NON-LINEARITY IN STATISTICAL SHORT RANGE FORECASTING

Ingemar Holmström

Swedish Meteorological and Hydrological Institute
Norrköping Sweden

I. INTRODUCTION

In 1973-74 Nils Gustafsson and I made at SMHI an experiment in order to see if, with linear regression the 24-hour error field of the numerical model could be predicted, using EOF (Empirical Orthogonal Functions) parameters of the initial fields as linear predictors. These fields consisted of three levels, 1000, 500 and 300 mb over an area considered as a probable domain of influence for height changes over Scandinavia.

Due to computer limitations only a very small sample could be processed. The results were promising, not only with respect to geographically fixed errors but also to some extent to model errors in flow development. However, these experiments were not continued at that time due to a number of intervening factors, SMHI to be moved from Stockholm to new premises in Norrköping, change of numerical model and of computer etc. and it is only now that the experiment can be taken up again.

It is then natural to consider also the possibility of introducing non-linearity in some simplified way and, if this turns out to be feasible, also the possibility of using a statistical model for forecasting directly atmospheric fields instead of error fields of a numerical model.

Of course the crucial point in any of these two undertakings is that the best data sample available only gives analyzed fields every twelve hours. With such a large timestep one would assume a very large area of influence and a high degree of non-linearity, all leading to a prohibitively large number of predictors.

However, the very important results obtained by Lorenz [1977] in his experiments with statistical barotropic forecasting using initial time derivatives as extra predictors, indicate that the effect of non-linearity may possibly be taken into account in a simplified way.

I have therefore found it of interest to try to investigate by means of a very simple equation what kind of approximations one can expect to be realistic in statistical prediction model. What I present here is not a final result but may still be of interest.

2. TIME DERIVATIVES

In order to discuss these problems it is sufficient to consider a simpliest possible non-linear equation

$$\frac{\partial u}{\partial t} = Au + Bu \frac{\partial u}{\partial x} \qquad (1)$$

In a finite difference scheme an integration over n timesteps will require an area of influence of $2n\Delta x$, which translated into atmospheric conditions means more than a hemisphere in less then 12 hours. This is really in drastic contrast to synoptic experience where one usually makes 12 hour extrapolations taking only a very limited area into consideration.

We shall now use eq. (1) in order to calculate terms in the Taylor expansion

$$u(x,t+\Theta) - u(x,t) = \Theta \frac{\partial u}{\partial x} + \frac{\Theta^2}{2!} \frac{\partial^2 u}{\partial t^2} + \frac{\Theta^3}{3!} \frac{\partial^3 u}{\partial t^3} + \ldots \qquad (2)$$

where all derivatives are taken at x and t and where Θ is large. It is of no interest in this connexion to use an expansion of $u(x,t-\Theta)$ in order to reduce the number of terms on the right hand side. Differentiating (1) with respect to time and eliminating lower time derivatives we obtain for second, third and fourth order derivatives.

$$\frac{\partial^2 u}{\partial t^2} = A^2 u + 3ABu \frac{\partial u}{\partial x} + B^2 \left[2u\left(\frac{\partial u}{\partial x}\right)^2 + u^2 \frac{\partial^2 u}{\partial x^2}\right] \qquad (3)$$

$$\frac{\partial^3 u}{\partial t^3} = A^3 u + 7A^2 Bu \frac{\partial u}{\partial x} + 6AB^2 \left[2u\left(\frac{\partial u}{\partial x}\right)^2 + u^2 \frac{\partial^2 u}{\partial x^2}\right] \tag{4}$$
$$+ B^3 \left[6u\left(\frac{\partial u}{\partial x}\right)^3 + 9u^2 \frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x^2} + u^3 \frac{\partial^3 u}{\partial x^3}\right]$$

$$\frac{\partial^4 u}{\partial t^4} = A^4 u + 15A^3 Bu \frac{\partial u}{\partial x} + 25A^2B^2 \left[2u\left(\frac{\partial u}{\partial x}\right)^2 + u^2 \frac{\partial^2 u}{\partial x^2}\right] \tag{5}$$
$$+ 10AB^3 \left[6u\left(\frac{\partial u}{\partial x}\right)^3 + 9u^2 \frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x^2} + u^3 \frac{\partial^3 u}{\partial x^3}\right] +$$
$$+ B^4 \left[24u\left(\frac{\partial u}{\partial x}\right)^4 + 72u^2\left(\frac{\partial u}{\partial x}\right)^2 \frac{\partial^2 u}{\partial x^2} + \right.$$
$$\left. + 16u^3 \frac{\partial u}{\partial x}\frac{\partial^3 u}{\partial x^3} + 12u\left(\frac{\partial^2 u}{\partial x^2}\right)^2 + u^4 \frac{\partial^4 u}{\partial x^4}\right]$$

In for instance (4) we immediately note that there is only one term, the last one in the last bracket, that requires five gridpoints in a finite difference approximation. For all the others it is sufficient with three gridpoints. Thus, even if the area of dependence increases linearly with the number of timesteps it is obvious that it carries very little weight at the boundaries in comparison with more central parts. The last term in (4) will have its largest values for large values of u and very small scales. Since this is something that one wants to suppress in atmospheric models beyond a certain limit, a neglect of such terms may be of advantage. Comparison with the atmosphere is facilitated if we non-dimensionalize for instance equation (4) and take

$$A = f = 10^{-4}s, \quad B = -1, \quad U \simeq 21\ ms^{-1}, \quad L \simeq 700\ km$$
$$u = U \cdot v, \qquad dx = L d\xi$$

We then obtain

$$\frac{\partial^3 u}{\partial t^3} = f^3 U\{v - 7Ro\ v \frac{\partial v}{\partial \xi} + 6Ro^2\left[2v\left(\frac{\partial v}{\partial \xi}\right)^2 + v^2 \frac{\partial^2 v}{\partial \xi^2}\right]$$
$$- Ro^3 \left[6v\left(\frac{\partial v}{\partial \xi}\right)^3 + 9v^2 \frac{\partial v}{\partial \xi}\frac{\partial^2 v}{\partial \xi^2} + v^3 \frac{\partial^3 v}{\partial \xi^3}\right]\}$$

where Ro is the Rossby number and with the choice given of U and L has the very large value of 0.3.

In this case third order terms are multiplied by 0.09 and fourth order terms by 0.027. Since Rossby numbers of 0.2 or 0.1 are more realistic for atmospheric conditions we find here a clear indication of truncation possibilities.

3. EOF-EXPANSION

In order to investigate another aspect of the non-linearity we shall assume that observed values of u(x,t) have been expanded into empirical orthogonal functions over a limited interval in x, (-L,L) and in time (0,T).

$$u(x,t) = \sigma\Sigma\mu_n \alpha_n(t) h_n(x) \qquad (6)$$

where we have the orthonormalization conditions

$$\frac{1}{2L}\int_{-L}^{L} h_n(x) h_m(x)dx = \delta_{nm}, \quad \frac{1}{T}\int_{o}^{T} \alpha_n(t) \alpha_m(t)dt = \delta_{nm}$$

and where σ^2 is the variance of u with a mean value subtracted out.

Under these conditions we have

$$\mu_n = \frac{\sqrt{\lambda_n}}{\sigma}$$

where λ_n are the eigenvalues of the covariance matrix of u from which the EOF:s are calculated. In order to present some typical values for μ_n I have taken data from an expansion of a year of 700 mb geopotential heights in a north-south direction over Sweden. With seven gridpoints (Δx = 150 km) the following μ_n-values were obtained

n =	1	2	3	4	5	6	7
μ_n =	0.8873	0.4399	0.1280	0.0460	0.0224	0.0133	0.0067

We now introduce the expansion (6) into the third order terms of eq. (3). We obtain

$$\sigma^3 B^2 \sum_k \sum_m \sum_n \mu_k \mu_m \mu_n \; \alpha_k \alpha_m \alpha_n \left[2h_k \frac{\partial h_m}{\partial x} \frac{\partial h_n}{\partial x} + h_k \, h_m \frac{\partial^2 h_n}{\partial x}\right]$$

With k, m, n = 1-7 we have here 84 third order terms and it is clearly not possible to have all these as predictors. However, considerable reduction can be made. Looking first at the third order product of the constants μ it is clear that μ_3^3 is much smaller than μ_1^3 their quotient being 0.003. Thus, it is easy to make a scheme by which small terms are neglected. In general small values of μ correspond to high frequency in the corresponding $\alpha(t)$-functions and small scale in the corresponding $h(x)$-functions and a truncation is therefore not only acceptable but also desirable.

With regard to the product $\alpha_k \; \alpha_m \; \alpha_n$ in (7) we shall here only consider the case of a second order product

$$z = xy$$

where we assume x and y to be independent and having a normal probability density distribution, a case that seems to be sufficiently realistic with regard to the α-functions.

With the distributions

$$\frac{a}{\sqrt{\Pi}} \, e^{-a^2 x^2} \quad \text{and} \quad \frac{b}{\sqrt{\Pi}} \, e^{-b^2 y^2}$$

we find the probability density distribution for z for the case $x \geqslant 0$, $y \geqslant 0$

$$f(z) = \frac{a \; b}{\Pi} \int_0^\infty e^{-a^2 x^2 - \frac{b^2 z^2}{x^2}} \frac{dx}{x} \tag{8}$$

This integral cannot (?) be solved directly but differentiating twice with respect to z and changing integration variable we find the following Bessel differential equation

$$\frac{d}{dz}\left(z\,\frac{df}{dt}\right) = 4a^2b^2fz$$

which has the modified Bessel function

$$f(z) = \frac{2ab}{\Pi}\,K_o(2abz)$$

as solution when integration constants have been determined so that the total probability becomes 1. For the case a=1 and b=2, the two distribution functions for x and y and the resulting distribution f(z) are shown in fig. 1.

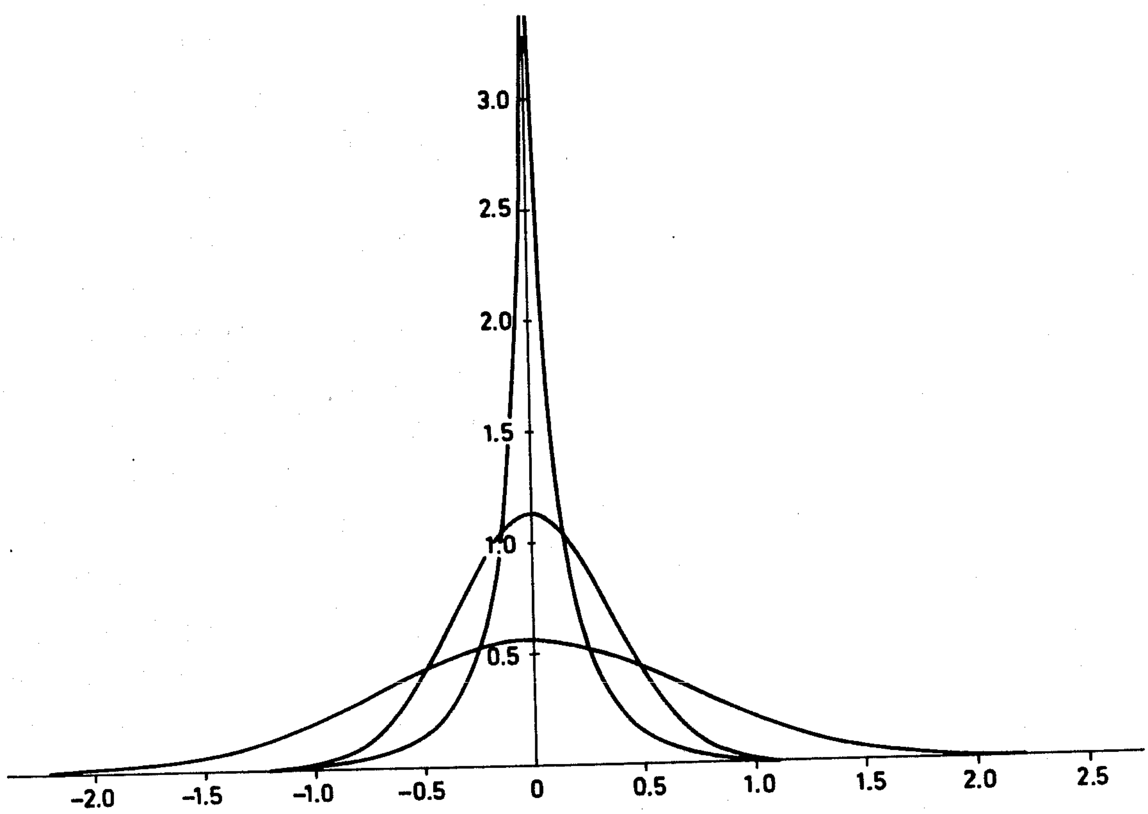

Fig. 1

It is seen that the possibility for small values of z is very large and we may conclude that this will be even more so in all cases where we have non-linear terms of higher order provided the factors are mutually independent. Therefore, terms of the type $\alpha_2^3(t)$ will dominate over terms of the type $\alpha_1(t)\ \alpha_2(t)\ \alpha_3(t)$ etc. and this will also provide guidance in the necessary truncations.

The same type of reasoning may also be applied to the last third order products of the functions h(x). Varying x they will also have a normal distribution and the probability for a third or fourth or higher order product of independent functions to be very different from zero becomes very small.

4. CONCLUSIONS

For a number of reasons it seems realistic to expect that the effective area of influence for a 12-hour timestep in a statistical forecasting model may be much more limited than indicated by conventional finite difference models. One seems also to have reason to believe that the high order non-linearity that would be natural to expect may be of limited importance and possibly restricted to only few important terms. However, the results shown here are too general to give sufficient guidance and further investigation is therefore needed. This may be carried out either in a theoretical way or by experiments. In one such experiment which is planned to be carried out in the near future, the 12-hour height change at one grid point will be statistically predicted using as predictors the amplitude functions (time-dependent) of an EOF-expansion of the height field over an area of 9x9 surrounding grid points. The EOF-expansion will be truncated and non-linearity introduced only stepwise. The results will be compared with conventional barotropic forecasts. One may here add that if the results turn out to be promising, then the real difficulties will start.

Acknowledgement

Thanks are due to Dr Lars Gottschalk at SMHI for indicating the method by which equation (8) may be derived.

Reference

Lorenz, E.N., 1977: An experiment in non-linear statistical weather forecasting. Monthly Weather Review. Vol. 105.