

# FORECAST DIVERGENCES OF A GLOBAL WAVE MODEL

D. J. M. Greenslade<sup>1</sup>, I. R. Young<sup>2</sup>

<sup>1</sup>Bureau of Meteorology Research Centre, Melbourne, Australia

<sup>2</sup>Swinburne University of Technology, Melbourne, Australia

## 1. INTRODUCTION

Several operational Numerical Weather Prediction (NWP) centres around the world run wave forecasting models that routinely assimilate satellite wave observations (e.g. Breivik and Reistad, 1994; NMOC, 1999). There has been a considerable amount of research into the development of wave data assimilation systems in recent years (e.g. Lionello *et al.*, 1992; Greenslade, 2001; Mastenbroek *et al.*, 1994). A major limitation to current assimilation systems is the specification of the model (or background) errors. This has not previously been explored to any great extent for wave models.

One method that may be used to investigate the structure of the background errors is the NMC method (Parrish and Derber, 1992; Rabier *et al.*, 1998). This method uses forecast divergence as a proxy for background error growth. Correlations and variances of the differences between forecasts of different ranges valid at the same time are determined and can be used to construct the background error covariances. In this work, forecasts of Significant Wave Height (SWH) from a global wave model are investigated. A general assumption is that most of the errors occurring in wave forecasts are due to errors in the surface winds (Cardone *et al.*, 1995). Therefore, the surface wind fields that are used to force the wave model are also examined.

A brief review of previous wave data assimilation research is given in Section 2 with an emphasis on specification of the wave model background errors. Previous work using the NMC method is also presented here. The wave model used in this work is a version of the WAM model (WAMDI Group, 1988; Komen *et al.*, 1994). This is described in Section 3. The details of the method used to calculate the correlations and variances are presented in Section 4. Results for isotropic and anisotropic correlations are presented and discussed in Section 5. Further issues are addressed in Section 6 and finally Section 7 provides a summary.

## 2. BACKGROUND

Current operational wave data assimilation systems at, for example, the European Centre for Medium-Range Weather Forecasts (ECMWF) (Lionello *et al.*, 1992) and the Australian Bureau of Meteorology (the Bureau) (Greenslade, 2001) use the sequential method of Statistical Interpolation (SI) to combine first-guess wave model fields with the observations to obtain analysed wave fields. Details of the SI algorithm can be found in Lionello *et al.* (1992) or Greenslade (2001). One of the requirements in the application of SI techniques is the specification of the model error correlation matrix,  $P$ . This is a symmetric  $N_{obs}$  by  $N_{obs}$  matrix ( $N_{obs}$  is the number of observations) whose element  $(k,j)$  is given by

$$P_{kj} = \frac{\langle (H_p^k - T^k)(H_p^j - T^j) \rangle}{\sigma_p^k \sigma_p^j} \quad (1)$$

where  $H_p$  is the model first-guess,  $T$  is the true field,  $\langle \dots \rangle$  is the expected value, and  $\sigma_p$  is the model prediction *rms* error, i.e.,

$$\sigma_p = \sqrt{\langle (H_p^j - T^j)^2 \rangle} \quad (2)$$

In other words, the value of element  $(k,j)$  of matrix  $P$  is the correlation between the model error at observation location  $k$  and the model error at observation location  $j$ .

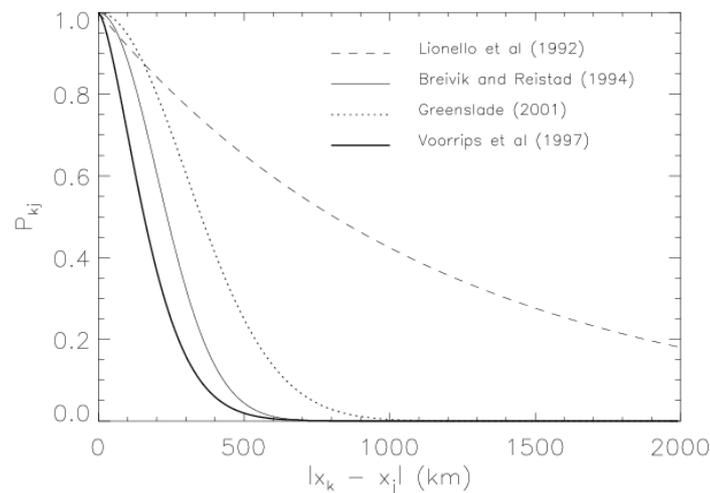
A wide range of structures has been used in the literature to describe  $P$ . A review of these can be found in Greenslade (2004). They generally have the form:

$$P_{kj} = \left( 1 + \frac{|x_k - x_j|}{L} \right)^a \exp \left[ -c \left( \frac{|x_k - x_j|}{L} \right)^b \right] \quad (3)$$

where  $L$  is the decorrelation length scale and  $|x_k - x_j|$  is the distance between the points  $k$  and  $j$ . The values of  $a$ ,  $b$ ,  $c$  and  $L$  used by various authors are listed in Table 1 and some of the resulting  $P_{kj}$  curves are plotted in Figure 1.

**Table 1. Values of the parameters in Equation 3 used by various authors.**

Author	a	b	c	L
Lionello et al (1992)	0	1	1	15°
Mastenbroek et al (1994)	1	1	1	~135 km
Breivik and Reistad (1994)	0	2	1/2	200 km to 40 km
Young and Glowacki (1996)	0	2	1/2	350 km
Greenslade (2001)	0	2	1/2	300 km
Voorrips et al (1997)	0	3/2	1	200 km



**Figure 1. Some example functions used for the background error correlations at mid-latitudes**

Some of these structures for  $P_{kj}$  have been *ad hoc* estimates, while others have provided some justification for the choice of background error structure. Most of the studies have assumed isotropy and homogeneity in the background error structure, with the exception of the studies in which  $L$  is defined in degrees of latitude and longitude. (This means that the model errors are defined to have a larger zonal spatial scale at low latitudes than at high latitudes, resulting in both anisotropy and inhomogeneity.) To date, there has been no extensive effort made to determine the spatial scale of the background errors in wave models on a global basis. A major limitation to current implementations of SI systems and indeed potential future systems is the lack of an accurate representation of the model (or background) errors.

There are several methods that have commonly been used to estimate the background error correlations in meteorology and oceanography. One of these is the observational method of Hollingsworth and Lönnberg (1986). This method uses observations from a long-term, dense, homogenous observational network and examines the difference between the observations and the background field. Greenslade and Young (2004) (henceforth referred to as GY) used altimeter data to determine background errors in SWH. Their results showed that the length scales of the background errors varied significantly over the globe, with the largest scales at low latitudes and shortest scales at high latitudes. Very little seasonal or year-to-year variability in the correlation length scales was detected. Conversely, the magnitude of the background error variance was found to have considerable seasonal and year-to-year variability. These results will be discussed further in Section 6.

In data assimilation, the background field is typically a model forecast from the previous assimilation step, i.e., the previous analysis. Therefore, the background error consists of two components: the model error and the analysis error. The model error describes the amount by which the error due to the model has grown since the previous analysis time. This could be due to either errors in the wind forcing, or errors internal to the wave model. The analysis (i.e. the wave field after the observations have been incorporated via data assimilation) will also include some error. This is because the length scales and error variances defined in the assimilation system represent expected values and are unlikely to be correct at all places and at all times. In addition, most areas of the ocean are not updated every assimilation period. Observational error also precludes a perfect analysis. The NMC method assumes that the background error growth is equivalent to the forecast divergence. There are obviously limitations to this method. Nevertheless, an examination of the forecast divergence can provide a guide to the structure of the background errors, and can also provide some insight into the error growth in wave models.

The forecast divergence can be estimated by examining the difference between pairs of model forecasts of different ranges (e.g. a 48-hour forecast and a 24-hour forecast) valid at the same time, using the shorter-range forecast as the validation. This method has been used to examine background errors in NWP systems at several forecasting centres. It was first given the name the NMC method in Rabier *et al.* (1998) in reference to a publication by Parrish and Derber (1992) who were based at the then U.S. National Meteorological Center (NMC) (now the National Centers for Environmental Prediction, NCEP).

One of the first studies to apply this technique to an atmospheric model was Baker *et al.* (1987). They considered the differences between analysed and 6-hour forecast geopotential height over a 60 day time period. Larger correlation scales were found in the tropics than in the mid-latitudes. It was suggested that the geographical variation could be due to deficiencies in the analyses due to the lack of data in some areas.

In Parrish and Derber (1992), 24-hour forecasts were compared with model analyses over a period of 30 days. The forecast error variance over 24 hours could therefore be determined. This was then empirically scaled to convert the variances to values representative of the 6-hourly interval at which data assimilation was performed. No attempt was made to examine directly the spatial structure of the forecast errors using this method.

Rabier *et al.* (1998) adapted and extended the method by considering the differences between 48-hour forecasts and 24-hour forecasts. It was suggested that although the choice of forecast fields used is fairly arbitrary, there are certain advantages in using the 24-hour and 48-hour fields. Firstly, the use of the 24-hour field as the verifying field avoids issues involved with using the analysis. For example, there may be problems in the analysis associated with the irregular spacing of observations. In addition, the 24-hour difference between the forecast periods is long enough so that the forecasts are not too similar and short enough so that the results will still be applicable to the desired 6-hourly time period, i.e., it will still be within the linear error growth regime. Both error variances and correlation scales were considered in Rabier *et al.* (1998). It was shown that results using this method compared reasonably well to results using other methods, for example, the observational method (Hollingsworth and Lönnberg, 1986).

Other authors to have used this method to determine background error statistics for atmospheric systems include Steinle *et al.* (1995), Desroziers *et al.* (1995) and Ingleby *et al.* (1996).

The method has been applied to SWH fields in Voorrips (1998). The spatial correlation structure of the background error for total wave energy was estimated by considering differences between the model forecast at 24 hours and the model analysis over a 2-month period in the North Sea. The resulting structure for the background error correlation matrix for SWH was that listed in Table 1 and shown in Figure 1.

There are some advantages to the NMC method over the observational method. These are mainly related to not having to rely on the availability of a high quality observational dataset. The method uses only model output, which is generally on a regular grid. The method is relatively straightforward and easy to apply and information on the background errors can be obtained in all directions, on all spatial scales (larger than the resolution of the model grid) and at all locations. In addition, for NWP systems, it provides the opportunity to examine the error structure of fields that are difficult to observe, for example, humidity.

However, there are limitations to its application. The inherent assumption that the structure and magnitude of the forecast divergence is similar to that of the forecast errors is difficult to justify. If the structure of the forecast errors varies little over the time range being examined, then this is a reasonable assumption, but in

the case of SWH, the long-term swell component of the wave field can complicate the situation substantially. This relates to the issue of needing to remain within the linear error growth regime. The SWH field may be better represented by the sum of two different linear error growth regimes - the wind-sea regime and the swell regime.

### 3. WAVE MODEL

The modelled wave fields used in this work are those produced operationally and archived at the Bureau. The wave model run at the Bureau is AUSWAM (NMOC, 1999), a version of the WAM model. The WAM model is a third generation wave model that solves the wave transport equation explicitly without assuming a form for the evolving spectrum. The wave transport equation is:

$$\frac{\partial F}{\partial t} + \nabla \cdot (\mathbf{c}_g F) = S_{ij} + S_{nl} + S_d \quad (4)$$

where  $F(f, \theta)$  is the wave spectrum as a function of frequency ( $f$ ) and direction ( $\theta$ ),  $\mathbf{c}_g$  is the group velocity and the terms on the right hand side represent the source terms:  $S_{in}$  is the energy input due to wind forcing,  $S_{nl}$  the non-linear energy transfer between groups of resonant waves and  $S_d$  the dissipation of energy due to whitecapping.

In the most recent version of the WAM model (Cycle 4, released in December 1991) the physics included a dynamic coupling between wind and waves (Janssen 1989, 1991). Bender (1996) compared the original Cycle 3 (Snyder *et al.* 1981; Komen *et al.* 1984) and Cycle 4 physics to independent waverider buoy data over a one-month period (July 1992) and concluded that a wave-forecasting model for the Australian region should be based on the Cycle 3 physics. Although this led to underprediction of wave heights, it was shown that by upgrading the first-order propagation numerics to third-order and increasing the magnitude of the dissipation term by approximately 40%, a good match with the buoy observations could be achieved. Therefore, the present implementation of the WAM model at the Bureau uses the Cycle 3 physics with increased dissipation and third-order upwinding numerics.

Wave spectra are discretised into 12 directional bins, centred at 15°, 45°, 75°, etc. This “staggering” of the directional bins is to avoid having spectral energy propagating directly along the axes of the north-south coordinate system (Bidlot *et al.*, 1997). There are 25 frequency bins ranging from 0.0418 Hz to 0.4114 Hz. Deep water physics only is used. The propagation and source term time steps are 20 minutes and 10 minutes respectively. For the global version of the model, the north-south extent of the domain is 78°N to 78°S. Surface parameters such as SWH are archived every 12 hours.

Forcing fields for the global wave model are wind velocities at 10 m above sea level. These are obtained from the Bureau's global atmospheric model: GASP (Seaman *et al.*, 1995). Surface winds are obtained from the lowest level of GASP via Monin-Obukhov theory with empirical stability functions (Garratt, 1992).

The time period considered in this work is a 2-year period from April 1999 to March 2001. This time period was chosen because there were no major upgrades or changes in resolution to either GASP or AUSWAM during this time. The 10 m wind fields are instantaneous snapshots of the surface fields and are provided to the wave model at 12-hourly intervals and 2.5° spatial resolution. These are linearly interpolated in time to 3-hourly intervals, and bi-linearly interpolated in space to the resolution of the wave model grid. The spatial resolution of the wave model is 3° and wave forecasts are made out to 96-hours. The only relevant change made to the operational system during this time period is that the assimilation of ERS-2 SWH data commenced in August 1999 (Greenslade, 2001). Thus data assimilation was being performed over most of the time period considered here.

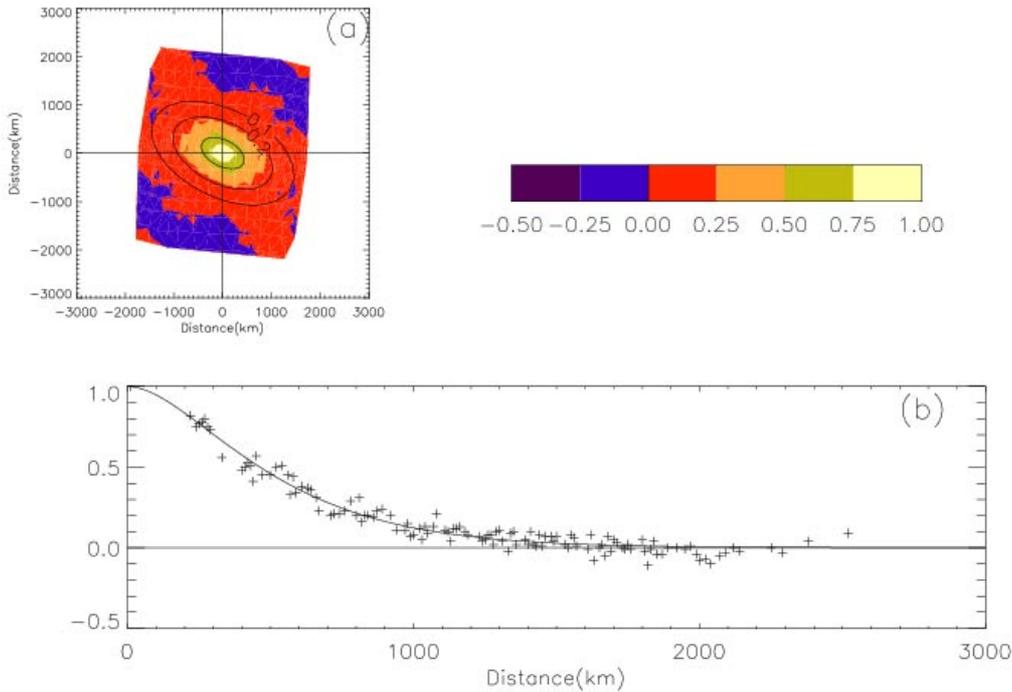
### 4. METHOD

In this work, 48-hour forecasts of SWH and 10 m wind speed (U10) are compared to 24-hour forecasts valid at the same time over the 2-year period to examine the magnitude, spatial scale and temporal and geographical variability of forecast divergence. In addition, over a period of 1 year, forecasts valid at all time periods greater than 24-hours (up to 96-hours) are compared to the 24-hour forecast in order to examine the variability of the forecast divergence with forecast period.

The background error correlation matrix in Equation 1 can be expressed as the spatial error correlation between two points,  $j$  and  $k$ , i.e., (Daley, 1991)

$$R_{jk} = \frac{\overline{(O_j - B_j)(O_k - B_k)}}{\sqrt{\overline{(O_j - B_j)^2} \overline{(O_k - B_k)^2}}} = \rho(r, \theta) \quad (5)$$

For the purposes of this work,  $B$  represents a forecast field (either SWH or U10) and  $O$  is the verifying field valid at the same time, i.e., the 24-hour forecast field. Correlations of the forecast divergence, i.e.,  $\rho(r, \theta)$  as a function of great circle distance  $r$  and angle  $\theta$  are computed for 3-monthly time periods and boxes of side length  $20^\circ$  in latitude and longitude globally. The average represented by the overbar is the time average over all matching pairs of model fields at 12-hourly intervals within the 3-month time period. The temporal mean SWH or U10 is removed at each grid-point to ensure that the difference fields are unbiased.



**Figure 2. (a) 48-hour - 24-hour SWH forecast divergence correlation for a  $20^\circ$  box centred at (110W, 40S) for the 3-month period April to June 2000. (b) Same as (a) but as a function of distance alone.**

An example of a 48-hour - 24-hour ( $t_{48} - t_{24}$ ) forecast divergence correlation for SWH is shown in Figure 2(a). This example is for a  $20^\circ$  box in the southeastern Pacific Ocean centred at (110W, 40S) for the 3-month period April to June 2000. Figure 2(b) shows the same correlation as a function of distance alone,  $\rho(r)$ . As expected, the correlations are highest at short distances and decay towards longer distances. The decay in the two-dimensional correlation function is slower in the southeast direction than the northeast direction. In other words, over the 24-hour forecast period the spatial scale of the forecast divergence is larger in the southeast direction than in the northeast direction. This behaviour can be interpreted as follows. Consider a model gridpoint within the domain. Assume that at this gridpoint, at a particular time within this 3-month period, the difference in SWH between the 48-hour forecast and the verifying 24-hour forecast is large. Then according to Figure 2(a) the SWH at a point  $x$  km to the southeast of this gridpoint is more likely to also differ between the 24-hour and 48-hour forecast than the SWH at a point  $x$  km to the northeast of this gridpoint. Note that the correlation functions in Figure 2 do not provide any information about the magnitude of the forecast divergence, but only the spatial scales and the directions in which the forecasts are most likely to diverge. To estimate the magnitude of the expected difference between 48-hour and 24-hour forecasts, the variance needs to be examined (see below).

Analytic descriptions of the correlations are obtained by applying curve-fitting procedures to the 3-monthly correlations. In addition, average correlations over longer time periods are obtained by averaging the 3-monthly correlations together bin-wise and then applying the curve-fitting procedure. The following forms were selected after testing many variations of Equation 3 (and its 2-dimensional equivalent) and selecting the curve that resulted in the lowest mean-square-error (mse) on average (Greenslade, 2004). The analytic form used for the 2-dimensional, or anisotropic case is:

$$\rho(r, \theta) = \exp\left[-\frac{d}{a_3}\right] \quad (6)$$

with

$$d^2 = r^2 \left( \frac{1}{a_1^2} \cos^2(\theta - a_2) + a_1^2 \sin^2(\theta - a_2) \right) \quad (7)$$

where  $a_1$ ,  $a_2$  and  $a_3$  are the parameters to be fitted. The best-fit surface for the example shown in Figure 2(a) is represented in the figure by the solid ellipses. In addition, the isotropic, or 1-dimensional case is considered. The analytic function that resulted in the lowest mse for the isotropic correlations is the 2nd-order auto-regressive function, i.e.,

$$\rho(r) = \left(1 + \frac{r}{a_1}\right) \exp\left[-\frac{r}{a_1}\right] \quad (8)$$

and the best-fit curve is again shown in Figure 2(b). For this example, the value of  $a_1$  for the isotropic case is 276 km. This is the isotropic length scale of the spatial correlation of the t48 - t24 forecast divergence. For the remainder of this paper, this will be referred to as  $L$ , the isotropic correlation length scale.

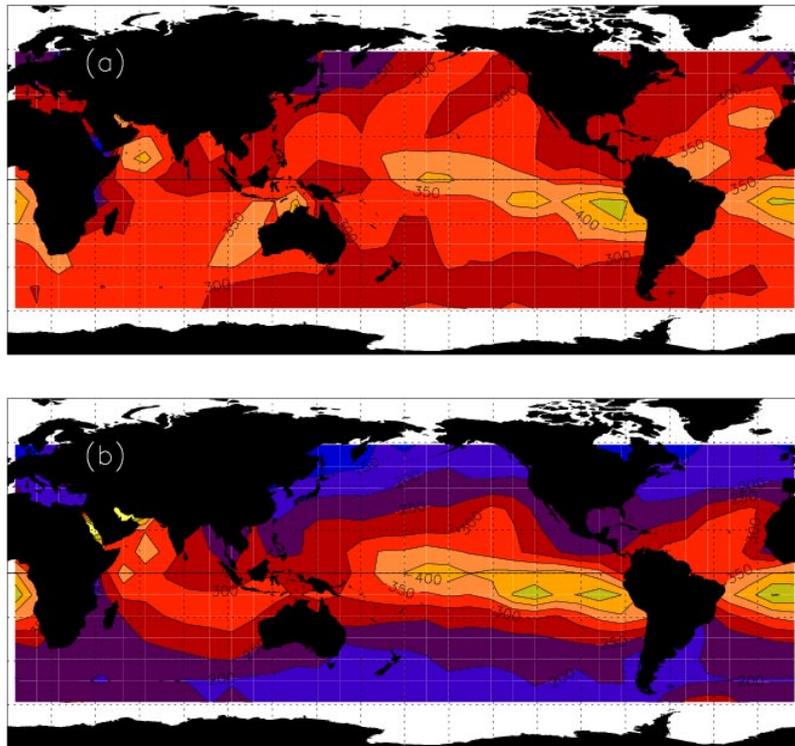
The variance of the forecast divergence is also calculated as follows:

$$E_B^2 = \frac{1}{K} \sum_{k=1}^K \overline{(B_k - O_k)^2} \quad (9)$$

where  $K$  is the number of model gridpoints within the  $20^\circ$  box.

## 5. RESULTS

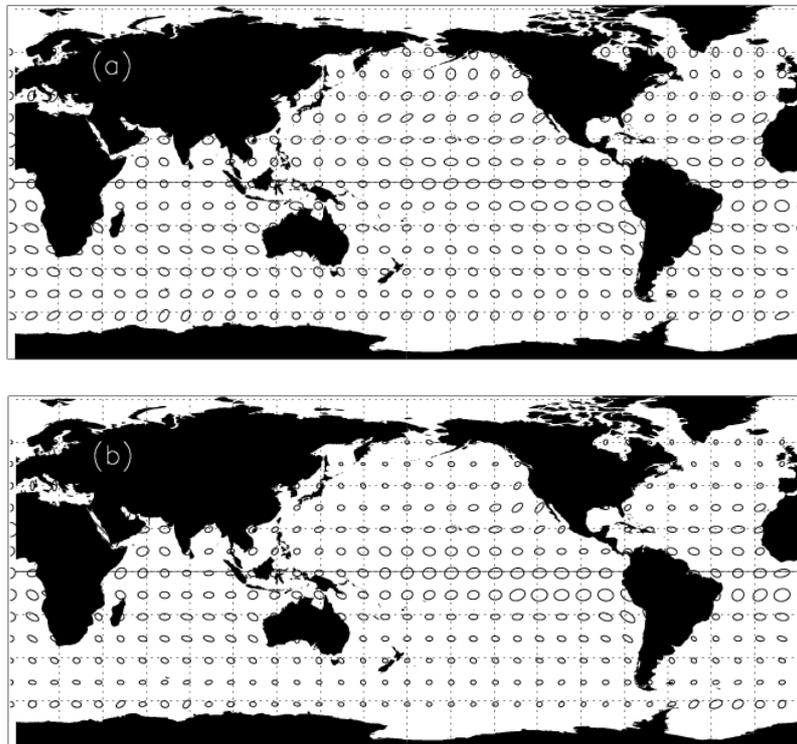
Firstly, the mean results over one year are considered - for t48 - t24 during the year April 2000 to March 2001. Note that wave data assimilation was included in the wave forecasting system throughout this time period. The operational Bureau data assimilation system has  $P_{kj}$  defined with a length scale of approximately 200 km. This means that variability with this spatial scale is being inserted into the analysis. It is difficult to know exactly what the impact of this would be on estimates of the spatial structure of forecast divergence.



**Figure 3. Average isotropic correlation length scale of the difference between 48-hour and 24-hour forecasts for the year April 2000 to March 2001 for (a) SWH and (b) U10.**

Maps of the isotropic correlation length scale over the globe are shown in Figure 3 for SWH ( $L_{SWH}$ ) and U10 ( $L_{U10}$ ). It can be seen that the geographical distribution of the length scales is similar for both parameters. The longest length scales occur just south of the equator in the eastern Pacific and Atlantic Oceans. It can be seen that there is in general a strong latitudinal dependence, with the smallest  $L$  at high latitudes. The regions in which  $L$  is large are those in which the two forecasts differ on a large spatial scale. The physical basis for these geographical distributions will be discussed further in Section 5.1.1.

To represent the anisotropic case, the 0.5 level contour line of the two-dimensional analytical correlation function,  $\rho(r, \theta)$  (see Figure 2(a)), is drawn at the centre of each  $20^\circ$  box. For the remainder of this work, these will be referred to as “error ellipses”. These are shown in Figure 4. The relative size of the error ellipses generally follows the pattern of the isotropic length scale. For SWH, the ellipses are generally anisotropic, with the longest scales in the east-west direction or close to it. There are some areas where there is less anisotropy, for example at high latitudes. The error ellipses for U10 follow a similar pattern, with perhaps a tendency to be less anisotropic at lower latitudes. The implications of the anisotropy will be discussed further in Section 5.1.2.



**Figure 4. 0.5-level contour line of the anisotropic correlation of the difference between 48-hour and 24-hour forecasts for the year April 2000 to March 2001 for (a) SWH and (b) U10.**

### 5.1 Forecast Period

The impact of the length of the forecast period is now considered. In particular, forecasts of different ranges are compared to the 24-hour forecast. Again, forecasts during the year April 2000 to March 2001 are examined.

#### 5.1.1 Isotropic case

Figure 5 shows how  $L_{SWH}$  varies with forecast period. A noticeable feature of Figure 5 is that the correlation length scales become larger everywhere as the forecast period increases. For example, in the central Indian Ocean, the length scale increases from approximately 300 km for  $t_{36} - t_{24}$  to more than 450 km for  $t_{96} - t_{24}$ . Note also the longest scales in the Eastern Pacific Ocean lengthening from 400 km to 650 km as the forecast period increases. This has been shown to be true for atmospheric systems also. For example, Bengtsson and Gustavsson (1971) show how the spatial scale of the forecast error autocorrelation of 500 mb geopotential height increases as the forecast period increases. In other words, for short-range forecasts, most of the error is in the short scales while longer-range forecasts have increasing errors in the larger scales. The trend towards larger spatial scales for longer forecast periods seen in Figure 5 may be due to

this trend occurring in the wind fields and being transferred into the wave fields. This can be investigated by examining the length scale of the forecast divergence of the winds used to force the wave model. This is shown in Figure 6.

It can be seen that  $L_{U10}$  does lengthen somewhat as the forecast period increases, but not to the same extent as in the SWH forecast divergence. For example, in the central Indian Ocean  $L_{U10}$  lengthens from approximately 300 km to more than 350 km and in the Eastern Pacific ocean, it lengthens from 450 km to 550 km. This suggests that the lengthening of scales in the surface wind forecast divergence does not fully explain the lengthening of scales seen in the SWH forecast divergences.

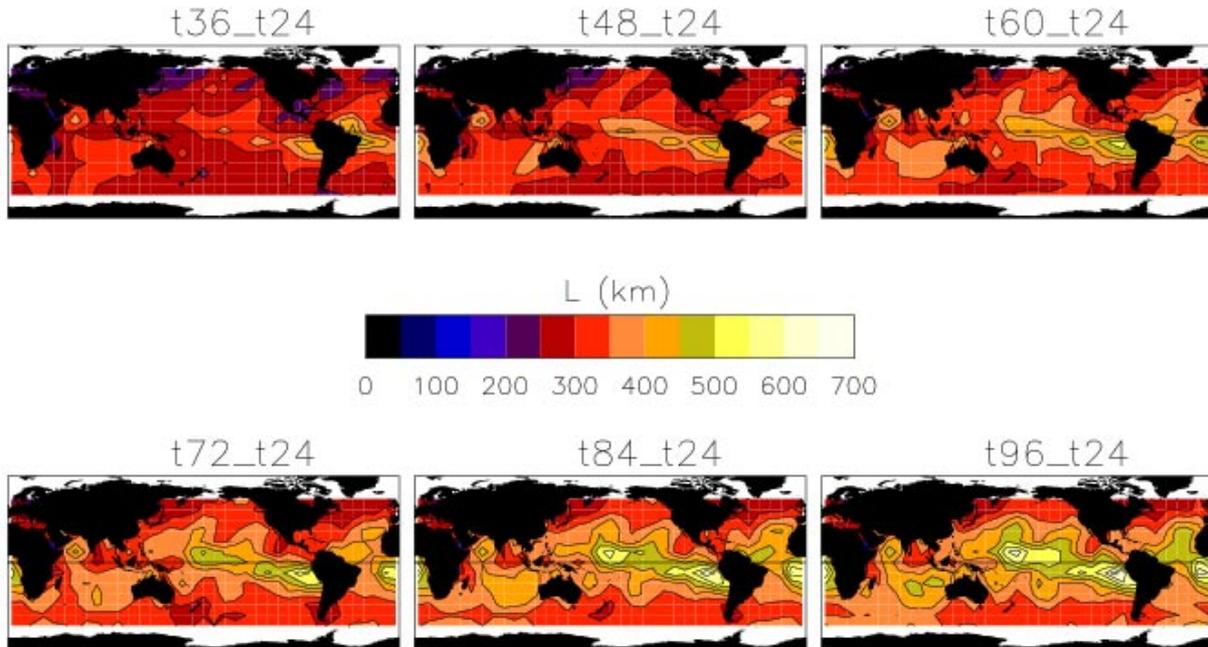


Figure 5.  $L_{SWH}$  for the year April 2000 to March 2001 for increasing forecast differences. t36 - t24 represents the difference between the 36-hour forecast and the 24-hour forecast, etc.

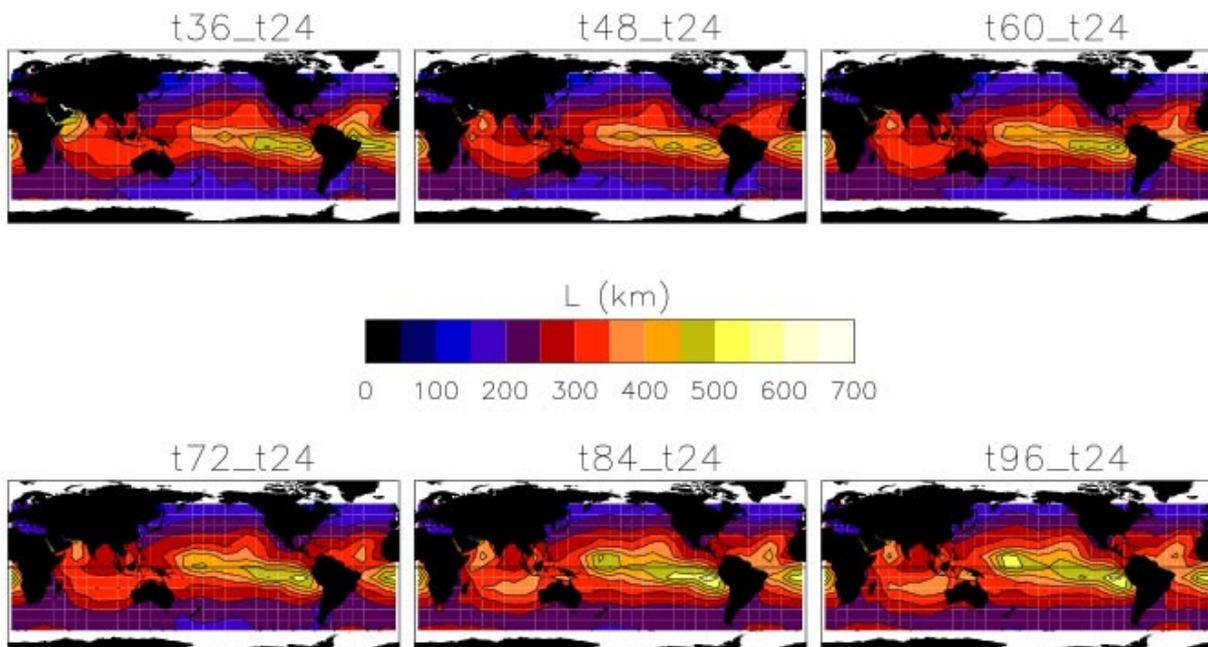


Figure 6. Same as Figure 5 but for U10.

For SWH, there is another simple physical explanation. The swell component of a wave field is that component of the wave spectrum that exists independently of the local surface winds. Errors in SWH that exist in the short-range forecasts (either due to errors in the analysis, errors in the wind forcing or errors internal to the wave model) can propagate forward in time, independent of the wind forcing in the longer term forecasts, similar to the way in which swell propagates. As these “swell errors” propagate, they disperse and their spatial scale becomes larger. Thus the differences between short-range and long-range forecasts will occur on larger scales. In other words, the errors existing in the SWH fields at longer forecast periods are more dominated by large-scale swell than the errors at shorter forecast periods. The lengthening of  $L_{SWH}$  seen in Figure 5 is likely to be due to a combination of this error propagation and the wind error pattern.

The latitudinal dependence of  $L_{U10}$  is more pronounced than that of  $L_{SWH}$ . This supports the idea that the forecast divergence of the SWH is made up of two components: one that is strongly coupled to the wind field and therefore strongly latitudinally dependent, and a component consisting of the propagation of swell errors, which superimposes a non-latitudinally dependent pattern.

### 5.1.2 Anisotropic case

The variation of the anisotropy of the forecast divergence with forecast period is now considered. Figure 7 shows plots of the error ellipses for increasing forecast periods for three different locations. Consider the box centred in the Indian Ocean at (80E, 30S) (panels (a) and (b)). For U10, the ellipses become slightly more anisotropic as the forecast period increases, but not to the extent that the SWH ellipses do. So, in this area, the increase of  $L_{SWH}$  with forecast period seen in Figure 5 is due only to the lengthening of the scales in one particular direction. Since the U10 ellipses do not vary much over the forecast period, this reinforces the hypotheses that a) the increase in scales in the SWH forecast divergence is due to swell errors propagating and dispersing, and b) the SWH error regime consists of two separate components - the swell error and the wind-sea error.

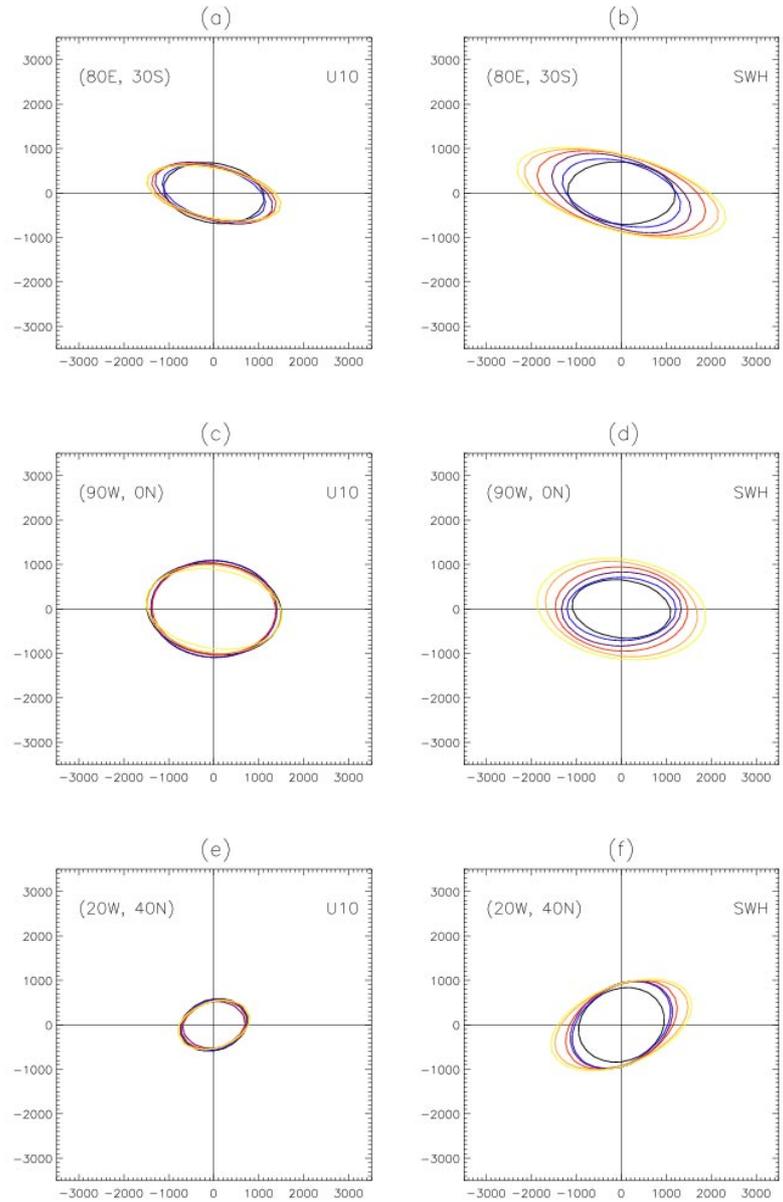
The SWH error ellipses in Figure 7 (b) lengthen predominantly in one direction, i.e., the spatial scale of the differences between long and short-term forecasts is larger in the southeast direction. In this region there are frequent swell systems propagating towards the northeast throughout the Indian Ocean. These are forced by the storm systems that move from west to east along 40°S to 50°S. The error ellipses are aligned with their short axes in the direction of propagation of the swell. It is likely then, that the structure of the correlations in this area is related to errors in the generation or propagation of swell within the wave model. Errors in the generation of swell are likely to be due to errors in the forcing. Potential propagation errors are the swell propagating too quickly, too slowly, in the wrong direction or the swell systems having the wrong amplitude. The group velocity for deep-water waves is well known and modelled, so it is unlikely that incorrect propagation speed is the cause. With the added knowledge that SWH is generally underestimated by the wave model in this region (Greenslade, 2004) other causes become more probable, for example, excessive numerical diffusion, or excess physical dissipation. Numerical diffusion is minimised in AUSWAM as it incorporates high order numerics in the propagation terms, so excessive physical dissipation is the most likely candidate.

In the eastern Pacific, (panels (c) and (d)), the opposite occurs: the U10 ellipses vary very little over the forecast period, but the SWH ellipses become larger in all directions. This suggests that the errors in the SWH field at this location are arriving from or are propagating in many directions. This is reasonable, as over a period of a year, this region of the ocean is exposed to swell arriving from both the Southern Ocean and the Northern Pacific Ocean.

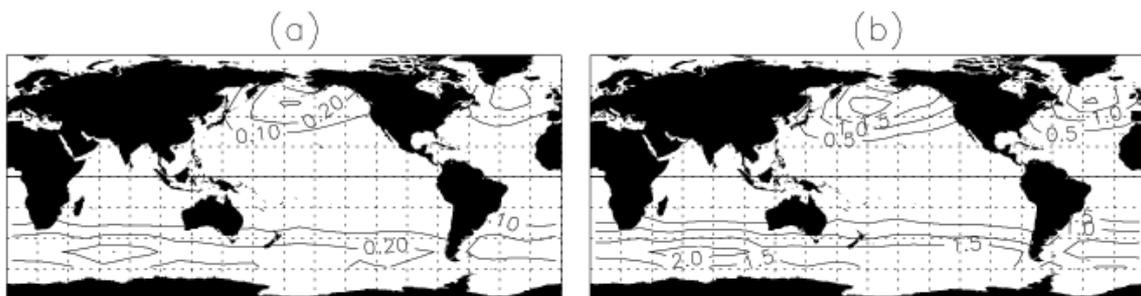
The North Atlantic location (panels (e) and (f)) is close to the European coast. A similar trend to that seen in the Indian Ocean location is seen here, but with the ellipses aligned towards the northeast rather than the southeast. This suggests that the errors in the SWH field at this location are mainly due to errors in swell systems propagating from the northwest.

### 5.1.3 Variance

The magnitude of the variance of the forecast divergence (see Equation 9) for two different forecast ranges is shown in Figure 8. These plots represent the average variance over the year April 2000 to March 2001. The distributions of the variance shown in Figure 8 for t36 - t24 and for t96 - t24 are similar. Maximum variances occur in the North Pacific and at around 50°S in the Indian Ocean sector of the Southern Ocean. These are the areas where SWH is generally highest (Young, 1999).

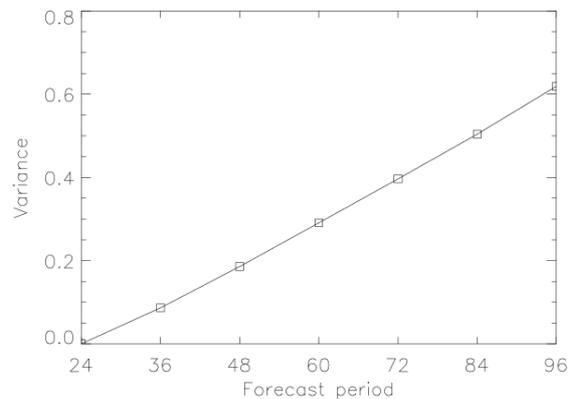


**Figure 7. Error ellipses as a function of forecast period for three separate locations for U10 (left column) and SWH (right column). In each case, the black line represents the error ellipse for t36 - t24. The other ellipses are for increasing forecast periods with the shade becoming lighter as the forecast difference increases up to t96 - t24.**



**Figure 8. Mean variance ( $m^2$ ) during the year April 2000 to March 2001 of (a) t36 - t24 (b) t96 - t24.**

As can be seen in Figure 8 the magnitude of the variance for  $t_{96} - t_{24}$  is considerably larger than that of  $t_{36} - t_{24}$ . This can also be seen in Figure 9 where the average variance over the globe of the difference between each forecast (at 12-hourly intervals) and the 24-hour forecast is shown. (Note that calculating the arithmetic mean of the variance over the globe assumes homogeneity, and Figure 8 demonstrates that this is clearly not the case). The background error variance should increase with time until it reaches the level of climatological error variance, where it should asymptote (Daley, 1991). It can be seen that while the variance of the forecast divergence does indeed increase with time, there are no signs of it asymptoting within this time period. If the forecast divergence is equivalent to the background error, then this implies that there is still significant skill (compared to climatology) in the 96-hour wave forecasts.



**Figure 9. Global spatial average of the variance ( $m^2$ ) of the forecast differences ranging from  $t_{36} - t_{24}$  up to  $t_{96} - t_{24}$ .**

As mentioned previously, a data assimilation system requires the value of  $E_B^2$  which is appropriate for the time interval at which data is inserted into the model. From Figure 9, the appropriate spatially averaged value at three hours (i.e.,  $t_{27} - t_{24}$ ) would be approximately  $0.02 m^2$ . However, as discussed in Section 2, the values shown here represent forecast error divergence - the appropriate value would be  $0.02 m^2$  only if the analysis were perfect.

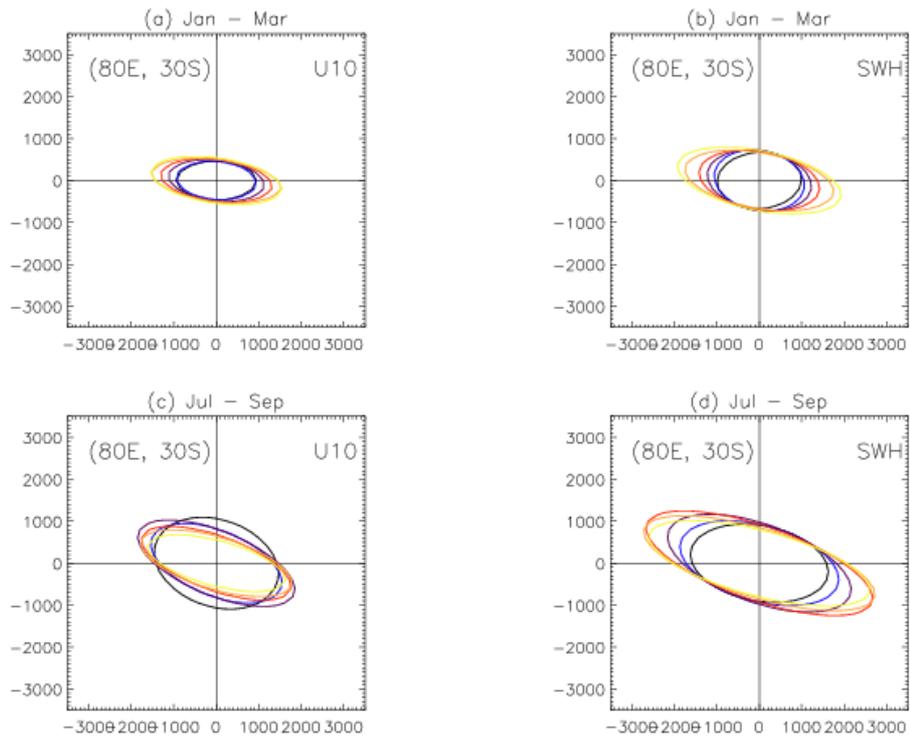
## 5.2 Seasonal Variability

In their study of atmospheric variables, Rabier *et al.*, (1998) found that the isotropic and homogenous component of the forecast error was very stable with season. The results of GY also showed that there was little seasonal variability in the isotropic length scale of the SWH error correlations. However, there are grounds for expecting to see seasonal variability in the forecast divergence of SWH. For example, during southern hemisphere (SH) winter, swell arriving in the central Pacific would mostly be from the SH, while during northern hemisphere (NH) winter, one would expect the central Pacific to be dominated by swell propagating from the NH. This is therefore best investigated by examining the seasonal variability of the anisotropic forecast divergence.

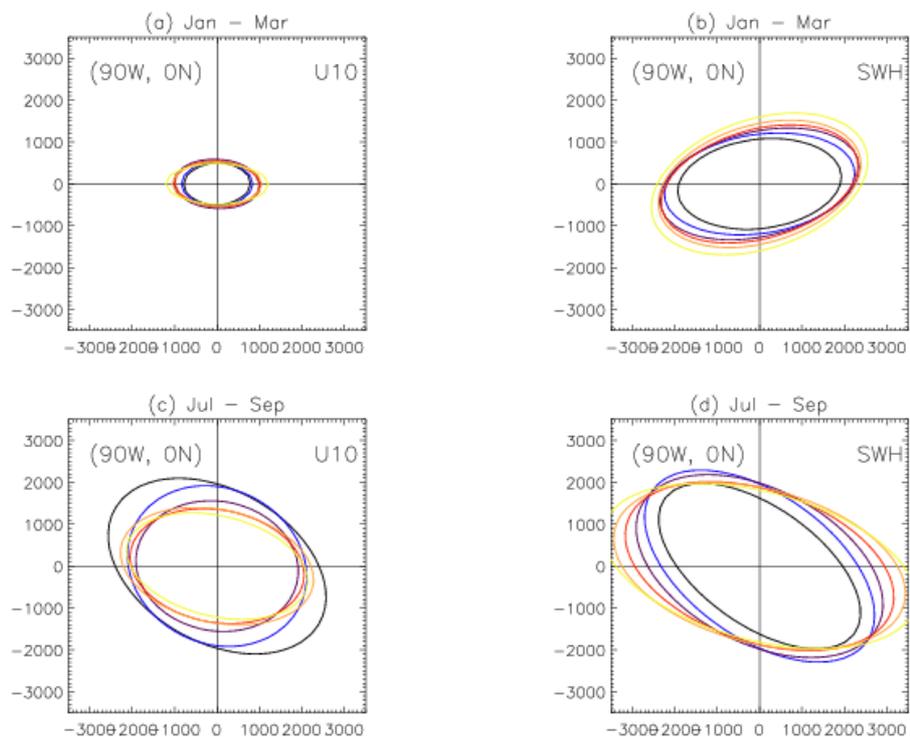
To obtain a time series longer than one year, correlations for specific three-month seasons over two years are averaged together, as described in Section 4. An even longer time series would be desirable, but the time period used here is limited to two years in order to avoid any major changes in the operational system. It has been shown (in GY) that system changes such as increases in the frequency of the forcing wind fields can have considerable impact on the SWH error correlation structure.

In this section, the seasonal variability of the error ellipses for two of the locations shown in Figure 7 is considered. First consider the box centred at (80E, 30S), in the central Indian Ocean. Seasonal variability is shown here in Figure 10. The top two panels are the error ellipses for U10 and SWH for January to March 2000 and 2001 and the bottom two panels are for July to September 1999 and 2000. It can be seen that the spatial scales of the correlations for both parameters are generally larger in SH winter and slightly more tilted, but overall, there are no major differences in angular distribution between the two seasons. This region experiences swell being generated in the Southern Ocean and propagating towards the Australian coast all year. The shape and size of the Indian Ocean (in particular its very small NH area) means that there is very little seasonal variability in the generation of swell in this region.

On the other hand, one would expect the Pacific Ocean to show a seasonal signal. As mentioned earlier, there should be more swell to be generated in the hemisphere experiencing winter and this would dominate



**Figure 10. Error ellipses of U10 (left column) and SWH (right column) as a function of forecast period for (a) and (b) Jan - Mar 2000 and 2001 and (c) and (d) Jul - Sep 1999 and 2000 for a box in the Indian Ocean centred at (80E, 30S).**



**Figure 11. Same as Figure 10 but for a box in the Eastern Pacific centred at (90W, 0N).**

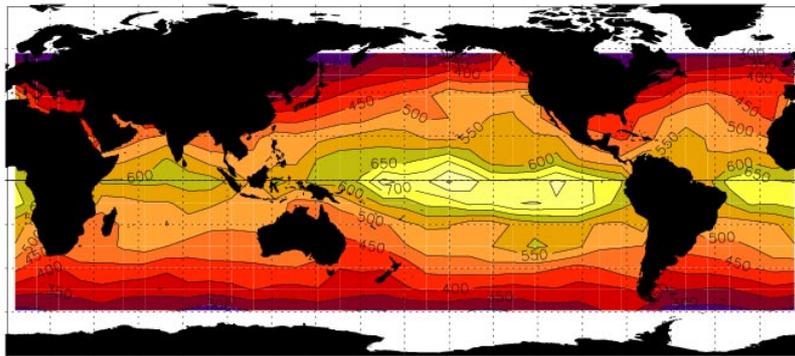
the error ellipses in forecasts in the central Pacific Ocean. Figure 11 shows the same as Figure 10 for the box centred in the eastern equatorial Pacific Ocean. It can indeed be seen that the SWH fields (right-hand panels) show error ellipses tilted perpendicular to the direction from which one expects swell to be propagating. The U10 error ellipses for SH summer are quite different to those for SH winter, and these are

not strongly related to the patterns in the SWH error ellipses. This demonstrates that the signal in the SWH fields is not related to the local wind fields and must be related to swell propagation.

## 6. DISCUSSION

These results can be compared to the results of GY, in which the same problem was addressed via a different method. In GY, the observational method of Hollingsworth and Lönnberg (1986) was used. Correlations of the differences between satellite altimeter observations of SWH and modelled SWH were examined globally. The modelled wave fields used were from AUSWAM at  $0.5^\circ$  spatial resolution over the four-year time period April 1998 to March 2002. The winds used to force AUSWAM were from the data assimilation cycle of the atmospheric model GASP, described in Section 3. Assimilation of wave observations was not included in the four-year wave model run.

The major results of GY are reproduced here in Figure 12. This figure can be compared to the various panels in Figure 5. It can be seen that the distribution of global correlation length scales obtained from the forecast divergences is in general quite similar to that obtained from the observational method.



**Figure 12. Isotropic SWH model error correlation length scale over the globe using the observational method (reproduced from GY).**

Note however, that the magnitude of the length scales in Figure 12 is larger than that of even the  $t_{96} - t_{24}$  forecast divergence in Figure 5. For the SWH fields considered in GY, it can be assumed that the wind-wave component of the wave spectrum is relatively well modelled, because the wind fields used were from the data assimilation cycle of the atmospheric model. However, there was no data assimilation in the wave model, so any component of the wave field that was incorrectly generated or modelled, would not be corrected. This component of the error would then propagate, disperse and become large-scale swell error. This suggests that the modelled wave fields in GY contain a large amount of long-term, or remotely-forced error. Therefore the length scales calculated from those modelled wave fields and shown in Figure 12 are likely to be overestimated.

As discussed earlier, the error structure that should be defined within a data assimilation system is the structure of the background error at the frequency of data assimilation (typically three hours). If the present results using the NMC method are used as a proxy for the background error, then this means that both the background error variance and the length scale must be scaled down to three hours. Figure 5 suggests that if the length scale grows linearly, then at a forecast range of three hours, the appropriate length scale is quite short. If a perfect analysis were produced then the appropriate error structure to use is indeed the forecast error structure found in this work, which describes the magnitude and spatial scale of the forecast divergence after 3 hours. However, as discussed earlier, the background error describes the difference between the background and the truth, and this must include analysis error, error in the forcing and forecast (or model) error. No analysis is perfect: the length scale and error variances defined in the assimilation system represent expected values and are unlikely to be correct at all times. In addition, most areas of the ocean are not updated every assimilation period. This means that there will always be some long-term or remotely-forced errors in the analysis and thus also in the forecast, and so the length scales and variances based on the forecast divergences alone are likely to be underestimated. One way to address this issue could be to use a combination of results from GY (in which there is likely to be too much remotely-forced error) and results from this work (in which there is too little remotely-forced error). Thus the results of GY provide an upper bound to the error structure of the background errors and error variance and the results of this work provide a lower bound. Note that there is quite a large difference between the length scales shown

in Figure 12 and those in Figure 5. This suggests that there still remains significant uncertainty in our knowledge of the structure of the background errors in wave models.

In Section 5.2 it was shown that the swell component of the error in forecast wave fields is likely to be anisotropic and to have a seasonal signal. The question of how to incorporate this information in a data assimilation system arises. To take advantage of this, one would need to know whether the error in the modelled wave field is in the wind-sea or swell component of the wave spectrum. Most current operational wave data assimilation systems assimilate satellite altimeter observations, which provide estimates only of SWH. This does not provide any information on whether the error is in the swell or wind-sea portion of the spectrum, so it is not easy to use this information. However data assimilation systems that use Synthetic Aperture Radar observations of the wave spectrum (e.g. Breivik *et al.*, 1998; Aouf *et al.*, submitted), or *in situ* buoy observations may be able to apply these results in updating the wave spectrum.

## 7. SUMMARY

In this work, the NMC method has been used to examine the structure and magnitude of background errors in an operational wind/wave forecasting system. The forecast error divergence is used as a proxy for the short-range forecast error. It is found that the isotropic correlation length scale of the SWH forecast divergence ( $L_{SWH}$ ) has considerable geographical variability, with the longest scales just to the south of the equator in the eastern Pacific Ocean, and shortest scales at high latitudes.

It was shown that  $L_{SWH}$  increases as the forecast period increases. This is partly due to the length scales of the surface wind forecast divergence increasing. Another explanation is that errors in the short-range SWH forecast propagate forward in time and disperse and their scale becomes larger. Examination of the anisotropy in the forecast divergences showed that longer-range forecasts are likely to be dominated by errors in swell propagation. This error had the largest spatial scales perpendicular to the expected direction of propagation of swell.

The results compared well, qualitatively, to those found using an alternate method in GY, but differ by a factor of more than 2. Limitations to both methods were discussed and it was suggested that the results in this work should be used as a lower bound and the results of GY as an upper bound to estimates of the structure of the background error in global wave models. This will be tested in a wave data assimilation system in future work.

## 8. ACKNOWLEDGEMENTS

The authors would like to thank Oscar Alves and Peter Steinle for reviewing early drafts of the manuscript.

## 9. REFERENCES

- Aouf, L., Lefevre, J-M., Hauser, D., Assimilation of directional wave spectra in the wave model WAM: an impact study from synthetic observations. *submitted to J. Atmos. Oc. Tech.*
- Baker, W.E., Bloom, S.C., Woollen, J.S., Nestler, M.S., Brin, E., Schlatter, T.W. and Branstator, G.W. 1987. Experiments with a three-dimensional statistical objective analysis scheme using FGGE Data, *Mon. Wea. Rev.*, **115**, 272 - 296.
- Bender, L.C. 1996. Modification of the Physics and Numerics in a Third-Generation Ocean Wave Model, *J. Atmos. Oc. Tech.*, **13**, 726 - 750.
- Bengtsson, L. and Gustavsson, N. 1971. An experiment in the assimilation of data in dynamical analysis. *Tellus*, **23**, 328 - 326
- Bidlot, J.R. Janssen, P.A.E.M., Hansen, B. and Günther, H. 1997. A modified set up of the advection scheme in the ECMWF wave model, *ECMWF Technical Memorandum*, **No. 237**, ECMWF, Reading.
- Breivik, L.A. and Reistad, M. 1994. Assimilation of ERS-1 altimeter wave heights in an operational numerical wave model, *Weather and Forecasting*, **9**, 440 - 451.
- Breivik, L.A., Reistad, M., Schyberg, H. and Sunde, J. 1998. Assimilation of ERS SAR wave spectra in an operational wave model, *J. Geophys. Res.*, **103**, **C4**, 7887 - 7900.
- Cardone, V.J., Graber, H.C., Jensen, R.E., Hasselman, S. and Caruso, M.J. 1995. In search of the true surface wind field in SWADE IOP-1: Ocean wave modelling perspective, *Glob. Atmos. Ocean Sys.*, **3**, 107 - 150.
- Daley, R. 1991. *Atmospheric Data Analysis*, Cambridge University Press, Cambridge, U.K., 457pp.
- Desroiziers, G., Mathiot, V. and Orain, F. 1995. A study of ARPEGE forecast error covariances. *Proceedings of the second WMO international symposium on assimilation of observations in meteorology and oceanography*, Tokyo, Japan, 263 - 268.

- Garratt, J.R. 1992. *The Atmospheric Boundary Layer*, Cambridge Univ. Press, Cambridge, U.K., 316pp.
- Greenslade, D.J.M. 2001. The assimilation of ERS-2 Significant Wave Height data in the Australian region, *J. Mar. Sys.*, **28**, 141 - 160.
- Greenslade, D.J.M. 2004. The Structure of the Background Errors in a Global Wave Model, *Ph.D. Thesis*, University of Adelaide.
- Greenslade, D.J.M. and Young, I.R. Background Errors in a Global Wave Model determined from Altimeter Data, *J. Geophys. Res.*, accepted.
- Hollingsworth, A. and Lönnberg, P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field, *Tellus*, **38A**, 111 - 136.
- Ingleby, N.B., Lorenc, A.C., Barker, D.M., Andrews, P.L.F. and Thurlow, S. 1996. Forecast error covariances for a variational analysis system, 11th Conference on Numerical Weather Prediction, Norfolk, Virginia, 246 - 248.
- Janssen, P.A.E.M. 1989. Wave-induced stress and the drag of airflow over sea waves, *J. Phys. Oceanogr.*, **19**, 645 - 754.
- Janssen, P.A.E.M. 1991. Quasi-linear theory of wind-wave generation applied to wave forecasting, *J. Phys., Oceanogr.*, **21**, 1631 - 1642.
- Komen, G.J., Hasselmann, S. and Hasselmann, K. 1984. On the existence of a fully developed wind-sea spectrum, *J. Phys. Oceanogr.*, **14**, 1271 - 1285.
- Komen, G.J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S. and Janssen, P.A.E.M. 1994. *Dynamics and Modelling of Ocean Waves*. Cambridge University Press, Cambridge, UK, 532pp.
- Lionello, P., Günther, H. and Janssen, P.A.E.M. 1992. Assimilation of altimeter data in a global third-generation wave model, *J. Geophys. Res.*, **97**, 14453 - 14474.
- Mastenbroek, C., Makin, V.K., Voorrips, A.C. and Komen, G.J. 1994. Validation of ERS-1 altimeter wave height measurements and assimilation in a North Sea wave model, *Glob. Atmos. Ocean Sys.* **2**, 143 - 161.
- National Meteorological Operations Centre. 1999. Changes to the Operational Sea State Forecast System, *Operations Bulletin No. 47*, Bur. Met., Australia.
- Parrish, D.F. and Derber, J.C. 1992. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, *Mon. Wea. Rev.*, **120**, 1747 - 1763.
- Rabier, F., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A., and Bouttier, F. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Q. J. R. Meteorol. Soc.*, **124**, 1809 - 1829.
- Seaman, R., Bourke, W., Steinle, P., Hart, T., Embery, G., Naughton, M. and Rikus, L. 1995. Evolution of the Bureau of Meteorology's Global Assimilation and Prediction System, Part 1: Analyses and Initialization, *Aust. Met. Mag.* **44**, 1 - 18.
- Snyder, R.L., Dobson, F.W., Elliot, J.A. and Long, R.B. 1981. Array measurements of atmospheric pressure fluctuations above surface gravity waves, *J. Fluid. Mech.*, **102**, 1 - 59.
- Steinle, P., Seaman, R., Bourke, W. and Hart, T. 1995. A generalized statistical interpolation scheme, Symposium on Assimilation of Observations in Meteorology and Oceanography, Tokyo, Japan, 205 - 208.
- Voorrips, A.C. 1998. Sequential data assimilation schemes for ocean wave models, *Ph.D. Thesis*, Delft University of Technology.
- Voorrips, A.C., Makin, K. and Hasselmann, S. 1997. Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model, *J. Geophys. Res.*, **102**, 5829 - 5849.
- WAMDI Group (S. Hasselmann, K. Hasselmann, E. Bauer, P.A.E.M. Janssen, G. Komen, L. Bertotti, P. Lionello, A. Guillaume, V.C. Cardone, J.A. Greenwood, M. Reistad, L. Zambresky and J.A. Ewing), 1988. The WAM model - A third generation wave prediction model. *J. Phys. Oceanogr.*, **18**, 1775 - 1810.
- Young, I.R. 1999. *Wind Generated Ocean Waves*. Elsevier Science, Oxford, U.K. 288pp.
- Young, I.R. and Glowacki, T.J. 1996. Assimilation of altimeter wave height data into a spectral wave model using statistical interpolation, *Ocean Engng.*, **23**, No. 8, 667 - 689.