METR 5303 – Lecture #6 - OBAN History

<u>Early Objective Analysis</u>

We will discuss some of the early objective analysis schemes, not for the intention of doing a thorough historical review, but for the purpose of introducing various principles or components of the analysis procedures that are in use today.

<u>Panofsky</u> (1949 *J. Meteor.*)

Panofsky's work was done in association with the first NWP project with Charney and Von-Neumann at Princeton, as he was assigned the task of providing the initial conditions on a uniform grid for the model.

He used 2-D cubic polynomials fitted over regional subsets of N. America. For example, the 500 mb height analysis was given by

$$h(x,y) = \sum_{i,j} a_{ij} x^i y^j, \qquad i + j \le 3$$

Verify that this causes 10 a_{ij} values to be determined. He realized that an over-determined analysis was best, so he used 12-14 height values per region, and, for the wind field (isotach) analysis, he used at least 20 values, since the wind field was assumed to be noisier.

Panofsky attempted to incorporate geostrophy in the contour analysis using a <u>direction constraint</u> only. That is, he sought to minimize

$$h(x,y) = \sum_{i,j} (\vec{\mathbf{V}} \cdot \nabla h)_{ij}^2$$

This quantity is small when $\nabla h \perp \vec{\mathbf{V}}$ or $h \parallel \vec{\mathbf{V}}$.

Note: To use both speed and direction as geostrophic constraints, one would have to minimize

$$E = b\sum_{i=1}^{M}(h-h_o)_i^2 + \sum_{i=1}^{N}(u_0-u_g)_i^2 + \sum_{i=1}^{N}(v_0-v_g)_i^2 \,,$$

where $b = \dfrac{\sigma_h^{-2}}{\sigma_u^{-2}} \sim 0.3\ \mathrm{s}^{-2}$ is a value that was used in the early days.

<u>Panofsky contributions include</u>:

1. Outlined many of the now familiar problems that need to be addressed in objective analysis.
2. Showed difficulties of polynomial fitting in meteorology – especially in data sparse regions. This technique was never used operationally.
3. Each kind of observation was weighted according to its accuracy.
4. Dynamic constraints (geostrophy) incorporated directly. This is also an advantage of a least squares minimization system where equation constraints can be incorporated.

<u>Gilchrist and Cressman</u> (1954 *Tellus*)

They fitted 2nd-order polynomials over a square region using least squares.

phere. The polynomial can be expressed in the form:

$$D = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2$$

or in condensed form:

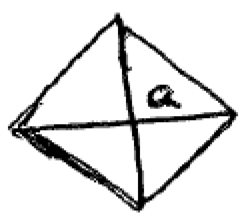
$$D = \sum_{i+j=0}^{i+j=2} a_{ij}\, x^i y^j \qquad (1)$$

Their contributions were:

1. Determined new polynomials for each grid point; that is, as in Daley's Sec. 2.1, they did a localized fit, where the data were fitted over a region to determine one grid point value at the center of the region.

2. Introduced the concept of an <u>influence area</u>; here the limits of the area was defined by $|x| + |y| = |a|$

   They did anticipate the use of circular areas.



3. Used multiple passes through the data, with changing influence radius

4. Devised procedure for determining the relative weights for σ_h , σ_v

5. Anticipated use of the first guess (background) field.

6. Stressed need for automatic data processing with quality control of the observations.

This is the last we will hear of localized function fitting. Recall that the advantages and disadvantages of function fitting include:

Advantages

1. Good for fixed no. of stations - with no large gaps

2. Best if have a relatively small number of observations

3. No first guess (background) field is required

4. Possible to account for observational error

Disadvantages

1. Has potential problems of over-fitting, under-fitting, and misspecification of the function basis.

2. Poor in data sparse regions and on the edges of the domain

3. Becomes expensive when N is large

4. Ignores meteorological knowledge except when constraints are imposed.

Bergthorsson and Doos (1955 *Tellus*)

Their contributions include:

1. First use of a background field (called a "preliminary field" in this paper).

   The background field $f_B(\mathbf{r_i})$ was obtained at each grid point from

$$f_B(\mathbf{r_i}) = \frac{\mu_F f_F(\mathbf{r_i}) + \mu_N f_N(\mathbf{r_i})}{\mu_F + \mu_N}$$

where $f_F(\mathbf{r_i})$ was a 12 hr forecast at grid point $\mathbf{r_i}$

$f_N(\mathbf{r_i})$ was a climatological normal at grid point $\mathbf{r_i}$

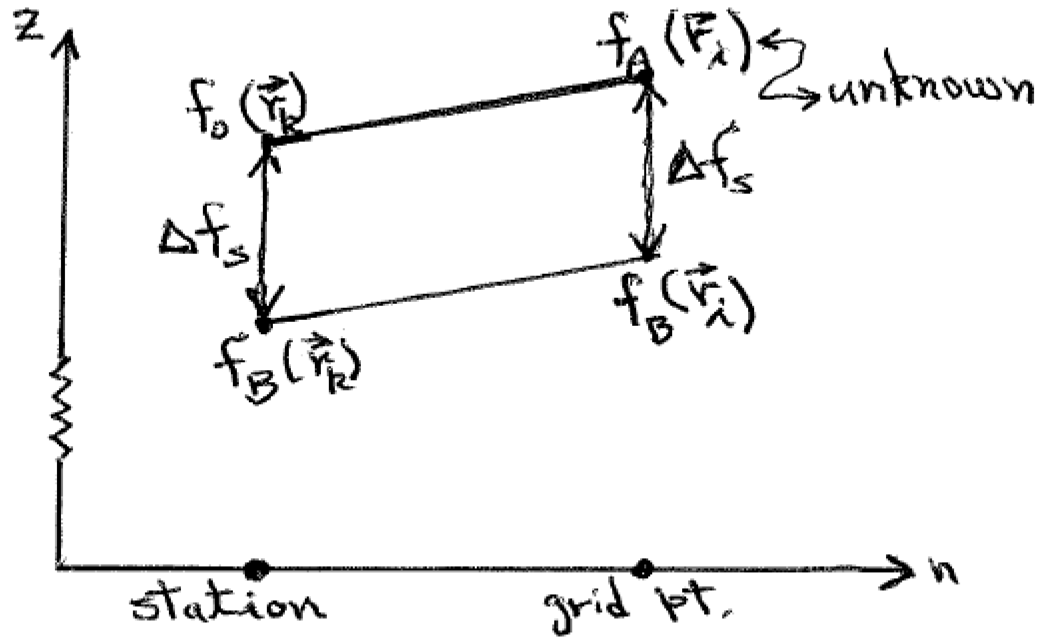
In B&D, $\mu_F = c_1 / \sigma_F$, $\mu_N = c_2 / \sigma_N$, where σ_F (σ_N) were root mean square differences between many forecasts (climatology) and observed values.

Note that we have seen in Daley that the most probable analysis results from using variance as the measure of error; that is, $E_B^2 = \langle \varepsilon_B^2(\mathbf{r}) \rangle$, the expected background error variance.

In B&D, there is a map of the ratio $\dfrac{\mu_F}{\mu_F + \mu_N}$ that shows the geographical distribution of how the forecast and climatological data were weighted relative to each other.

2. First use of a discrepancy (observation increment) analysis

a. Consider heights only first. We illustrate this concept using the *n*-*z* cross-section diagram to the right:



We know 3 of the values shown in this figure:

$f_o(\mathbf{r_k})$ - observed value at the station

$f_B(\mathbf{r_k})$ – background value at the station (obtained via interpolation)

$f_B(\mathbf{r_i})$ - background value at the grid pt.

We wish to obtain $f_A(\mathbf{r_i})$ - the analysis value at the grid point. We assume that the difference

$\Delta f_s = f_o(\mathbf{r_k}) - f_B(\mathbf{r_k})$ at the station is the same at nearby grid points

(within a reasonable distance). Thus

$$f_A'(\mathbf{r_i}) = f_B(\mathbf{r_i}) + [f_o(\mathbf{r_k}) - f_B(\mathbf{r_k})] \qquad (1)$$

Now try to weight f_B and f_A' in an optimal way:

$$f_A(\mathbf{r_i}) = \frac{E_B^{-2} f_B(\mathbf{r_i}) + E_O^{-2}(k)\ w(\mathbf{r_k} - \mathbf{r_i})\ f_A'(\mathbf{r_i})}{E_B^{-2} + E_O^{-2}(k)\ w(\mathbf{r_k} - \mathbf{r_i})} \qquad (2)$$

where $w(\mathbf{r_k} - \mathbf{r_i})$ is a generalized distant-dependent weight function.

If $w(\mathbf{r_k} - \mathbf{r_i}) = 1$, $\mathbf{r_k} = \mathbf{r_i}$, meaning that the ob. and grid pt. are co-located, and (2) becomes

$$f_A(\mathbf{r_i}) = \frac{E_B^{-2} f_B(\mathbf{r_i}) + E_O^{-2}(k)\ f_O(\mathbf{r_i})}{E_B^{-2} + E_O^{-2}} \qquad (3)$$

Recall that E_o^{-2} is the expected observational error variance.

When $\mathbf{r_k} - \mathbf{r_i} \rightarrow \infty$, the weight function $w(\mathbf{r_k} - \mathbf{r_i})$ should go to 0 and $f_A(r_i) \rightarrow f_B(r_i)$. That is, the background field becomes the analysis value when they are no nearby observations.

Inserting eq. (1) into (2), and rearranging we have

$$f_A(\mathbf{r_i}) - f_B(\mathbf{r_i}) = \frac{E_O^{-2}(k) w(\mathbf{r}_k - \mathbf{r}_i)[f_o(\mathbf{r_k}) - f_B(\mathbf{r_k})]}{E_B^{-2} + E_O^{-2}(k)\ w(\mathbf{r}_k - \mathbf{r}_i)} \qquad (4)$$

Recall that the terms on the left represent the <u>analysis increment</u>, and the term on the right is a <u>weighted observation increment</u> (or “correction” or “discrepancy”).

Note: $w(\mathbf{r_k} - \mathbf{r_i})$ is a specified *“a priori”* weight. However, if we write eq. (4) as

$$f_A(\mathbf{r_i}) - f_B(\mathbf{r_i}) = W_{ik}\ [f_o(\mathbf{r_k}) - f_B(\mathbf{r_k})]$$

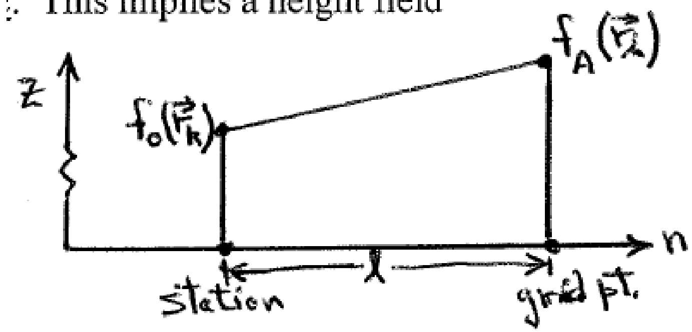
then we have the *"a posterori"* weight W_{ik} given by

$$W_{ik} = \frac{E_B{}^2(k)\,w(\mathbf{r_k} - \mathbf{r_i})}{E_B{}^2(k)\,w(\mathbf{r_k} - \mathbf{r_i}) + E_o{}^2(k)}$$

b. Now consider the case when heights and winds are used together (multivariate)

B&D made 2 additional estimates of f_A, $\tilde{f}_A$ and $\tilde{\tilde{f}}_A$:

(i) Assume $\mathbf{V}_o(\mathbf{r_k})$ at the station is geostrophic. ∴ This implies a height field gradient $\dfrac{\partial Z}{\partial n}$ which is used in the following way:

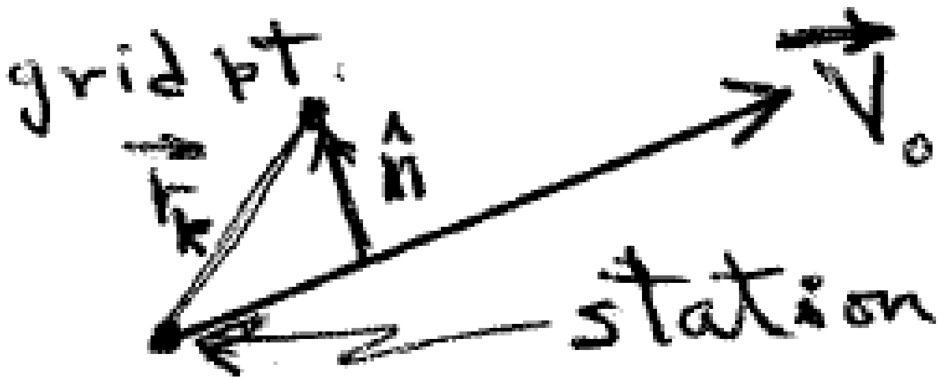


$$\tilde{f}_A(\mathbf{r_i}) = f_O(\mathbf{r_k}) + l \cdot \frac{\partial Z(\mathbf{r}_k)}{\partial n}. \qquad (5)$$

Class exercise: Show that for the Cartesian coordinate case that, for $f_A(\mathbf{r_i})$ = the height field $Z_A(\mathbf{r_i})$, we can write

$$Z_A(\mathbf{r_i}) = Z_O(\mathbf{r_k}) + (\mathbf{r}_i - \mathbf{r}_k)\cdot\nabla Z = \mathrm{Z}_o(\mathbf{r}_k) + \frac{f}{g}\,[(y_i - y_k)u_0 - (x_i - x_k)v_0]$$

where the diagram for this problem is



with f the Coriolis parameter, g is gravity, and the observed wind $\mathbf{V_o}$ is assumed geostrophic.

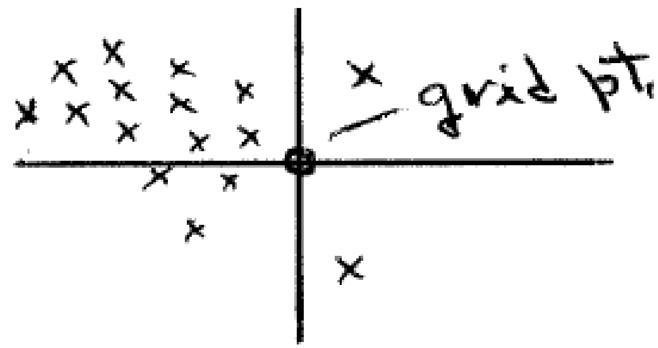
(ii) The third estimate of the analysis value $\tilde{\tilde{f}}_A$ is made by assuming that the background height gradient at the grid point is representative of the gradient between the grid point and the station:

$$\tilde{\tilde{f}}_A(\mathbf{r_i}) = f_O(\mathbf{r_k}) + l\cdot\frac{\partial Z_B(\mathbf{r}_k)}{\partial n} \qquad (6)$$

B&D used regression techniques to determine the best combination of f_B and f_A, $\tilde{f}_A$ and $\tilde{\tilde{f}}_A$. They found that the weights $w(\mathbf{r_k} - \mathbf{r_i})$ were primarily a function of the distance between the grid point and the stations, and gave the most weight to the closest stations (see Fig. 3.1 in Daley).

B&D also found that $E^2_o / E^2_b = 1/9$ at 500 mb, indicating that the observations were much more accurate than the 12-hr forecast.

3. A third contribution from B&D was that they tried to account for station density, although in an *ad hoc* way. The problem is that clustered observations are highly correlated with each other, and thus less independent; - they each should receive less weight than an isolated or "lonely" observation of the same accuracy.



B&D addressed this issue by modifying the weights by $1/\rho_i$ where ρ_i is the station density in the local area. This will be done more properly in the statistical methods.

Cressman Method (Cressman, 1959, *MWR*)

This is often called the method of successive corrections because the analysis was iterated. Cressman is most well known for introducing the Cressman (distance-dependent) weight function given by

$$w(r) = \begin{cases} \dfrac{R^2 - \tilde{r}^2}{R^2 + \tilde{r}^2}, & \tilde{r} \le R \\ 0 & \tilde{r} > R \end{cases} \tag{7}$$

where $\tilde{r}$ is the distance from the grid point to the station, and R is the influence radius. Eq. (7) was chosen simply because it was very fast and economical to use. Note that this weight function is isotropic, meaning that it is independent of direction. The weight function (7) has the following shape as a function of radial distance *R*:

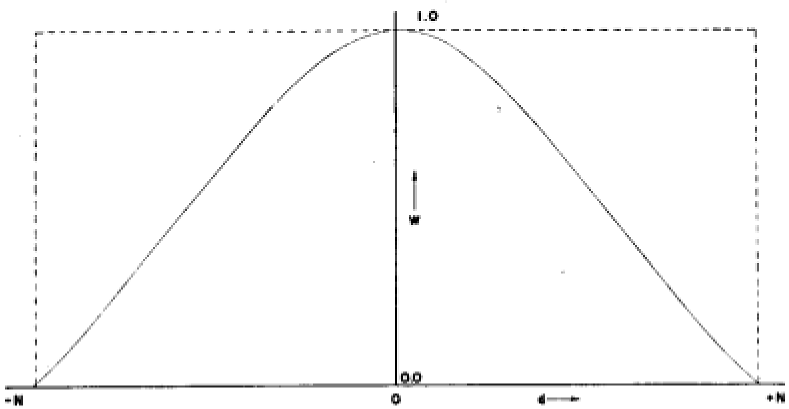


FIGURE 3.—Curve of the weighting function *W* vs. distance *d*. Solid line refers to equation (2). Dashed line refers to recent changes for scan 4 (see text).

As before, the corrections (obs. increments) were determined using a background field from a forecast. A height analysis increment using a height observation at *k* point is given by

$$C_h\left(\tilde{r}_i\right)=w_k(\tilde{r}_i)\left[Z_o(\tilde{r}_k) - Z_B(\tilde{r}_k)\right] \quad . \tag{8}$$

A height analysis increment using height and wind observations is given by:

$$C_{hv}(\tilde{r}_i)=w_k(\tilde{r}_i)\left[Z_o(\mathbf{r}_k)+ \frac{kf}{mg}(u\Delta y - v\Delta x) - Z_B(\mathbf{r}_\mathrm{i})\right]. \tag{9}$$

Note that in (9) the background height at the observation point is estimated using the background height at the grid point and the height gradient estimated from the wind measurement.

Here, the observed station winds *u* and *v* are assumed nearly geostrophic, *m* is the map factor (contains both our image and map scale factors), and *k* is the ratio V_g / V_{ob} which was set to 1.08. Thus final analysis equation used was, for all *i* grid points:

$$Z_A(\mathbf{r}_i) = Z_B(\mathbf{r}_i) + \frac{A\sum_{m=1}^{M} w(\tilde{r}_m)\left[Z_o(\tilde{r}_k) - Z_B(\tilde{r}_k)\right] + \sum_{n=1}^{N} w(\tilde{r}_n)\left[Z_o(\mathbf{r}_k) + \frac{kf}{mg}(u\Delta y - v\Delta x) - Z_B(\mathbf{r}_k)\right]}{A\sum_{m=1}^{M} w(\tilde{r}_m) + \sum_{n=1}^{N} w(\tilde{r}_n)} \tag{10}$$

where *M*, *N* were the number of height and wind obs. respectively, within the influence radius *R*, and *A* was set to ¼ which means less weight was given to an estimate from use of heights alone compared to the estimate using heights and winds together.

Cressman's second main contribution was successive scans (an iterated analysis).

Once the analysis Z_A was available at all grid points, the values are interpolated back to the stations – thus Z_A becomes the new background field, and the analysis procedure is repeated, usually using a different value of R

R should be large enough on the first "scan" (or "pass" or "iteration") so that all grid points have "sufficient" observations to produce an analysis. R is deceased in subsequent scans to allow more detail in the analysis where the observations are dense.

Cressman used:

scan no	R	
1	4.75 Δx	(heights only)
2	3.60 Δx	
3	2.20 Δx	
4	1.80 Δx	(here, w = 1 for all $\tilde{r}$ within R)

This yielded less smoothing (fitted the obs more closely) for each successive pass.

In the earlier days, Δx = 381 km at 60°N (known as "1 bedient").

Note: Decisions on R should always be made on the basis of average station separation distance d, not a particular grid spacing.

That is, for the same data base (e.g., raobs), an analysis using Δx = 400 vs Δx = 50 km should not be different because Δx = 50 km does *not* add new information to the analysis. This "mistake" above was not fatal for Cressman because his Δx is very close to the average distance between raob stations.

It is important to know d for any analysis you do. Here are two methods to estimate it:

(a) Compute the average distance between all the stations (as you did in Computer Prob. #1 where you got d ~ 345 km.)

(b) If know the area A of the analysis domain and assume that the N stations are randomly spaced, then

$$A = N d^2 \quad \text{and thus} \quad d = (A/N)^{1/2}$$