General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

NASA

Technical Memorandum 83851

THE GEMPAK BARNES OBJECTIVE ANALYSIS SCHEME

S. E. Koch, M. des Jardins,
and P. J. Kocin

NOVEMBER 1981

THE GEMPAK BARNES OBJECTIVE ANALYSIS SCHEME

S. E. Koch, M. des Jardins, and P. J. Kocin

THE GEMPAK BARNES OBJECTIVE ANALYSIS SCHEME

S. E. Koch, M. des Jardins, and P. J. Kocin

ABSTRACT

GEMPAK is an interactive computer software system developed at the Severe Storms Branch of NASA's Goddard Laboratory for Atmospheric Sciences for the purpose of assimilating, analyzing, and displaying various conventional and satellite meteorological data types. The Barnes (1973) objective map analysis scheme possesses certain characteristics that allowed it to be adapted to meet the analysis needs of GEMPAK. These characteristics and the specific adaptation of the scheme to GEMPAK are described here. A step-by-step guide for using the GEMPAK Barnes scheme on an interactive computer (in "real-time") to analyze various types of meteorological datasets is also presented.

The GEMPAK Barnes scheme is unique in the way in which it achieves a balance of objectivity, versatility, and practicality. Demonstration of these qualities is accomplished by applying the scheme to both a high quality, uniformly distributed radiosonde data set, and to a non-uniformly distributed data set of undetermined quality composed of satellite-derived cloud motion vectors.

CONTENTS

FIGURES

THE GEMPAK BARNES OBJECTIVE ANALYSIS SCHEME

1. AN OBJECTIVE ANALYSIS SCHEME FOR GEMPAK

One of the basic program objectives of the Severe Storms Branch (Code 914) of the Goddard Laboratory for Atmospheric Sciences (GLAS) is "to study the subsynoptic to mesoscale processes which create the favorable environment for the development and evolution of severe convective storms by using the recently acquired SESAME data sets, TIROS-N (and VAS) temperature and moisture soundings and satellite derived cloud tracked winds" (p. 44 of Simpson et al., 1980). In order to assess the impact that merging of the various data sources has upon the dynamic computations and subsequent interpretation of storm-environment processes, the GEMPAK (GEneral Meteorological data assimilation, analysis, and display software PACKage) was created. This system is designed to (1) convert the various types of data to a standard file structure, (2) vertically and horizontally interpolate the standardized data to uniformly spaced grid points, (3) perform data assimilations to initialize numerical models, (4) compute diagnostic quantities, and (5) display the products onto various maps. The GEMPAK has been developed on the AOIPS (Atmospheric and Oceanographic Information Processing System) at NASA/Goddard. The adaptation of the Barnes (1973) objective analysis scheme to GEMPAK for the purpose of performing the interpolation function (step 2) is discussed in this report.

The term "objective analysis" refers to a process by which data observed at irregularly spaced points in space (and/or time) are numerically processed to retrieve the two- or three-dimensional structure in the spatial distribution of some meteorological field parameter. Usually, the observed parameter values are computed at a regularly spaced array of grid points in order to provide the basis for mapping isopleths of the parameter distribution, and to allow quantitative diagnostic and/or predictive calculations to be made.

The process of objective analysis results in smoothing of both high-frequency, short-wavelength features commonly referred to as "noise" and of gross data errors in the data distribution. It is

necessary to suppress "noise" and reduce error amplitudes significantly, particularly before computation of meaningful estimates of such derivative quantities as divergence and vorticity from the grid point values.

The degree of smoothing is governed by the selected weights applied to the data. The Barnes scheme uses a Gaussian weighted-averaging method to assign a non-zero weight to each observation according to its distance S from a mesh point relative to the number and distance of all other observations within some "cutoff radius" R_c, beyond which all weights are assigned a value of exactly zero. Given the Gaussian nature to the scheme, an "influence radius" R can be defined as that S where the weight falls to a value of exp (–1) = 0.37 (as discussed in section 2). However, the weight function is not tied to the chosen value of R_c, and in fact *all* the data points in a given dataset can be incorporated in determining each grid point value if desired, by simply setting $R_c = \infty$ (normally, for reasons of computational economics, R_c is set to a finite value with no loss in analysis accuracy as long as R_c is significantly larger than R, as discussed in section 3e).

It is of some interest to provide some comparisons between the Barnes (1973) and Cressman (1959) objective analysis schemes, as they both are weighted-averaging techniques in common usage. Perhaps the most important difference is that the Cressman weight function is determined by the choice of R, since the weights do not asymptotically approach zero with increasing S as they do in the Barnes technique, but instead abruptly become zero at $S = R(= R_c)$. This nature of the Cressman scheme can present serious difficulties when the data distribution is non-uniform.

All weighted-averaging, "successive correction" (see below) objective analysis methods like those developed by Barnes and Cressman are designed to work best on two-dimensional data fields of rather uniform data distribution. Those fields that have very irregular data densities over the domain can present a problem known as "ballooning," characterized by large amplitude and phase distortions in the neighborhood of any grid point whose value is determined primarily by the value of only one inaccurate datum. In general, this problem can be reduced by requiring that at least several observations be used in the calculation of each grid point value. In the Barnes technique,

this is accomplished by extending the distance R_c (with no resulting effect upon the weight function). In some applications of the Cressman technique, the current scan R is locally increased to insure that a sufficient number of data influence each grid point value (Inman, 1970); however, "noise" introduced by such a locally varying weight function must be suppressed with additional numerical filters, thus producing an unknown final response and requiring additional computer time. The other correctional method used in the Cressman technique is to increase R sufficiently on the first pass throughout the entire domain, yet such a procedure means more passes are required to achieve the desired final response.

The Cressman and Barnes techniques both employ the method of successive corrections whereby an adjustment is made to the first pass analysis by decreasing R in the second pass through the data to restore the amplitude of large wavenumber components suppressed in the first interpolation-filtering pass. In the Cressman case, neither the number of additional passes nor the value of the second pass R are governed explicitly by the data distribution, thus neither is the filter response.[1] An advantage of the Barnes technique is that (in the 1973 version) only two passes are required to achieve the desired pattern recognition whereas typically (but rather arbitrarily) four to six scans are employed with the Cressman technique.

Classical sampling theory (Peterson and Middleton, 1963) dictates that a wave whose horizontal wavelength does not exceed at least twice the average observation spacing ($2\Delta n$) cannot be resolved since five data points are required to describe a wave and its derivatives. Random errors in the observations generate fictitious $2\Delta n$ waves (Barnes, 1964); therefore, it is desirable to filter these from the analysis as much as possible. The Barnes (1973) technique has the highly desirable property that the weight function constants can be chosen upon the basis of the data distribution prior to the analysis so that pattern scales resolvable by the data will be revealed to a known response amplitude.

---

[1] Stephens and Stitt (1970) show that an optimum choice for R on the first pass can be made in terms of Δn for a uniform data distribution. However, the optimum R is not well defined theoretically on the second pass, so the final filter response is rather arbitrary.

To summarize, the Barnes (1973) technique is selected because of the following advantages (listed by him).

1. The scheme is a computationally simple, Gaussian weighted-averaging technique which assigns a weight solely as a known function of distance between datum and grid point.

2. Since the weighting function decreases to zero asymptotically, the influence of data may be extended any distance to insure that a sufficient number of observations influence each grid point value without having any effect upon either the weight function or the response characteristics.

3. The low-pass filter response characteristics can be determined prior to the analysis so that pattern scales resolvable by the data distribution will be revealed.

4. Only two passes through the data are required to achieve the desired scale resolution because of the rapidity with which convergence is reached (see appendix). Even when a large influence radius (weight factor κ in the next section) is chosen to reduce noise due to variations in observation density, convergence is attainable (because of the numerical convergence parameter γ discussed below).

5. Small scale "noise" is sufficiently filtered from the analysis after only two passes.

The reader interested only in applying the GEMPAK Barnes scheme to a particular problem can skip the next section without loss of continuity.

2. CHARACTERISTICS OF THE BARNES OBJECTIVE ANALYSIS SCHEME

Barnes (1964) has developed an objective analysis scheme that utilizes a Gaussian weight function in the spatial domain based on the supposition that two-dimensional data distributions can be represented by Fourier integrals. Let r_m represent the distance between the (i,j) grid point and an observed datum $f(x_m, y_m)$. Further, let κ be the weight parameter that determines the shape of the filter response function. Then, the weight function w_m is expressed as:

$$w_m = \exp(-r_m^2/\kappa).\tag{1}$$

For a given choice of κ, an "influence radius" may be thought of as that r_m where $w_m = e^{-1}$. It is shown later that κ can be chosen to reveal those scales resolvable with the particular data distribution.

Barnes (1973) has modified the earlier method in order to decrease the amount of computer time necessary to achieve "desired response" at small wavelengths. This modification consists of applying only a single correction pass through (iteration upon) the initial interpolated field $g_o(i,j)$, rather than making several iterations as before. To accomplish this, parameter κ is decreased from its first pass value (κ_0) to its correction pass value of

$$\kappa_1 = \gamma\kappa_0,\tag{2}$$

by using a "numerical convergence parameter" $\gamma(0<\gamma<1)$ that forces a high degree of convergence (agreement) between the observation field $f(x,y)$ and the correction (second) pass interpolated field $g_1(x,y)$. Of course, it is not always desirable to have the interpolation field fit the data exactly; by manipulating the value of γ, the analyst obtains the desired response following the second data pass.

The first pass yields an interpolated field given by:

$$g_0 = f(x,y)\,D_0,\tag{3}$$

where the spectral response function

$$D_0 = \exp\left[-\kappa_0(\pi/\lambda)^2\right]\tag{4}$$

derived by Barnes (1964) is of the form of a low-pass filter that attenuates signals at small horizontal wavelength λ. Figure 1 shows that *use of smaller values of the dimensionless weight parameter $\kappa_0^\star$ results in greater filter response during the first pass, particularly for the short waves.*

5
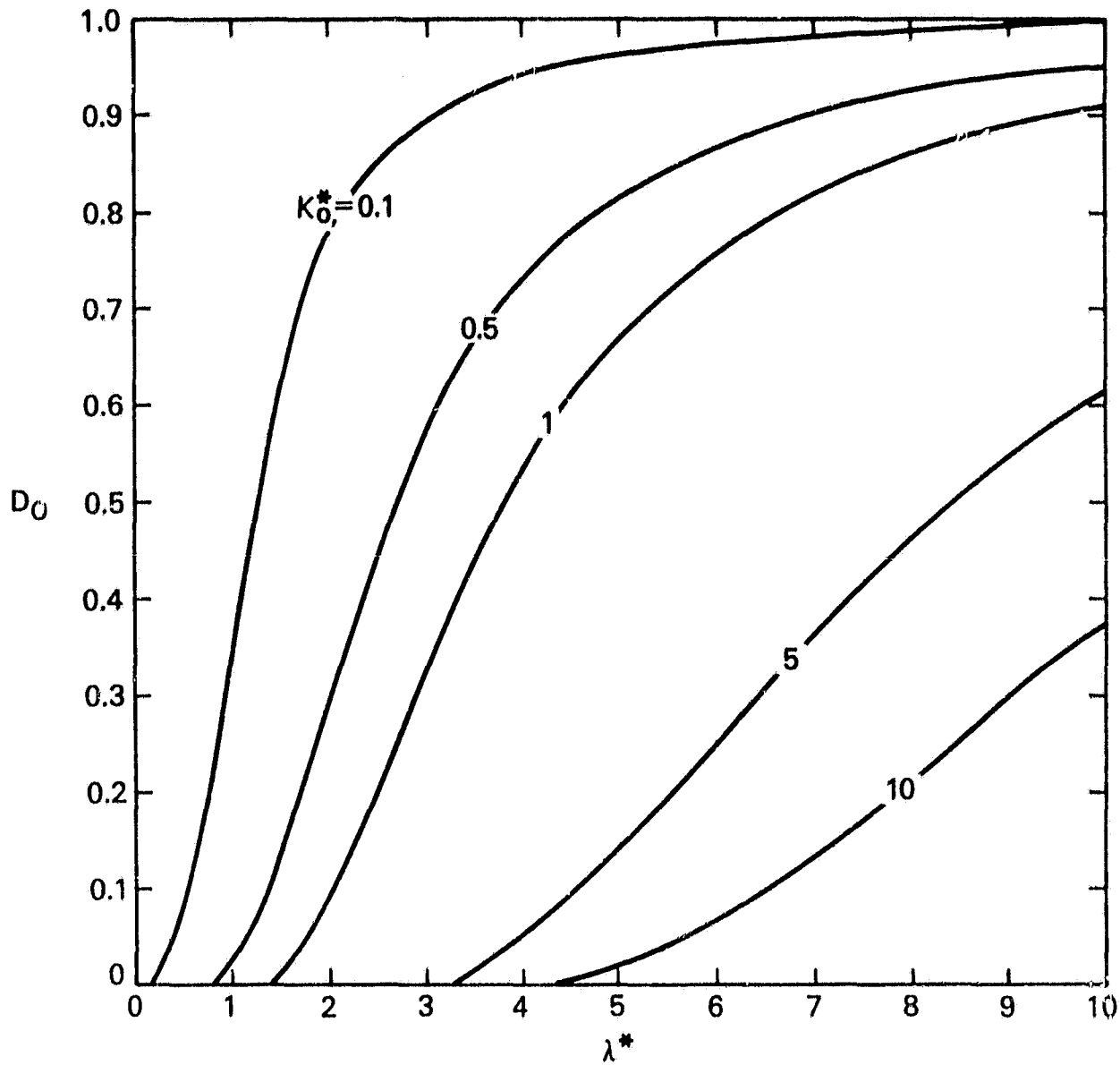
Figure 1. Relationship of first pass response D_0 to dimensionless wavelength $\lambda^{\star} = \lambda/L$ for various values of dimensionless weight parameter $\kappa_0^{\star} = \kappa_0/L^2$, in which L is an arbitrary scale length based upon the data distribution.

6

The "correction (second) pass" grid point field $g_1(i,j)$ is the result of adding to the first pass field the smoothed residual difference between the observed data values and the first pass estimated values "at" the data locations $g_0(x,y)$, or

$$g_1(i,j) = g_0(i,j) + \left[f(x,y) - g_0(x,y) \right] D_1 , \qquad (5)$$

where (after Barnes, 1973),

$$
\begin{aligned}
D_1 &= \exp\left[-\kappa_1 (\pi/\lambda)^2 \right] \\
&= \exp\left[-\gamma \kappa_0 (\pi/\lambda)^2 \right] \qquad (6) \\
&= D_0^{\gamma}
\end{aligned}
$$

is the response function corresponding to the weight function

$$w_m' = \exp(-r_m^2/\gamma\kappa_0) . \qquad (7)$$

A simple bilinear interpolation between the values of $g_0(i,j)$ at the four surrounding grid points can be used to obtain an estimate for $g_0(x,y)$ at each data location. The actual correction pass value "at" each grid point is computed as the sum of the weighted averages from the two passes with M observations according to (after (5)):

$$g_1(i,j) = \frac{\sum\limits_{m=1}^{M} w_m \, f(x_m,y_m)}{\sum\limits_{m=1}^{M} w_m} + \frac{\sum\limits_{m=1}^{M} w_m' \left[f(x_m,y_m) - g_0(x_m,y_m) \right]}{\sum\limits_{m=1}^{M} w_m'} . \qquad (8)$$

Since it follows from (5) that (omitting arguments for convenience)

$$D_1 = \frac{g_1 - g_0}{f - g_0} , \qquad (9)$$

7

the function (6) might be referred to as "a difference field response function." The true correction pass response function (i.e., that one which corresponds to the first pass response function given by (3)) is:

$$D_1^\star = \frac{g_1}{f},$$ (10a)

which upon substitution from (5) gives

$$D_1^\star = D_0 + (1 - D_0)D_1,$$ (10b)

and upon further substitution from (6) we have (Barnes, 1973)

$$D_1^\star = D_0(1 + D_0^{\gamma-1} - D_0^\gamma).$$ (11)

Function $D_1^\star$ is the proper measure of the degree of analysis convergence, or in other words, how closely the interpolated values agree with the observed ones after a second pass through the data. *Making a second pass will increase the degree of convergence when $0 < \gamma < 1$, and particularly so when $\gamma < 0.5$* (Figure 2). Notice that the greatest increase in response occurs at the shorter wavelengths (small D_0 for a given choice of κ_0, according to Figure 1).

A mathematical analysis of the effect of making additional passes ($N \geqslant 2$ iterations) through the data appears in the appendix. It is proven there that no real benefit can be gained in making more than one correction pass because of the rapidity with which convergence is approached when γ is chosen small enough (Figure 2). It is also proven that the 1973 version of the Barnes objective analysis technique is absolutely convergent (although the 1964 version has been shown by Barnes (1964) to be convergent also, several more passes are required to reach the same degree of convergence as with the 1973 version). This fact enables the analyst to control the amount of small-scale detail to be revealed in the analyzed data fields.
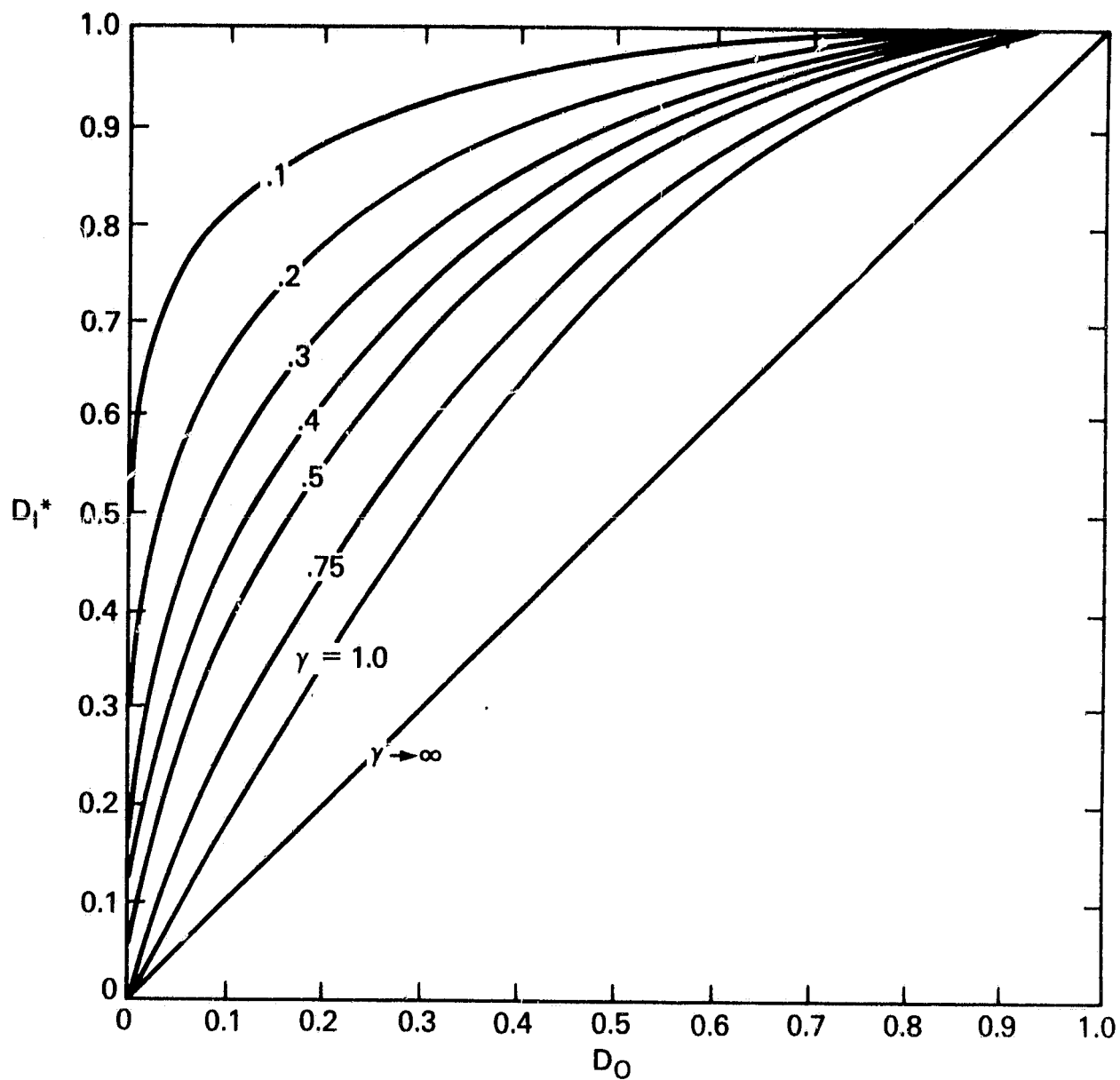
Figure 2. Correction pass response function $D_1^\star$ as a function of first pass response D_0 and numerical convergence parameter γ (from Figure 3 of Barnes, 1973).

These properties of the Barnes analysis scheme, (a) weight selection based wholly upon the data spacing, (b) filtering of the $2\Delta n$ wave from the analysis, (c) adequacy of only two data passes to achieve desired fit to the data, and (d) insurance that a sufficient number of observations influence each grid point value, are all incorporated into the GEMPAK Barnes analysis scheme described in the next section. Other factors, such as objective determination of grid size and accounting for variable observation density, are unique features of the GEMPAK version of the Barnes scheme.

3. THE BARNES SCHEME TAILORED TO GEMPAK NEEDS

Certain factors were considered in accommodating the Barnes (1973) objective analysis scheme to the specific needs of GEMPAK:

1. An agreeable balance between user manipulability and the objectivity inherent to the Barnes scheme must be found.

2. Differences in the nature of the various conventional and satellite data must be taken into account. In particular, satellite-derived cloud motion wind data tend to occur in clumps separated by varying distances because of the problem in finding a sufficient number of suitable, unobscured cloud targets for tracking purposes (Negri and Vonder Haar (1980), Peslen(1980), Maddox and Vonder Haar (1979)). Missing swaths of data are common with satellite-retrieved temperature and moisture soundings (Hillger and Vonder Haar, 1981). On the other hand, the Barnes scheme assumes uniform distribution of data.

3. The maximum allowable detail in the interpolated fields is to be governed by the data spacing.

4. The results of the mathematical analysis of the Barnes convergence properties (appendix) are to be incorporated by limiting the number of passes through the data.

5. Bounds to the grid box size should be wholly determined by the data spacing, since the magnitude of derivative fields like divergence and vorticity is highly sensitive to grid size.

10

6. Measures of analysis quality should be made available to the GEMPAK user to allow objective determination of the effect of user manipulation of the analysis filter control parameters upon the resulting objective analysis.

7. Finally, it is essential that the GEMPAK Barnes scheme be easily and quickly usable by one who does not possess a thorough understanding of the mathematics describing the Barnes scheme filter response, convergence properties, etc. appearing in this report, but who can judge the quality of the resulting objective analysis in terms of its physical content, consistency with other meteorological fields, and temporal continuity. The uniqueness of the GEMPAK version of the Barnes scheme rests in the manner in which these seven considerations are incorporated into an easy-to-use objective analysis package. The unique features are discussed below.

a. Domain Definitions

Unique data and mesh domain definitions were devised in order to permit easy manipulation of data file structures and to obtain a uniformly reliable analysis throughout the entire grid area displayed (Figure 3). One of these domains is termed the *data file*, which consists of the entire data set to be considered for the Barnes objective analysis, e.g., cloud motions in the lower troposphere determined over a subsynoptic region for a specified time interval.

That subset of the data file in which grid point values are computed from the data by the GEMPAK Barnes scheme is termed the *data area*. Those observations which lie outside of the data area do not influence any grid point value. One conceivable data area is that in which the distribution of the data is more uniform than that within the data file. Another choice for a data area might be one which contains a sufficient number of observations from one data type needed to make a comparison with the objective analysis of data by some other type. For example, if there is a wish to study the impact of satellite cloud winds upon an analysis of conventional rawinsonde winds, then the data area should be chosen large enough to include rawinsonde data locations outside of the area of satellite data.

11
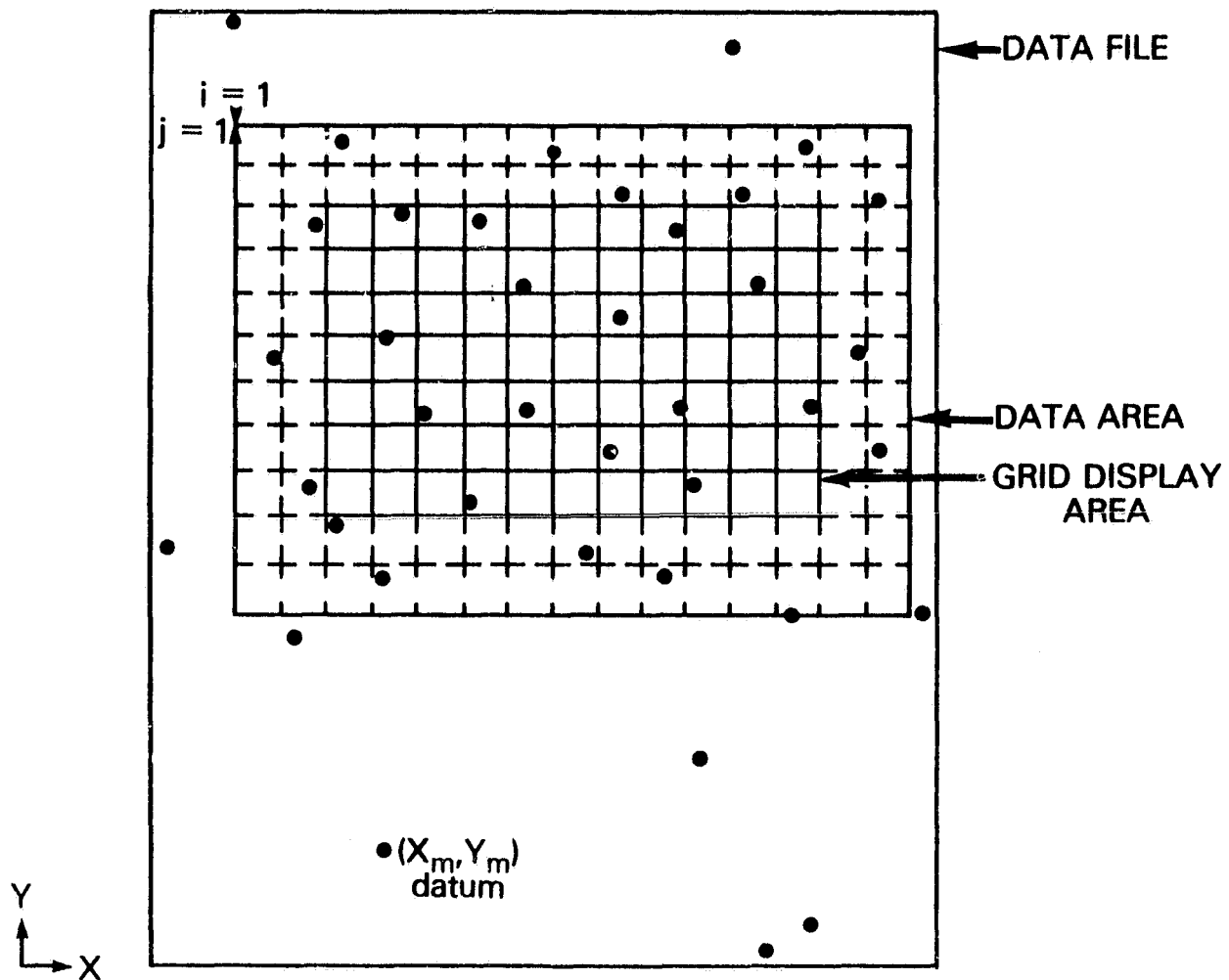
Figure 3. Schematic illustration of data and mesh domains used within GEMPAK Barnes scheme.

The *grid display area* is that portion of the entire gridded domain (i.e., data area) which is displayed as "the objective analysis" to the analyst. The grid display area should lie entirely within the data area to avoid attempted interpretation of the analysis near the data-sparse boundaries of the data area (the dashed region in Fig. 3). In this region, the data are simply extrapolated by the Barnes (or similar) scheme because there are not enough data nearby to give reliable grid point estimates. *A good approach to take is to have at least one or two observations lie within the data area existing beyond each side of the grid display area.*

b. Non-Uniform Data Distributions

The problem of handling non-uniform distributions within the data area is presently dealt with in the following manner. Within any data area input to the computer, the distance between each observation and the nearest observation to it is determined, regardless of how uniformly distributed the observations are. The *average of these distances over the entire data area is termed the "computed data spacing"* Δn_c. *Its value determines the maximum datail permitted in the objective analysis.* However, a smoother analysis can be obtained by manually inputing any $\Delta n > \Delta n_c$, as should be done, for example, when examining the *synoptic-scale* impact of non-uniformly distributed, satellite-derived *mesoscale* wind, temperature, or moisture data upon a *synoptic* field obtained from conventional data. An appropriate choice in this case is the Δn computed from the data area comprised only of the conventional data. In other words, the purpose of the analysis should govern the choice for Δn, under the constraint that $\Delta n \geqslant \Delta n_c$.

c. Control of Detail in a Two-Pass Barnes Analysis

Once the data spacing has been defined, the analyst "fine tunes" the degree of analysis detail (or convergence of the interpolated field towards the observed field) by choosing a value for the numerical convergence parameter[2] (γ). In the GEMPAK version of the Barnes scheme, *the maximum*

---

[2] A slight degree of analysis control is also made possible by manipulation of the mesh size. However, the mesh size, which is allowed to vary within the limits discussed in Section 3d, should not be manipulated for the purpose of achieving a desired degree of analysis convergence, but rather for reasons discussed later.

detail is obtainable with $\gamma = 0.2$, *whereas the least detail results with* $\gamma = 1.0$, *for a given value of* Δn. The analyst must decide for himself how closely the final analysis should be made to fit the data. If one desires to see the greatest resolvable detail in the field, it is justifiable to use a small γ value when errors in the data are only a small fraction of the signal present over the field and the observations are not substantially contaminated by subgrid-scale atmospheric processes.

A limit of two passes through the data is imposed in the GEMPAK Barnes scheme (see appendix). Under this constraint, and the further constraint that $0.2 \leqslant \gamma \leqslant 1.0$, the range of analysis responses permitted can be found (Figure 4). The first pass values $D_0(\lambda)$ are arrived at by inserting the value $D_0(2\Delta n) = 0.0064$ for integral values of $(\lambda/2\Delta n)$ in (A18); the second pass values $D_1^{\star}(\lambda)$ are then calculated from (11). The value of 0.0064 for $D_0(2\Delta n)$ gives a second pass response of $D_1^{\star} = e^{-1}$ at the $2\Delta n$ wavelength when $\gamma = 0.2$. Since the maximum response and best fidelity characteristics of the Barnes low-pass filter are arrived at by choosing this γ value, use of $D_0(2\Delta n) = 0.0064$ lets us obtain a baseline value by which responses at multiples of the $2\Delta n$ wave can be calculated. It is important to realize that under these conditions, the weight parameter κ_0 is fixed by the data spacing, since when $\lambda = 2\Delta n$ is inserted into (4) with $D_0(2\Delta n) = 0.0064$,

$$
\begin{aligned}
\kappa_0 &= -\left(\frac{2\Delta n}{\pi}\right)^2 \ln D_0(2\Delta n) \\
&= 5.052 \left(\frac{2\Delta n}{\pi}\right)^2 .
\end{aligned}
\tag{12}
$$

In other words, *the weight parameter* κ_0 *is fixed by the data spacing to give maximum response of* e^{-1} *at the* $2\Delta n$ *scale* $(\gamma = 0.2)$. *The user of the GEMPAK Barnes scheme has the option of either accepting this default analysis or of making a smoother, less detailed one by inputting a larger* γ *value* $(0.2 \leqslant \gamma \leqslant 1.0)$. Increasing γ affects only the second interpolation field, which is the one displayed, such that the amplitude of the shorter wavelength features is reduced.
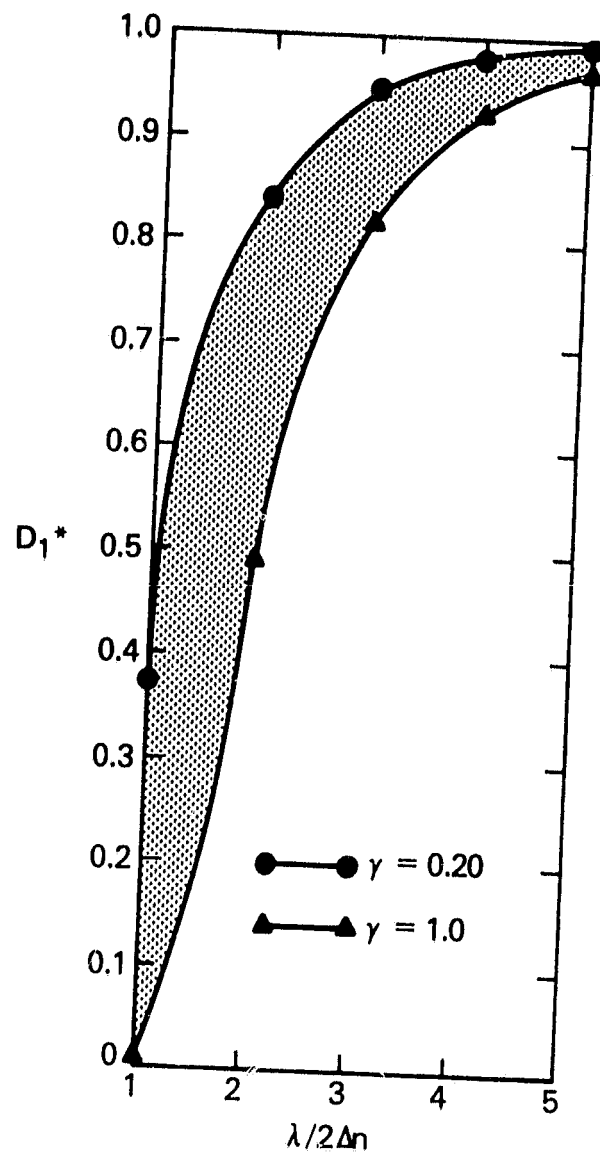
Figure 4. Permissible second (correction) pass filter responses as functions of multiples of the $2\Delta n$ wavelength in the GEMPAK Barnes scheme.

d. Objective Determination of Grid Size

An abundant literature exists illustrating bounds on the ratio between the grid box size Δx and the data spacing Δn, such that $\Delta x / \Delta n$ lies in the range of about 0.3 to 0.5 (Barnes 1964, 1973; Doswell, 1977; Maddox, 1980; and others). The apparent reason for the upper limit is that since five grid points are required to *represent* a wave, and the minimum *resolvable* wave is of $2\Delta n$ scale, then Δx must be no larger than one-half of Δn (as shown in Figure 5). As for the lower limit, since calculations of derivative quantities like divergence and vorticity are highly sensitive to grid length, then if such derivative fields are to represent resolvable features, one must not use a grid length that is much smaller than the data spacing.

Accordingly, the GEMPAK Barnes scheme imposes the constraint that

$$1/3 \leqslant \Delta x / \Delta n \leqslant 1/2. \tag{13}$$

The number of grid points (KX by KY) are then determined by

$$KX(KY) = 1 + (x(y) \text{ data area dimension})/\Delta x. \tag{14}$$

Even though the GEMPAK Barnes scheme places stringent limits on the mesh size, it remains versatile enough that it can accommodate round-off of the computed mesh size to convenient whole numbers (such as $1.0°$ latitude, or 10.0 km, etc.). This versatility may be necessary when making comparisons of objectively analyzed data to numerical model output, for instance.

e. Quality Control Indicators

Serious "ballooning" problems can occur when an insufficiently small number of data determine the value at a grid point (as discussed in section 1). In order to avoid such problems, a warning flag appears to the user of the GEMPAK Barnes scheme whenever less than 3 data values determine any grid point value. Such a problem is particularly likely to occur when the grid display area is chosen so large that it extends into data-void regions. It is desirable to caution the analyst that a better choice for the grid display area should have been made.

16

DATA POINTS

GRID POINTS
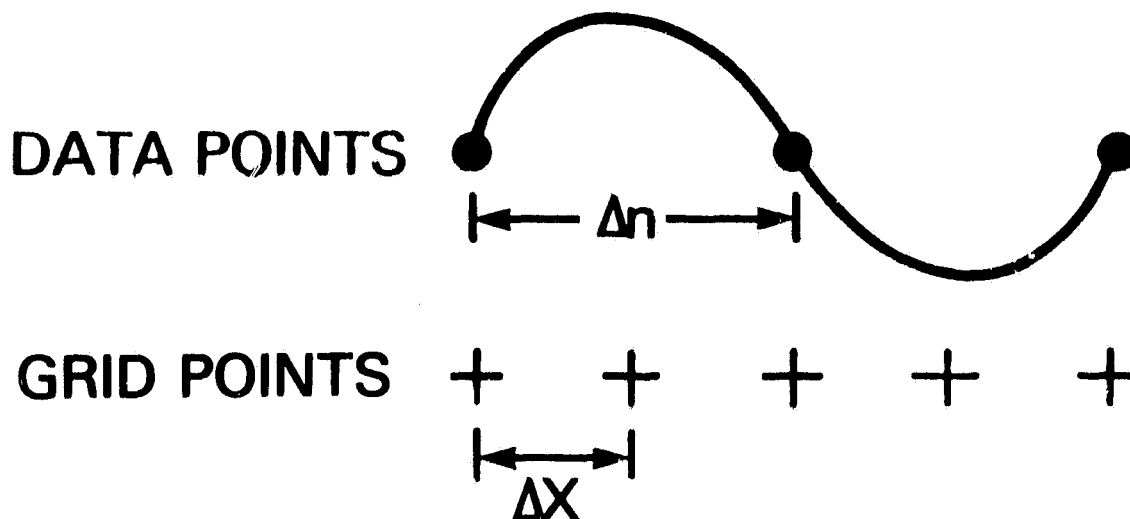
Figure 5. Determining the proper relationship between grid size
and data spacing from elementary considerations.

The warning flag appears slightly more often than it would otherwise because of the inclusion
of a delimiter on the computation of extremely small weights at very large distances between a grid
point and a datum. Its insertion into the program increases the efficiency of its operation without
noticeable loss of any accuracy. Presently, this cutoff distance has the value $R_c = \kappa_o \sqrt{20}$, which is
4.5 times as large as the "radius of influence" R at which the weight $w_m = e^{-1}$ (see (1)). At distances
$r > R_c$, the weights are assigned zero values, whereas in actuality it can be shown that the weight
value at R_c is about 1.5×10^{-8} when $\Delta x = \Delta n/2$.

Another measure of analysis quality is the computed rms difference between the interpolated
and observed fields. This number is displayed to the analyst after both the first and second passes
through the data to enable one to see the amount of rms reduction between passes and how much
greater this reduction is with the use of a smaller γ value. This number can then be compared to
the known observational errors.

Finally, computer limitations (section 4) dictate that KX and KY be no larger than 40.
Problems arise when an analysis is attempted of clumps of closely spaced data separated by large
distances from other clumps over a sizable data file. If one wishes to make a detailed analysis using
$\Delta n = \Delta n_c$, then several separate analyses over subsets of the data file must be made.

17

Despite the last limitation, the GEMPAK Barnes scheme remains a versatile and easy-to-use objective analysis package. By incorporating a weight selection procedure based purely upon the data spacing, filtering the $2\Delta n$ wave from the analysis, and determining the grid size from the data spacing, the scheme is objective in nature. Yet, by permitting user input of arbitrarily larger data spacing than actually exists, user manipulation of the various spatial domains, and user selection of the γ and Δx values within bounds, the scheme retains its necessary flexibility. In the following section, it is shown how use of an interactive, menu-type format and various cursor/display controls on a computer CRT enhances the ease with which the GEMPAK Barnes scheme can be used.

4. OPERATION OF THE GEMPAK BARNES SCHEME ON AN INTERACTIVE COMPUTER

A CRT computer graphics terminal on AOIPS is utilized to permit the user to step through a series of menus (questions) in order to select appropriate files and parameters for the GEMPAK Barnes objective analysis. For each analysis the user must select 1) the vertical level for the analysis, 2) the data sets to be included in the analysis, 3) the data and grid display areas, 4) values for the observation spacing Δn, grid spacing Δx, and γ, and 5) the meteorological parameters to be analyzed. When appropriate, the values chosen most recently will be displayed as the default values.

Unless default values from a previous analysis are available the following defaults are used and may be chosen by entering 0 on the terminal.

(1) The default vertical level is the earth's surface.

(2) The default data set is the sounding data set used most recently in GEMPAK.

(3) The data and grid display areas both correspond to the data file area. Areas are defined by the lower left and upper right coordinates. Since it is necessary to align the areas to lie on grid points, the lower left corner will be fixed and the upper right corner moved up and right until it coincides with a grid point. ΔN is set to ΔN_c; $\Delta X = \Delta N/2$.

(4) Default values are $\Delta N = \Delta N_c$ (value computed within default data area), $\Delta X = \Delta N_c/2$, and $\gamma = 0.30$.

(5) The meteorological parameters are not defaulted; they must be selected by the user.

When a user works in latitude/longitude coordinates, he has the capability to display the data and grid display areas on the CRT. Then, by manipulation of a joystick on the console, these analysis boxes can be easily reshaped and moved until, by visual inspection, the desired result is obtained. After selecting which files and parameters to use, both the actual data and the objective analysis are displayed. If the analysis is unsatisfactory to the user, he may then alter the input parameters to the scheme by stepping through the menus once more, and subsequently see another analysis. Of course, instead a different parameter may be analyzed at this point.

Following are the menus in the order in which they appear to the user and a description of the required responses.

```
┌─────────────────────────────────────────┐
│  Enter Vertical Level (Default = Surface):│
└─────────────────────────────────────────┘
```

A single vertical level to be analyzed, in the units of the vertical coordinate system of the sounding dataset name (see below), must be entered (with a decimal point included). A ⟨CR⟩ (carriage return) or 0.0 will result in default to the surface level.

If the surface level is selected, the data set file labels will be checked to ensure that surface data is available in each dataset. If it is not, an error message is printed and the user may select another level.

```
┌──────────────────────────────────────────────────┐
│  Sounding Data Set Number 1                        │
│                                                    │
│  Enter Sounding Dataset Name (Default Disk  =    ) │
│                                                    │
│  (0              ) or (No Default):                │
└──────────────────────────────────────────────────┘
```

Up to *four* sounding datasets[3] can be entered by name. The default disk and user's UIC (User Identification Code) will be added to the name if not specified; also ".SND" will be appended to the name if a file type is not entered. When a default file name appears, the file may be selected by entering a "0".

An incorrect file selection will result in an error message and the file prompt will be redisplayed. The file labels will be checked to ensure that they have a common vertical coordinate system.

A (CR) will end the file selection process.

```
Average minimum station spacing     0 = old values
default = compute:
```

This prompt is displayed only if default parameters from a previous objective analysis are available. Entering a "0" will restore the previous parameters, including the average minimum station spacing, ΔN_c, previously computed from the present *data file*.

```
Enter 2 bit planes for data and grid display areas
Default = no display:
```

This prompt appears only if the user has allocated a CRT display terminal and has previously plotted sounding data.[4]

---

[3] A GEMPAK sounding dataset is a collection of soundings formatted through one of the options in the sounding input program.

[4] Each display terminal is presently a Hazeltine terminal which has one memory of 512 X 512 X 8 bits, addressable as 8 bit (graphic) planes.

The two bit planes (1-8) must be entered separated by a comma. The first bit plane will be used to show the data area; the second will display the grid analysis area. The same bit plane may be entered for both displays. Note that any previous displays on those bit planes will be erased.

The menu shown in Figure 6 lists information about the data and grid analysis areas as well as the current values of ΔN, ΔX, and γ. For reference, the latitude/longitude coordinates of the lower left and upper right corners of the *data file* area, and the number of stations reporting data at the vertical level selected, are displayed at the top. Additionally, the number of stations within the *data area* are displayed to the right of the station spacing.

In 1., the selected value of ΔN is shown. Also listed are the last computed station spacing and the random station spacing. The average minimum station spacing, ΔN_c, within a *data area* is computed only if a "-1" is entered in response to this menu. The uniform station spacing is the average station spacing within the *data area* assuming that the stations are evenly distributed. This number should be used only for reference. *If this number is different by a factor of 2 or more from the computed minimum station spacing, then the data set is not really uniform in nature.*

2. and 3. show the lower left and upper right corners of the data area, in degrees latitude, longitude.

4. is the grid spacing, ΔX. Note that the grid spacing and the station (observation) spacing are both in degrees latitude. The grid spacing is initially set equal to one-half of the station spacing.

5. displays the number of grid points covering the data area in both X and Y directions.

6. shows the grid display coordinates in units of grid points, i.e., the subset of the data area grid which will become the grid display area. The latitude, longitude corners of the grid display area, which cannot be altered directly, are shown under the coordinates.

```
                    ***BARNES ANALYSIS MENU***

            (28.8          -106.0       41.4        -86.6)      34 STATIONS

    1. STATION SPACING                                1.93      32 STATIONS

           LAST COMPUTED SPACING      1.93

           UNIFORM STATION SPACING    2.51

           -1 TO COMPUTE STATION SPACING


    2. DATA AREA (LOWER LEFT)         28.82              -106.37


    3.              (UPPER RIGHT)     41.36              -86.11


    4. GRID SPACING                   0.965


    5. # GRID POINTS (KX, KY)         21                 14


    6. GRID DISPLAY COORDINATES       1,21               1,14

           (LOWER LEFT)               28.82              -106.37

           (UPPER RIGHT)              41.36              -86.11
```

Figure 6. Example of GEMPAK Barnes "menu" of input parameters.

22

```
Enter number and new values

0 = initialize    P = PRINT    D = Data Box

Default = accept:
```

This prompt appears following the menu shown in Figure 6. The user may enter a menu selection number (1-6) followed by a comma and new parameter value(s) separated by commas. The values entered are checked for obvious errors, changes are made to the required parameters and the menu is redisplayed. A station spacing $\Delta N > \Delta N_c$ can be entered at this step to account for non-uniform data distributions. Entering a new value for one parameter can affect values for the other parameters (e.g., new station spacing alters data area slightly).

A response of "0" will reset the menu to the values initially displayed. "P" will print the menu on the line printer. "D" will display a yellow box on the display screen. The box may be moved and shaped to redefine the data area using the "MOVE" and "SHAPE" buttons on the terminal console. The "DEFINE" button is used to accept position of the box. The user may enter ":XF" to turn the box off.

```
ENTER GAMMA BETWEEN 0.2 and 1.0

DEFAULT GAMMA = 0.3:
```

Any value for γ within the limits $0.2 \leq \gamma \leq 1.0$ is acceptable. If any other value is entered, the menu is displayed again. A ⟨CR⟩ will default to $\gamma = 0.3$.

```
Enter list of parameters to be analyzed (< = N):

   .

   .

   .
```

Up to N parameters (e.g., temperature, wind speed, etc.) may be selected by number from the list. N is limited by computer considerations. Since the most time-consuming part of the objective analysis is the exponential computation, large grids or data areas containing a large number of stations will limit the number of parameters that can be analyzed at one time. The number of exponentials computed is approximately $2 \times M \times KX \times KY$, where M is the number of stations. Since this number is not a function of the number of parameters being analyzed, *it is desirable to compute as many parameters as possible during a single analysis.*

PASS 1	PARAMETER 2	RMS = 2.38	#STATIONS = 32
PASS 2	PARAMETER 2	RMS = 0.74	#STATIONS = 32

This message prints the RMS after each pass. No user response is required. The reduction in RMS error between the two passes is greater when a small γ value is used. The RMS at the second pass can be compared to known observational uncertainties to help the user make a judgment of analysis quality.

Following each pass, parameter values are computed at each station using bilinear interpolation with the four surrounding grid points. The RMS is computed using the differences between the interpolated station value and the actual station value (see (8) in Section 2). The parameter number refers to the GEMPAK number for the parameter being analyzed. The number of stations actually used in the analysis is also shown.

THERE ARE_____ POINTS IN THE GRID DISPLAY AREA WITH INSUFFICIENT DATA. A LISTING WILL BE SENT DIRECTLY TO THE LINE PRINTER UNLESS AN "N" IS ENTERED:

24

The prompt is displayed only if fewer than three data values have influenced any grid point, as mentioned in Section 3e.

```
ENTER GRIDDED DATASET NAME (DEFAULT DISK = )
(0 =                    ) OR (NO DEFAULT)
```

This gridded dataset, a file which stores up to 20 grid analysis arrays, and has the form "NAME GRD," must be entered. Again, the default disk and the users UIC will be added if not specified here; "GRD" will be appended if a file type is not entered. When a default file name is given, the file may be selected by entering a "0".

```
OUTPUT TEMP IN GRID NUMBER 10
ENTER NAME FOR GRID:
```

This prompt asks the user to specify a name for grid 10; in this case, the gridded data resulting from an objective analysis of temperature (TEMP) is stored in grid 10. A 12-character name may be entered to identify the grid in the file. The grid number is the location of the grid in the file. The grid is accessible for purposes of algebraic manipulation (e.g., pressure tendency calculations made by subtracting two gridded pressure datasets) and contouring.

5. METEOROLOGICAL APPLICATIONS OF THE GEMPAK BARNES SCHEME

In this section, station and grid spacings (Δn and Δx), data and grid display areas, and the numerical convergence parameter (γ) are each varied to examine the resulting effect upon the objective analysis of subsynoptic-scale wind data. Each of the resulting GEMPAK Barnes analyses, except for the last one to be discussed, will be of a uniformly distributed wind data set obtained from rawinsonde data collected during the SESAME (Severe Environmental Storms and Mesoscale Experiment). The last analysis will be that of non-uniformly distributed wind data composed of both the rawinsonde observations and satellite-derived cloud motion estimates. The purpose of the latter analysis is to
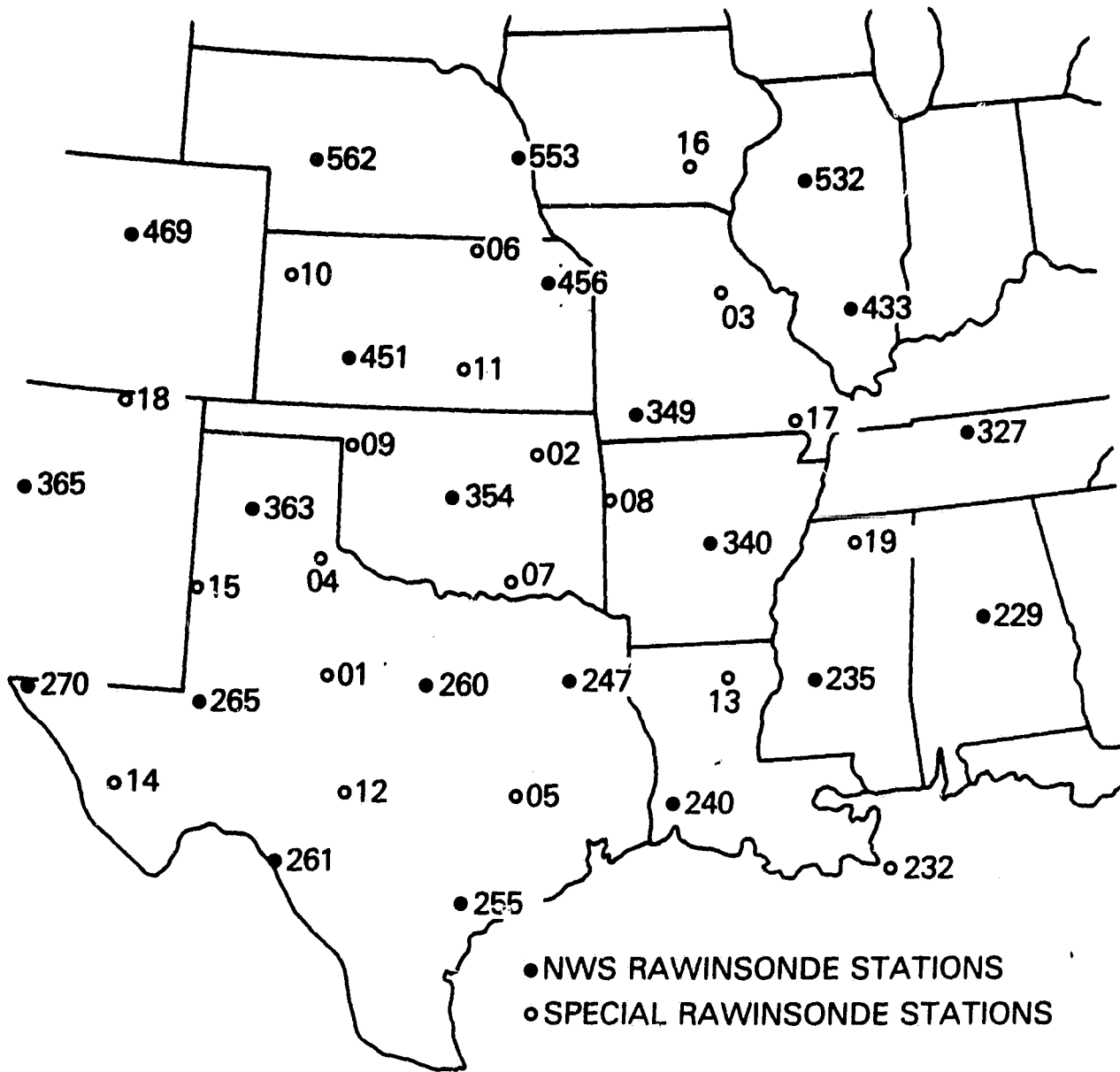
Figure 7. Location of rawinsonde stations participating in the AVE-SESAME I (April 10, 1979) experiment.

show how Δn can be chosen larger than the computed value Δn_c so as to obtain an analysis more representative of the scale resolvable by the uniformly distributed, conventional data set.

The data file used for all but the last analysis is derived from the SESAME regional-scale network, a uniformly distributed array of rawinsonde observations including 23 National Weather Service stations and 19 supplemental sites over the south central United States (Figure 7). In particular, an analysis of wind speed at the 310 K isentropic surface at 2100 GMT on April 10, 1979 is made from data existing at 32 of the 42 possible stations. These 32 stations constitute the *data file*, which form the base for the objective analysis. The latitude and longitude bounds of the data file are:

	Latitude Bounds	Longitude Bounds	
	upper: 41.36	western: −106.37	(15)
	lower: 28.82	eastern: −86.11	

a. Map Generation

The first procedure in generating a Barnes analysis is to display a map of the observations within the *data area* (the only real requirement is for the analyst to display just a map, but the inclusion of the observations could minimize errors related to the placement of data and display areas). A map containing the SESAME wind speed observations (Figure 8) is generated and displayed by GEMPAK's sounding plotting program. For this example, the coordinate bounds of the map were chosen as follows:

	Latitude Bounds	Longitude Bounds	
	upper: 43.4	western: −108.4	(16)
	lower: 25.4	eastern: −83.4	

These coordinate choices produce a map covering 25° of longitude and 18° of latitude, which includes all the observations within the data file plus an additional surrounding data-free area of about 3°.
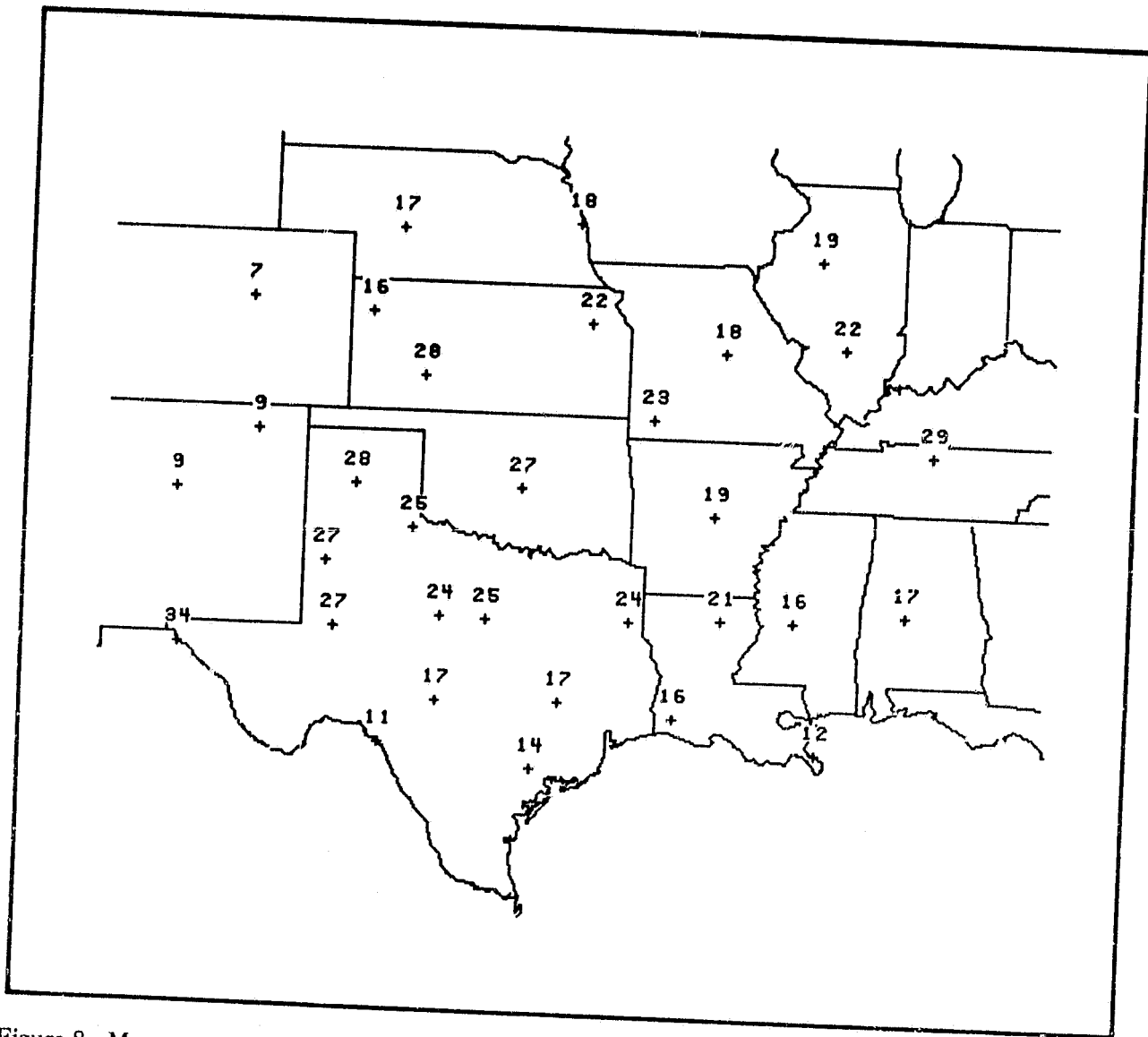
Figure 8. Map generated by GEMPAK sounding plot program when SESAME data file consisting of wind speeds (shown in units of m s^{-1}) at the $\theta =310$K level at 2100 GMT on April 10, 1979 is chosen.

As long as no further changes are made to the map's coordinates within the sounding plot program, the latitude/longitude coordinates of the map area will be retained throughout the use of the Barnes program, so that the data and display areas can be accurately placed over the map.

b. Station/Grid Spacings and Data/Grid Display Areas

Once the map has been generated, the GEMPAK Barnes menus are utilized. Following the procedure outlined in Section 4, the user first specifies his desire to analyze the wind speed at vertical level = 310 (in this case, 310°K). He is afterwards faced with the task of generating suitable data and grid display areas.

The interactive CRT terminal allows the analyst to create and alter two "boxes" on the CRT display screen, one box enclosing the data area, and the other enclosing the grid display area. During the initial internal computation of Δx and Δn, only one box will appear since the program will assign the same coordinates to both the data and grid display area. Unless old values for Δn are available, the average minimum station spacing within the *data file* will be computed, which in the present case (Figure 8) is $\Delta n_c = 1.93°$. Under the constraint (13), the grid spacing is initially set to its maximum allowable value of $\Delta x = \Delta n_c/2 = 0.965°$. Upon input of the data file bounds (15), the values of KX and KY are computed internally from (14), and hence the Barnes Analysis Menu will appear on the CRT display screen with the parameter values (shown earlier) in Figure 6.

The calculated value of Δn_c may be modified now to meet the needs of the analyst. It is essential to realize the importance that such a modification has upon the analysis (see Section 3c), within the limit that any input Δn must be no smaller than Δn_c. For discussion purposes, the value is increased slightly to $\Delta n = 2.00°$ in order to allow an internally calculated grid spacing of $\Delta x = 1.00°$. *In general, the value of Δn should not be arbitrarily increased beyond a few percent of its calculated value unless a non-uniform data area is to be analyzed, for otherwise details resolvable by the data will be lost.*

The change of Δn from 1.93° to 2.00° results in a slight enlargement of the box representing the data area. While holding the lower left corner of the original coordinates (15) fixed and accounting for the grid spacing Δx of 1.0 degree, the GEMPAK Barnes program computed new coordinates for the *data area*:

Latitude Bounds	Longitude Bounds	
upper: 41.82	left: −106.37	(17)
lower: 28.82	right: −85.37	

Note that the coordinates of the upper right corner of the data area have changed to allow the box to increase in size while including all observations from the data area. The degree differences between the latitude bounds and the longitude bounds are now multiples of the grid spacing Δx of 1.0 degree.

Following the choice of Δx, the data area is chosen by either typing in the coordinates or by using a joystick-controlled function that shapes a rectangular box around the data area. It is easier to position a box manually than to determine the coordinates of the lower left or upper right edges of the data area. In either case, once the data area has been entered, the program may alter the coordinates and the box slightly to accommodate the grid spacing. The resulting data area box is the larger of the two boxes in Figure 9.

The grid display area should be entered as a *subset* of the data area. A good approach to take (see Section 3a) is to allow at least one or two observations to lie beyond each side of the grid display area. Since the slight increase of Δn from 1.93° to 2.00° results in an increase of the grid array from 21 X 14 (Figure 6) to 22 X 14 when $\Delta x = 1.00°$ (see (17)), a grid display area imbedded two grid points (or 2°) within each side of the data area requires the following sequence of values to be input in item #6 of the Barnes menu:

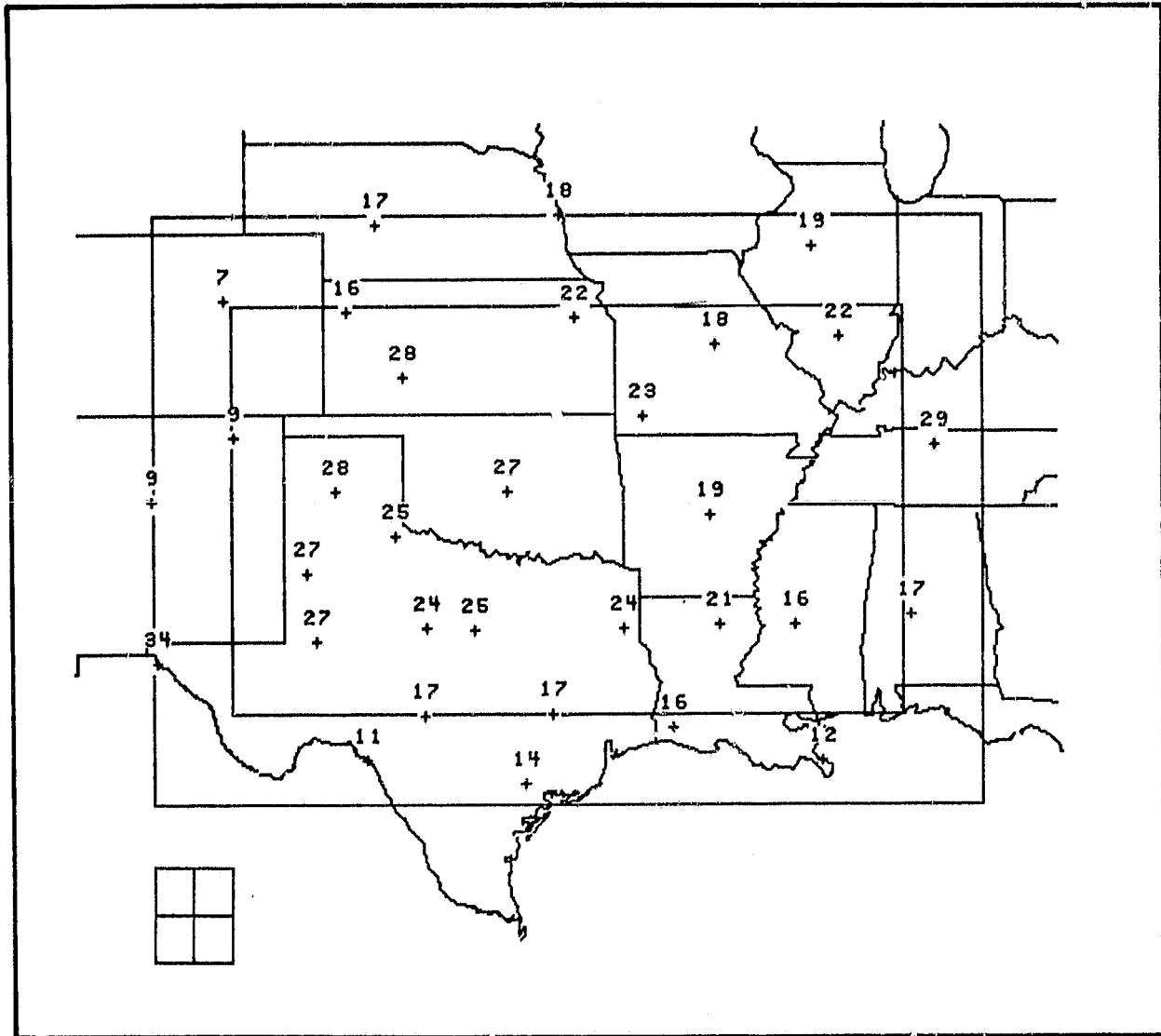$$3 \qquad 20 \qquad 3 \qquad 12,$$

30

Figure 9. Data and grid display areas selected for analysis of SESAME data set (see Figure 8). The boxes were obtained both by typing in the coordinates and by using joystick-controlled functions on the AOIPS computer console. Note size of grid mesh in lower left corner.

and results in an 18 X 10 array of grid points within a *grid display area* with the following coordinates:

Latitude Bounds	Longitude Bounds
upper: 39.82	west: −104.37
lower: 30.82	east: −87.37

The grid display area is shown enclosed by the smaller of the two boxes in Figure 9.

A comparison of two objective analyses of the SESAME data set generated by (1) having a grid display area smaller than the data area (Figure 10a) and by (2) having the two areas represented by one box (Figure 10b) demonstrates the rationale behind choosing the former option. A less reliable analysis results when the data and grid display areas coincide (Figure 10b), especially near the east and south sides of the box. Notice that the observations at Nashville, Tennessee and throughout southern Texas are not incorporated into that analysis, resulting in loss of valuable information on the strong wind speed gradients actually there. On the other hand, the analysis displayed in Figure 10a, which was generated from values interpolated to grid points enclosed by the outer box (the data area) but displayed only within the inner box (the grid display area) is more reliable because observations exist in the region between the grid display area and the data area. Therefore, each grid point within the grid display area has been influenced by values at station locations which totally *surround* it.

c. Numerical Convergence Parameter

In the case of the two analyses just described, the numerical convergence parameter γ was assigned a value of 0.3. The GEMPAK Barnes program allows γ to range in value between 0.2 and 1.0. Analyses of the SESAME wind speed data set with four different values of γ (0.2, 0.3, 0.5, and 1.0) are displayed in Figure 11 to illustrate the impact of γ variations upon analysis detail (recall Section 3c). The analysis produced with $\gamma = 0.2$ exhibits significant detail at the $2\Delta n_c$ scale. As
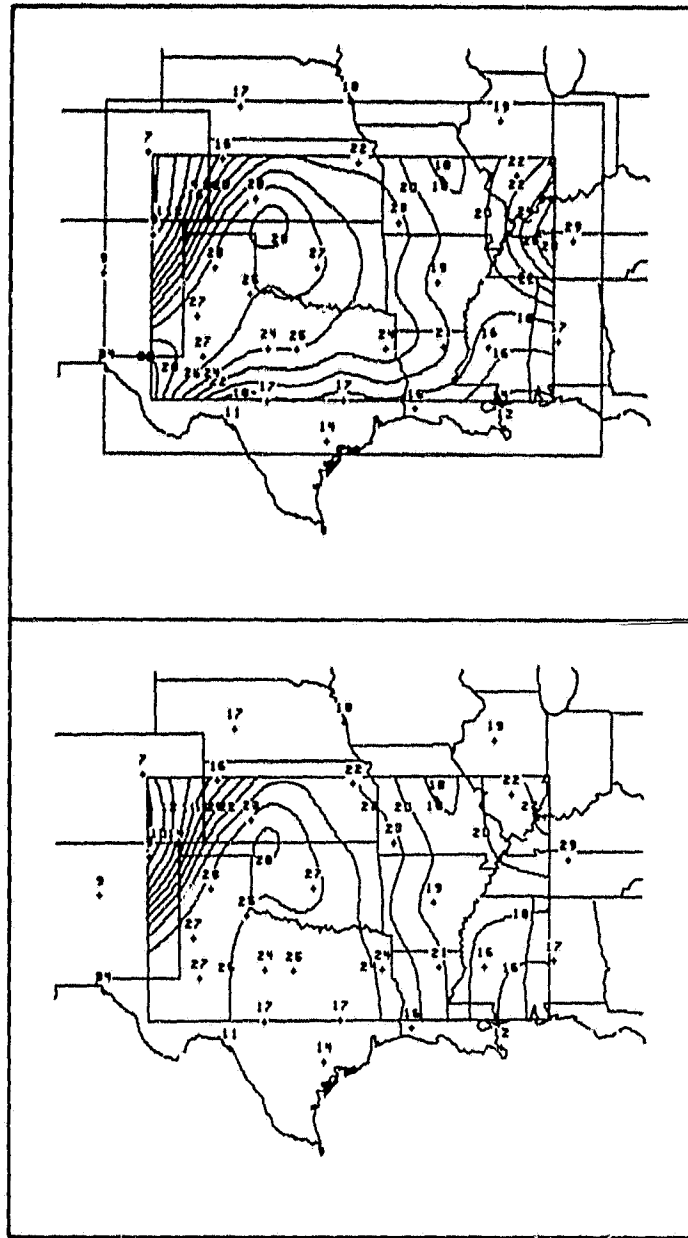
32

Figure 10. Comparison of two analyses of SESAME data set by GEMPAK Barnes scheme (isotachs
in m s^{-1}), with following values for the input parameters: $\Delta n = 2.0°$, $\Delta x = 1.0°$, and
$\gamma = 0.1$. Top figure shows analysis generated by having grid display area nested 2°
within the data area. The two areas are coincident in the bottom figure.
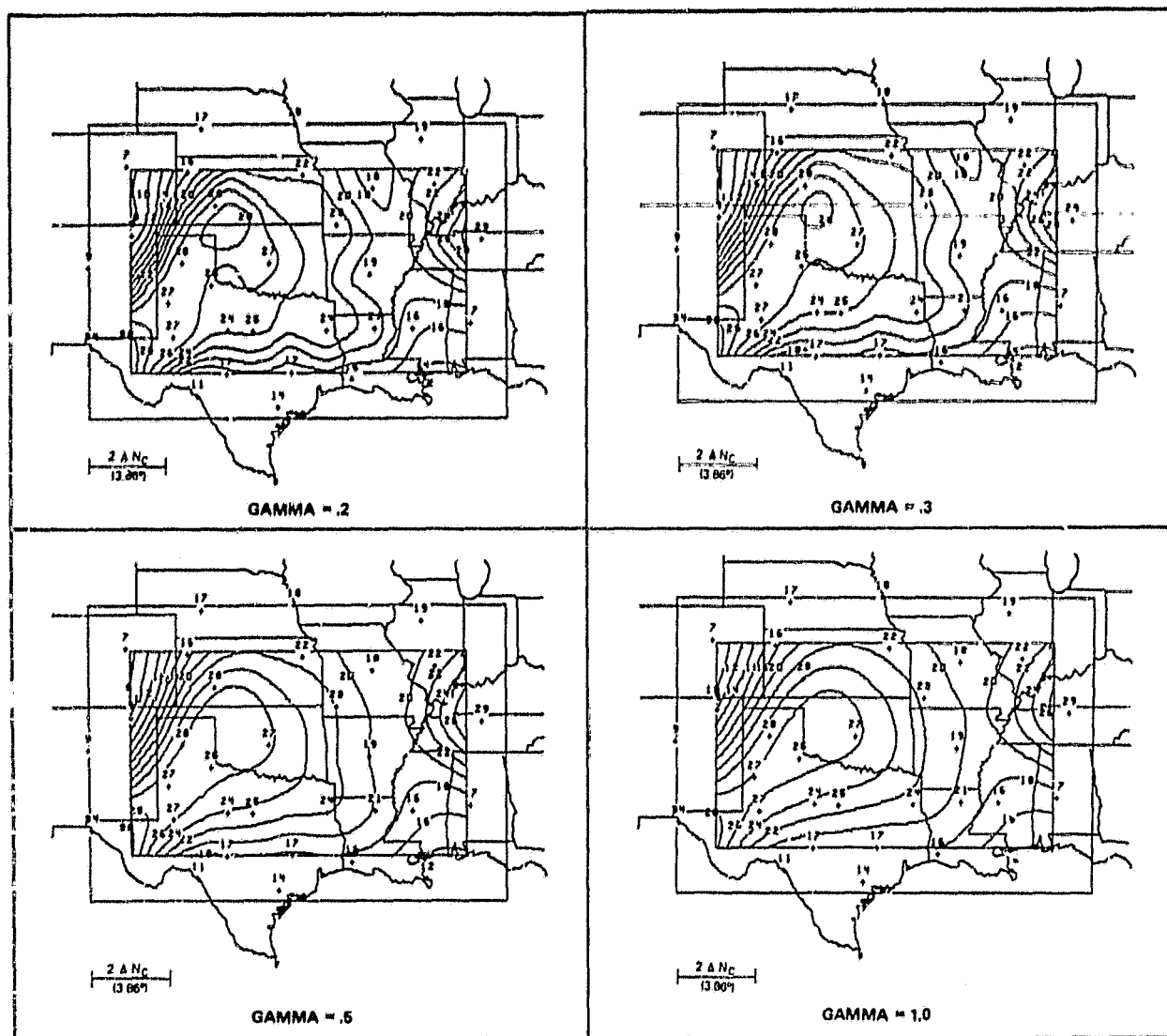
Figure 11. Changes in the objective analysis of the SESAME wind speed data set (see Figure 10) brought about by variations in the value input for γ.

discussed in Section B of the Appendix, this detail is likely to be the undesirable result of aliasing and of random errors in the data. Larger values of γ lower the response to that scale (as given by (11) and Figure 2) and accordingly dampen the analysis detail.

The analyst must judge which value of γ produces the best analysis for his purposes. Although there is no intention in this report to cover all aspects of the general factors involved in objective analysis quality, a few points should be kept in mind. First, the *rms difference* between the interpolated and observed fields should be much less than the estimated rms error in the observations; this guideline should help serve as an upper limit on the choice of γ. Second, the meteorological features appearing in the analyses should exhibit acceptable temporal and spatial *continuity*. The lower limit on γ is determined by the continuity considerations, which is best made by examining derivative fields like divergence and vorticity. Third, the analyst should examine the interpolated fields resulting from the largest γ value first, and then *reduce γ in steps* until an unacceptably small signal-to-noise ratio becomes evident.

Application of the rms quality criterion to the objectively analyzed wind speed fields in Figure 11 is illustrated first. An observational rms error of 3 to 6 m s^{-1} can be expected for winds between the 700 and 400 mb pressure surfaces, through which the 310 K isentrope passes in this case (Fuelberg, 1974). The computed rms differences between observed and objectively analyzed fields exceeded 1.5 m s^{-1} for γ values equal to or greater than 0.5. Thus, the $\gamma = 0.2$ or $\gamma = 0.3$ analysis would be selected upon this basis.

The other measures of analysis quality can be derived from the calculated velocity divergence fields at 3-hourly intervals from 1500 GMT to 2100 GMT on this date. The fields shown in Figures 12, 13, and 14 are generated from the objectively analyzed u and v wind component fields using γ values of 0.2, 0.3, and 1.0, respectively. Examination of these fields for spatial and temporal continuity is made, bearing in mind that, although the archived SESAME data have been

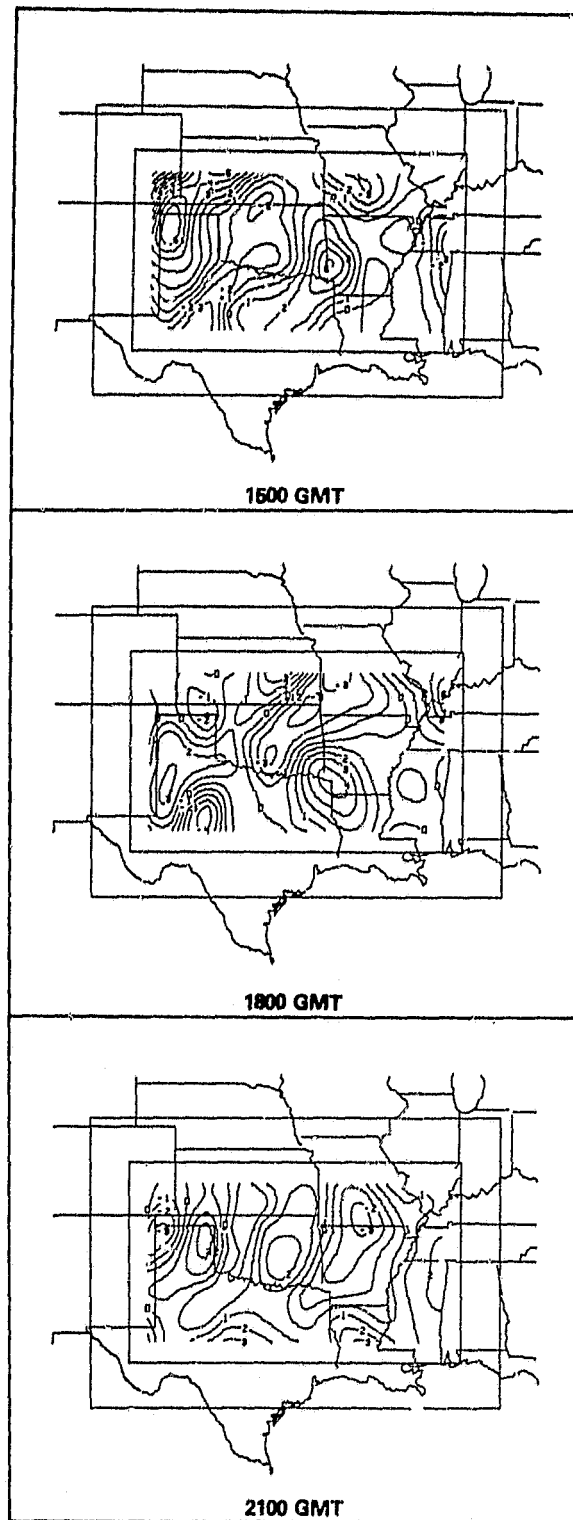Figure 12. Three-hourly analysis of velocity divergence ($\times 10^{-5} s^{-1}$) fields derived from objectively

analyzed u and v wind component fields with $\gamma = 0.2$.

1500 GMT
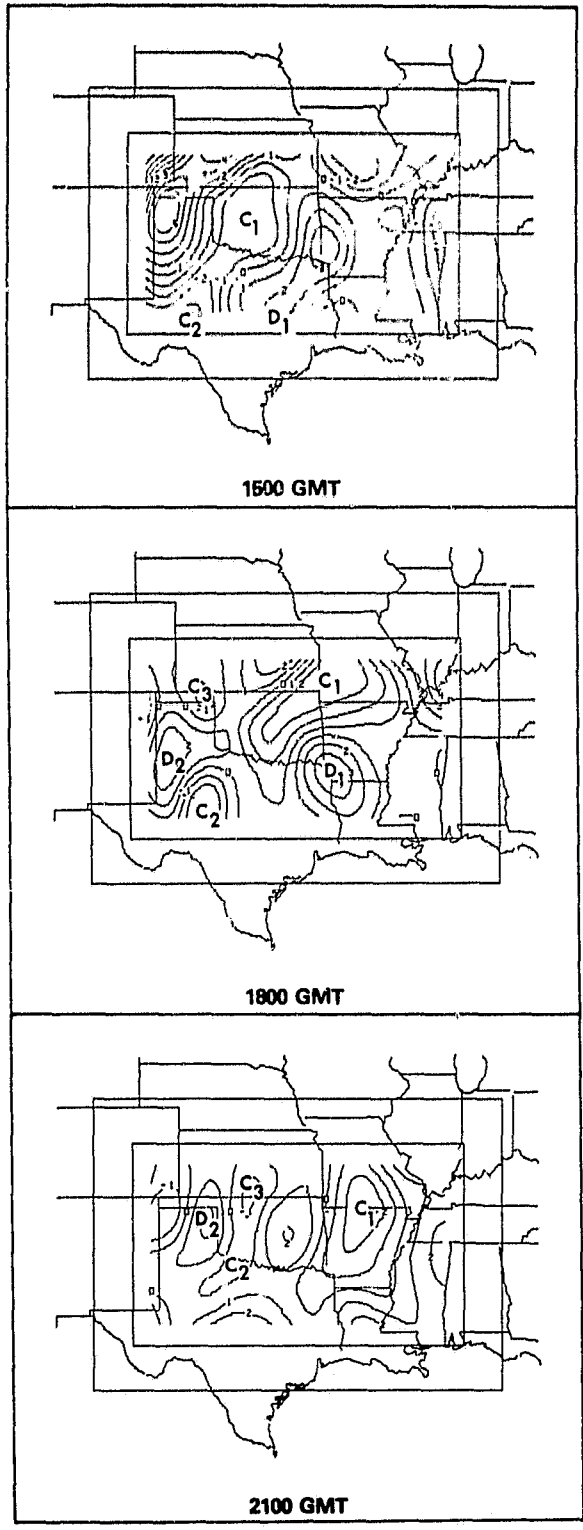
1800 GMT

2100 GMT

Figure 13. Same as in Figure 12, except $\gamma = 0.3$. Labelled features were followed from one analysis to the next.

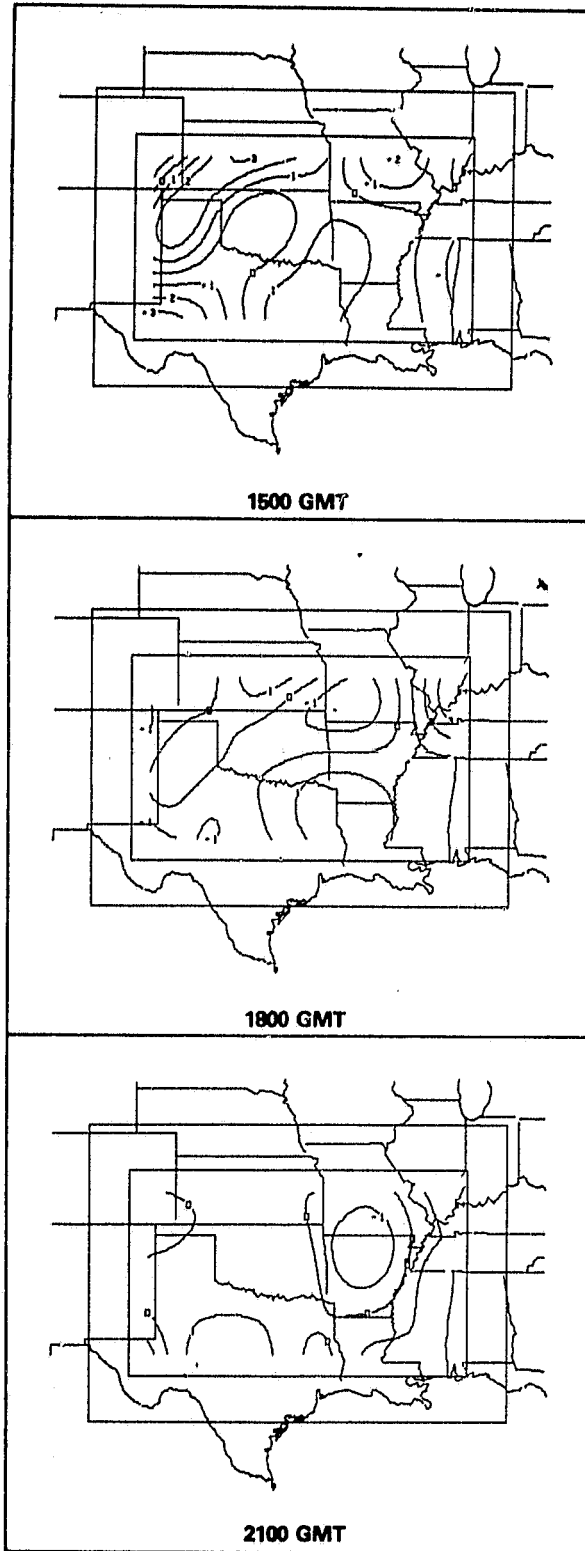**1500 GMT**

1800 GMT

2100 GMT

Figure 14. Same as in Figure 12, except $\gamma = 1.0$.

checked for gross errors, stricter quality control measures were not applied to the data set to account for any possible bias errors.

Each set of analyses exhibits at least some continuity, with the degree of continuity decreasing as γ is reduced. Features in the $\gamma = 1.0$ analysis are easy to follow, but lack detail, which is consistent with its large rms error value. Many more features appear in the $\gamma = 0.2$ analysis, but the fact that their amplitudes and propagation velocities fluctuate greatly indicates an unacceptably low signal-to-noise ratio in the analysis. This observation is clearly evident in the case of the isolated convergence maxima $(-3 \times 10^{-5} \text{ s}^{-1})$ over eastern Oklahoma and western Missouri at 1800 GMT (Figure 12b). Such small-scale features are evidently not resolvable with this particular data set.

In contrast, the alphabetically labelled features in the $\gamma = 0.3$ analysis can all be traced with acceptable temporal continuity over at least two time periods. Furthermore, these features move at velocities (20-30 m s^{-1}) characteristic of the observed wind speeds (see Figure 8) and in a northeasterly direction consistent with the prevailing advecting flow. Other unlabelled centers either dissipate between periods (as with the divergence maximum over southwestern Arkansas at 1500 GMT) or move out of the gridded display area too quickly to be followed (as with the divergence maximum over the Texas Panhandle at 1500 GMT). Thus, the $\gamma = 0.3$ analysis provides the best results in terms of rms error and temporal/spatial continuity criteria.

d. Effect of Grid Size on the Objective Analysis

Recalling that for a given data spacing Δn, the GEMPAK Barnes program constrains Δx to range between $\Delta n/3$ and $\Delta n/2$, then for the SESAME wind example whereby $\Delta n = 2.0°$, Δx must lie between 0.65° and 1.0°. The latter case appeared in Figure 10a, and for comparison the case $\Delta x = 0.65°$ appears in Figure 15. The differences between the two analyses are insignificant. This comparison emphasizes the fact that *variations of the grid size within the allowable limits has a negligible effect upon detail in the objective analysis,* and that it should not be varied with that purpose in mind. Rather, it should be varied for such purposes as to ease the discussion (as here,
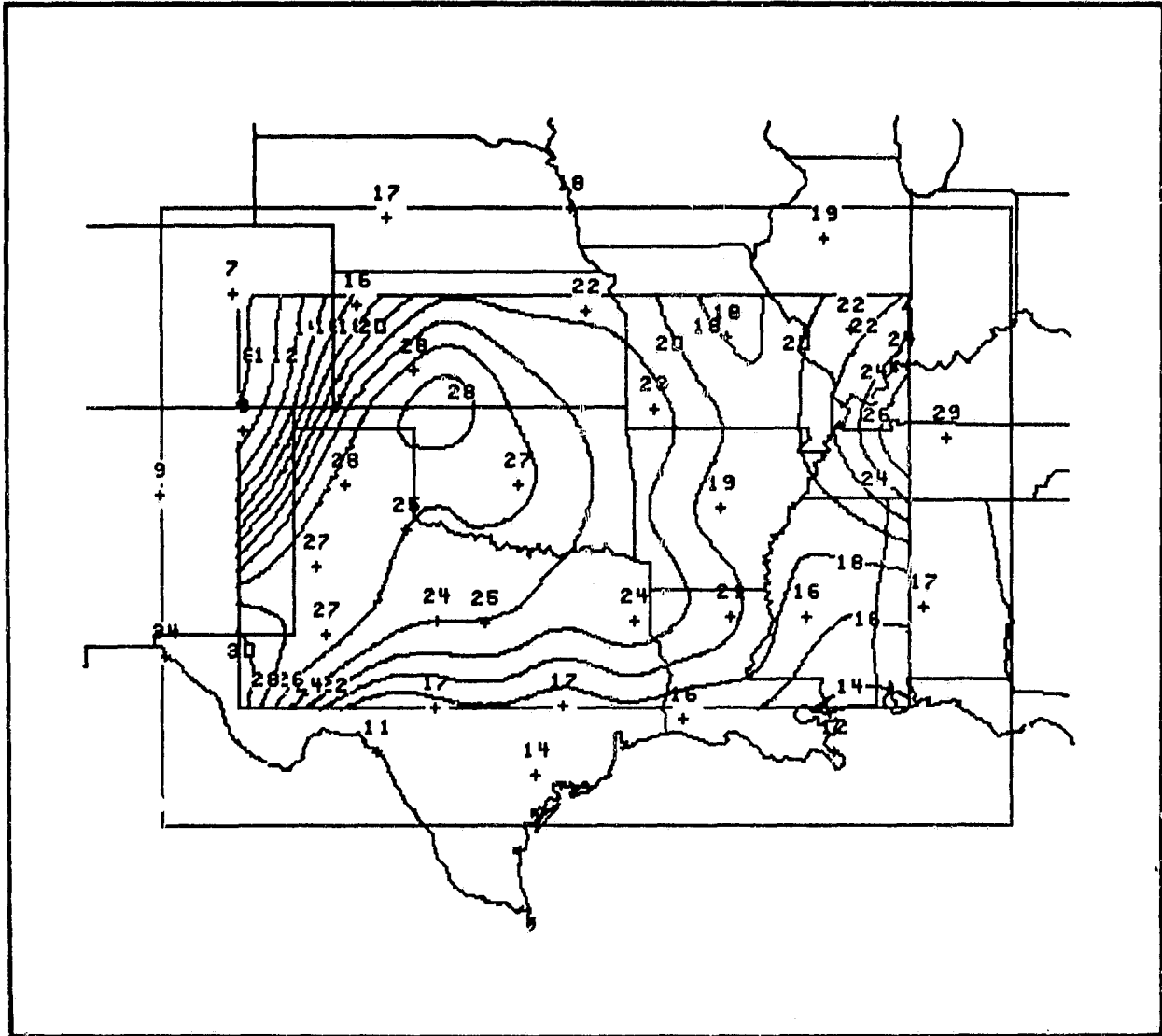
Figure 15. Effect upon objective analysis of reducing grid size to its smallest value permitted within the GEMPAK Barnes scheme ($\Delta x = \Delta n/3$). This analysis should be compared to that in the upper portion of Figure 10.

Δn was increased from $1.93°$ to $2.00°$ to make Δx = $1.00°$ a convenient round number), or to make a one-to-one comparison between objectively analyzed data and numerical model data.

e. Objective Analysis of Non-Uniformly Distributed Data

A comparison is made between an objective analysis of SESAME rawinsonde reported winds at 825 mb at 1800 GMT on April 10, 1979 (Figure 16a) and an objective analysis of those same winds when combined with non-uniformly distributed satellite data (Figure 16b). These analyses are being used in an on-going investigation at NASA/GLAS to determine to what extent assignment of the cloud motion vectors to an incorrect height will degrade a conventional analysis of winds at the SESAME regional scale (Peslen, et al., 1982). With this purpose in mind, the grid display area is adjusted to just barely cover the area of satellite data. Likewise, for this purpose, the value of Δn input to the computer is that of the uniformly distributed, conventional rawinsonde data of the SESAME regional scale (Δn = $2.0°$), rather than the computed minimum observation spacing Δn_c, which has a very small value because of the small distances between adjacent satellite observations. Moreover, use of Δn_c would have been impossible on the AOIPS computer (see Section 3e), because Δx would also have been too small and hence KX (KY) too large.

Clearly, the satellite data do exert an impact upon the conventional wind analysis on the scale resolvable by the SESAME regional rawinsonde network. The question remains, however, whether the satellite data can provide useful meteorological information on scales smaller than is resolvable by these rawinsondes. At least part of the difference between Figures 16a and 16b can be explained by a systematic mis-assignment of all the cloud vectors to one isobaric level, when in actuality the clouds exist at different heights above the local terrain (Peslen, et al., 1982). The problem of mis-assignment of heights is most crucial in a vertically sheared environment. The following questions need to be unravelled: (1) what coordinate system to assign cloud vectors to with minimum error; (2) what the smallest scale is at which satellite wind data provides reliable information; and (3) what observation spacing Δn is the proper one to employ with such non-uniformly
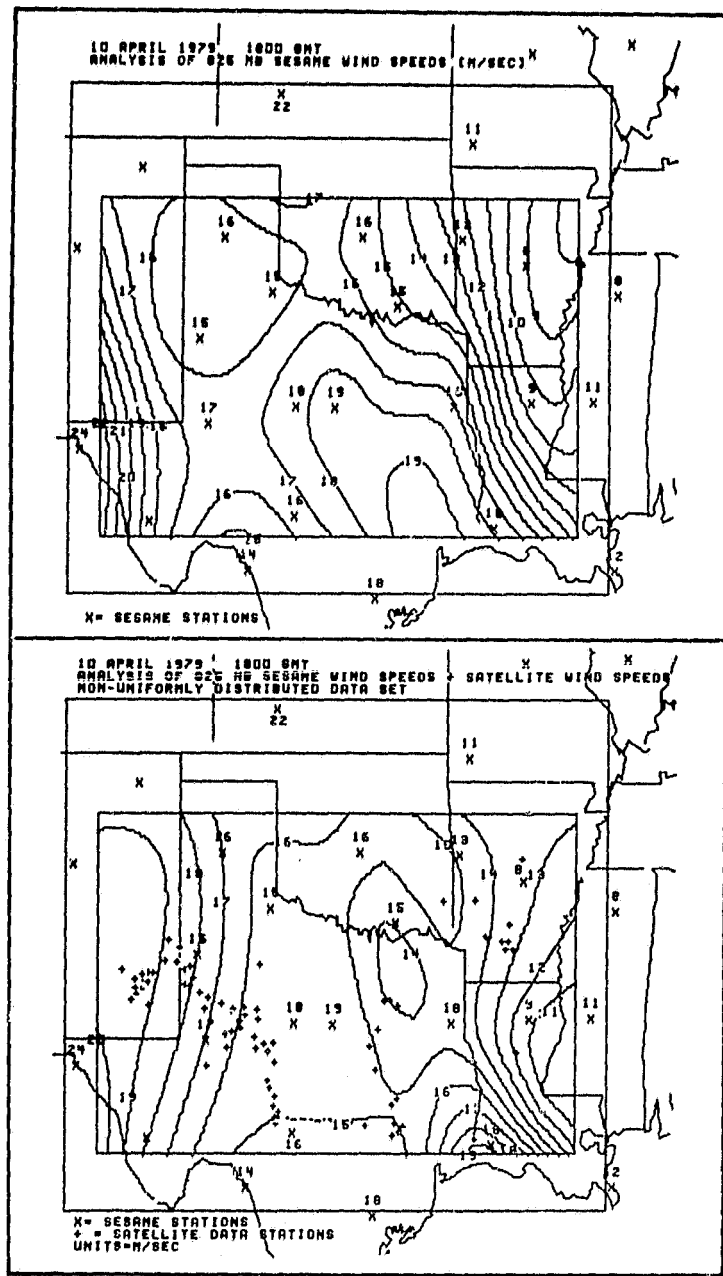
41

Figure 16. Comparison of objective analyses of uniformly (upper figure) and non-uniformly (lower figure) distributed wind data sets. Data spacing in lower analysis is set equal to that in the upper analysis in order to examine impact of satellite-derived cloud motion data upon scales resolvable by the rawinsondes.

distributed data. Further refinement of such analyses as that in Figure 16b can only be made following the unraveling of these questions.

6. SUMMARY

The specific adaptation of the Barnes (1973) objective analysis scheme to GEMPAK on the AOIPS computer at NASA/Goddard has been described in this report. The Barnes scheme was selected because it is computationally efficient, has filter response characteristics that are known functions only of the data, and adequately recovers details on the second pass through the data even when a large influence radius is used to insure that sufficient data influence is exerted at all grid points. A mathematical analysis proves that the scheme is absolutely convergent, and that only two passes through the data are needed to enable the analyst to control resolvable small-scale details in the interpolated fields.

The GEMPAK version of the Barnes scheme is unique in that it makes possible a sound objective analysis of any two-dimensional scalar data field by incorporating several objective constraints upon the analysis over which the user-analyst has no control, while still providing some user input via an interactive computer video terminal. Thus, the GEMPAK Barnes scheme is altogether *versatile, objective, and practical* (Figure 17).

By permitting user selection and alteration of the following input parameters, the scheme retains its necessary *versatility*:

(1) Data area as subset of a data file.

(2) Gridded display area as subset of the data area.

(3) Data spacing Δn.
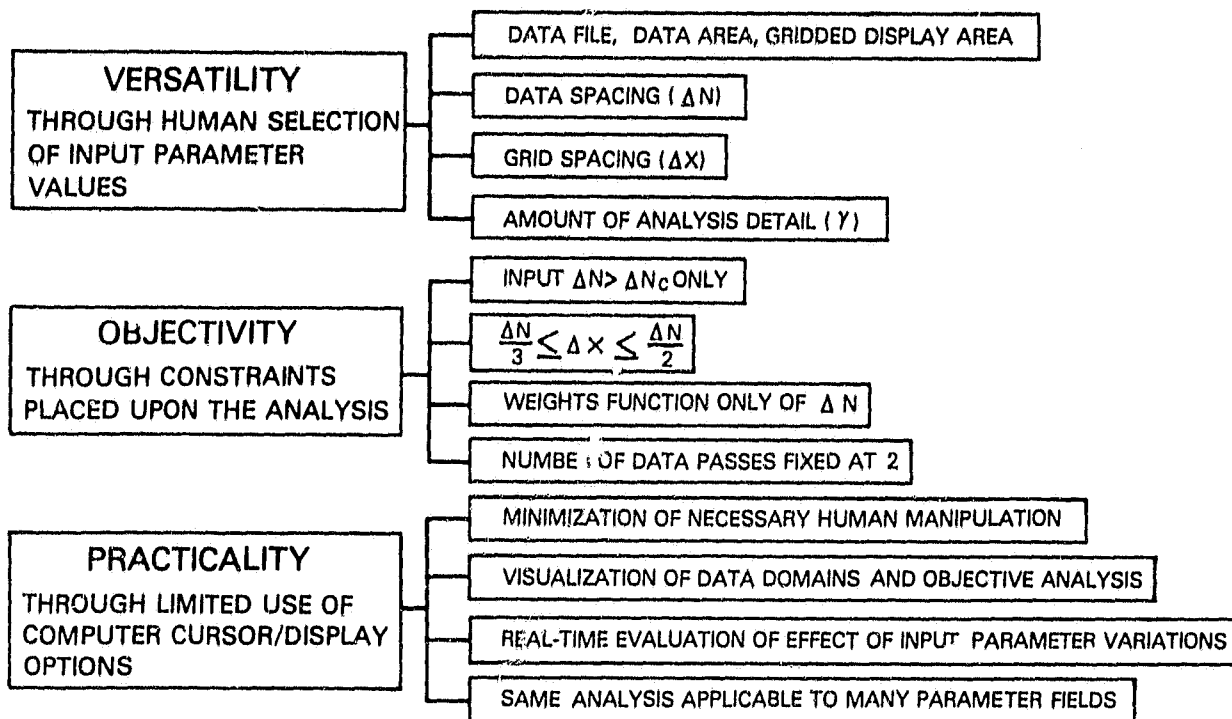
(4) Grid spacing Δx.

Figure 17. A summary description of the unique features of the GEMPAK Barnes scheme

(5)   The numerical convergence parameter γ, within the limits $0.2 \leqslant \gamma \leqslant 1.0$ imposed

(determines the degree to which the interpolated field converges to the observed data

field, i.e., the amount of analysis detail).

By retaining the following *objective features*, which are unique features of the GEMPAK Barnes

program, a reliable objective analysis can be made without unnecessary subjective human intervention:

(1)   The detail in the analysis is constrained by the data distribution as the only permitted Δn

that can be inserted is one larger than Δn_c (computed minimum spacing between data

points, averaged over the data area).

(2)   Bounds are placed upon Δx by the data spacing, namely $\Delta n/3 \leqslant \Delta x \leqslant \Delta n/2$, to insure

proper representation of resolvable features.

(3)   The weights of the Barnes low-pass filter are determined solely by the data spacing, and

are calculated internally with no user control once Δn has been input. The most detailed

44

analysis permitted is that one resulting from input of $\Delta n = \Delta n_c$ and $\gamma = 0.2$, in which case the $2 \Delta n_c$ wave is assigned a mere $1/e$ response amplitude.

(4) No user control of the number of passes through the data is permitted, as this number is fixed at two based upon a mathematical analysis of the convergence properties of the Barnes scheme (presented in the Appendix).

Use of an interactive, menu-type format and various cursor/display controls on a computer video screen makes the GEMPAK Barnes scheme an *easy and practical* operation. The user steps through a series of menus (questions) displayed on the screen to help in the selection of input parameters, quickly sees the data and grid display areas and the contoured, interpolated fields presented on a map, and can then alter the values of the parameters to test the effects upon the fields. In addition, several quality control indicators are presented to him (such as rms difference between the interpolated and observed fields). Practicality is augmented by being able to use the same calculated weights for many parameter fields. Thus, a person with only a superficial knowledge of the Barnes scheme can quickly obtain sound objective analyses that faithfully represent the data.

Two cases were presented to illustrate the versatility, objectivity, and practicality of the GEMPAK Barnes scheme. The first, that of a uniformly distributed rawinsonde data set, is one in which Δn was assigned its calculated value Δn_c, so that the $2 \Delta n_c$ wave was essentially filtered from the gridded fields. It was demonstrated that when 0.2 is selected for the γ value, the maximum detail for a given data spacing is realized in the gridded fields. Such a small γ value is justifiable only when the data are not substantially contaminated by errors and subgrid-scale atmospheric processes (which may unrealistically alias energy to larger wavelengths). In the case presented, the data were rather good and $\gamma = 0.3$ was chosen as the best analysis. The effects of varying the other input parameters upon the gridded analysis were also shown.

In the other case, inclusion of satellite-derived cloud motion vectors with the rawinsonde data resulted in a non-uniform data set. Use of Δn_c would be inappropriate here, because it has yet to

be determined what the smallest scale is at which such satellite data provide reliable information. It is for such a reason as this that the user is allowed to intervene (by inputting a larger Δn than the calculated one), with the result that a less detailed analysis is made.

The practicality of the GEMPAK Barnes scheme is demonstrated by the fact that each of the analyzed maps in the cases just presented were produced in hard copy form at intervals of 3-4 minutes once the set of input parameters had been decided upon. For an experienced AOIPS computer user, it takes one 1-2 hour session to generate the map display and go through the variations leading to the final decision of which parameter values to use in future map analysis generation with the given dataset.

7. FUTURE PLANS

Several modifications to the GEMPAK Barnes scheme are both anticipated and possible with present AOIPS computer resources. Upstream-downstream enhancement of the weighting parameter κ_0 would be preferable in jet-like regimes where isolines of the analyzed scalar tend to align with the wind direction (Endlich and Mancuso, 1968). Use of the following modification

$$\kappa_0' = \kappa_0 (1 + \beta \cos^2 \phi) \tag{19}$$

has been suggested by Barnes (1973), where $\beta = V/V^\star$, ϕ is the angle between the wind vector V and the vector that points between grid point and observation, and $V^\star$ is a scaled wind speed with value of from 1 to about 3. In the case of hurricane observations, a modification of κ_0' to account for its circular flow (as attempted by Bergman and Carlson (1975)) is desirable, perhaps based upon the Rankine vortex model.

One highly desirable addition to the present package is a comprehensive, automated, interactive data editing/quality control routine that would allow the analyst the power to selectively alter or omit troublesome observations. Displays of data time series, time tendencies, map plots, hydrostatic and superadiabatic checks, etc. are envisioned for this addition.

46

Additional work is needed to more objectively determine what data spacing to use with non-uniformly distributed data sets. It is likely that the optimum Δn will be a function of the representativeness, or error magnitude, characteristic of the particular data type as well as the nature or degree of non-uniformity of the data distribution. Map displays of both the influence radius and the difference between observed and interpolated values at actual station locations would help one see the spatial variation in accuracy of the analysis due to non-uniformity of the data distribution.

Access to a larger computer can make possible several other extensions to the core GEMPAK Barnes scheme: (1) use of time series data can enhance detail in the analysis by adding more data in space through a time-to-space conversion process (Barnes, 1973); (2) a three-dimensional Barnes analysis scheme would better serve the needs of dual Doppler radar meteorologists than a two-dimensional one;[5] (3) finally, it is envisioned that two-dimensional gridded fields produced by the GEMPAK Barnes scheme can be used as initial fields for three-dimensional mesoscale analysis schemes that require "data at grid points," such as variational schemes that incorporate various physical constraints into the analysis.

---

[5]According to Dr. Gerald Heymsfield (NASA/GSFC), personal communication.

REFERENCES

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, *3*, 396-409.

_____, 1973: Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069, (NTIS COM-73-10781), 60 pp.

Bergman, K. H., and T. N. Carlson, 1975: Objective analysis of aircraft data in tropical cyclones. *Mon. Wea. Rev.*, *103*, 431-444.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, *87*, 367-374.

Doswell, C. A., 1977: Obtaining meteorologically significant surface divergence fields through the filtering property of objective analysis. *Mon. Wea. Rev.*, *105*, 885-892.

Endlich, R. M., and R. L. Mancuso, 1968: Objective analysis of environmental conditions associated with severe thunderstorms and tornadoes. *Mon. Wea. Rev.*, *96*, 342-350.

Fuelberg, H. E., 1974: Reduction and error analysis of the AVE II pilot experiment data. NASA Contractor Report CR-120496. Marshall Space Flight Center, Alabama, 140 pp.

Gerhard, M. L., H. E. Fuelberg, S. F. Williams, and R. E. Turner, 1979: AVE-SESAME I: 25-mb sounding data. NASA Technical Memorandum TM-78256. Marshall Space Flight Center, Alabama, 361 pp.

Hillger, D. W., and T. H. Vonder Haar, 1981: Retrieval and use of high-resolution moisture and stability fields from Nimbus 6 HIRS Radiances in pre-convective situations. *Mon. Wea. Rev.*, *109*, 1788-1806.

Inman, R. L., 1970: Papers on operational objective analysis schemes at the National Severe Storms Forecast Center. NOAA Tech. Memo. ERL NSSL-51, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069, (NTIS COM-71-00136), 91 pp.

Maddox, R. A., 1980: An objective technique for separating macroscale and mesoscale features in meteorological data. *Mon. Wea. Rev.*, *108*, 1108-1121.

_____, and T. H. VonderHaar, 1979: Covariance analyses of satellite-derived mesoscale wind fields. *J. Appl. Meteor.*, *18*, 1327-1334.

Negri, A. J., and T. H. VonderHaar, 1980: Moisture convergence using satellite-derived wind fields: a severe local storm case study. *Mon. Wea. Rev.*, *108*, 1170-1182.

Peslen, C. A., 1980: Short-interval SMS wind vector determinations for a severe local storm area. *Mon. Wea. Rev.*, *108*, 1407-1418.

_____, S. E. Koch, and L. W. Uccellini, 1982: A study of the height specification problem for satellite-derived wind vectors within a severe storm environment. *Preprints, 12th Conf. on Severe Local Storms*, San Antonio, Amer. Meteor. Soc., 489-492.

Peterson, D. P., and D. Middleton, 1963: On representative observations. *Tellus, 15*, 387-405.

Simpson, J., R. Adler, D. Atlas, L. Dubach, A. Hasler, E. Rodgers, W. Shenk, and L. Uccellini, 1980: Guidelines for the Severe Storms Program of the Goddard Laboratory for Atmospheric Sciences in the 1980's. *NASA Report X-910-80-22*, NASA/GLAS, Greenbelt, MD 20771, 86 pp.

Stephens, J. J., and J. M. Stitt, 1970: Optimum influence radii for interpolation with the method of successive corrections. *Mon. Wea. Rev.*, *98*, 680-687.

APPENDIX

The mathematical-numerical analysis presented here shows the Barnes (1973) scheme is abso-
lutely convergent, but that no significant improvement in filter response fidelity (steepness of
response curve) is obtained by making more than two passes through the data when γ is chosen
sufficiently small. It is concluded that only one correction pass is sufficient to achieve acceptable
analysis convergence at resolvable wavelengths.

A. Convergence Proof

Convergence in an absolute sense occurs when the difference between the observed and analyzed
data fields vanishes as the number of passes through the data approaches infinity, or stated mathe-
matically

$$\lim_{N \to \infty} [f(x,y) - g_N(x,y)] = 0, \qquad (A1)$$

where $g_N(x,y)$ is the interpolated field obtained after N *iterations* (N+1 passes). It was shown in
the main text that a single application of the reduced weight parameter κ_1 (2) upon the weight function
$w_m(1)$ results in the true response function $D_1^\star(11)$ and the "difference field response function"
D_1 (6) at the second ("correction") pass. Weight functions at additional passes using successively
decreasing values of κ_N, assuming that γ is kept constant through all passes, are defined by:

$$\kappa_N = \gamma \kappa_{N-1} = \gamma^N \kappa_0 \qquad (A2)$$

Application of these additional filter functions result in "difference response functions" at each
pass given by

$$D_N = D_0^{(\gamma^N)} . \qquad (A3)$$

When the value chosen for γ is less than unity, the response is further accentuated at each additional
pass, particularly at short wavelengths where the initial response D_0 is small.

It follows from (5) that the third pass (N=2) interpolated field is:

50

$$g_2 = g_1 + (f - g_1) D_2, \tag{A4}$$

or upon substitution of the expression for g_1 from (5):

$$g_2 = g_0 + (f - g_0) [D_1 + D_2 (1 - D_1)]. \tag{A5}$$

Following this procedure one step further, one can easily show that on the fourth pass:

$$\begin{aligned}
g_3 &= g_2 + (f - g_2) D_3 \\
&= g_0 + (f - g_0) [D_1 + D_2 (1 - D_1) + D_3 (1 - D_2) (1 - D_1)]. \tag{A6}
\end{aligned}$$

Accordingly, the general form of the equation describing the Nth pass interpolated field is:

$$g_N = g_0 + (f - g_0) \left\{ D_1 + \sum_{i=2}^{N} [D_i \prod_{j=2}^{i} (1 - D_{j-1})] \right\}. \tag{A7}$$

To obtain the condition for convergence, this equation is substituted into (A1), resulting in:

$$\lim_{N \to \infty} (f - g_0) \left\{ 1 - \left\{ D_1 + \sum_{i=2}^{N} [D_i \prod_{j=2}^{i} (1 - D_{j-1})] \right\} \right\} = 0. \tag{A8}$$

Finally, because $(f - g_0)$ is a fixed constant, convergence is attained when

$$\psi_N = 1 - D_1, \tag{A9}$$

where D_1 is given by (6) and

$$\psi_N = \sum_{i=2}^{N} [D_i \prod_{j=2}^{i} (1 - D_{j-1})]. \qquad 2 \leqslant N \leqslant \infty \tag{A10}$$

Before proceeding with the analysis of the convergence criterion, it is important to understand the relationship between it and the actual response function at the Nth iterative pass $D_N^\star$, defined by (see ref.)

$$D_N^\star = \frac{g_N}{f}. \tag{A11}$$

51

Use of (A7) in (A11) results in the simple relationship:

$$D_N^\star = D_0 + (1 - D_0)(D_1 + \psi_N). \qquad 2 \leqslant N \leqslant \infty. \qquad (A12)$$

Thus, convergence defined in the equivalent sense to (A1), namely

$$\lim_{N \to \infty} g_N = f, \qquad (A13)$$

demands that $D_N^\star$, as defined by (A12), approach unity in light of criterion (A9).

The nature of the convergence properties of the Barnes (1973) objective analysis scheme can be understood from an analysis of (A9), (A10), and (A12). It can be shown analytically that (A10) is a convergent power series, and numerically that it converges to the value given in (A9). Applying the ratio test for convergence to (A10), we have

$$\frac{D_{N+1} \prod_{j=2}^{N+1} (1 - D_{j-1})}{D_N \prod_{j=2}^{N} (1 - D_{j-1})} = \frac{D_{N+1}}{D_N} (1 - D_N) \equiv L. \qquad (A14)$$

The series (A10) is absolutely convergently only if $L < 1$. Substitution of (A3) into (A14) results in:

$$L = D_0^{[\gamma^{N+1} - \gamma^N]} (1 - D_0^{\gamma^N}). \qquad (A15)$$

Recalling that D_0 is constant for any given choice of weight factor κ_0 and has a value $0 \leqslant D_0 \leqslant 1$, then L must be a constant, as required for convergence. In the limiting case as $N \to \infty$, both $\gamma^{N+1} \to 0$ and $\gamma^N \to 0$ for the range $0 < \gamma < 1$, so that $L \to 0$. Thus, there is absolute convergence in the limiting sense as $N \to \infty$.

The results of calculating $D_N^\star(D_0, \gamma, N)$ on a programmable desk computer (Figure A1) indicate that it converges rapidly to unity, as required. Barnes (1973) noted that the fastest restoration of small wavelength amplitude suppressed in the first pass filtering-interpolation (D_0) is obtained with

the smallest values of γ (Figure 2). For a choice of $\gamma = 0.20$, the value of $D_N^\star$ approaches 1.0 to within ten decimal places by the $N = 6$th iteration, whereas the same degree of convergence is reached by $N = 9$ when $\gamma = 0.45$ ($D_0 = 0.0064$). Convergence is obtained relatively quickly even for $\gamma = 0.8$ and is attainable in a finite time as long as $\gamma < 1$. Thus, the Barnes (1973) objective analysis scheme converges to the value specified in (A9) when $0 < \gamma < 1$. It is noted that this scheme forces the interpolated field to converge much more rapidly to the observed field than does the Barnes (1964) scheme, as can be easily seen by comparing (A12) to equation (20) of that paper.

B.   Effect of Multiple Passes Upon Filter Fidelity

Having thus verified that convergence is attainable, the question of whether making more than two passes through the data can effect a significant enhancement of the small, but *resolvable*, waves must be answered. For most purposes, it is desirable to suppress the response to a wave whose scale does not exceed twice the average minimum data spacing ($\lambda < 2\Delta n$). Considering the Gaussian nature to the response function at the first pass given by (4) and illustrated at additional passes by Figures 2 and A1, then the final response at this minimum resolvable scale should be limited by $D_N^\star(\lambda = 2\Delta n) \leqslant e^{-1}$. Under this constraint, high frequency "noise" generated by random errors and energy aliased from shorter wavelengths to larger wavelengths will be effectively filtered from the analysis. The aliased energy can result from both the data discretization process and the sampling of the atmosphere when it exhibits such subgrid-scale events as thunderstorms.

Thus, the question can be rephrased as, for a given fixed weight parameter κ_0, whether the steepness of the filter response curve at wavelengths larger than $2\Delta n$ can be appreciably increased (filter fidelity enhanced) by making more than two passes under the constraint that the final response at $\lambda = 2\Delta n$ does not exceed e^{-1}. A comparison was made between the response curves generated by a two pass, small γ filter and those generated by multiple pass, larger γ filters, since the best filter fidelity is obtained when γ is smallest (Figure 2).

The results of the comparison (Figure A2) were obtained by employing the following procedure. First, the $2\Delta n$ wavelength was defined in terms of known quantities. Since the weight parameter
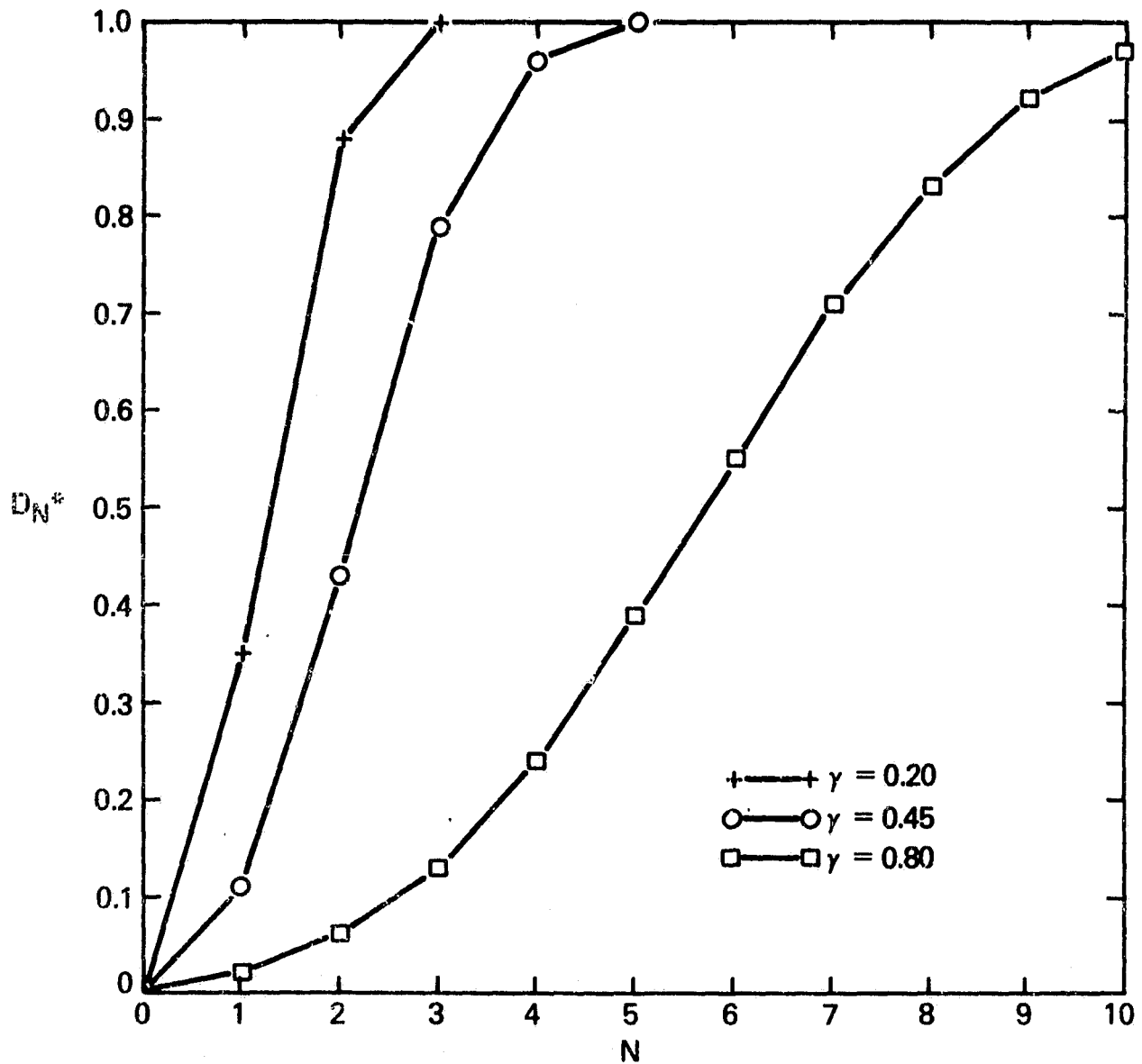
Figure A1. Nature of Barnes analysis convergence properties. The variation of analysis response at

the Nth iterative pass $D_N^\star$, given by (A12), with the numerical convergence parameter γ

is shown in the case of $D_0 = 0.0064$ (used to correspond to a $2\Delta n$ wave when $\gamma = 0.2$).
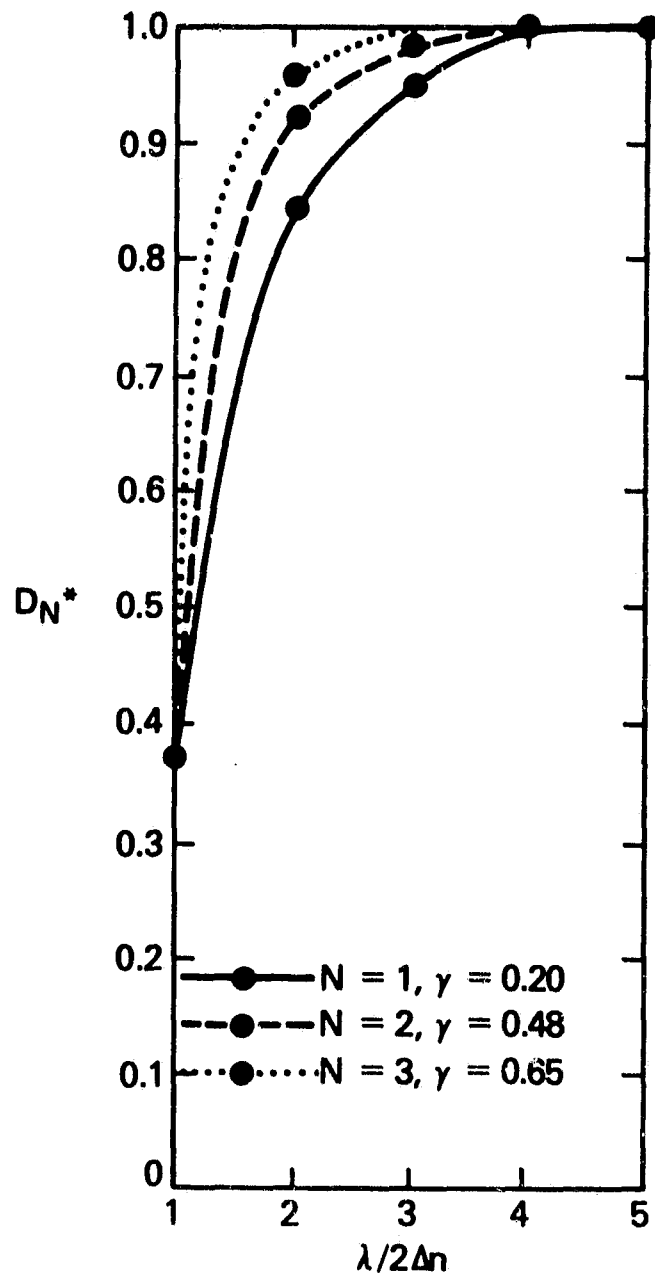
Figure A2. Effective enhancement of filter response $D_N^\star$ produced by making multiple passes, as function of resolvable wavelengths $\lambda > 2\Delta n$. Each response curve results from holding γ constant for N iterative passes under the constraint that $D_N^\star(2\Delta n) = e^{-1}$, from (A12), where $\lambda = 2\Delta n$ is defined by (A16) with $D_o(\lambda = 2\Delta n) = 0.0064$, and $D_o(\lambda \neq 2\Delta n)$ is calculated from (A18).

κ_0 *must be held constant* for the purpose of the comparison, and under this condition a singular relationship exists between any wavelength λ and the first pass response at that wavelength $D_0(\lambda)$, then a useful definition according to (4) would be:

$$2\Delta n = \pi\sqrt{-\kappa_0 / \ln\ D_0(2\Delta n)} \ . \tag{A16}$$

The actual value of κ_0 need not be known to use this definition for $2\Delta n$, since when κ_0 is constant

$$\frac{\lambda}{2\Delta n} = \sqrt{\frac{\ln D_0(2\Delta n)}{\ln D_0(\lambda)}} \ . \tag{A17}$$

Thus, the effectiveness of making additional passes at larger γ is examined as a function of multiples of the $2\Delta n$ wavelength. Each curve in Figure A2 is the result of the search for that value of γ which gives $D_N^\star(\lambda=2\Delta n)=e^{-1}$ ($N>1$), and then finding $D_N^\star$ at other wavelengths by the following formula derived from (A16),

$$D_0(\lambda) = [D_0(2\Delta n)]^{(2\Delta n/\lambda)^2} \tag{A18}$$

for integral values of $(\lambda/2\Delta n)$, inserting the result in (A12) for the chosen γ value, and plotting $D_N^\star$ against multiples of the $2\Delta n$ wave according to (A17). The selected value of $D_0(2\Delta n)=0.0064$ used in (A18) is that one which gives a second pass response of $D_1^\star(2\Delta n)=0.37=e^{-1}$ when $\gamma=0.2$ (see Figure A1). The value of $\gamma=0.2$ is chosen to represent the two pass, small γ case.

The results in Figure A2 show that when $N=1,2$, and 3 iterations through the data are made, the responses are 0.84, 0.92, and 0.96, respectively, at twice the minimum resolvable scale. Differences between these responses are no more than 12% and decrease at larger wavelengths. These differences are further attenuated when an even smaller γ value case is chosen as the basis for comparison with multiple pass filters. Thus, one correction pass using a small γ value provides for a highly acceptable degree of filter fidelity. Therefore, if one wishes to make the final objective analysis fit the data as exactly as possible, the same result can be obtained by making a greater number of interpolation passes (which is computationally wasteful and can cause greater "ballooning" effects in data-sparse areas), or by using a small γ value to reduce the correction pass influence radius.