Technical Memorandum No. 15

U.S. Joint Numerical Weather Prediction Unit

LIBRARY
MAY 13 1959
WEATHER BUREAU

LIBRARY
FEB 25 2004
National Oceanic &
Atmospheric Administration
U.S. Dept. of Commerce

Multiple Linear Regression Equations Expressing Heights of Certain Pressure Surfaces

RAREBOOK
QC
996
.T33
no. 15

John A. Brown, Jr.
Capt., U.S. Air Force

April 7, 1959

USCOMM-WB-DC

104 478

National Oceanic and Atmospheric Administration

U.S. Joint Numerical Weather Prediction Unit

ERRATA NOTICE

One or more conditions of the original document may affect the quality of the image, such as:

Discolored pages
Faded or light ink
Binding intrudes into the text

This has been a co-operative project between the NOAA Central Library, National Center for Environmental Prediction and the U.S. Air Force. This project includes the imaging of the full text of each document. To view the original documents, please contact the NOAA Central Library in Silver Spring, MD at (301) 713-2607 x124 or www.reference@nodc.noaa.gov.

LASON
Imaging Contractor
12200 Kiln Court
Beltsville, MD 20704-1387
April 13, 2004

At the present time the JNWP machine analysis scheme, constructed by Dr. Cressman, requires an estimate (first guess) of the height of the pressure surface being analyzed. In the operational analyses the first guess allows maximum height errors of 213 meters and maximum wind errors of 53.6 meters per second. These tolerances can easily be adjusted according to the quality of the first guess, and with a certain amount of smoothing a final analysis is obtained. This means, of course, that in regions of the octagonal grid containing no data reports the first guess determines the final analysis. Hence a good first guess is desirable. In cases of analyzing more than two levels it is also desirable to have a convenient means of height interpolation when the height at an intermediate level is missing or garbled.

It was the purpose of this study to obtain an expedient means of interpolation and extrapolation to the 300-mb, 250-mb, and 200-mb levels given data at other pressure surfaces.

Multiple linear regression equations based on rawinsonde data have been obtained with the aid of JNWP 704 library program WP214 written by Mr. Frank Lewis, Chief, Electronic Computer Branch, Air Weather Service. All data samples were obtained from randomly chosen reporting stations over the northern hemisphere which are included in the JNWP octagonal grid. The annual data samples were obtained from observations during the months of May and July of 1958 and January, March, September, and November of 1956. Reports from both observation times were used. The dates considered in each month were the 5th, 10th, 19th, 25th, and 29th. The November sample for equations 9, 11, and 12 was obtained from observations made in 1958 on the 2nd, 10th, 16th, 20th, 27th, and 28th day of the month. The October sample was obtained from reports from all days of the month in 1956 and 1958.

In each case the multiple correlation coefficient ($r = r_{xyz}$) and the standard error of estimate (S_e) are given. Here

$$r_{xyz} = \left(\frac{r_{xy}^2 + r_{xz}^2 + r_{yz}^2 - 2r_{xy}\, r_{xz}\, r_{yz}}{1 - r_{yz}^2} \right)^{\frac{1}{2}}$$

where $$r_{xy} = \frac{\sum (x_i - \bar{x})\ (y_i - \bar{y})}{N \sigma_x\ \sigma_y},$$

x = dependent variable,

y and z = independent variables,

$\bar{x}$ = mean of x,

N = number of sets of observations, and

σ = standard deviation.

For large samples

$$S_e = \sigma_x (1 - r^2)^{\frac{1}{2}}$$

The following notations apply:

Z = Height of the pressure surface in meters

T = Temperature of the pressure surface in tenths of degrees Centigrade

V = Wind speed at the pressure surface in meters per second

η = 500-mb absolute stream function vorticity times 10^6 sec^{-1}.

ξ = 500-mb relative stream function vorticity times 10^6 sec^{-1}.

The subscripts refer to the pressure surface in decibars.

The multiple linear regression equations obtained are listed below:

$$Z_3 = 346 + 1.7832 Z_5 - 0.7683 Z_{8.5} \qquad (1)$$

$r = 0.99218$

$S_e = 50$ meters

(From annual data. $N = 146$.)

$$Z_{2.5} = 870 + 1.9873Z_5 - 1.0844Z_{8.5} \quad (2)$$

$r = 0.98992$

$S_e = 61$ meters

(From annual data. $N = 146.$)

$$Z_{2.5} = 1077 + 1.8378Z_5 - 0.6521Z_{8.5} \quad (3)$$

$r = 0.98850$

$S_e = 58$ meters

(From October data. $N = 300.$)

$$Z_{2.5} = 526 + 1.1989Z_3 - 0.2056Z_5 \quad (4)$$

$r = 0.99768$

$S_e = 29$ meters

(From annual data. $N = 146.$)

$$Z_{2.5} = -262 + 0.5646Z_3 + 0.4620Z_2 \quad (5)$$

$r = 0.99933$

$S_e = 16$ meters

(From annual data. $N = 146.$)

$$Z_{2.5} = 1677 + 0.9746Z_3 + 0.5383T_3 \quad (6)$$

$r = 0.99962$

$S_e = 12$ meters

(From annual data. $N = 140.$)

$$V_{2.5} = 3.3775 + 0.9803V_3 \quad (7)$$

$r = 0.90327$

$S_e = 6.5$ meters per second

(From annual data. $N = 94$)

$$z_2 = 2023 + 2.1371z_5 - 1.4618z_{8.5} \quad (8)$$

$r = 0.98478$

$S_e = 76$ meters

(From annual data. N = 146.)

$$z_2 = 2966 + 1.8171z_5 - 0.8758z_{8.5} \quad (9)$$

$r = 0.97941$

$S_e = 81$ meters

(From November data. N = 101.)

$$z_2 = 1666 + 1.4552z_3 - 0.5738z_5 \quad (10)$$

$r = 0.99325$

$S_e = 51$ meters

(From annual data. N = 146.)

$$z_2 = 2894 + 1.8298z_{8.5} - 0.8809_{8.5} + 0.0848\eta \quad (11)$$

$r = 0.97942$

$S_e = 81$ meters

(From November data. N = 101.)

$$z_2 = 2715 + 1.8704z_5 - 0.9094z_{8.5} + 0.8581\xi \quad (12)$$

$r = 0.98092$

$S_e = 78$ meters

(From November data. N = 101.)

$$V_2 = 4.6104 + 0.9141V_3 \quad (13)$$

$r = 0.83207$

$S_e = 8.6$ meters per second

(From annual data. N = 86.)

Through subjective comparisons of analyses and extrapolations from 850 mbs and 500 mbs to 300 mbs, 250 mbs, and 200 mbs it appears that wind magnitudes in the vicinity of jet streams are underestimated to some extent. At the 200-mb level in particular the heights in the southwestern Pacific and Caribbean areas are usually estimated too low.

It seems worthwhile to mention here that in the cases considered the absolute vorticity and relative vorticity expressed little of the variation of the 200-mb level not already expressed by the 500-mb and 850-mb heights.

As can be seen from the multiple correlation coefficients and the standard errors of estimate, heights which were extrapolated through the tropopause gave considerable scatter. It is conceivable that linear regression equations obtained from height space-means and time-means might deviate considerably from those obtained from particular geographical locations and from particular seasons of the year.

Only linear equations were obtained in this study. Of course this is not meant to suggest that these should give the best estimates. Non-linear relationships of the independent variables and the dependent variables may in some cases give better results. Considering other independent variables may also give improved estimates of the dependent variables.

104 478