itself. The ocean can act as an integrator of noise and in fact the signal is quite large in the late 80's early 90's even though this is not the time of greatest wind error. Data assimilation acts to reduce these differences (not shown). Further illustration of the extent to which data assimilation can act to reduce the impact of wind error is shown in Vialard et al 2005.

4. Weaknesses in the ocean assimilation strategy

The hypothesis underlying the assimilation strategy is that the system is unbiased. This hypothesis is definitely not true in the case of ocean data assimilation. This can be seen by evaluating the mean increment applied in the assimilation. Fig 6 shows this for a section along the equator in the upper 400m. The upper panel shows the mean increment in temperature, averaged over eight years. Far from this being zero, one can see that the assimilation acts to warm the ocean in the west and to cool it in the east Pacific i.e. to strengthen the gradient along the equator. There is a strong systematic effect in the Atlantic too but not much in the Indian ocean (which might just reflect the fact that there are few observations in the Indian ocean).

If one makes the velocity correction mentioned earlier, this bias is reduced somewhat. Including a salinity correction also acts to reduce the bias. Nonetheless, regardless of these changes the bias remains substantial. It could result from error in the wind, in the ocean model physics, in the way that momentum is transferred from atmosphere to ocean, or in the assimilation system itself. Whatever the reason, the assimilation system will operate at reduced efficiency since it is not designed to deal with bias. In fact most assimilation schemes assume the system is unbiased. Can we somehow adapt the assimilation system to take account of the bias? This has been considered by Dee and Da Silva 1998, Dee and Todling 2000, Vidard et al 2004, and Bell et al 2004. One approach is to correct the pressure gradient as suggested by Bell et al 2004. One might think that it would be better to correct the temperature since that is the field which seems to be biased, but this is not the case. The vertical velocity along the equator is distorted (not shown). It is not possible to show what this field should really look like since it is hard to measure, but a descending circulation in the east Pacific is generated which looks very unlikely. It is thought that this arises from the assimilation cycle itself. (See also Vialard et al 2003, Huddleston et al 2004). Correcting bias in the pressure gradient greatly reduces this spurious circulation and the mean temperature increment (not shown). On the other hand correcting the bias in T actually aggravates the situation and leads to an enhanced spurious vertical circulation, although the T increment is reduced.

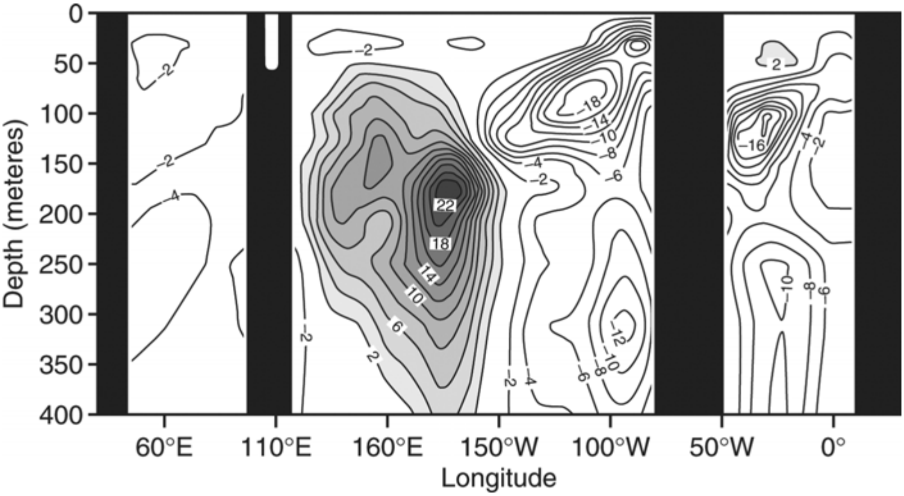

Figure 6. Assimilation increment averaged over an 8-year period in a vertical section along the equator. The left panel is the Indian ocean, the middle the Pacific and the right is the Atlantic. In a well-balanced system the mean increment should be close to zero. In practice most ocean assimilation systems have considerable bias, as shown here. The assimilation is acting to warm the west Pacific but to cool the east. Some of the bias is caused by the assimilation system itself. Warming regions, i.e. where the assimilation increment is positive, are shaded. Contour interval is 2K/year.

5. **Ocean observing system experiments**

As funding is always limited, the question of the relative merit of each observational system arises. This can be estimated through observation system experiments (OSEs), well known to meteorologists and described in fig 1. See Daley 1992 for general discussion and Anderson et al 1991 and Isaksen and Stoffelen 2000 for discussion relevant to surface wind field from scatterometers. In oceanography, this is a relatively new field, as observations have always been sparse.

The fairest way to evaluate the potential of an observing system is to selectively withdraw components of the system starting with the full array. So we take the system with TAO/PIRATA/TRITON moorings, the XBT network and the ARGO array as the standard, and then remove, TAO/PIRATA/TRITON. The impact of this component of the observing system can then be assessed. The XBTs can be assessed by removing them and comparing with the standard system. This strategy is different to starting with a zero system and then adding components. The approach used here allows for redundancy. So, it might be that a component of the array can be withdrawn at little overall effect because more or less the same information is available from another component of the observing system. Some redundancy is not necessarily bad, however, as it allows an evaluation of the different parts of the array. It should also be remembered that even though there is some redundancy in some average sense, there might be occasions when this is not so. There are good examples of this in weather forecasting though we are not aware of this in ocean analyses for climate forecasting. Evaluation of observing systems through their impact on forecasts is standard practice in meteorology and should generally be so in oceanography, even if it has not been so in the past.

It is important to realise that results from OSEs are dependent on the analysis system used and on the weight given to the data. The results will be application-dependent. We are interested mainly in seasonal forecasts; this emphasizes the tropics over middle latitudes. For other forecast horizons, or other objectives, different areas may be important and different conclusions might be drawn.

Various experiments to assess the impact of the TAO and XBT networks are given in Vidard et al 2005. To assess the importance of the observing systems on forecasts, 200 six-month forecasts are made, spanning the period 1993-2002 using the ocean analyses as initial conditions. The importance of the equatorial moorings is demonstrated. The exciting new aspect of the observing system is the development of the ARGO network. Can this system replace the XBT network? To consider this, the impact that can be derived from ARGO is discussed. The ARGO experiments cover only the two years 2002-3, but show considerable impact even when the array is only partially developed and optimum techniques for using the data have not been developed (for example, salinity data from the ARGO floats are not yet assimilated).

In addition to OSEs, one can conduct Observing System Simulation Experiments (OSSEs) to assess the potential impact of a proposed observing system, or to assess the relative merits of a given array design. It is difficult to gauge the error characteristics of such observing systems and the experience from meteorology is that results are often too optimistic because errors or difficulties in using the data are underestimated.

6. Summary and conclusions

In this paper we have considered data assimilation methods and issues arising, as related to operational analysis of the atmosphere and ocean. At ECMWF, the primary purpose is, and has been, to provide initial conditions for various forecasts made. These range from forecasts out to 10 days at resolution of ~20km to seasonal forecasts at atmospheric resolution of ~200km. Forecasts for the monthly and seasonal timescales are made with coupled atmosphere-ocean models and so require initial conditions for the coupled system. In real-time, these are made by taking the atmospheric analysis performed for the highest resolution forecast and truncating to the appropriate resolution needed for the forecast. In the case of the ocean, they are obtained by running the ocean analysis system, assimilating all in situ thermal data and relaxing strongly to observed sea surface temperature. The ocean analysis systems for the monthly and seasonal forecast systems are essentially the same, though they do differ slightly as the analysis for the monthly forecast has to be available within one day of real-time, whereas the seasonal forecasts start 11 days behind real-time. The analysis for the monthly system starts from the last analysis for the seasonal system but then is accelerated to real time using the surface forcing from the atmospheric analyses and assimilating what ocean data are available.

Because model error is significant at longer forecast range, it is necessary to take this into account when preparing forecast products. This is done by calculating the model drift and climate over a calibration period which is currently 15 years but will be longer in future. The forecast products are anomalies relative to the model climate. Hindcasts over the period 1987-2004 are possible, largely because ocean forcing fields are available from the reanalyses ERA-15, ERA-40. These atmospheric reanalyses are not without their problems, as shown in the paper, but mark a significant improvement over previously available products, such as the analyses carried out at the time. It is hoped that atmospheric reanalyses will continue to be performed, perhaps every 5 to 10 years, and covering perhaps the last 50 years or so. These are major undertakings using advances in model development and assimilation techniques and some recovery of old data. They should be viewed as part of the effort to produce monthly, seasonal and even multi-annual forecasts. Ocean reanalyses spanning the period of the atmospheric reanalyses are also made on a routine basis. These are currently done every year or so but as model resolution increases and/or assimilation techniques become more sophisticated, they will be undertaken less frequently. Currently, several ocean groups are active in this area, in contrast to meteorology where only two or three are performed.

At present it is unclear how best to initialize the coupled model. Initialising the atmosphere and ocean separately, as is done currently, may not be the best. One would like to do a more coupled assimilation but this is some way off at present. It is not straight forward because of the disparate time scales of the atmosphere and ocean.

The ocean observing system has advanced rapidly over the last decade. In the tropics it is clear that there are systematic differences between the ocean model state and the observations. This could be because of deficient forcing fields, ocean model physics or the way the surface fluxes are transferred from atmosphere to ocean. ECMWF uses a wave model as part of all its forecast systems, but this model is not fully tied into the ocean as yet. Most other groups using coupled models do not even include a wave model and do not pass the fluxes through the wave field. The assimilation system itself may contribute to the bias increments noted here. Finally, we are just beginning to evaluate the observing system and full use is not yet made of all the data available. For example, salinity from the ARGO floats is only just beginning to be used (Haines et al 2005). The combination of an expanding ocean observing system, new strategies for assimilation, different techniques for initializing coupled models combined with better models suggest a busy time ahead.

Acknowledgements

The authors thank the members of the ECMWF analysis and medium range and seasonal forecast groups for constructive discussions, and Graeme Kelly and Federico Grazzini for providing figs 1 and 2 respectively. This work was supported in part by ENACT EU-contract EVK2-2001-00077.

References

Anderson D.L.T., A. Hollingsworth, S. Uppala, and P. Worceshyn, 1991: A study of the use of scatterometer data in the ECMWF Operational Analysis - Forecast model. Part II Data impact. J. Geophys. Res., 96, C2, 2635-2649.

Andersson E., and H. Jarvinen, 1999: Variational quality control. Quart. J. Roy. Met. Soc., 125, 697-722.

Alves O., M.A. Balmaseda, D.L.T. Anderson, and T. Stockdale, 2004: Sensitivity of dynamical seasonal forecasts to ocean initial conditions. Quart. J. Roy. Met. Soc., 130, 647-668.

Balmaseda M.A., 2003: Ocean data assimilation for seasonal forecasts. In Recent developments in data assimilation for atmosphere and ocean. Seminar Proceedings ECMWF.

Bell M., M.J. Martin, and N.K.Nichols, 2004: Assimilation of data into an ocean model with systematic errors near the equator. Quart. J. Roy. Met. Soc., 130, 873-893.

Burgers G., M.A. Balmaseda, F. Vossepoel, G.J van Oldenburgh, and P.J. van Leeuwen, 2002: Balanced ocean data assimilation near the equator. J. Phys. Ocean., 32, 2509-2519.

Courtier P., J-N Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4DVAR, using an incremental approach. Quart. J. Roy. Met. Soc., 120, 1367-1388.

Daley R., 1991: *Atmospheric Data Analysis*, Cambridge University Press, 457pp

Dee D.P., and A. Da Silva, 1998: Data assimilation in the presence of forecast bias. Quart. J. Roy. Met. Soc., 124, A, 269-295.

Dee D.P., and R Todling, 2000: Data assimilation in the presence of forecast bias: the GEOS moisture analysis. Mon. Wea. Rev., 128, 3268-3282.

Derber J., and F. Bouttier, 1998: A reformulation of the background error covariance in the ECMWF global data assimilation system. Tellus 51A, 195-221.

Fisher M., 2005: On the equivalence between Kalman Smoothing and weak-constraint four-dimensional variational assimilation. Quart. J. Roy. Met. Soc., to appear. See also ECMWF Tech Memo 447 available on line.

Haines K., J. Blower, J-P Drecourt, C. Liu, A. Vidard, I. Astin, and X. Zhou, 2005: Salinity assimilation using S(T) relationships. Part 1: theory. Mon Wea. Rev., to appear.

Huddleston M.R., M.J. Bell, M.J. Martin, and N.K.Nichols, 2004: Assessment of wind stress errors using bias-corrected ocean data assimilation. Quart. J. Roy. Met. Soc., 130, 853-871.

Isaksen L., and A.C.M. Stoffelen, 2000: ERS scatterometer wind data impact on ECMWF's tropical cyclone forecasts. IEEE Trans. Geosci. Remote Sens., 38, 4, 1,885-1,892.

Kelly G., 2004: The proceedings of the third WMO Workshop on the Impact of Various Observing Systems on Numerical Weather Prediction (Alpabach, Austria, 9-12 March 2004).

Lalaurette F., 2003: Early detection of abnormal weather conditions using a probabilistic extreme forecast index. Quart J. Roy. Met. Soc., 129, 3037-3057.

Palmer T., F.J. Doblas-Reyes, R. Hagedorn, and 22 others, 2004: Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER). BAMS, 85, 853-872.

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.

Stockdale T., D.L.T. Anderson, M.A. Balmaseda, and O. Alves, 1998: Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. Nature 392, 370-373.

Troccoli A., M. Balmaseda, J. Segschneider, J. Vialard, D. Anderson, K. Haines, T. Stockdale, F. Vitart, and A Fox, 2002: Salinity adjustments in the presence of temperature data assimilation. Mon. Wea. Rev., 130, 89-102.

Uppala S.M., P. Kallberg, A.J. Simmons, and forty-three others, 2005: The ERA-40 reanalysis. Submitted. See also ECMWF Newsletter 101. Available from ECMWF or from the web: http://www.ecmwf.int/publications/newsletters/

Vialard J., A. Weaver, D.L.T. Anderson, and P. Delecluse, 2003: Three and four dimensional variational data assimilation with a general circulation model of the tropical Pacific ocean. Part 2: Physical validation. Mon. Wea. Rev., 131, 1379-1395.

Vialard J., F. Vitart, M.A. Balmaseda, T. Stockdale, and D.L.T. Anderson, 2005: An ensemble generation method for seasonal forecasting with an ocean-atmosphere coupled model. Mon. Wea. Rev., in press.

Vidard A., A. Piacentini, and F.-X. Le Dimet, 2004: Variational data analysis with control of the forecast bias. Tellus A, 56, 177-188.

Vidard A., D.L.T. Anderson, and M.A. Balmaseda, 2005: Impact of ocean observation systems on seasonal forecasts. ECMWF Techical Memorandum 460.

Weaver A., J. Vialard, and D.L.T. Anderson, 2003: Three and four dimensional variational data assimilation with a general circulation model of the tropical Pacific ocean. Part 1: Formulation, internal diagnostics and consistency checks. Mon. Wea. Rev., 131, 1379-1395.