

EVAP MEMO NUMBER 10

PRINCIPAL COMPONENT ANALYSIS - PROGRAMS  
by P B WrightINTRODUCTION

This memo describes a set of programs in Metocode available for performing Principal Component Analysis on any set of data of maximum dimensions 31 x 90 or 10 x 279.

PROGRAMS F24 and F25

These require settings as follows:

- a. Number of variables V (maximum 31)
- b. Number of cases (usually days) on which a set of values is available (maximum 90)
- c. Number of eigenvectors wanted (maximum = number of variables)
- d. The value of  $V \times (V + 1) / 2$

F25 also requires:

- e. Identifier of Magtape
- f. Number(s) of block(s) of Magtape containing data.

F24 uses data punched on the end of the Directory Tape.

F252 uses one block of data in the format of  
Block 5 on MLZR0006 (daily May temperatures).

F253 uses 3 blocks of data each in the format of  
Block 5 on MLZR0006.

F254 uses 1 block of data in the format of  
Block 59 on MEPW0001 (evaporation by 4 methods).

Program performs PCA on the covariance matrix then on the correlation matrix.

An example of the results is shown as Annex 1 (available on request). Some versions also produce means of the eigenvector terms. Some versions print out the correlation matrix in triangular form.

PROGRAM F26

This program uses a set of V "artificial means" provided as a separate paper tape. It performs PCA on the (modified) covariance matrix. Data as for F252.

PROGRAM F27

This does the same as F26 except that a set of "artificial linear means" is calculated such that the first and last items are about the same as the true means. This is suitable for use with a time series, such as temperature, in which a linear variation is "expected" from beginning to end of the period. F270 uses data as for F252; F273 uses data as for F253.

/PROGRAM F28



## PROGRAM F28

This does the same as F25 except that:

maximum number of variables = 10  
maximum number of cases (days) = 279  
F280 uses data as for F254.

## MISSING OBSERVATIONS

It is assumed that the set of data is complete and does not include any missing observations. If missing observations are present, provision must be made for replacing them by suitable artificial data - see Meto programming manual instructions for Option M11.

## RESULTS

Evap Memo 9 describes the results of analyses of evaporation data.

Singularities Memo 4 describes the results of analyses of daily May temperatures.

Met 0 8  
Bracknell  
November 1969