LONDON, METEOROLOGICAL OFFICE.

Met.O.3 Technical Note No.20.

Root-mean-square interpolation errors for wind speeds in the United Kingdom. By SMITH, S.G.

London, Met.Off., Met.O.3 Tech.Note No.20, 1982, 31cm. Pp.7, 4 pls. 3 Refs.

An unofficial document - restriction on first page to be observed.

MET O 3 TECHNICAL NOTE NO. 20

ROOT-MEAN-SQUARE INTERPOLATION ERRORS FOR WIND SPEEDS IN THE UNITED KINGDOM

by

S G Smith

Climatological Services (Met O 3)
Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ

September 1982

1. Introduction

The work to be described was carried out during 1978 and 1979 but not formally completed. Two interim hand-written reports were produced during that period which exist in loose covers in room 128A. This report, written in September 1982, summarises the salient points but it is stressed that it is only a summary.

Requests are often made to the Meteorological Office for estimates of wind speed at a particular place and time in the past which are usually answered by subjective assessments of relevent data. The study here examines methods of objective estimation, namely interpolation, in order to

- provide a means by which the adequacy of the existing anemograph network for answering these enquiries can be gauged.
- investigate the effects of removing or adding stations at specified points in the network
- determine the relative merits of different forms of interpolation in the estimation procedure.

The work follows that of Hopkins (1977) who analysed the spatial variability of maximum and minimum temperatures and daily sunshine over East Anglia. In the present study hourly mean and maximum gust speeds in the hour are considered.

2. Theory

It was ultimately intended to produce a map showing expected root-mean-square (r.m.s.) errors of interpolation for any point in the U.K. based on estimates from the nearest n anemograph stations. An obvious way to achieve this would have been to use results for the observed r.m.s. error at each anemograph station. However this direct approach suffers from the fact that errors would fluctuate irregularly from station to station, since they would be heavily dependent on local site factors and larger scale topography. An indirect approach was therefore preferred, the theory of which is now presented.

Let $x_o' = x_o - \bar{x}_o$ be the deviation from the mean $\bar{x}_o$ at an interpolated point.

x_i' be the deviations from their respective means at the stations i used to estimate x_o' by interpolation ($i = 1$ to n)

p_i the weights applied to each of the stations

ϵ the observational error, inherent in all speed observations,

and N the number of cases.

Then the m.s. error of interpolation is given by

$$R^2 = \left(\sum_{i=1}^{n} p_i (x_i' + \epsilon) - x_o' \right)^2 / N \qquad - (1)$$

with summation over N cases implied.

Assume that the standard errors of observation are constant, that they are uncorrelated with the magnitude of the wind speed and mutually independent between stations.

Then, expanding (1),

$$R^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j c_{ij} + \sum_{i=1}^{n} p_i^2 \sigma_F^2 - 2 \sum_{i=1}^{n} p_i c_{io} + \sigma^2 \qquad - (2)$$

where c_{ij} is the covariance of stations i and j

c_{io} is the covariance of station i and the interpolated point

σ_F^2 is the squared standard error of observation

and σ^2 the variance, assumed constant for a given region.

Factorizing by σ^2,

$$R^2 = \sigma^2 \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j \mu_{ij} + \sum_{i=1}^{n} p_i^2 \frac{\sigma_F^2}{\sigma^2} - 2 \sum_{i=1}^{n} p_i \mu_{io} + 1 \right\} \qquad - (3)$$

where μ_{ij} is the correlation function i.e. correlation against site separation. The correlation field is assumed to be homogeneous and isotropic, so that the correlation depends only on distance apart.

Two forms of interpolation were considered

- non-optimal interpolation. The weights ρ_i were simply set inversely proportional to the distance from the interpolated point.

- optimal interpolation. The ρ_i are chosen to minimise R^2. This form of interpolation has been popularised, amongst others, by Gandin (1970) and has been applied within the Meteorological Office to rainfall amounts by Nicholass et al (1981). From (3), mininizing R^2 w.r.t. ρ_i gives

$$R^2_{opt} = \sigma^2 \left(1 - \sum_{i=1}^{n} \rho_i \mu_{io}\right) \qquad \text{-- (4)}$$

where the ρ_i satisfy $$\sum_{j=1}^{n} \mu_{ij} \rho_j + \frac{\sigma_F^2}{\sigma^2} \rho_i = \mu_{io} \quad (i = 1, n) \qquad \text{-- (5)}$$

3. Data

Development work was carried out on data from 13 stations in East Anglia and, separately, from 12 stations in the Central Lowlands of Scotland. Root-mean-square errors of interpolation were computed for all magnitudes of hourly means and of maximum gust speeds in the hour by extracting these variables for each station in the region at 12Z every other day in 1977 and 1978. Errors were also computed for 'strong' hourly means and for 'strong' gust speeds but using data for 1970-78. 'Strong' mean occasions were selected corresponding to the hour of day when the maximum gust speed of the day occurred at a specified reference station for the region, provided this speed exceeded 25 Kn. 'Strong' gust speed occasions were similarly defined but for hours when the maximum hourly mean exceeded 16 Kn. Interpolation errors were subsequently calculated for 12 more regions for 'strong' gust speeds only. These regions are displayed in fig 1; there was an average of about 10 stations in each region. Some areas of the U.K.

could not be assigned to one of these regions either because there were insufficient stations or because the topography was much too variable to assume homogeneity.

4. Analysis and Results

For East Anglia and the Central Lowlands of Scotland the correlation functions for gust and hourly mean speeds were determined using correlations from all possible station pairings in the region. In practice the Fisher Z transform was employed in preference to the correlation since it has more satisfactory statistical properties as correlations approach unity. All functions were fitted by expressions of the form

$Z(x) = a e^{-bx}$ where Z represents the Fisher Z transform

x is the site separation in Km

and a, b are constants.

The observed Z transform values and the fitted curve (given by the line of 'predicted' values) for East Anglia are shown in fig 2.

For σ^2 a mean value was determined using the stations in the region. σ_F^2 was computed using the approximation

$$\sigma_F^2 \approx \sigma^2 \left(\frac{1 - \rho_0}{\rho_0} \right) \qquad - (6)$$

where ρ_0 is the extrapolated value of the correlation function to zero distance.

For 'strong' gusts in East Anglia, $\sigma^2 = 40.07$ Kn^2 and $\sigma_F^2 = 11.29$ Kn^2.

The r.m.s. interpolation errors were then computed under a variety of assumptions to compare, for example, the effects of increasing the number of stations used in the interpolation and to compare optimal with non-optimal interpolation. Equation (3) was applied to derive R for non-optimal interpolation and equations (4) and (5) for the optimal case. Some findings were:

- insignificant decreases in R were observed as n was increased beyond 3
- of the four variables studied (viz all magnitudes of hourly mean and gust speeds and 'strong' means and gusts), the largest errors occurred for 'strong' gusts.
- errors for optimal interpolation were about 5-10% less than for non-optimal interpolation for the same number of stations.
- correlations of speeds were higher, and thus r.m.s. interpolation errors less, if the stations used in the interpolation and the interpolated point were coincident with the wind direction.

Root-mean-square errors for 'strong' gusts were calculated on a 10 Km grid for each of the 14 regions separately, using the three nearest stations at each grid point and optimal interpolation. Errors were determined for the network as it existed in 1979. An example of the output produced is shown for S.W. England in fig 3. (Contours over the sea should be ignored). Results for all regions were combined subjectively by eye and are displayed in fig 4. It is observed that R ranges from about 2 Kn in areas of the south-east of the U.K. to over 4 Kn in parts of Wales and Scotland. The lowest errors occur in areas where the network is most dense and the terrain is most uniform. Conversely in mountainous areas, where the correlation between speeds is lowest, the network is sparse, enhancing the r.m.s. errors.

5. Concluding remarks

Qualitatively the three aims of this study listed in the introduction have been fulfilled. However quantatively there are doubts about the validity of the results. These arise because

- assumptions required for the estimation of r.m.s. errors are that within each region the wind field is homogeneous and that each station observes speeds representative of its region. Only in inland East Anglia are these conditions likely to be met

- a comparison of estimated r.m.s. errors derived in this study with those calculated directly from the data at each anemograph station showed the latter to be consistently greater, particularly as the number of stations in the interpolation increased.
- for some areas, such as inland Scotland and Wales, the topography is too complex and the network too sparse to derive meaningful values for the r.m.s. interpolation errors.

6. References

Gandin, L. S. 1970 The planning of meteorological station networks. W.M.O. Tech. Note No. 111.

Hopkins, J. S. 1977 Spatial variation of daily temperatures and sunshine over uniform terrain. Met. Mag., 106, pp. 278-292.

Nicholass, C. A., O'Connell, P. E. Senior, M. R. 1981 Rain-gauge network rationalisation and its advantages. Met. Mag., 110, pp. 92-102.

FIG 1

Areas used to determine correlation functions and stations ▲ used in mapping of interpolation errors

FIG 2

PLOT OF Z TRANSFORM AGAINST DISTANCE

FOR 'STRONG' GUSTS IN EAST ANGLIA

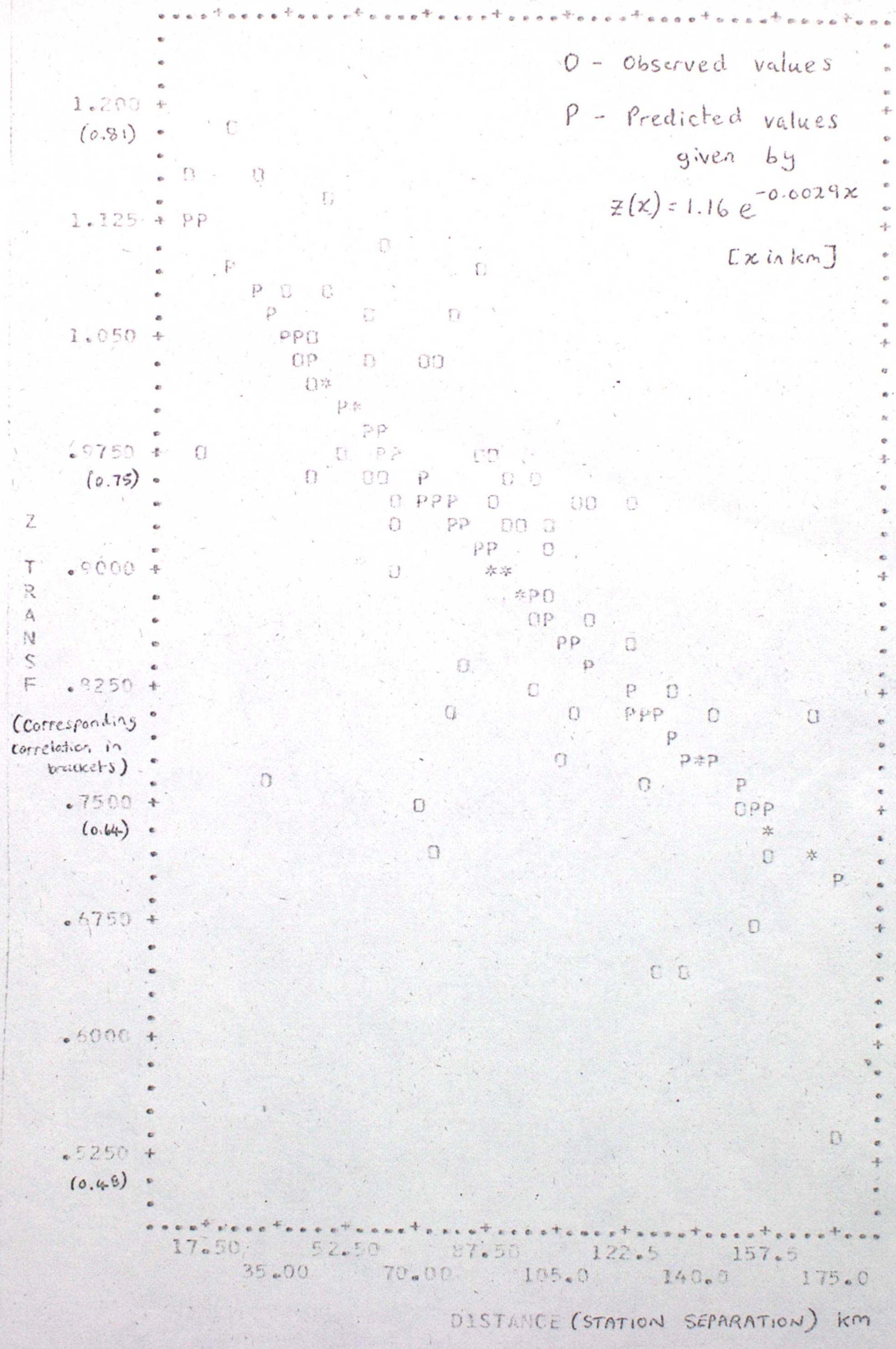
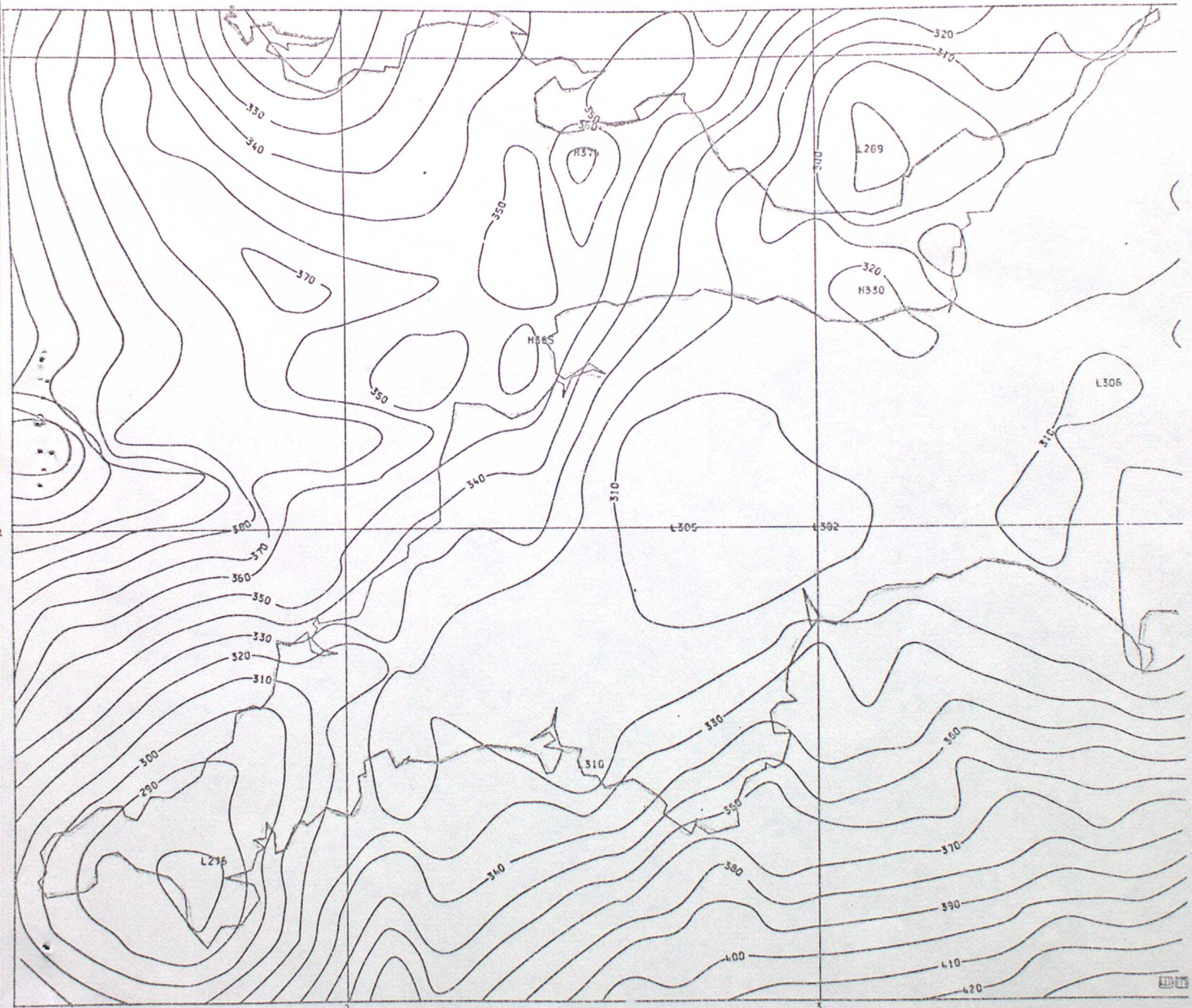


Fig 3

Root-mean-square interpolation errors for 'strong' gusts using nearest 3 stations and optimal interpolation.

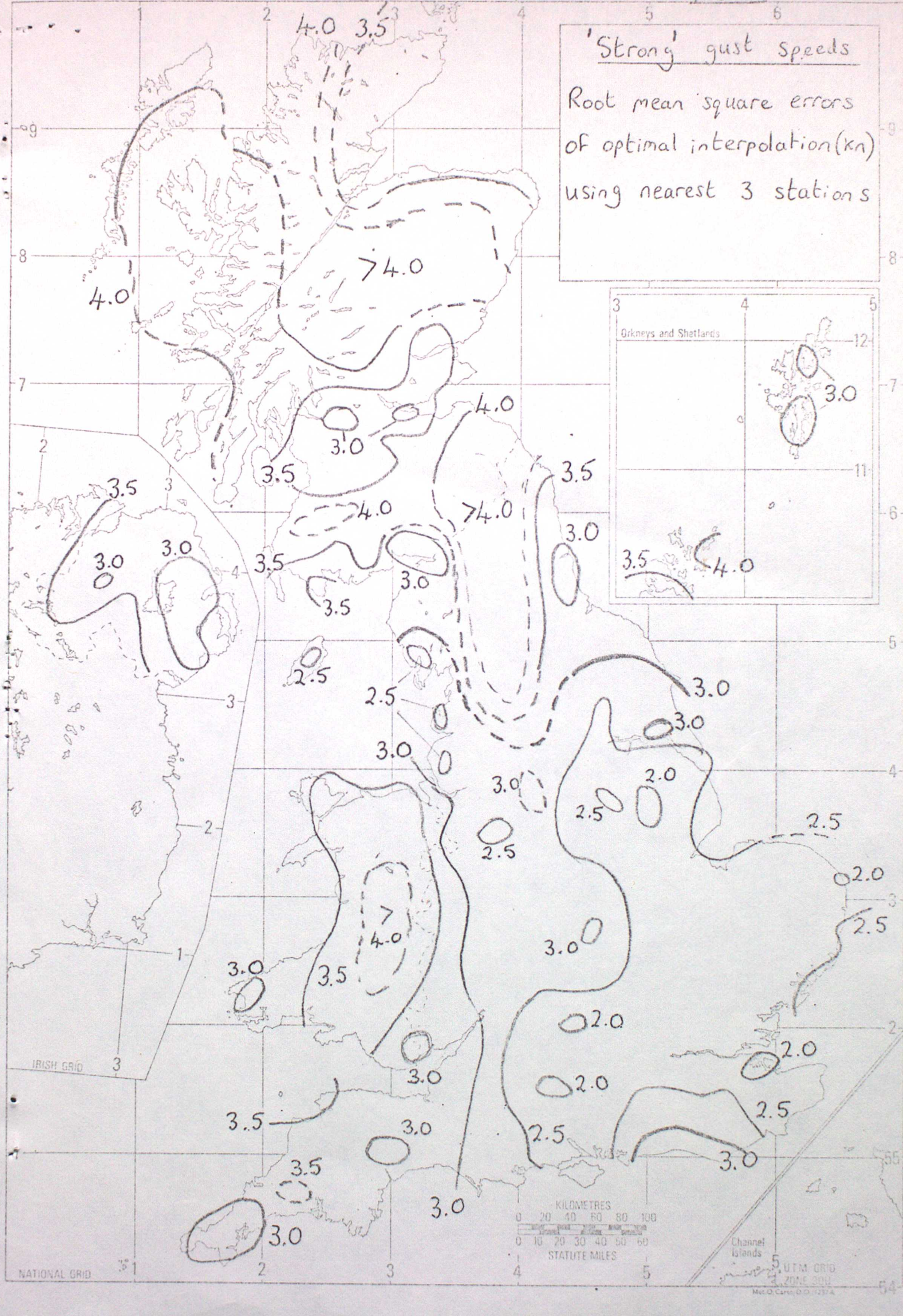
S.W. England

(values are in hundredths of knots)



RMS INTERPOLATION ERRORS - STRONG GUSTS
USING NEAREST 3 STATIONS OPTIMALLY

PRODUCED BY JOB MO3NSGES ON 25 OCT 79 AT 000



'Strong' gust speeds
Root mean square errors of optimal interpolation (kn) using nearest 3 stations
Orkneys and Shetlands
>4.0
4.0
3.5
3.0
2.5
2.0
IRISH GRID
NATIONAL GRID
KILOMETRES
0 20 40 60 80 100
0 10 20 30 40 50 60
STATUTE MILES
Channel Islands
UTM GRID ZONE 30U