U. S. DEPARTMENT OF COMMERCE
WEATHER BUREAU
NATIONAL METEOROLOGICAL CENTER

September 1964

A PROPOSED ANALYSIS SCHEME

OFFICE NOTE NO. 21

Lt. Col. Harold A. Bedient
United States Air Force

A PROPOSED ANALYSIS SCHEME

The analysis scheme in use at Suitland has not changed basically since the commencement of operation of the IBM 7090 machine. Such changes as have been made are improvements to the first guess (for a parameter) by extrapolation or interpolation so as to improve the results. This note considers what might be done to improve the treatment of large numbers of aircraft reports, the extension of the analysis to the tropics, and the improvement of wind analysis.

The Suitland scheme is straight-forward in the treatment of single height observations. An observation of a height and a wind imposes a linear wind field about the observation point, but this is not destructive. A wind observation alone is not handled very well. The procedure of taking the guess height value at the wind point and treating this as an observation of height and wind results in the wind "pinching" the stream field here and there but not in a uniform way. A large number of wind observations result in cancellation since the guess heights inserted at the wind observation points may not change. Too many wind only reports, therefore, are harmful. This is why this program has not been successful in direct stream-function analysis.

The u and v analysis have been done regularly with this code. Usually, the guess has been the geostrophic wind. It has been found that the divergence calculated from this analysis is somewhat too high, and large-scale control has to be supplied by the height analysis to prevent erroneous flows.

To correct the excessive divergence, the Honolulu version is smoothed by computing a stream function after an analysis pass then substituting the resulting stream-function wind as a new guess. This is done after the first two scans to eliminate the large-scale spurious divergence caused by the data weighting function. The stream function derived in this way has a freedom that the Suitland analysis does not have as the stream function can rise or fall to any value since the guess values of the stream function do not enter as data. Thus, there is no difficulty with a large volume of data.

On the other hand, this scheme lacks large-scale control and must be controlled by manual insertion of bogus data rather than by its own mechanisms. The stream function could be used to generate a height field derived from wind only plus an arbitrary constant. This height field could be generated from the geostrophic law or the balanced wind law or any other wind law. The resulting height field could be used as "offset differences" to combine the height and wind observations. The resulting height analysis could feed back to the stream function and hence to the wind analysis.

This is a form of analysis similar to that proposed by Sasaki, but is not applied point-wise. The total field is integrated rather than a series of overlapping Green functions. The following procedure describes exactly how this procedure could be carried out.

First, the u and v analysis is done. Some further remarks will be made about this step. Second, the stream function is found. Experience in Honolulu shows that this must be done carefully. A "Neuman boundary procedure" is recommended. The following steps find the stream function

$$\psi = \psi_1 + \psi_2$$

$$(1) \qquad \nabla^2 \psi_1 = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad \text{interior}$$

$$\psi_1 = \text{arbitrary on the boundary}$$

$$\nabla^2 \psi_2 = 0 \quad \text{interior}$$

$$(2) \qquad \nabla^2 \Phi = 0$$

$$\frac{\partial \psi_2}{\partial n} = \frac{\partial \Phi}{\partial s} = V_s - \frac{\partial \psi_1}{\partial n} \quad \text{boundary}$$

this is used to find Φ which is the conjugate function of ψ_2

then

$$(3) \qquad \frac{\partial \psi_2}{\partial x} = -\frac{\partial \Phi}{\partial y} \qquad \frac{\partial \psi_2}{\partial y} = \frac{\partial \Phi}{\partial x}$$

these steps are only used to find the boundary values of ψ as at this point we do not need ψ itself. Now the balance equation provides

$$(4) \qquad \nabla^2 gz = f\nabla^2 \psi + 2J(u,v) - \beta u + \ldots$$

Note that higher order corrections to the geostrophic vorticity will not substantially increase running time as they do not increase the number of relaxations required. Another higher order correction that might profitably be used is the divergence due to the advection of the low-level wind over the mountains. This would account for the tendency of the low-level wind to blow across the contours on the great plains in a cold outbreak, and also for "lee-side" low pressure. If this were done it might be necessary for consistency sake to introduce the χ wind component due to this in the reanalysis of the wind. If the initial ψ tendency for the analysis time were available from a dynamic model it could be used to make the turning of the wind correction of the gradient wind. The value of including these higher order times would have to be demonstrated by experiment. The boundary values of z would be obtained by

$$(5) \qquad \frac{\partial gz}{\partial s} = \frac{f \partial \psi}{\partial s}$$

If this boundary procedure is applied over the equator, caution is necessary. It is necessary to take a height value in the southern hemisphere and one in the northern hemisphere and to make proportional changes on the north-south sides as peculiar results can occur if an integration is made all around the boundary and the error of closure is uniformly distributed. (This will be further explained.)

From this integration a height field specified by the wind alone is available. The height analysis is made thus wise. The height value at a point is

$$(6) \qquad H = \frac{\Sigma\, [H_o + (H_I - H_G)]\, \omega_i}{\Sigma\, \omega_i}$$

where H_o is an observed height H_I is the height interpolated at the observation point in the "wind-height" field, H_G is the grid point height in the "wind-height" field, and ω_i is the weighting function.

It should be noted that formula (6) is essentially the same as the "wind" correction used in the Suitland analysis. However, $H_I - H_G$ is not restricted to being taken from linearized wind measurements. Guess values of height are not used to utilize wind-only observations. Wind-only observations appear in the detail of the wind analysis and are essentially tied laterally to a true measured height value. This is the fundamental difference from the present analysis scheme used at Suitland.

Since the "wind-height" field has been integrated over the field with due consideration to variation of the Coriolis parameter and map-factor, the relative weighting functions of the height analysis and the wind analysis can be made truly independent to produce desired results.

The conventional weighting function at Suitland is

$$(7) \qquad \omega = \frac{N^2 - R^2}{N^2 + R^2}$$

(where R is the grid distance between an observation point and N is an arbitrary parameter which expresses the weighting distance). When the weight is used as in formula (6) a grid point will receive the correction due to the weighted mean of the observation error.

If N is a large number so that many data points lie within N/2 of the data point then the analysis is heavily smoothed. If N is small then the analysis tends to draw each grid point so as to agree with its nearest neighbor. In a region where aircraft wind measurements are thin, if the winds are analyzed with N small and the heights with N large then the details in the analyses will be due to the winds and yet the height measurements will be satisfied in the large. An isolated height will be drawn for exactly. Its effect will spread smoothly over large areas so it will not obscure the details due to the wind observations. Conversely, if N is small in the height analysis then the measure height gradients will be imposed on the gradients found by the wind measurements. Thus, it becomes possible to independently control the relative effect of height and wind measurements which is not possible in the present system.

This type of analysis makes it possible to use heights measured in the tropics. The gradients of heights in the tropics are notoriously poor as the uncertainty of the height measurement is as large as the geostrophic differences of two nearby stations. However, the heights are important in controlling the large-scale flow patterns so that the total flows over large regions are correct. The patterns from one hemisphere to another controls the position of the intertropical low. This means that N for the winds should be small and N for the heights should be large in areas where the winds are numerous relative to the heights, and in low latitudes. This will force heavy reliance on the winds. Conversely, N for the winds and N for the heights should be nearly equal where heights are reliable and dense.

After the procedure (1) through (4) has been carried out and the height has been analyzed, then the procedure must be reversed so that the effects of the height analysis are returned to the stream function. Caution must be exercised with equation (4) and (5) inversion because f approaches zero as the equator is approached. The corrections can be set up in the following form:

$$(8) \qquad \frac{\nabla^2 \psi'}{\nabla^2 \psi} = \frac{1/f\,(\nabla^2 g z' - 2 J_1 (u,v) - \ldots)}{1/f\,(\nabla^2 g z - 2 J_1 (u,v) - \ldots)}$$

where prime refers to the desired stream function from the revised heights and unprimed refers to the original stream and wind height.

Except at the equator itself $1/f$ can be cancelled out on the right and the values of $\nabla^2 \psi$ changed to the value of $\nabla^2 \psi'$ (the corrected stream vorticity) by the ratio on the right. It is proposed that the second order terms be left unchanged as this is only an approximate solution which is to be reanalyzed with data. The equatorial problem can be avoided by regarding the atmosphere as two regions, separated by the equator, which have to approach in the limit common values along a line at the equator in the manner of integration theory. Similarly

$$(9) \qquad \frac{\partial \psi' / \partial s}{\partial \psi / \partial s} = \frac{1/f(\partial g z' / \partial s)}{1/f(\partial g z / \partial s)}$$

Note that the normal boundary condition is not used here as hopefully changes will be small and will not disturb the general nature of the solution of the original stream function. Now it is possible to relax for ψ and thus get a non-divergent u and v guess which contains information consistently arrived at from both the first analysis of the wind and height data.

Steps (1), (2), and (3) return to u, v and analysis are already employed in Honolulu and produce a stable analysis which fits the observed wind to the same closeness as an uncontrolled u and v analysis. Height fields have been produced from the wind and look reasonable so, there is no reason to believe that the above procedure would not produce an improvement in all analyses fields if iterated several times. The small neglected times might preclude convergence in a mathematical sense (that is, it might not be possible to prove that it could be iterated enough times so that the difference $\nabla^2 \psi' - \nabla^2 \psi$ was everywhere less than a criterion that could be set arbitrarily small.)

Note that this procedure is free of arbitrary restrictions on the magnitude of the vorticity. It might be desirable to smooth regions of excessive anticyclonic vorticity relative to the hemisphere because spurious values could arise from the shape of the weighting functions and the accidental placement of data or erroneous observations. Since the wind and height fields are going to be independently controlled by the data, the ellipicity criterion (which essentially selects one of two possible solutions) is not necessary.

The running time of this procedure is going to be greater than the present analysis scheme. Since the intention is to make consistent analysis in areas of thin data coverage, this scheme could be applied with a minimum number of levels to get the big picture, followed by a finer scale treatment which the relaxations of more layers in areas where data coverage warranted.

Data coverage is not going to substantially increase over the present. However, exotic data from satellites may eventually be of assistance. The present conventional data coverage is reasonably dense over continental areas, and perhaps in Europe is even a little overly dense. Conventional data has internal temperature and wind consistency and gives values at all levels to nearly 100 millibars. Improvements could be made in checking, collecting, and transmitting these data so as to eliminate errors and ensure timely arrival. Surface data are excessively dense at least for large-scale analysis and if the ocean ships are properly collected and evaluated, reasonable surface analysis cannot be done over most of the earth. Satellite pictures can probably benefit the surface analysis most quickly by filling in some holes.

The other large source of data which is hardly touched for machine purposes is aircraft reports. Advent of large range jet transport has suddenly filled the ocean areas with strings of doppler wind reports between 300 mb and 200 mb. Because of the range of the aircraft they are not channeled to narrow "island hopping" routes but are striking out into areas where observations have never before been made.

This distribution of data leads to the following recommended analysis levels for large-scale analysis. The surface should be analyzed using primarily surface data with a guess constructed from a consistent forecast model. The analysis of u and v should be done at 700, 300, 250, and 200 mbs. Airplane reports should be interpolated vertically as well as horizontally. There should be a vertical weighting function so that each level gets the heaviest weight from aircraft nearest its level and a broken line interpolation should be used in the 300 to 200 mb layers to allow for a sharp maximum wind layer in these levels. The greatest densities of aircraft data would be expected around 700 mbs and in the 300 to 200 mb layers. One scan of the analysis should be made at each level. The value of N should be about three grid lengths. An elliptical weighting function could be used but probably it would be better used in Scan 2. An elliptical weighting function is one of the form

$$\omega_2 = \left(\frac{N^2 - R^2}{N^2 + R^2}\right)\left(1 - \frac{\psi_i \psi_G}{K}\right)$$

where ψ_i is the interpolated value of the stream function at the observation and ψ_G is the grid point value. K is an arbitrary parameter and neither quantity in parenthesis is allowed to be negative.

Since the aircraft data are of variable quality and quantity and should not be allowed to submerge checked observations by sheer volume, the following scheme is being prepared in Honolulu.

$$\Delta u = \frac{\omega_{mm}}{\omega_{mf} + \omega_{ma}} \left[\omega_{mf} \frac{\Sigma \omega_i (u_f - u_i)}{\Sigma \omega_i} + \omega_{ma} \frac{\Sigma \omega_i (u_a - u_i)}{\Sigma \omega_i} \right]$$

where Δu is the correction to the guess of element u, ω_i is the weight of an individual observation, u_f is a fixed observation of an element, u_i is an interpolation in the guess field at the observation point. An aircraft observation of the element is u_a, ω_{ma} is the largest aircraft weight (i.e. nearest aircraft) to a grid point, and ω_{mf} is the largest fixed weight (i.e. nearest). This effectively treats all airplane reports in a region as one composite observation and all fixed reports as one in the region likewise. The effect of an isolated observation dies off smoothly with distance due to ω_{mm} which is the maximum weight for either an aircraft or fixed observation.

The height observations and extrapolations are assumed to be thermally consistent and available at each level; so if a sounding is incomplete, it should be completed if possible by extrapolation and interpolation. The described procedure is then followed to force the wind analysis to be consistent with the height observations and extrapolations. This can be followed through two cycles. Then, guess values for any other levels can be interpolated and given two cycles at a scan radius of two to fill in internal detail where there are sufficient reliable observations.

The resulting fields then could be reconciled height to stream or stream to height and used as balanced initial data for a primitive equation model without further treatment. The extra running time could be absorbed from the initialization time of the forecast models.

Pieces of this procedure are already in existence. It is intended to carry this procedure out completely on Pacific data at Honolulu. Start should be made at Suitland to do this for the total area from the North Pole at 30° South where data presently exist to carry this out. Arrangements for collecting these tropical data have to be made.