Tellus (1980), *32*, 189–197

On the use of a time sequence of surface pressures in four-dimensional data assimilation

By LENNART BENGTSSON, *European Centre for Medium Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, England*

(Manuscript received May 31; in final form October 8, 1979)

ABSTRACT

The possibility of using a time sequence of surface pressure observations in four-dimensional data assimilation is being investigated. It is shown that a linear multilevel quasi-geostrophic model can be updated successfully with surface data alone, provided the number of time levels are at least as many as the number of vertical levels. It is further demonstrated that current statistical analysis procedures are very inefficient to assimilate surface observations, and it is shown by numerical experiments that the vertical interpolation must be carried out using the structure of the most dominating baroclinic mode in order to obtain a satisfactory updating. Different possible ways towards finding a practical solution are being discussed.

1. Introduction

In the synoptic analysis of weather maps, originally developed by the Bergen school, the surface pressure tendencies or a time sequence of pressure observations are of primary importance and essential when using the weather maps in short range forecasting. The surface pressure tendencies give the change over the last 3 hours and they are included as a part of the SYNOP and the SHIP observations reported and exchanged globally every 6 hours (in some regions even every 3rd hour). Consequently we have a global coverage of surface pressure information which is available every 3rd hour with a relative accuracy of less than 2 millibars practically everywhere. Moreover in areas where observations still are sparse, surface pressure information can with a reasonable cost be obtained from floating buoys or from other automatically recording instruments. Data collection can, as has been demonstrated during FGGE, successfully be carried out from interrogating polar orbiting satellites.

It is natural to ask why the surface pressure tendencies have not been used in numerical weather prediction. Surface pressure is recorded by instruments which due to their relative inertia exclude high frequency gravity waves. The observations therefore describe pressure changes related to the Rossby waves except for large scale gravity oscillations due to tidal effects.

However, these large scale fluctuations have no dynamic significance and are easy to eliminate. The observed surface tendencies are thus providing useful information about the state of development of baroclinic waves in particular, and as will be shown in section 2 can be used to compensate for the lack of information in the free atmosphere.

The use of surface pressure tendencies in numerical prediction has previously been very difficult due to an unsatisfactory initialization of the numerical models. This has created erroneous gravity waves which usually have taken around a day to be dampened to an acceptably low level. These computational gravity modes are characterized by large surface tendencies of periods from less than an hour and up to a day, and there are no simple ways to separate them from Rossby modes except by computing each of the individual Rossby and gravity modes *per se*. A technique to do this, normal mode initialization, has recently been developed into a practical solution and there are now satisfactory methods to exclude gravity waves from the initial state and produce an almost

 0040-2826/80/030189-09$02.50/0

completely noisefree prediction from the very beginning of the forecast (Machenhauer, 1977; Temperton and Williamson, 1979). For this reason it is now feasible to use tendencies to initialize a numerical integration.

In the following we will address some problems related to the question of how surface pressure tendencies can be used practically in numerical weather prediction. The most straightforward way seems to be to make use of a continuous data assimilation or an intermittent assimilation with a 3-hour cycle. Hereby the tendencies will implicitly be incorporated. The surface observations can then either replace the predicted values or be analysed using a 3-hour prediction as a first guess (background field).

Considerable interest has been devoted, during the last decade, to the adjustment problem in relation to a 4-dimensional data assimilation. It is obvious than an updating of the wind field will have very little effect on the large scale mass field unless a balance correction is built into the assimilation scheme and, vice versa, updating of the mass field will have very little effect on the small scale wind field if no balance procedure is incorporated.

However, there is another equally fundamental problem in 4-dimensional data assimilation, which so far has not been discussed very much. This is a problem which occurs when we have a data distribution which is very unevenly distributed in the vertical. The most extreme case is when we have observations at only one level in the atmosphere, as for instance surface pressure observations. We may consider this as a *second kind of adjustment* between a dynamical process described by the model and a statistical process described by the data.

We will concentrate this study on this particular problem. For the sake of reasoning we will set up the simplest possible system and investigate the updating of a baroclinic wave on a β-plane. We will make use of the quasi-geostrophic equations and we will disregard non-linear effects. We realize naturally that such a model is very different from reality, but nevertheless, as will be shown, it can be used to illustrate some fundamental problems when we have information on one level only.

2. Theoretical considerations

The mathematical framework for this study will be based on the quasi-geostrophic equation on a β-plane. This is a system where the relation between wind and mass field is expressed by the linear part of the balance equation assuming a constant Coriolis parameter, f_0, and where the static stability is allowed to vary only in the vertical. Under these assumptions the model is governed by the vorticity equation and the thermal equation in the following form. In the p-system these equations read

$$\nabla^2 \frac{\partial \psi}{\partial t} + J(\psi; \nabla^2 \psi) + \beta \frac{\partial \psi}{\partial x} - f_0 \frac{\partial \omega}{\partial p} = 0 \qquad (2.1)$$

$$\frac{\partial^2 \psi}{\partial p \partial t} + J\left(\psi; \frac{\partial \psi}{\partial p}\right) + \frac{\sigma \omega}{f_0} = 0 \qquad (2.2)$$

where ψ is the stream function, ω the vertical motion and $\sigma = \sigma(p)$ the static stability. This system of equations has been extensively investigated during the last 25 years. For a description relevant to this investigation see Bengtsson (1978). In the following we will refer to this paper as B1.

The solutions of (2.1) and (2.2) can be found by assuming solutions of the particular form

$$\begin{aligned} \psi(x, y, p, t) &= -U(p)y + \Psi(p, t)e^{i(kx+my)} \\ \omega(x, y, p, t) &= \Omega(p, t)e^{i(kx+my)} \end{aligned} \qquad (2.3)$$

where $U(p)$ is the zonal wind and k and m are the wave numbers for the perturbation.

Inserting (2.3) into (2.1) and (2.2) and using finite differences in the vertical we obtain a set of ordinary differential equations in the perturbation Ψ. In the case of one vertical separation we obtain a 2-level model in the following form

$$\begin{aligned} \dot{\Psi}_1 &= k_{11}\Psi_1 + k_{12}\Psi_2 \\ \dot{\Psi}_2 &= k_{21}\Psi_1 + k_{22}\Psi_2 \end{aligned} \qquad (2.4)$$

where Ψ_i is the complex perturbation at each level and k_{ij} are complex coefficients. Ψ_1 represents here the streamfunction at the upper level which we can think of as 250 mb and Ψ_2 as the lower level at 750 mb. If Ψ is known initially, eqs. (2.4) gives the value of Ψ_i at all future times.

As was shown in B1 the analytic solution of eqs. (2.4) can be written

$$\Psi_1(t) = e^{i\nu_1 t}\{A\Psi_1(0) - B\Psi_2(0)\} + e^{i\nu_2 t}\{C\Psi_1(0) + D\Psi_2(0)\} \qquad (2.5)$$

$$\Psi_2(t) = e^{i\nu_1 t}\{C\Psi_2(0) - D\Psi_1(0)\}$$

$$+ e^{i\nu_2 t}\{A\Psi_2(0) + D\Psi_1(0)\} \quad (2.6)$$

A, B, C, D, ν_1 and ν_2 are all functions of the basic large scale flow and the wave number. In particular ν_1 and ν_2 represent the frequencies of the wave.

Formally we can write

$$\Psi_1(t) = f_1(t)\Psi_1(0) + f_2(t)\Psi_2(0) \quad (2.7)$$

$$\Psi_2(t) = f_3(t)\Psi_1(0) + f_4(t)\Psi_2(0) \quad (2.8)$$

where $f_i(t)$ depends on the large scale flow which is supposed to be known at all times.

If now Ψ_1 has an error, this error will grow and as has been shown in B1 penetrates very fast to the lower level.

We will now assume that Ψ_1, which represents the stream function at the upper level, has an initial error, $\delta\Psi_1$, while the stream function at the lower level, Ψ_2, is correct. Moreover, the initial tendency at the lower level $\delta_t\Psi_2$ is also assumed to be known exactly.

We can use this tendency to calculate $\Psi_2(\delta t)$ which is also given by (2.8) evaluated at (δt). Equating the two we can obtain $\Psi_1(0)$ as follows:

$$\Psi_2(0) + \delta_t\Psi_2(0) = \Psi_2(\delta t)$$

$$= f_3(\delta t)\Psi_1(0) + f_4(\delta t)\Psi_2(0)$$

or

$$\Psi_1(0) = -\frac{f_4(\delta t)}{f_3(\delta t)}\Psi_2(0) + \frac{1}{f_3(\delta t)}\Psi_2(\delta t) \quad (2.9)$$

where $\Psi_2(\delta t)$ is the value of the stream function Ψ_2 at the time δt. After insertion in eqs. (2.7) and (2.8) we formally obtain

$$\Psi_1(t) = h_1(t, \delta t)\,\Psi_2(0) + h_2(t, \delta t)\,\Psi_2(\delta t) \quad (2.10)$$

$$\Psi_2(t) = h_3(t, \delta t)\,\Psi_2(0) + h_4(t, \delta t)\,\Psi_2(\delta t) \quad (2.11)$$

where $h_i(t, \delta t)$ are functions of the large scale flow (linear functions of $f_i(t)$) and as assumed above supposed to be known.

Assuming next that we have an N-level model the solution of which we formally could write

$$\begin{aligned}\Psi_1(t) &= \sum_{l=1}^{N} k_{1l}\,\Psi_l(0)\\ &\vdots\\ \Psi_N(t) &= \sum_{l=1}^{N} k_{Nl}\,\Psi_l(0)\end{aligned} \quad (2.12)$$

where k_{ij} are functions of time and of the large scale flow.

Assuming next that we know Ψ_N at N succcessive times $\Psi_N(0)$, $\Psi_N(\delta t) \ldots \Psi_n((N-1)\,\delta t)$ we can then write in analogy with (2.10) and (2.11)

$$\begin{aligned}\Psi_1(t) &= \sum_{l=1}^{N} h_{1l}\,\Psi_N((i-1)\,\delta t)\\ &\vdots\\ \Psi_N(t) &= \sum_{l=1}^{N} h_{Nl}\,\Psi_N((i-1)\,\delta t)\end{aligned} \quad (2.13)$$

Consequently, if we know a *time sequence of observations* at one level in the atmosphere, we can then determine the "initial state" for a numerical model for all *vertical levels* if we disregard the non-linear effect.

3. A numerical study

It is clear from the discussion in the preceding section that it is possible to initialize a baroclinic perturbation knowing the structure (phase and amplitude) at a particular level for a given length of time. We can of course solve this by using eqs (2.13) and we have done so using the analytic solution given in B1. However, in order to obtain some guidelines which can be applied under more realistic conditions, we will use a numerical procedure.

In the experiment which we are going to describe, we have been using a 10-level model. The wind profile as well as the specified static stability are given in Fig. 1. As was pointed out in B1, the short waves and the very long waves turned out to be very sensitive to the initial error. In particular the short waves with wavelengths around 3–4000 km turned out to be very sensitive to the initial error. In this experiment we have, therefore, for the sake of argument selected a wave with a wavelength of $L = M = 4000$ km. L represents the wavelength in the east–west direction and M the wavelength in the north–south direction. This wave is slightly unstable for a vertical stratification profile and wind profile given in Fig. 1. The most unstable mode given in Fig. 1 has an e-folding time of 4–5 days.

The updating experiments have been carried out as an observing simulation experiment. A given initial state which will represent the "true state" has been integrated for a period of 5 days. This initial state is specified in Table 1. The initial error is characterized by an initial phase error. While the

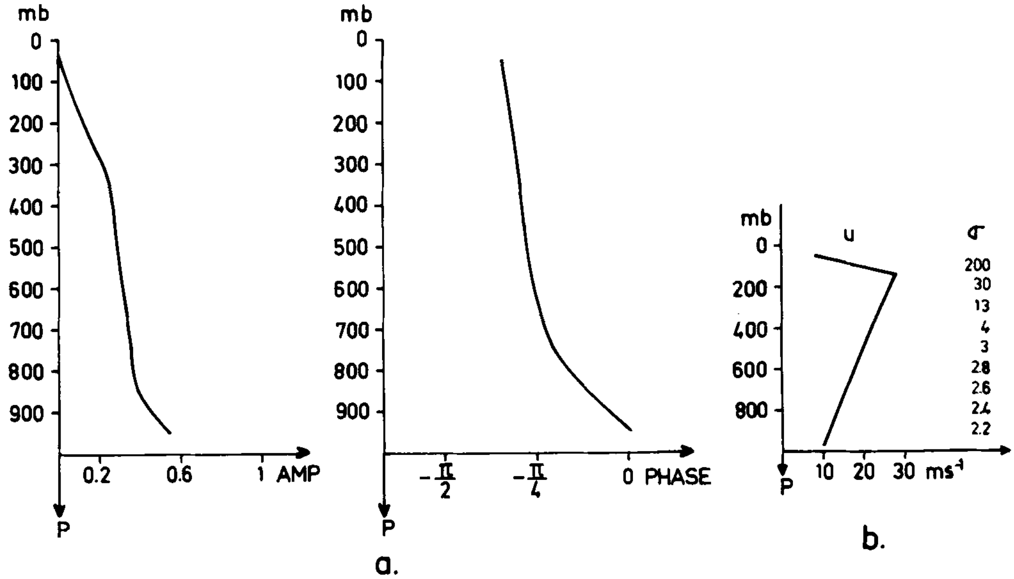

Fig. 1. (a) Amplitude and phase-structure for the most unstable mode given by the vertical profile of $\bar{U}$ and σ. (b) The amplification factor is 0.2135/day and the phase-speed 13.87 ms^{-1}.

true state has a small phase tilt, as can be seen from Table 1, the reference state has no phase tilt at all. This represents a kind of error which is rather common in operational numerical weather prediction where, if at all, surface observations are modifying upper levels only along a vertical axis. The initial "reference state" has been integrated for a period of 5 days as well. As can be expected, this wave amplifies much slower than the true state and a substantial "error" in the amplitude can be seen at all levels.

Table 1. *Initial condition for the true case and the reference case*

Pressure level (mb)	True state: Amplitude	True state: Phase in degrees	Reference state: Amplitude	Reference state: Phase in degrees
50	100	30	100	0
150	100	30	100	0
250	100	30	100	0
350	100	28	100	0
450	100	25	100	0
550	100	20	100	0
650	100	15	100	0
750	100	7.5	100	0
850	100	3	100	0
950	100	0	100	0

Fig. 2 shows the percentage increase of the amplitude at level 10 (the lowest level) in time for the true state and the reference state, respectively. Fig. 2 also shows the vertical structure of the wave after 48 hours.

The actual updating is carried out in such a way that data from level 10, the lowest level of the model, are taken from the true state and inserted in the reference state. The insertions are done during the 10 first time steps of 1 hour and the experiment is evaluated by measuring to what degree this information is being assimilated and to what extent the reference state is being modified towards the true state. According to section 2, if this insertion happens to be perfect, the reference state will be changed into the true state by just 10 insertions.

We will now carry out two different updating experiments: In the first experiment we will try to simulate what is done in operational numerical weather prediction. For everyone of the first 10 time steps we replace the values at level 10 by the true values, and we further modify the lowest 5 levels according to a vertical function given in Fig. 3.

In order to illustrate the procedure, let us first separate the complex variable Ψ in an amplitude part, A and a phase part, D. We can thus write

$$\Psi = Ae^{iD} \tag{3.1}$$

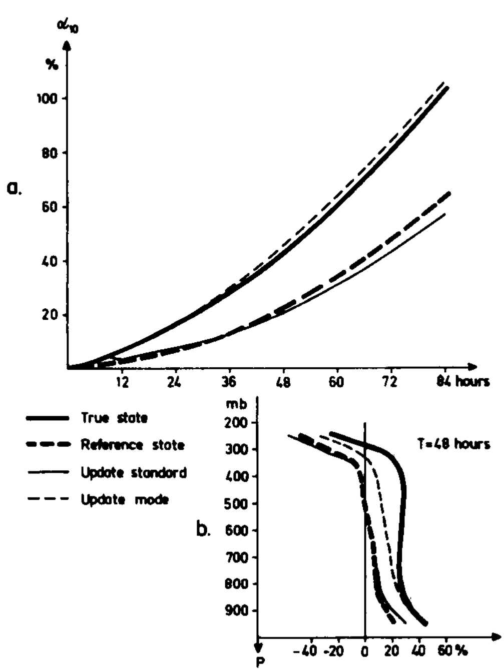

Fig. 2. (a) Percentage change in the amplitude at level 10 (A_{10}) or 950 mb for the true state (full heavy line), the reference state (dashed heavy line), standard updating (full thin line) and mode updating (dashed thin line). (b) Vertical amplitude structure for the same cases at $t = 48$ h. Wavelength $L = M = 4000$ km.

Here A and D are both functions of p and t. Numerically we will use the notations

$$A = A(k, \tau)$$
$$D = D(k, \tau) \tag{3.2}$$

where k represents the vertical levels $k = 1, \dots N$ and τ the time step $\tau = 1, 2 \dots$

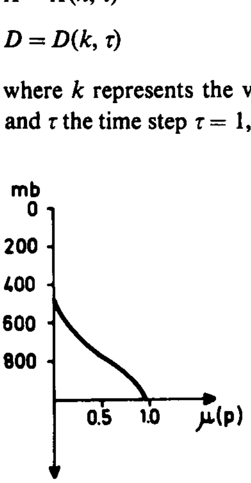

Fig. 3. Weighting factor μ as a function of p.

In the first updating experiment we have used the relations

$$\Psi(N, \tau) = \Psi(N, \tau)^{\text{true}}$$
$$\Psi(k, \tau) = \Psi(k, \tau)^{\text{guess}} + a\mu(k)\,\{\Psi(N, \tau) - \Psi(N, 0)\}, \quad k = 1, \dots, N-1 \tag{3.3}$$

where a is an empirical constant and $\mu(k)$ a vertical weighting function.

Updating of experiment I is very unsuccessful, and as can be seen from Fig. 2 the inserted information is dispersed and the solution develops towards the reference state.

In the second updating experiment we are assuming that the flow is developing towards a state dominated by the most unstable mode, that is, the fastest amplifying baroclinic wave. We are therefore simultaneously modifying the phase of the wave towards a vertical structure of the most unstable mode.

We have thereby used the following expressions

$$\Psi(N, \tau) = \Psi(N, \tau)^{\text{true}}$$
$$A(k, \tau) = A(k, \tau)^{\text{guess}} + a\mu(k)\,\{(A(N, \tau)^{\text{true}} - A(N, 0)^{\text{true}}\} \tag{3.4}$$
$$D(k, \tau) = D(k, \tau)^{\text{guess}} + a\{(D(N, \tau)^{\text{true}} + D_1(k, 0) - D(k, \tau))\}$$

$D_1(k, 0)$ is here the phase difference between level k and level N for the most unstable mode.

After some experimentation we found that $a = 0.05$ was a suitable value, although $a = 0.1$ gave very similar results.

The results shown in Fig. 2 are very interesting. It is found that when we try to update stream functions analogous to operational weather prediction, we can see hardly any effect at all. The amplitude is forced to the correct one, but the fundamental error, which is the error in the phase, is not corrected. When we correct the error in the phase as well, the result is improved dramatically.

An additional experiment was carried out to investigate the effect of incorporating surface friction. A value being the equivalent of Ekman friction (10° cross isobar flow and the wind of 70% of the geostrophic wind) was used for the reference model as well as for the update model. Fig. 4 shows the result of this experiment. The same relative improvement can be observed. The same normal mode (most unstable adiabatic mode) was used.

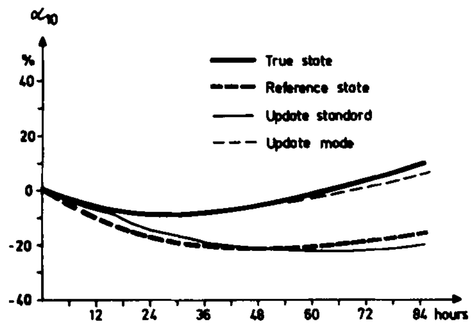

Fig. 4. The same as Fig. 2a but Ekman friction included.

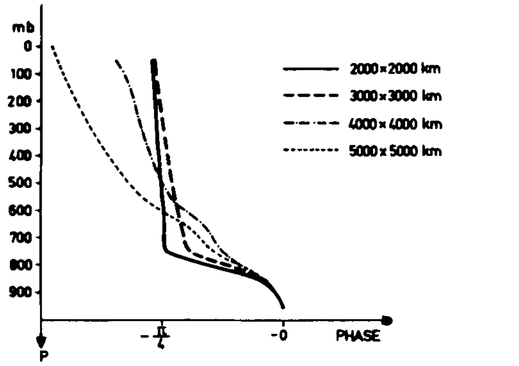

Fig. 6. The phase structure for the waves $L = M =$ 2000 km (amplification 0.508/day) $L = M =$ 0.428/day), $L = M =$ 4000 km (amplification 0.473/day) $L = M =$ 5000 km (amplification 0.570/day). At the wavelengths 3000 km and 4000 km there is a second unstable mode with an amplification factor of around 0.1/day. The outlined structure of this mode is very similar to the mode given in the figure.

We have extended the present study to a case which is characterized by a higher degree of baroclinic instability. The wind profile and the vertical stratification profile are given in Fig. 5. For this particular case we have calculated the most unstable eigenvalues and the average structure for most unstable waves in the range 2000–5000 km is also given in Fig. 5. The phases for the individual (most unstable) wave components are given in Fig. 6. These are waves which are characterized by rapid developments in the westerlies with corresponding large changes in the surface pressure. Although the structure of the modes naturally varies by the wavelength, they are rather similar with characteristically large tilts between the surface and 700 mb.

This fact has been recognized in the subjective analysis of upper area maps but is very difficult to incorporate in an objective analysis scheme. A small positive impact has been reported by N. Gustafsson (private communication) using an average tilt obtained from climatology.

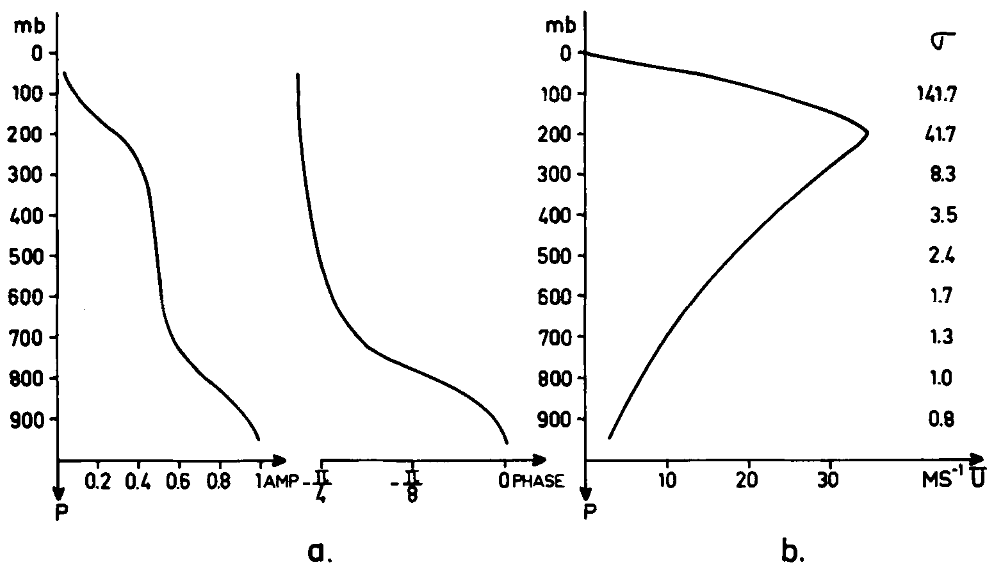

Fig. 5. (a) Average amplitude and phase structure for the most unstable mode with a wavelength between 2000 and 5000 km given by the vertical profile of U and σ. (b) The average amplification factor is around 0.5/day.

Table 2. *Phase and amplitude at 48 h for 4 different wavelengths. Update I is the standard update and update II is the mode update*

	True		Refer		Update I		Update II		
mb	Ampl.	Phase	Ampl.	Phase	Ampl.	Phase	Ampl.	Phase	
950	183	81	147	81	147	87	185	85	$L = M =$
550	48	95	29	100	29	103	44	99	2000 km
950	260	123	189	117	185	122	241	127	$L = M =$
550	148	160	102	149	103	152	131	163	3000 km
950	318	163	236	152	233	156	285	165	$L = M =$
550	222	−149	164	−163	165	−161	193	−149	4000 km
950	316	−160	245	−171	244	−168	286	−159	$L = M =$
550	245	−107	190	−121	191	−120	215	−107	5000 km

To test the idea outlined above for several wave components, we have carried out a data assimilation experiment for all the 4 waves using the mean structure function in Fig. 5 for all waves. The result is presented in Table 2. It is evident that method II is by far superior to method I.

Naturally, we have simplified this very much, and we have so far mainly investigated waves where the true state has a positive tilt. The problem has to be treated differently when we have initially a negative tilt. In this case the wave is dominated by the decaying mode which has a tilt which is exactly opposite to the amplifying mode. The amplitude of the wave is being reduced for a short time before the development is overtaken by the fastest growing unstable mode. We have carried out some experiments with such waves as well. We have therein made use of the fact that such a wave must have a negative tilt proportional to the decay of the wave. Fig. 7 shows an example of such an updating.

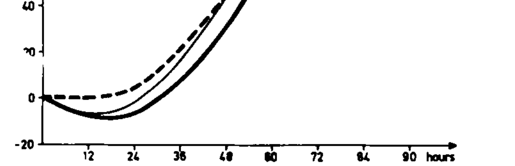

Fig. 7. The same as Fig. 2a but the true state is initially eastwards (in the direction of the basic flow). The amount of the tilt is the same as for Fig. 2.

4. Conclusions

The present study has indicated a problem in the data assimilation of the surface pressure tendencies or information available at one pressure surface only but during a sequence of time. In numerical weather prediction statistical structure functions are used to analyse surface observations and to let these observations influence other levels where data may not be available. The structure functions can be calculated from the prediction error covariance Hollett (1975) and have a form similar to the weighting curve shown in Fig. 3. The statistical structure functions cannot describe the characteristic variation of the vertical phase which a depression is undergoing during its lifetime. If there is no vertical tilt in the forecast (first guess), the insertion of the surface data will maintain this state even if it is clear from the surface tendencies that the depression is developing and consequently must have a positive tilt. We would in general expect a numerical model to *underpredict the development of the depressions* in their early phase since we often have to rely upon surface observations to detect them and because they are then inserted in the model without any positive phase change with height.

We may conclude from this simple study if surface observations, or in fact any observations which are valid on one level only, are going to be used efficiently in numerical weather prediction, the vertical extrapolation should be done using the structure function of the actual weather system.

We realize that there are difficulties in implementing this technique in practice. However, we may suggest a practical method along the following lines. In operational numerical weather prediction we normally analyse the difference between a forecast valid at the time, and the observations. Naturally, this forecast should be as accurate as possible and when the forecast happens to be identical to the observations they are not included. Let us now consider a case where we observe a difference between an observation and the first guess (the valid forecast) which is greater than the accuracy of the observation. At the same time we will have available $\partial p_s/\partial t$ which has been analysed separately.

The technique outlined is applicable only to well-defined depressions and therefore only cases when the tendencies are large enough will be considered. At middle and high latitudes this means tendencies equal to or larger than 3 mb/3 h.

The second step is to find out if the depression is amplifying or decaying. This can easily be done by calculating the difference between the maximum tendency and the minimum tendency on each side of the depression or by comparing the amplitude of the depression at the preceding observation time.

We will then suggest that if the depression is amplifying by more than a factor 0.5/day we will carry out the extrapolation along the axis of the most unstable mode (fastest amplifying). If on the other hand the depression is decaying by more than a factor of 0.5/day we extrapolate along the axis of the fastest decaying mode. If the change falls between these values we simply propose a linear relation. As the direction of the large scale flow, the mean wind direction between 700 and 300 mb can be used. The structure of the mode is naturally a function of the wavelength and of the large scale flow itself. Unless the weather situation is very extreme we suggest that one may use a similar structure to the one given in Fig. 5. An alternative method would be to organize the data set which is used to calculate the vertical error covariances into say 3 groups representing amplifying, neutral or decaying systems. Any of these covariances could then be selected in the 3-dimensional analysis.

It is important to stress that the proposed method is based on the assumption that the dominating process for cyclone development is caused by adiabatic processes. This is certainly not the case everywhere, and the release of latent and sensible heat are essential processes to incorporate. However, at middle and high latitudes these effects could possibly be neglected in this context.

The proposal outlined is thought to be of practical value in areas at middle and high latitudes where great interest is connected to the accuracy of short range forecasts (12–36 hours) and where the observing network upstream has serious limitations in the free atmosphere, but where surface observations are abundant. This is a situation which is typical for the west coast of United States and Canada and for Western Europe and to an even higher degree at the Southern Hemisphere.

5. Acknowledgement

This investigation was carried out mainly during a shorter stay by the author at the International Meteorological Institute in Stockholm. The stimulating discussions with the staff members and the technical assistance are very much appreciated.

REFERENCES

Bengtsson, L. 1978. Growth rate and vertical propagation of the initial error in baroclinic models. *Tellus* *30*, 324–334.

Hollett, S. R. 1975. Three-dimensional spatial correlation of P.E. forecast errors. M.Sc. thesis. McGill University, Quebec, Canada, 73 p.

Machenhauer, B. 1977. On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialization. *Contr. Atmos. Phys. 50*, 253–271.

Temperton, C. and Williamson, D. 1979. Normal mode initialization for a multi-level grid point model. European Centre for Medium Range Weather Forecasts Technical Report No. 11, 91 p. Available from ECMWF.

ИСПОЛЬЗОВАНИЕ ВРЕМЕННОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ПРИЗЕМНОГО ДАВЛЕНИЯ В 4-МЕРНОМ УСВОЕНИИ ДАННЫХ

Исследуется возможность использования последовательности наблюдений за приземным давлением в 4-мерном усвоении данных. Показано, что линейная многоуровенная квазигеострофическая модель может успешно усваивать текущие данные только по приземному давлению при условии, что число временны́х уровней, по крайней мере, столь же велико, как и число вертикальных уровней. Далее продемонстрировано, что обычный статистический анализ весьма не эффективен для усвоения приземных наблюдений, и с помощью численных экспериментов показана необходимость вертикальной интерполяции, использующей структуру доминирующей бароклинной моды, для удовлетворительного усвоения текущих данных. Обсуждаются различные возможные пути отыскания практического решения.