1. Introduction

Present observing systems and even more so present computer systems produce very large amounts of data in the form of analyses and forecasts. It is necessary to find ways to compress this data into sets which can be handled in a simple and rational way and thereby eliminate information which is redundant for practical purposes.

In statistics so-called principal components, orthogonal functions derived from data through a process of successive minimisation, have been applied since the 1930s. Twenty years later principal components were first applied in meteorology by Fukuoka (1951). Later, Lorenz (1956), Obukhov (1960) and Holmström (1963) independently arrived at the same method and the notation "empirical orthogonal functions" (eof) was introduced into the meteorological nomenclature by Lorenz. The method of eof's has therefore been known for a long time, having been introduced into meteorology by a pioneer whose work had been forgotten.

Eof's are not only a method of compressing data sets in an optimum way. It is also possible to identify some of the individual functions and, at least in a tentative way, associate them with a particular physical process. Some eigen-functions of simplified atmospheric models are similar to eof's calculated from atmospheric observations. During the last 15 years a substantial development has taken place in the field of eof's and the methodology has been applied in many meteorological fields.

In order to review recent work and to investigate to what degree eof's can be used in the work by the Centre, a workshop was held from 2 - 4 November 1977. It is our hope that this workshop, which is the first international meeting in this field, should stimulate the interest in the application of eof's and activate the international cooperation in this field.

The present volume contains a further short discussion on the nature of eof's, a number of examples of applications as well as a list of references, and the scientific lectures which were given during the workshop. The following scientists took part in the meeting:

H. Fechner	Institut für Meereskunde an der Universität Kiel Düsternbrooker Weg 20 D-2300 Kiel 1
I. Holmström	The Swedish Meteorological and Hydrological Institute Fack S-601 01 Norrköping
S. Järvenoja	University of Helsinki Department of Meteorology Hallituskatu 11-13 00100 Helsinki 10
V. Karhila	University of Helsinki Department of Meteorology (as above)
A. Kasahara	National Centre for Atmospheric Research P.O. Box 1470 Boulder, Colorado 80302 U.S.A.
J. Rinne	University of Helsinki Department of Meteorology (as for S.Järvenoja)
H. Törnevik	The Swedish Meteorological and Hydrological Institute (as for I.Holmström)
A. Törnvall	The Swedish Meteorological and Hydrological Institute (as above)

The following ECMWF staff members participated in the meeting:

K. Arpe
L. Bengtsson
D. Burridge
B. Machenhauer
H. Savijärvi
D. Williamson

2. The nature of eof's

There are at least three different approaches to this problem:

1) In general the functions can be found as eigenvectors of a covariance matrix. In this "mathematical" approach the functions are termed "eigenvectors".

2) In the statistical approach the functions are derived by minimizing the mean square residual. The term "principal components" is then normally used. Also the terminology "factor analysis" has also been used.

3) The third approach to eof's is that of applying variational analyses to a series expansion in order to determine the functions.

The eof's are a convenient way to represent data. They are empirical functions determined by the data rather than an arbitrarily chosen set of mathematical functions and hence the term "empirical orthogonal functions" introduced by Lorenz.

Mathematical functions are usually derived as eigensolutions of simplified differential equations. However, Holmström has shown that by applying the variational analysis to non-linear equations, one can derive mathematical functions resembling empirical orthogonal functions. The eof's could, therefore, be understood as eigensolutions of some equations which we know only approximately. Mathematical functions are defined once and for all, whereas different sets of eof's may be obtained from different limited data samples of the same variable. In most cases these differences are not important with regard to the convergence of the series. In their use the mathematical functions are usually ordered by scale, whereas the eof's often lack such distinct ordering. One of the advantages of mathematical functions is that derivatives and integrals can be obtained analytically or calculated very accurately. This is not always the case with eof's where approximate methods must be used. However, this limitation has so far not been found to be serious.

3. Areas of potential use

With increasing cost of instrumentation and labour it has become necessary to analyse critically the needs for meteorological, hydrological and oceanographical networks. In order to determine redundancy and to introduce a measure of optimization into network planning, investigations using eof analysis seem to be very promising. An example of such an investigation shedding new light on a mareograph network is given later in this report.

Expansion into eof of geopotential and wind data shows characteristic features of the structure of the atmosphere and of the energy distribution between the different modes. It may be possible to check if models reproduce the same modes and distribute energy in the same way as the real atmosphere. Comparisons of this kind could be an efficient tool, not only for testing models but also for analysing and determining points of weakness in the model construction. In the same way eof provide a basis for forecast verification.

One of the advantages with statistical methods is that we can relate a predictand to a number of predictors directly and thereby avoid an explicit treatment of intermediate physical processes which are incompletely known. This may possibly eliminate errors introduced by approximate physical-numerical models. A typical example is given later, where sea level changes are directly related to surface pressure changes, thus eliminating the transformation of pressure to wind and from wind to stress. It is believed that the same approach may be used in order to transform satellite radiance data into amplitudes of geopotential modes. The same approach may be used for local forecasts.

The amount of data in the form of observations as well as analysed and forecast fields that must be stored or transmitted between various meteorological centres continues to increase at a very fast rate, due to the introduction of models with higher resolution and the extension of forecasts to longer periods. Steps will have to be taken to compress the volume of data. In this process we must retain as much information as possible, or, in other words, one should eliminate insignificant information. An eof truncated expansion is a very efficient method as is shown later in this report.

The expansion of observed or analysed data into empirical orthogonal functions also provides time dependent amplitudes. Since they reflect variations at many points, they should in general show a rather smooth behaviour in time. Sudden deviations therefore are likely to indicate errors and the method may therefore be used as a means for data control. An example of this is given later in this report. The same approach could provide a method for filling in missing data.

For any numerical forecast there is a scale of motion below which the flow is unpredictable. This scale increases with the length of the forecast. In medium range forecasts therefore it is advisable to remove the detailed synoptic features before interpretation and verification.
There is no a priori method to define such a scale and filtering based on Fourier expansion or expansion using other analytic functions which have been used so far are not particularly satisfactory. Eof's can here be worth while to try for such a filtering procedure. Physically this will imply the assumption that lower modes are successively more predictable than higher modes.

It is also possible to formulate numerical models in eof's. This can be done by carrying out the expansion in the vertical or in the horizontal or in all three dimensions. For detailed weather forecasting such an approach can be questioned,since an expansion in empirical functions is unsuitable to predict extreme and rare events. Experimentation is necessary in order to determine the feasibility to construct efficient models utilizing eof expansions and to find out methods by which important physical processes may be incorporated in such a system.

One may here add the possibility to use statistical methods of this kind in order to construct models where the basic dynamical equations are utilized but where statistics are used to provide a mechanism in the model to account for processes about which our knowledge is insufficient for a correct deterministic formulation.