U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
OFFICE OF SYSTEMS DEVELOPMENT
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 92-11

ON THE PACKING OF GRIDPOINT DATA FOR EFFICIENT TRANSMISSION

Harry R. Glahn

September 1992

ON THE PACKING OF GRIDPOINT DATA FOR EFFICIENT TRANSMISSION

Harry R. Glahn

1. INTRODUCTION

It has been a goal of the National Weather Service (NWS) for many years to transmit gridded fields of data to field forecasters. However, the concentration to date has been to provide graphical products, primarily because the Automation of Field Operations and Services (AFOS) system is oriented toward displaying graphics and not to providing enough computational power to rapidly produce large numbers of graphics on-site for subsequent display, or for providing a graphic "on-the-fly." Since AFOS communications was saturated with graphic and alphanumeric products, very few gridpoint products have been transmitted.[1] Now, with the availability and capability of PC's on-station, it is feasible to deal with gridded products locally. In addition, the upgrading of AFOS with System Z creates the communication capacity that will handle an expanded suite products for display on the AFOS Graphic Display Module (GDM) and in addition handle a significant set of gridpoint products.[2]

However, the ability of the National Meteorological Center (NMC) to create gridpoint products still far exceeds AFOS communications capacity. The NWS regions, acting through the NWS Office of Meteorology, has already provided a list of desired products exceeding AFOS capacity. Therefore, it is of considerable interest to transmit the gridpoint fields efficiently, and thereby allow a larger number of fields to be transmitted than would be possible under less efficient schemes. This is, of course, not a new subject, and methods already exist for packing meteorological data before transmission--methods which take advantage of the redundancy of the meteorological information within any given field. Two general methods that have been used are:

(1) Minimum removal methods--The World Meteorological Organization (WMO) has incorporated this method into their GRIdded Binary (GRIB) storage and transmission standard. The current documentation (WMO, 1988;

---

[1] A probability of thunderstorms product has been transmitted to the Western Region since 1986 (Reap, 1986), and similar products for other areas have since been added (Reap, 1990). Also, a few fields have been transmitted to the NWS Headquarters Test Facility and to Topeka, Kansas, to support risk reduction activities.

[2] There has been some hesitancy to send "binary" products over AFOS because of the belief that the end of message bit pattern (203_8) that can occur within the message will be misinterpreted. At the introduction of AFOS, this was a hardware limitation, but has long ago been removed. Some software associated with GDM display still has that limitation. A "work around" was devised and used whereby a 203_8 was transformed to a 20_8 14_8 and a 20_8 was transformed to a 20_8 20_8 within the message. These patterns were then decoded at the appropriate time. However, this problem (and solution) does not apply to data that are transmitted and stored but not displayed. Therefore, binary data can be transmitted and stored in the AFOS database and transmitted to PC's with no 203_8-related problem.

Stackpole, 1992) allows for the removal (subtraction) of the minimum value of the entire field, thereby reducing in many cases the number of bits required to adequately represent the individual field values. The minimum must also be transmitted in order for it to be restored to the field upon receipt. A second-order method is also described by the WMO (1988) in which minima of smaller parts of the field are removed after the overall minimum removal.[3]

(2) Differencing methods--The early AFOS documentation incorporated a method for transmission of gridded binary data.[4] In this method, first-order difference values are formed, each being the difference between a gridpoint value and a neighbor. A complete field can be recreated by using the first value and the N-1 difference values, where N is the number of points in the field. Alternatively, and this is what the AFOS documentation describes, second-order differences can be found, each being the difference between a first-order difference and a neighbor's difference. A complete field can be recreated by using the first value in the field, the first first-order difference, and the N-2 second-order differences. In order to enhance the algorithm and to take full advantage of the point-to-point correlation, differencing proceeds along a row (say, left to right), then at the right edge, the next difference is between the rightmost points on the two adjacent rows. Differencing then proceeds right-to-left, etc.

In order to study the packing efficiency of certain algorithms, software has been written to calculate statistics on real meteorological fields for four versions of the "minimum removal" method and for two versions of the "differencing" method. Finally, a procedure is described which takes advantage of the strengths of both methods.

2. GENERAL PROCEDURE

In each case, a real (floating point) array was input. The first step was to scale the values in it, usually by a power of 10, and to round them to integers. A specific number of bits in which to pack the data was not specified; rather, the data were processed to retain exactly the scaled and rounded values. That way, if accuracy of sea level pressure to tenths of millibars (mb) is desired, no information is lost other than that of rounding to tenths of mb.[5] For each method, the fields were processed according to an

---

[3] While this is sometimes thought of as a removal of the mean, it is the range of numbers that is important, and if a measure of central tendency were used, it should be the median. However, it is convenient to deal with all positive values; hence, the removal of the minimum value.

[4] A document "Universal Transmission Format" dating back to January 15, 1978, (Davis, 1978) describes this method. It has also been incorporated into the document "Standard Formats for Weather Data Exchange among Automated Weather Information Systems" (Chapter 10) published by the Office of the Federal Coordinator for Meteorological Services and Supporting Research (OFCM, 1990).

[5] This also means that no more accuracy than that specified by scaling and rounding is retained. This tack was taken for two reasons. First, it guaranteed that each method tried would retain exactly the same accuracy, and

algorithm, then reprocessed to recover the original values. In some cases actual packed "messages" were formed, then unpacked to assure correctness.

After processing the data according to one of the algorithms, the number of bytes (octets) necessary to transmit the data was calculated. Note that this does not include the field (product) description information such as is contained in the GRIB Product Definition Section (PDS)--a minimum of 28 bytes--(Stackpole, 1992, pp. 4-5), the Indicator Section (IS) of 8 bytes (op. cit., p. 4), or other optional information such as actual grid definition.

In each case, 32 bits were allocated to the first value transmitted. This could be an overall mean, the smallest value in the field, the first value in the field, etc. Although this may be larger than necessary, it is what is provided for in GRIB. In those cases where it is necessary to carry in the transmission the number of bits needed for individual values, 5 bits were used; this allows individual values of up to 31 bits. The size of the grid is represented as NX (left to right) by NY (bottom to top).

3. METHOD 1--ZERO-ORDER

Purely for comparison and not counted as one of the four minimum removal methods mentioned in the Introduction, a field was packed without removal of the minimum except when necessary to make all values positive. Each value was reduced by the minimum value only if the minimum were negative so that values transmitted (other than the minimum itself) would be positive. For large values, such as 500-mb height, not subtracting the minimum is very wasteful. For other fields which range down to zero, such as vertical velocity, there is no inefficiency. The values transmitted and associated bits are:

32 bits	the minimum value--will be zero for a positive minimum value,
5 bits	IBIT, the number of bits required for each value, and
IBIT*NX*NY bits[6]	the gridpoint values.

4. METHOD 2--FIRST-ORDER MINIMUM REMOVAL

This is the method that is currently implemented at NMC and will be called in this document "basic GRIB" [WMO, 1988, p. I-Bi-6, 92.6.3, Note (2)]. The overall minimum of the field is removed (subtracted out). The values transmitted and associated bits are:

32 bits	the minimum value,
5 bits	IBIT, the number of bits required for each value, and
IBIT*NX*NY bits	the gridpoint values.

This GRIB method, other than providing a standard "framework" or message format, does (only) two things--it provides for transmitting the "values"

---

the message sizes would have that as a common denominator. Secondly, with the more elaborate methods which involve choosing groups based on whether or not a particular accuracy can be maintained, the accuracy has to be specified.

[6] An asterisk represents multiplication.

(differences from the minimum) in as few bytes as possible *with a constant "word width"* and still maintain desired accuracy, and in case the *median* of the original values is not zero, the removal of the minimum, by making all values to be transmitted positive, *may* reduce the number of bits needed.[7]

5. METHOD 3--SECOND-ORDER MINIMA REMOVAL (1)

Many variations of second-order minima removal can be imagined. Three were programmed and are reported here. The first, designated as (1), was to subtract the minimum of each *row* of the grid after subtracting the overall minimum. The values transmitted and associated bits are:

32 bits	the overall minimum value,
5 bits	IBIT, the number of bits required for the row minimum values,
IBIT*NY bits	the row minimum value for each row,
5*NY bits	JBIT(J), the number of bits required for each value in the row (J=1,NY), and
JBIT(J)*NX bits	the row gridpoint values after the second removal for each row J (J=1,NY).

This method can be expected to have some advantage over the previous method when there is a change in the actual range (minimum and maximum, not necessarily just maximum minus minimum) when proceeding from bottom to top of the grid.[8]

6. METHOD 4--SECOND-ORDER MINIMA REMOVAL (2)

This method is similar to Method 3, except that the second-order minima are determined in groups of size MINPK, MINPK being an input parameter to the packing routine. Each group will be exactly K*MINPK in size for determining the second-order minima, where K is an integer (except for the last group in case NX*NY is not evenly divisible by MINPK). The number of bits necessary to pack the values in the first group after subtracting the second-order minimum is found. Then, the same is done for the second group *individually* and also *combined* with the first. The second group is merged with the first only if it can be done without increasing the number of bits required by either the first or second group. If the first and second groups combine (only one second-order minimum is required for both together), then a third group is considered

---

[7] Note that when both negative and positive values are present (or have to be assumed to be present), the removal of the minimum may be more than a convenience. The number of bits J necessary to accommodate numbers between X and Y (inclusive) is the smallest value of J for which the maximum of the absolute values of X and Y is less than 2^J, plus 1 for the sign. Even when X and Y are both positive, if this is not known (i.e., the presence of negative numbers cannot be ruled out), the sign bit must be allocated. For instance, suppose X = -2 and Y = 200. X requires three bits for its representation. Y requires 8 bits, yet 9 bits have to be allocated to the field to accommodate the measly -2. Removal of the minimum (-2) gives a maximum of 202 to pack, which still requires 8 bits, and a sign does not have to be provided for.

[8] The bottom is assumed to be generally nearer the equator than the top. For other orientations of the grid, left to right might be more appropriate.

for combining with the first and second groups according to the same pattern. The values transmitted for JG groups and associated bits are:

32 bits	the overall minimum value,
5 bits	IBIT, the number of bits required for the group minimum values,
5 bits	KBIT, the number of bits required for the number of values in each group,
IBIT*JG bits	the group minimum value for each of the JG groups (J=1,JG),
5*JG bits	JBIT(J), the number of bits required for each value in group J (J=1,JG),
KBIT*JG bits	NOV(J), the number of values in the group for each group J (J=1,JG), and
JBIT(J)*NOV(J) bits	the group gridpoint values after the second removal for each group J (J=1,JG).

The advantage of this method over Method 3 is that the IBIT+5+KBIT bits necessary for transmitting information about a group are not needed for each row in the grid unless the algorithm indicates such is desirable. Also, the size of the groups need not be based on the number of values in a row. The parameter MINPK can be adjusted based on experience. The disadvantage is that the number of values in each group needs to be transmitted for each group.[9]

7. METHOD 5--SECOND-ORDER MINIMA REMOVAL (3)

This method is similar to Method 4, except that the group sizes are not fixed to multiples of MINPK. Rather, a minimum group size is specified, MINPK, then an attempt is made to add values at increments specified by another input parameter, INC. Consider that a group of size A exists, which is at least MINPK in size. An attempt is made to add INC values to it. If this does not increase the number of bits necessary to pack values in group A, and the number of bits required for a group, group B, of size MINPK immediately following group A is not less than those required for group A, then the INC values are added and the new group is now called group A. If the addition of the INC values increases the bits necessary to pack values in group A, the values are not added and the group A is complete. The process starts over with the formation of a new group A of size MINPK. Special consideration keeps the last group from being very small. The values transmitted are the same as for Method 4.

A disadvantage of this method compared to Method 4 is that the packing process (actually the definition of the groups) is slightly more computationally intensive. We note that as a more concerted effort is made to use as few bits as possible to transmit the individual values, more overhead is incurred in terms of the parameters that must be sent of effect the unpacking. For instance, Methods 4 and 5 require that the size of each group be sent, while for Method 3 each group is the number of values in a row, which is assumed known from header information. (In any case, only one value would be required, not one per row.)

---

[9] Variations on how to transmit the group sizes are possible, but likely to not be of importance in saving bits.

As stated previously, GRIB also provides for second-order minima removal [WMO, 1988, p. I-Bi-6, 92.6.3, Note(3)]. Sections 5, 6, and 7 in this document specify algorithms for determining the groups on which to define the second-order minima. These algorithms allow for zero bits to be transmitted for groups in which all values are equal, (differences are all zero). GRIB also provides for this possibility. A subroutine which can be used for Methods 4 and 5 is given in the Appendix.

8. METHOD 6--FIRST-ORDER DIFFERENCES

This method has been essentially described in the Introduction. The minimum first-order difference is subtracted to make all differences positive. The values transmitted and associated bits are:

32 bits	the "first" value in the field,
5 bits	MBIT, the number of bits required for the absolute value of the minimum first-order difference,
MBIT bits	the absolute value of the minimum first-order difference,
1 bit	the sign of the minimum first-order difference,[10]
5 bits	IBIT, the number of bits required for each first-order difference, and
IBIT*(NX*NY-1) bits	the first-order differences.

One would expect that for realistic free-air meteorological fields, the differences between adjacent gridpoints would be less than the differences between gridpoints and some overall or group minimum, unless the groups were so small that the overhead of sending group minima, sizes, etc. would be prohibitive. However, for highly variable fields, this may not be true.

9. METHOD 7--SECOND-ORDER DIFFERENCES

This method has also been described sufficiently in the Introduction. The minimum second-order difference is subtracted to make all differences positive. The values transmitted and associated bits are:

32 bits	the "first" value in the field,
5 bits	MBIT, the number of bits required for the absolute value of the first first-order difference,
MBIT bits	the absolute value of the first first-order difference,
1 bit	the sign of the first first-order difference,
5 bits	NBIT, the number of bits required for the absolute value of the minimum second-order difference,
NBIT bits	the absolute value of the minimum second-order difference,
1 bit	the sign of the minimum second-order difference,
5 bits	LBIT, the number of bits required for each second-order difference, and
LBIT*(NX*NY-2) bits	the second-order differences.

---

[10] The minimum first order difference could, of course, be treated as a signed number of MBIT+1 bits.

Note, again, that all differences are being treated as positive; the sign is sent as a separate bit only for the first first-order difference and the minimum second-order difference.

A relevant question is, "Are second-order differences smaller than first-order differences?" They are not guaranteed to be, but experience has shown that for free-air meteorological variables the range about zero is usually less for the second-order differences. The disadvantage of the differencing methods described is that only one "word size" is provided for (as distinct from Methods 3, 4, and 5, where word size can vary over the field). Even though the magnitudes of the second-order differences may be less than those of the first-order differences, they may not be sufficiently so to be accommodated in less bits. For instance, numbers in the range -120 to +120 require 8 bits, but so do those in the range -70 to +70.

The second-order minima removal methods are based on the assumption that the field values vary in an organized manner along rows or from bottom to top of the grid. Similarly, one could assume that first- or second-order differences would vary in an organized manner over the grid. The next method combines the desirable aspects of both differencing (point-to-point redundancy is exploited) and of minima removal ("word width" can be reduced by group minima removal and can vary over the field).

10. METHOD 8--MINIMA REMOVAL FROM SECOND-ORDER DIFFERENCES

This MRSOD (Minima Removal from Second-Order Differences) method is a combination of Methods 7 and 5. First, the second-order differences are found, then those NX*NY-2 differences are grouped and packed according to Method 5. The values transmitted and associated bits are:

32 bits	the "first" value in the field,
5 bits	MBIT, the number of bits required for the absolute value of the first first-order difference,
MBIT bits	the absolute value of the first first-order difference,
1 bit	the sign of the first first-order difference,
5 bits	NBIT, the number of bits required for the absolute value of the minimum second-order difference,
NBIT bits	the absolute value of the minimum second-order difference,
1 bit	the sign of the minimum second-order difference,
5 bits	IBIT, the number of bits required for the group minima,
5 bits	KBIT, the number of bits required for the number of values in each group,
IBIT*JG bits	the group minimum value for each of the JG groups (J=1,JG),
5*JG bits	JBIT(J), the number of bits required for each value in group J (J=1,JG),
KBIT*JG bits	NOV(L), the number of values in the group for each group J (J=1,JG), and
JBIT(J)*NOV(J) bits	the group second-order differences after removal of the group minimum for each group J (J=1,JG).

11. RESULTS ON LFM 33 X 29 GRIDS FOR JUNE 16, 1992, 0000 UTC

A number of 33 X 29 Limited-area Fine Mesh (LFM) model (Gerrity, 1977) fields were readily available. Each field was processed by each of the algorithms described above and statistics calculated. Three were processed with more than one scaling. The fields processed and some of the statistics are shown in Table 1. The resolution for each variable before scaling and rounding (see Section 2) was to the fourth decimal place (hundred thousandths) in the units indicated in Table 1, except for 500-mb height which had been rounded to quarter meters. A map of 500-mb height is shown in Fig. 1. When values of MINPK and/or INC were required, 33 and 3, respectively, were used. Some points to note from Table 1 are:

- In those cases where all values are far above zero--e.g., 500-mb height and 1000-mb temperature--the basic GRIB (Method 2) provides for a significant reduction in bytes required, in addition to using a word size appropriate to the accuracy to be retained (compare Methods 1 and 2). In those cases where the values range down to zero or below, basic GRIB provides only the minimum word width plus the possible advantage of using only positive numbers. For instance, compare Methods 1 and 2 for 1000-mb height and precipitable water.

- The removal of row minima (Method 3) generally improves on the basic GRIB (Method 2) by 5 to 15% with an overall improvement for all entries in Table 1 of 9%. The more elaborate group minima removals (Methods 4 and 5) gave some additional improvement. Method 4 was always better than Method 3, and Method 5 was always better than Method 4. Overall, Method 5 improved over Method 2 by 12%.

- The second-order difference method (Method 7) was always as good or better than the first-order difference method (Method 6).

- The second-order difference method was always better than the best of the minima removal methods (Method 5) and by the overall margin of 11%. One variable, 1000-mb height with the accuracy carried to tenths of meters, showed a 20% improvement.

- MRSOD (Method 8) offered further substantial improvement, improving on Method 7 overall by 14%.

- MRSOD improved on the basic GRIB overall by 33%, and over the best minima removal method (Method 5) by 24%.

The algorithm for determining the groups in Method 5 (given in the Appendix), which is also used in MRSOD, Method 8, is relatively insensitive to the parameters MINPK and INC. As stated previously, the values in Table 1 were computed with MINPK = 33 (the number of values in a row) and INC = 3. Table 2 shows some results of varying MINPK for Methods 4, 5, and 8, for INC = 3.

The values in the table show that the number of bytes required is surprisingly stable for values of MINPK between about 17 and 99, although there is an ill-defined minimum at 17 to 33. This is probably related to the grid width of 33. Groups of size 4 and 8 are too small--the overhead of sending group parameters is greater than the gain of 1 or possibly 2 bits per value by so doing.

Table 2 also indicates that for the three variables 500-mb height, 1000-mb temperature, and even boundary layer U-wind, the range of second-order differences is less than the range of first-order differences, and each is less than the range of the original values.

Table 3 shows some results of varying INC with a constant MINPK of 33. The number of bytes required is remarkably independent of INC between 1 and 33. Note that the results of INC = 33 for Method 5 match those for Method 4 which implicitly uses INC = 33 because that is the grid width.

Although there are some differences due to MINPK and/or INC, they are relatively minor and not consistent enough to strongly prefer a particular combination of MINPK and INC. The 33,3 combination seems about right.

12. DECREASING THE MESH SIZE FOR LFM GRIDS

The LFM fields have a gridspacing (mesh size) of 190.5 km at 60° North latitude, and fields from other models may be transmitted with more definition than that. In particular, a mesh size of 1/2 or 1/4 of the LFM may be used in the not distant future. To see how this might affect the results, the 33 X 29 fields were linearly interpolated to 1/2 of their original spacing, giving fields 65 X 57. While it is true that these interpolated fields carry no more information than the originals, and smaller-mesh models may be more accurate in terms of feature placement and amplitude, the overall features will still represent the future analyses just as the LFM fields do. 500-mb amplitudes would not be expected to be much different overall, although individual cases might vary. On the other hand, some fields, such as vertical velocity, might have more amplitude and, more importantly, have more small-scale spatial detail. For this test to be meaningful, it has to be assumed that the fields would not have a large component of high spatial detail. With the group sizes used, the redundancy of change due to linear interpolation should not give an advantage to the differencing methods. Maps of 1000-mb temperature and boundary layer U-wind are shown in Figs. 2 and 3. The spatial variation of the boundary layer wind in Fig. 3 is considerable, although the patterns are relatively large scale.

Table 4 shows similar statistics for the 95.25-km grid as Table 2 shows for the 190.5-km grid. The number of points is approximately quadrupled (the more exact factor is 3.87). The average message sizes for Methods 4, 5, and 8 increased by factors of 3.53, 3.43, and 2.90, respectively, where these factors were computed from the totals of 500-mb height, 1000-mb temperature, and boundary layer U-wind for MINPK = 33 and INC = 3. However, one can note from Table 4 that the optimum MINPK is no longer 17 to 33, but is more like 12 to 17. The basic GRIB gave message sizes of 4636, 4173, and 4173, respectively, for the three variables mentioned above.

The number of groups used to achieve a near minimum message size is surprisingly high, generally about 200. For instance, for the 33,3 combination for 1000-mb temperature, Methods 4, 5, and 8 gave 8, 22, and 11, respectively for number of groups for the 190.5-km grid (see Table 2). Correspondingly, the 12,3 combination for the 95.25-km grid gave values of 274, 256, and 230 (see Table 4). Also, the full software output indicates that for the 190.5-km grid the change in the number of bits required to pack a group almost always changed by at most one for adjacent groups, while for the 95.25-km grid, the bits required sometimes changed by more than one from group to group. As an example, for 1000-mb temperature and Method 8, one group of 15 required 3 bits for each value, the next group of 12 required 5, the next group of 21 required 3, the next group of 12 required 2, and the next group of 12 required 4.

Table 5 shows similar statistics for the 95.25-km grid as Table 3 shows for the 190.5-km grid. Since smaller groups tend to be better for the 95.25-km grid, small values in INC are also better, although the variation in message size for INC = 1 to 6 is only about 2%. It's interesting to note that having a high INC may actually increase the number of groups; this is because the minimum group size is almost always used with a large INC.

If we use values from Table 4 for the 12,3 combination, we find that for the three variables combined, Method 8 is 34% better than Method 5 (the best of the minima removal methods), and better than the basic GRIB by 50%. These percentages compare to 24% and 33% for the 190.5-km grid, emphasizing the greater superiority of MRSOD over the basic GRIB as the mesh size is decreased when spatial detail is not increased.

13. RESULTS ON ETA MODEL 83 X 59 GRIDS FOR AUGUST 19, 1992, 0000 UTC

Four fields produced by the NMC eta (or mesoscale) model (Mesinger, et al., 1990) were made available from the August 19, 1992, 0000 UTC run. These are forecasts from the 40-km version of the model interpolated to the southern portion of an 83 X 83 NGM (Hoke, et al., 1985) "C-grid" (Grid 105 in NMC parlance). Values above row 59 were not available, so the fields used were 83 X 59. Note that this interpolation roughly doubled the mesh length. The four fields are the 500-mb height, 700-mb omega, surface temperature, and accumulated precipitation--all 24-h forecasts, the precipitation being accumulated for 12 hours. These fields are shown in Figs. 4 through 7. Three of the four fields were chosen for study because they were expected to have smaller scale information than the LFM fields, and also smaller scale information than temperature, height, and winds well above the earth's surface.

Fig. 5 indicates a very detailed omega pattern, with the sign of these instantaneous values changing every few gridpoints. The maximum upward motion is reached in central Canada--a value of -17.8 μbars/s. The four closest values are -16.4, -0.6, -14.9, and +1.7, indicating the rapid variation in space. The second highest value in magnitude is just north of Idaho. This value of -9.8 has a neighboring value of zero the next gridpoint to the west and a +3.4 just two gridpoints to the east. Table 6 indicates that Method 4 improved on Method 3, and Method 5 improved on Method 4. However, the first- and second-order differencing methods by themselves were not competitive with the group minima removal methods because of the low spatial correlation in the fields. Method 8 was in the right ballpark, being better than Method 3, but not as good as Methods 4 and 5. In this case, the first- and second-order differences actually required a greater range of values to pack than the original values, as indicated in Table 7.

Surface temperature, shown in Fig. 6, was also quite detailed, but didn't have as much small scale variation, compared to the total range of values, as did 700-mb omega. Again, Method 4 was better than 3, and 5 better than 4. Even here, the large spatial differences did not favor the differencing methods alone, but MRSOD improved on the best of the minimal removal methods (Method 5).

A number of realistic temperature features are identifiable--the lower temperatures hugging the east and west coasts, low temperatures over Hudson Bay, warm ahead and cool behind a probable cold front nearing the U.S. east coast, warm temperatures in the San Joaquin and Sacramento Valleys and the far Southwest, and warm temperatures over Cuba.

The precipitation field also exhibited very small scale features. Fig. 7 shows two pronounced maxima, one of 6.35 inches south of Arizona and one of 1.64 inches near northern Florida. One of the neighboring gridpoints to the 6.35-inch maximum has a zero value. When differences between adjacent gridpoints are of the same order as the total variation over the field, first- and second-order differences have nothing to offer for packing possibilities, and for this precipitation field were worse than the basic GRIB. However, MRSOD improved on the basic GRIB by 51%. The group minima removal Methods 4 and 5 were even better, Method 5 improving on the basic GRIB by 65%. This means 2.89 (2.16) fields packed by Method 5 (Method 8) could be sent in the same space as 1 packed by basic GRIB. It should be remembered that the gridpoint values used here were not the actual model values, but rather those values interpolated to a coarser grid. It is likely that MRSOD would improve on the group minima removal methods if applied to the original values.

It is worth noting again that the second-order minima removal methods--including MRSOD--allow for individual values to not be transmitted for a group when they are all equal (differences equal zero). This is a form of run-length encoding that is useful on the precipitation fields. For instance, there were several groups of sizes of 30 to 40 that fit this description.

The patterns in Fig. 7 show general precipitation over much of the Atlantic and Pacific Oceans covered by the grid, general precipitation associated with the 500-mb trough in eastern Canada, precipitation along the Ohio Valley probably associated with a front, and precipitation associated with the strong upward motion at 700-mb in central Canada. The temperature field also indicates a frontal boundary associated with the precipitation and 500-mb trough in the area north of Montana and North Dakota.

If the totals in Table 6 are used as indicators, Method 5 improves on the basic GRIB by 37% and MRSOD improves on Method 5 by 7%; MRSOD improves on the basic GRIB by 41%. This value compares with 33% for the 190.5-km LFM gridpoint fields and 50% for the interpolated 95.25-km LFM fields. For the eta model, most of the improvement of MRSOD over Method 5 came from the 500-mb field, as would be expected.

Table 7 for the eta model is similar to Tables 2 and 4 for the LFM. It gives the results for Methods 4, 5, and 8 for a constant INC = 3 and varying values of MINPK for each of the four fields. The results generally agree with those of Table 4--a group size of 15 is about right, and the variation in bytes required is not great between groups of size 10, 15, and 21. Other results, not shown, indicate that varying INC from 1 to 9 for MINPK = 15 on the precipitation field changes the bytes required by at most only about 4%. It is noted that for each of the fields listed in Table 7 except 500-mb height, the range of second-order differences is greater than the range of the original values. This would not be true for many of the fields that would be of interest to forecasters; these three fields were chosen for study because of their potential for high spatial variability.

14. GENERAL COMPRESSION SOFTWARE

Commercial software is available that attempts to compress a file provided to it. One such product was readily available and was tried. The output of the MRSOD method was given to the commercial package PKZIP.[11] Except for the unrealistic situation of 60 duplicate grids in one file (on which the compression ratio was 30 to 1), we got the compression we expected--zip.

15. SUMMARY AND CONCLUSIONS

It has been shown that a significant reduction in message size can be achieved over the basic GRIB method presently in use at NMC by three versions of a group minima removal method, and a structuring of the message to retain the data resolution desired. Removing the minimum of each row reduced the size by 17%, where this value is computed on the combined totals in Tables 1 and 6.[12] Removing group minima where the groups were made up of multiple groups of constant size (Method 4) reduced the basic GRIB message size by 26%. Removing group minima where the groups were of varied size improved on basic GRIB by 28%. First-order and second-order differencing each improved on GRIB by 10%. The combined second-order differencing and group minima removal (MRSOD) improved on basic GRIB by 38%. These numbers do not include the results on interpolated LFM fields, where the improvement of MRSOD over basic GRIB was 50%.

One can argue for various reasons that these values are not (or are) representative of what could be achieved when transmitting gridpoint data over AFOS or AWIPS. Better values could be obtained if we knew the exact product suite to be transmitted. A number of decisions regarding the products desired by the field forecasters over AFOS have yet to be made, besides the basic decision of which fields will be transmitted. For instance, will the instantaneous omega fields have a time and/or space smoother applied? If the 40-km eta model output is used, will interpolation to the NGM grid without smoothing provide the best product? As has been demonstrated, the packing of model fields that are highly variable benefits considerably by the removal of group minima, differencing methods alone are not competitive, and the MRSOD combination always improves on basic GRIB. For smoother fields, MRSOD is decidedly best, and when one considers that the AWIPS plans are to send many such fields at 80-km resolution (e.g., height, temperature, wind, and relative humidity at 10 levels and 17 projections to support the Regional Areas at Weather Forecast Offices), the saving by using MRSOD rather than the basic GRIB should approach 50%; about half of this improvement would come from using second-order spatial differences and half from using second-order minima removal.

As Table 7 indicates, Method 8 is not always better than Method 5. When Method 5 is better, an indicator of this is the range of the original values compared to the range of the second-order differences. For example, for the precipitation field, the second-order differences have a range of 1350, while the original values have a range of 635. On implementation, these ranges could be computed and compared, and either Method 5 or Method 8 (MRSOD) be

---

[11] No indorsement or otherwise of the commercial product PKZIP produced by PKWARE, Inc. is implied.

[12] $[(12{,}265 + 22{,}668) - (11{,}160 + 17{,}821)]/(12{,}265 + 22{,}668) = 0.17$.

used. The packing routine could do this very easily (a "flag" in the message would be required, of course), and the unpacking routine could easily deal with this option.

The minima-removal algorithm in Methods 5 and 8 requires the two parameters MINPK and INC. These could be made to vary depending on the field being packed, but the message size is not highly dependent on them within a range for MINPK of about 10 to 20 and for INC of 1 to 10. Small values of INC are slightly better, and the difference in computer time involved was not noticeable on limited tests.

The bit arrangements in the "message" used in this study were devised for efficiency, and are not meant to be a final transmitted arrangement. However, the efficiencies quoted do depend on the bits allowed for, as stated in Sections 4 through 10, and the associated algorithms for arriving at those values.

GRIB provides a framework for a standardized message for transmitting gridpoint data. The basic minimum removal method is not very efficient for many meteorological fields produced by NMC models. Especially when we consider the limited AFOS capacity--although AWIPS-era transmission load is also important--we need to pack the data in a way that allows us to meet as high a percentage of the field requirements as we can. Second-order differencing combined with group minima removal offers a solution which will permit us to send approximately 60% to 100% more products than with the basic GRIB.

Acknowledgments. I thank Ralph Petersen, Mark Pecnick, and Tom Black for furnishing the eta model fields; Tim Chambers for dragging them over the ethernet to the TDL DAR^{3}E system, where this work was done; and Joe Lang and Matt Peroutka for running PKZIP.

REFERENCES

Davis, R. A., 1978: Universal transmission format, Version 2.1. Unpublished manuscript, National Weather Service, NOAA, U.S. Department of Commerce, 22 pp.

Gerrity, J. P., Jr., 1977: The LFM model--A documentation. NOAA Technical Memorandum NWS NMC-60, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 68 pp.

Hoke, J. E., N. A. Phillips, G. J. DiMego, J. G. Sela, and D. G. Deaven, 1985: The Regional Analysis and Forecast System of the National Meteorological Center. Wea. Forecasting, 4, 323-334.

Mesinger, R., T. L. Black, D. W. Plummer, and J. H. Ward, 1990: Eta model precipitation forecasts for a period including tropical storm Allison. Wea Forecasting, 5, 483-493.

OFCM, 1990: Standard formats for weather data exchange among automated weather information systems FCM-S2-1990. National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 10 chapters plus appendices.

Reap, R. M., 1986: New 6-hr thunderstorm probability forecasts for the West. Technical Procedures Bulletin No. 362, National Weather Service, NOAA, U.S. Department of Commerce, 6 pp.

_____, 1990: Six-hour thunderstorm probability forecasts for Kansas-Oklahoma and the Northeast United States. Technical Procedures Bulletin No. 385, National Weather Service, NOAA, U.S. Department of Commerce, 6 pp.

Stackpole, J. D., 1992: GRIB, Edition 1, the WMO format for the storage of weather product information and the exchange of weather product messages in gridded binary form. National Meteorological Center Office Note 388, National Weather Service, NOAA, U.S. Department of Commerce, 41 pp. plus appendices.

WMO, 1988: Manual on Codes, Vol. 1, Part B--Binary Codes. WMO No. 306, World Meteorological Organization, Geneva.

Table 1. The bits required per packed value and the total bytes for the packed message for the methods described in the text for several LFM fields. For some methods, a range of bits is shown. The variables are in order: 500-mb height, 1000-mb height, 1000-mb temperature, 1000-mb dew point, boundary layer U- and V- wind components, and precipitable water. An asterisk on the units means the scaling factor was 10 (e.g., for wind the units were tenths of m/sec). The projections are in hours.

Field/Projection/Units			Method 1		Method 2		Method 3		Method 4		Method 5		Method 6		Method 7		Method 8	
			Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes
500-mb Hgt	0	m	13	1560	10	1201	7-9	979	7-9	949	7-9	945	7	843	7	843	4-7	737
500-mb Hgt	36	m	13	1560	9	1082	6-9	914	6-9	890	5-8	879	7	843	7	843	2-7	578
1000-mb Hgt	6	m	9	1082	9	1082	7-8	985	7-8	950	7-8	947	7	843	6	724	4-6	649
1000-mb Hgt	6	m*	12	1441	12	1441	10-12	1400	10-12	1364	10-12	1361	10	1202	9	1083	7-9	1048
1000-mb Temp	12	°K	9	1082	6	723	3-5	573	3-5	557	3-5	543	4	484	4	485	2-3	353
1000-mb Temp	12	°K*	12	1441	9	1082	7-9	980	7-9	956	6-9	944	8	963	7	843	5-6	725
1000-mb DewP	12	°K	9	1082	6	723	4-5	627	4-5	606	4-5	594	5	604	4	485	2-4	457
1000-mb DewP	12	°K*	12	1441	9	1082	7-9	1022	7-9	1001	7-9	989	8	963	8	963	5-8	840
BL U-Wind	6	ms^{-1*}	9	1082	9	1082	7-8	984	7-8	954	6-8	950	7	843	7	843	5-7	768
BL V-Wind	6	ms^{-1*}	8	962	8	962	7-8	973	7-8	946	6-8	944	8	963	7	843	5-7	736
Precip W	12	kgm^{-2}	6	723	6	723	4-6	664	4-5	640	4-6	639	5	604	5	604	3-5	475
Precip W	12	kgm^{-2*}	9	1082	9	1082	7-9	1059	7-9	1031	7-9	1021	8	963	8	963	5-8	862
Total				14538		12265		11160		10844		10756		10118		9522		8228

Table 2. The number of groups (Grp.) in the packed message and the total bytes required for the methods described in the text which use the parameters MINPK and/or INC. INC was held constant at 3. The ranges for the scaled values and for the first-order differences (FOD) and second-order differences (SOD) are shown for each variable. See the caption of Table 1 for further explanation.

Field/Projection/Units			MINPK	INC	Method 4		Method 5		Method 8	
					Grp.	Bytes	Grp.	Bytes	Grp.	Bytes
500-mb Hgt	0	m	4	3	222	1126	218	1117	207	878
			8		105	998	97	970	95	739
Ranges			17		43	980	40	949	40	724
Original	615		33		8	949	12	945	22	737
FOD	116		49		5	956	6	943	11	739
SOD	92		66		6	952	4	952	5	745
			82		5	964	4	952	5	745
			99		5	961	4	952	3	751
1000-mb Temp	12	^{o}K*	4	3	224	1070	220	1058	209	852
			8		111	975	99	937	96	747
Ranges			17		49	958	42	924	37	720
Original	398		33		8	956	22	944	11	725
FOD	140		49		7	958	7	945	4	727
SOD	68		66		6	963	6	945	2	729
			82		6	966	6	951	2	729
			99		4	982	5	959	2	729
BL U-Wind	6	ms^{-1*}	4	3	217	1096	218	1068	192	908
			8		102	990	101	973	89	809
Ranges			17		41	968	38	936	34	775
Original	263		33		6	954	15	950	15	768
FOD	123		49		3	953	5	945	6	776
SOD	84		66		3	954	5	945	4	781
			82		3	961	3	950	4	781
			99		2	968	3	950	4	781

Table 3. The number of groups (Grp.) in the packed message and the total bytes required for the methods described in the text which use the parameter MINPK and/or INC. MINPK was held constant at 33. Although Method 4 did not vary by INC, it is listed for comparison. See the caption of Table 1 for further explanation.

Field/Projection/Units			MINPK	INC	Method 4		Method 5		Method 8	
					Grp.	Bytes	Grp.	Bytes	Grp.	Bytes
500-mb Hgt	0	m	33	1	8	949	14	945	22	737
				2			11	945	22	737
				3			12	945	22	737
				6			12	945	22	738
				9			13	945	21	739
				12			13	946	22	737
				17			9	948	21	741
				33			8	949	16	748
1000-mb Temp	12	$^{\circ}K^{*}$	33	1	8	956	20	941	13	724
				2			20	941	13	724
				3			22	944	11	725
				6			22	944	10	721
				9			21	946	12	723
				12			22	959	10	725
				17			20	951	10	722
				33			8	956	8	731
BL U-Wind	6	ms^{-1*}	33	1	6	954	17	951	16	770
				2			17	951	16	770
				3			15	950	15	768
				6			15	948	14	778
				9			15	950	13	779
				12			13	955	10	781
				17			9	952	12	778
				33			6	954	7	785

Table 4. The same as Table 2, except for the interpolated 65 X 57 LFM grids.

Field/Projection/Units			MINPK	INC	Method 4		Method 5		Method 8	
					Grp.	Bytes	Grp.	Bytes	Grp.	Bytes
500-mb Hgt	0	m	8	3	414	3499	401	3386	339	2071
			12		262	3380	248	3289	242	2068
Ranges			17		178	3377	167	3258	169	2072
Original	615		25		114	3445	107	3308	106	2114
FOD	60		33		84	3522	78	3419	75	2156
SOD	48		49		33	3548	54	3463	51	2195
			65		10	3596	21	3544	40	2268
1000-mb Temp	12	°K*	8	3	430	3321	397	3142	360	2173
			12		274	3240	256	3169	230	2066
Ranges			17		196	3285	179	3178	162	2066
Original	398		25		128	3357	117	3194	108	2086
FOD	71		33		89	3453	78	3365	75	2167
SOD	35		49		44	3552	55	3368	46	2206
			65		9	3616	35	3537	27	2236
BL U-Wind	6	ms^{-1*}	8	3	423	3484	396	3404	367	2373
			12		266	3412	241	3296	229	2327
Ranges			17		181	3422	169	3292	152	2329
Original	263		25		111	3454	110	3340	97	2332
FOD	63		33		77	3510	74	3446	67	2352
SOD	42		49		34	3590	55	3481	43	2371
			65		6	3639	30	3555	30	2436

Table 5. The same as Table 3, except for the interpolated 65 X 57 LFM grids.

Field/Projection/Units			MINPK	INC	Method 4		Method 5		Method 8	
					Grp.	Bytes	Grp.	Bytes	Grp.	Bytes
500-mb Hgt	0	m	17	1	178	3377	167	3249	165	2060
				3			167	3258	169	2072
				6			172	3298	163	2088
				9			174	3332	165	2097
				17			178	3377	174	2140
1000-mb Temp	12	°K*	17	1	196	3285	180	3150	162	2037
				3			179	3178	162	2066
				6			179	3201	165	2088
				9			186	3246	165	2095
				17			196	3285	177	2138
BL U-Wind	6	ms^{-1}*	17	1	181	3422	171	3239	154	2287
				3			169	3292	152	2329
				6			179	3338	154	2335
				9			177	3382	148	2340
				17			181	3422	164	2380

Table 6. The bits required per packed value and the total bytes for the packed message for the methods described in the text for four eta model fields. For some methods, a range of bits is shown. The variables are in order: 500-mb height, 700-mb omega, surface temperature, and accumulated precipitation. All fields are 24-h forecasts. 700-mb omega was packed to retain tenths of μbar/s, surface temperature was packed to retain tenths of degrees Kelvin, and precipitation retained hundredths of inches.

Field	Units	Method 1		Method 2		Method 3		Method 4		Method 5		Method 6		Method 7		Method 8	
		Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes	Bits	Bytes
500-mb Hgt	m	13	7963	9	5514	5-9	4642	2-9	4239	2-9	4059	6	3678	6	3679	2-6	2269
700-mb Omega	μbar/s*10^{-1}	9	5514	9	5514	6-8	4282	4-8	3999	3-8	3919	9	5515	9	5515	2-9	4042
Sfc Temp	^{o}K*10^{-1}	12	7351	9	5514	7-9	4905	2-9	4385	2-9	4276	9	5515	10	6127	2-9	4203
Precip	in*10^{-2}	10	6126	10	6126	5-10	3992	0-10	2239	0-10	2121	11	6739	11	6739	0-11	2832
Total			26954		22668		17821		14862		14375		21447		22060		13346

Table 7. The number of groups (Grp.) in the packed message and the total bytes required for the methods described in the text which use the parameters MINPK and/or INC. INC was held constant at 3. The ranges for the scaled values and for the first-order differences (FOD) and second-order differences (SOD) are shown for each variable. See the caption of Table 6 for further explanation.

Field	Units	MINPK	INC	Method 4		Method 5		Method 8	
				Grp.	Bytes	Grp.	Bytes	Grp.	Bytes
500-mb Hgt	m	10	3	439	4191	393	4069	309	2338
		15		281	4239	261	4059	214	2269
Ranges		21		189	4290	184	4109	141	2251
Original	511	41		81	4433	82	4331	64	2299
FOD	61	83		9	4568	27	4480	35	2441
SOD	46								
700-mb Omega	μbar/s*10^{-1}	10	3	355	4086	377	3994	368	4069
		15		253	3999	235	3919	250	4042
Ranges		21		174	4003	177	3898	171	4035
Original	271	41		77	4098	79	4008	84	4252
FOD	302	83		22	4250	37	4164	38	4481
SOD	348								
Sfc Temp	°K*10^{-1}	10	3	415	4475	398	4338	360	4193
		15		262	4385	269	4276	249	4203
Ranges		21		181	4486	184	4290	175	4283
Original	447	41		70	4703	76	4608	85	4718
FOD	406	83		11	4842	31	4733	36	4894
SOD	546								
Precip	in*10^{-2}	10	3	329	2181	318	2051	331	2818
		15		233	2239	224	2121	227	2832
Ranges		21		186	2445	169	2275	171	2945
Original	635	41		93	3003	93	2718	96	3402
FOD	1075	83		26	3992	46	3501	44	4125
SOD	1350								

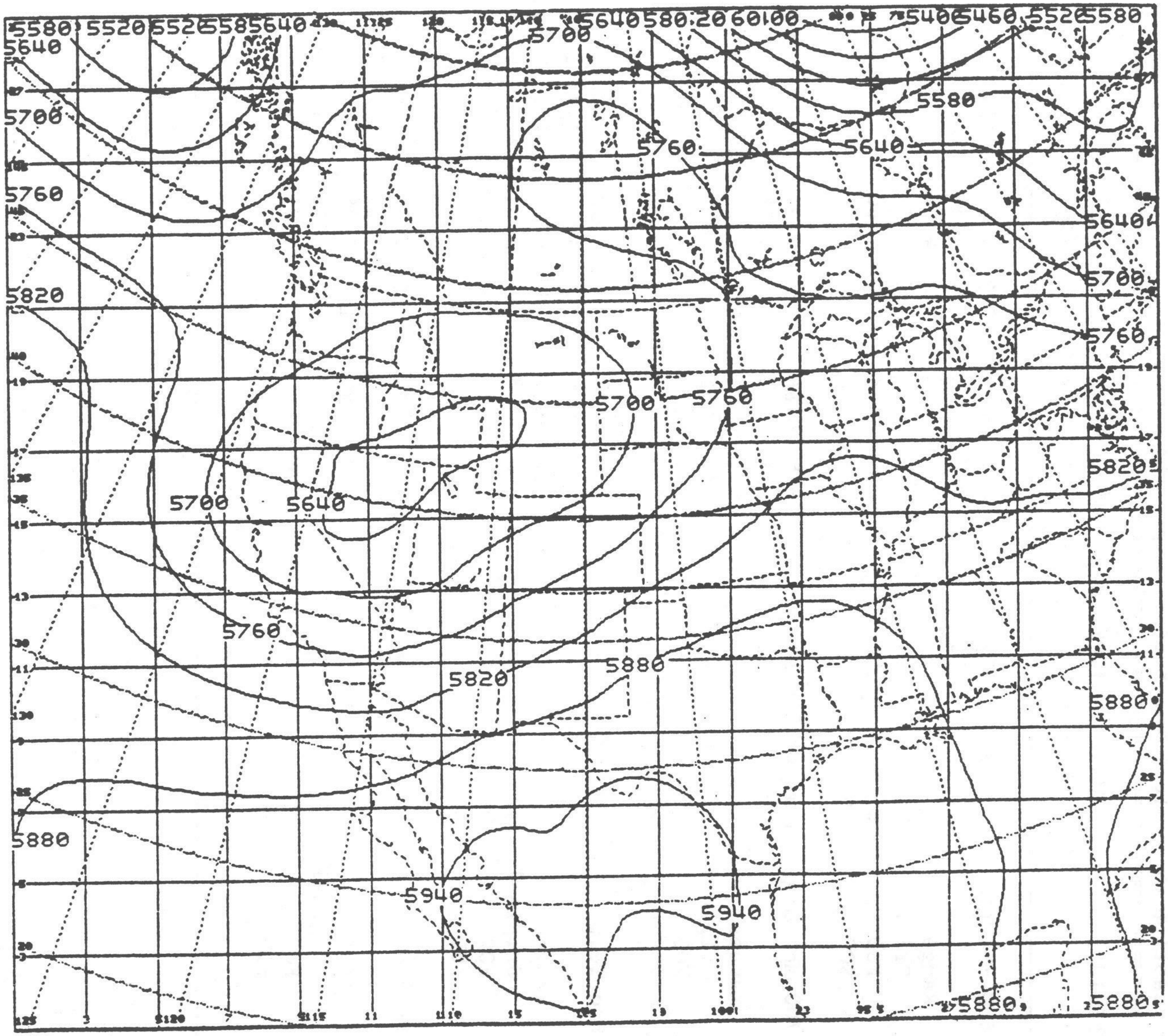

Figure 1. 500-mb height for June 16, 1992, 0000 UTC. This is from the LFM dataset for projection 0 hours. The grid is 33 X 29; alternate rows and columns are shown. Units are meters.

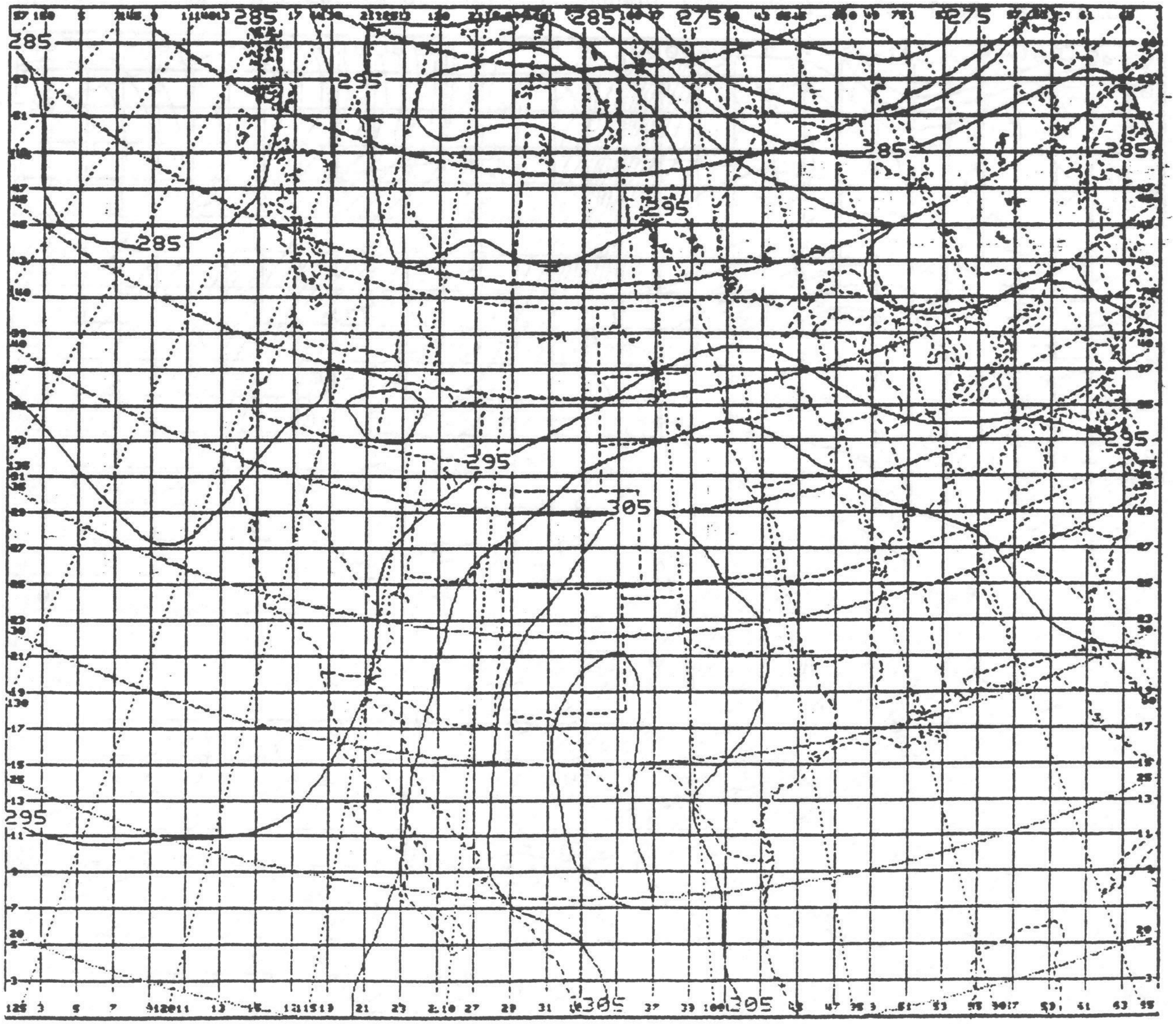

Figure 2. LFM 12-h forecast of 1000-mb temperature from the June 16, 1992, 0000 UTC cycle interpolated to a 95.25-km, 65 X 57 grid. Alternate rows and columns are shown. Units are degrees Kelvin.

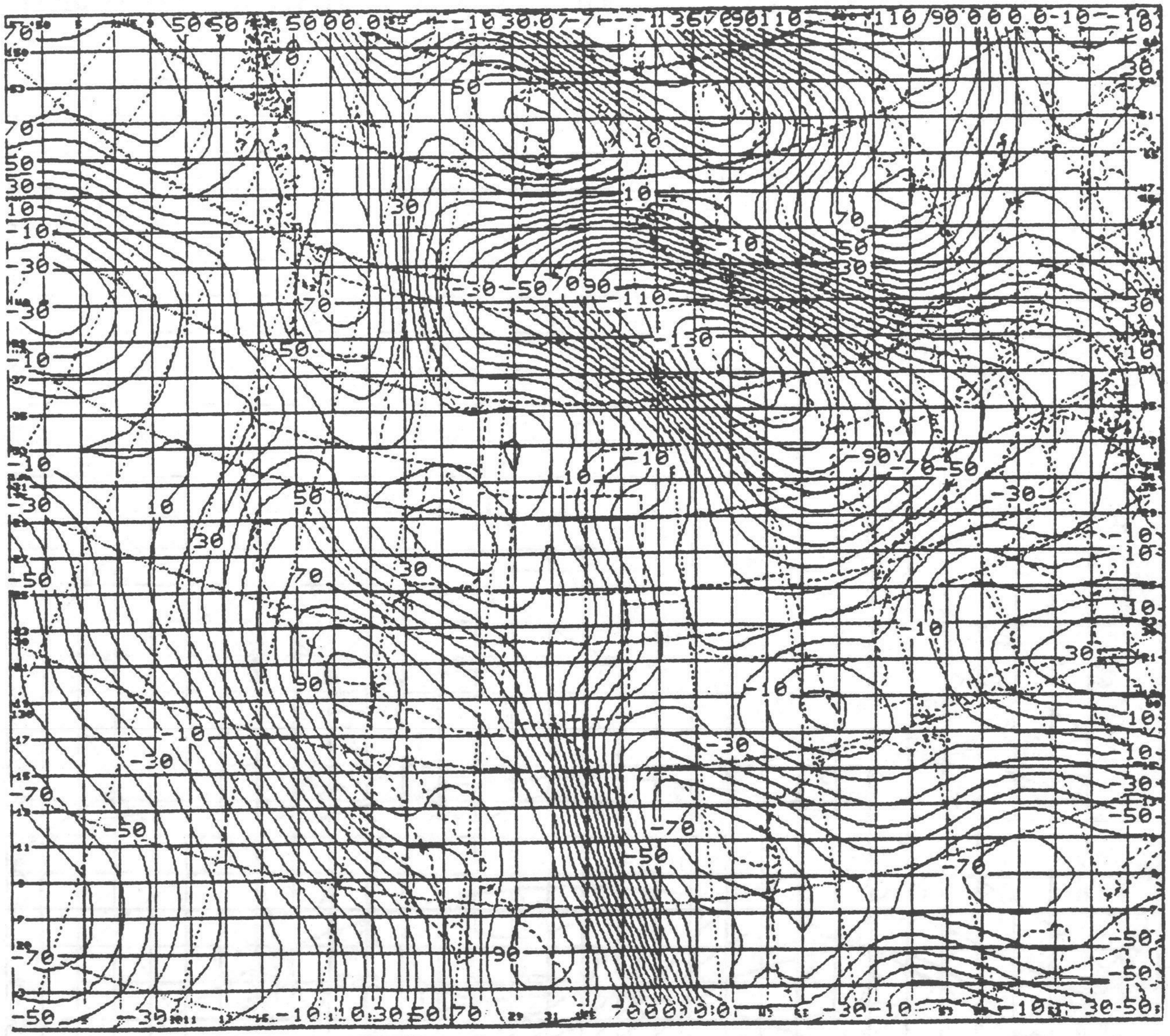

Figure 3. LFM 6-h forecast of U-wind component from the June 16, 1992, 0000 UTC cycle interpolated to a 95.25-km, 65 X 57 grid. Alternate rows and columns are shown. Units are tenths of ms^{-1}.

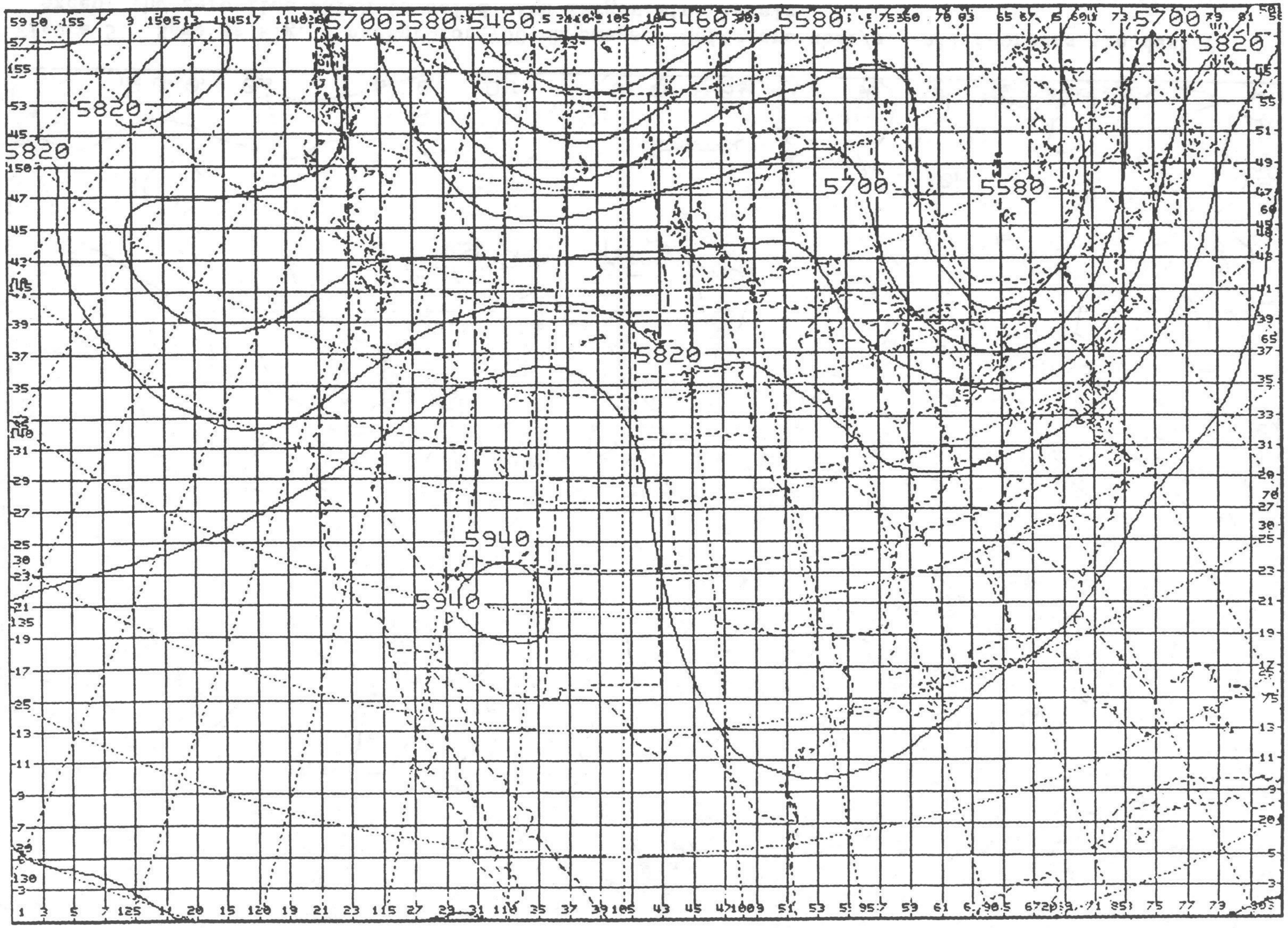

Figure 4. Eta model 24-h forecast of 500-mb height from the August 19, 1992, 0000 UTC cycle interpolated to the southern 83 X 59 portion of an NGM C-grid. Alternate rows and columns are shown. Units are meters. Alternate contours are labeled.

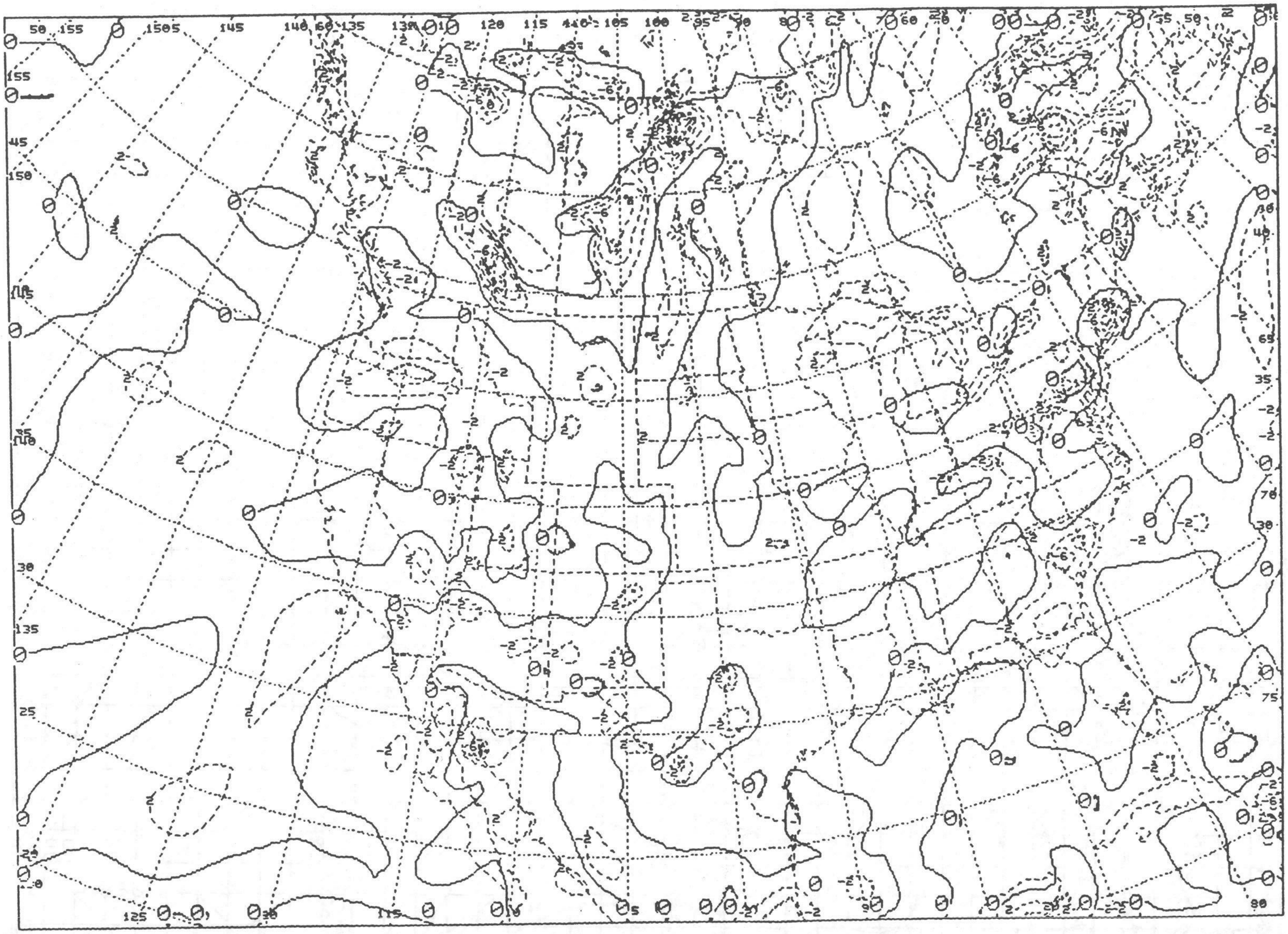

Figure 5. Same as Fig. 4 except for 700-mb omega. Units are μbar/s. All contours are dashed except the zero line which is solid with larger line labels.

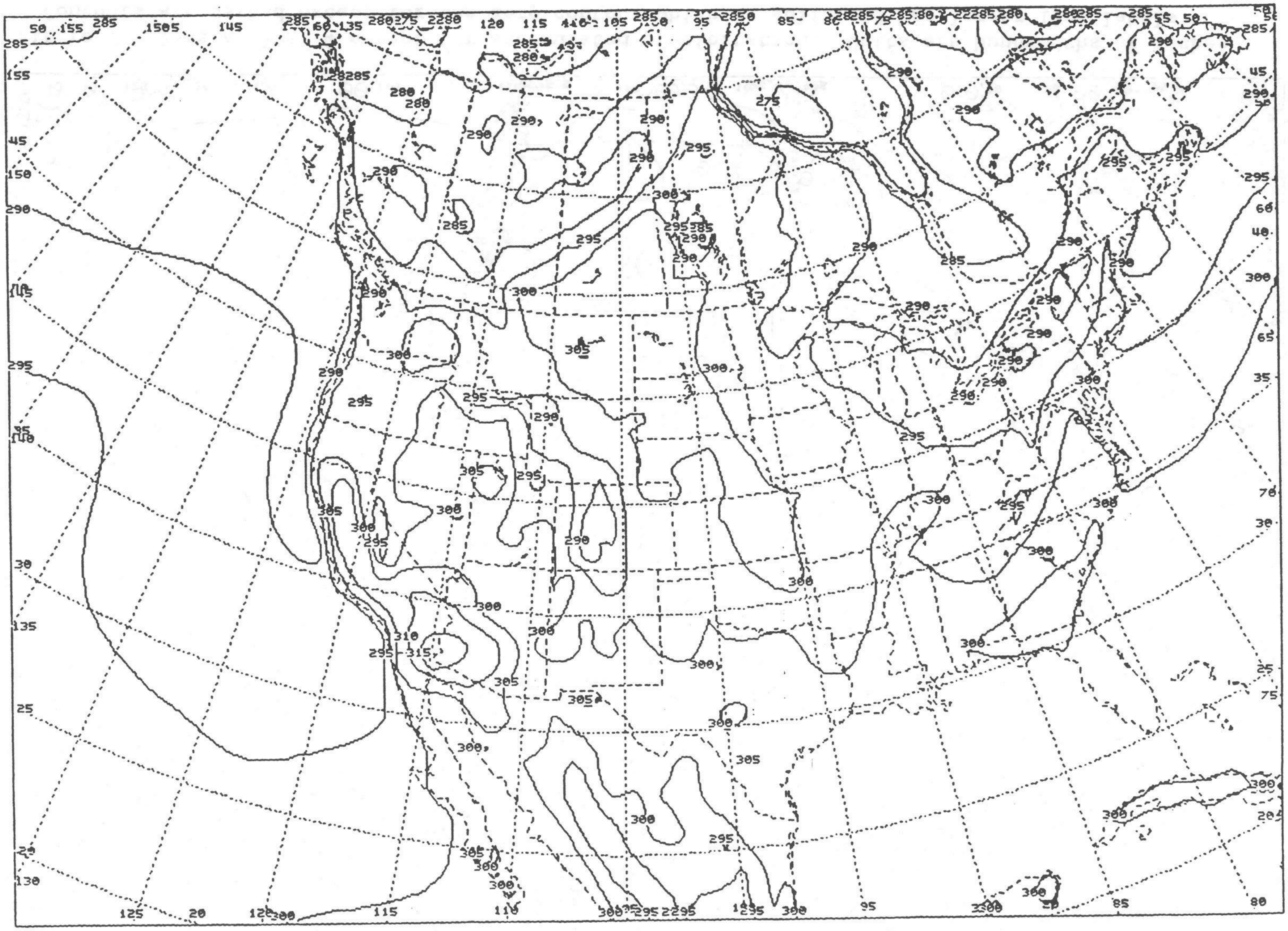

Figure 6. Same as Fig. 4 except for surface temperature. Units are degrees Kelvin.

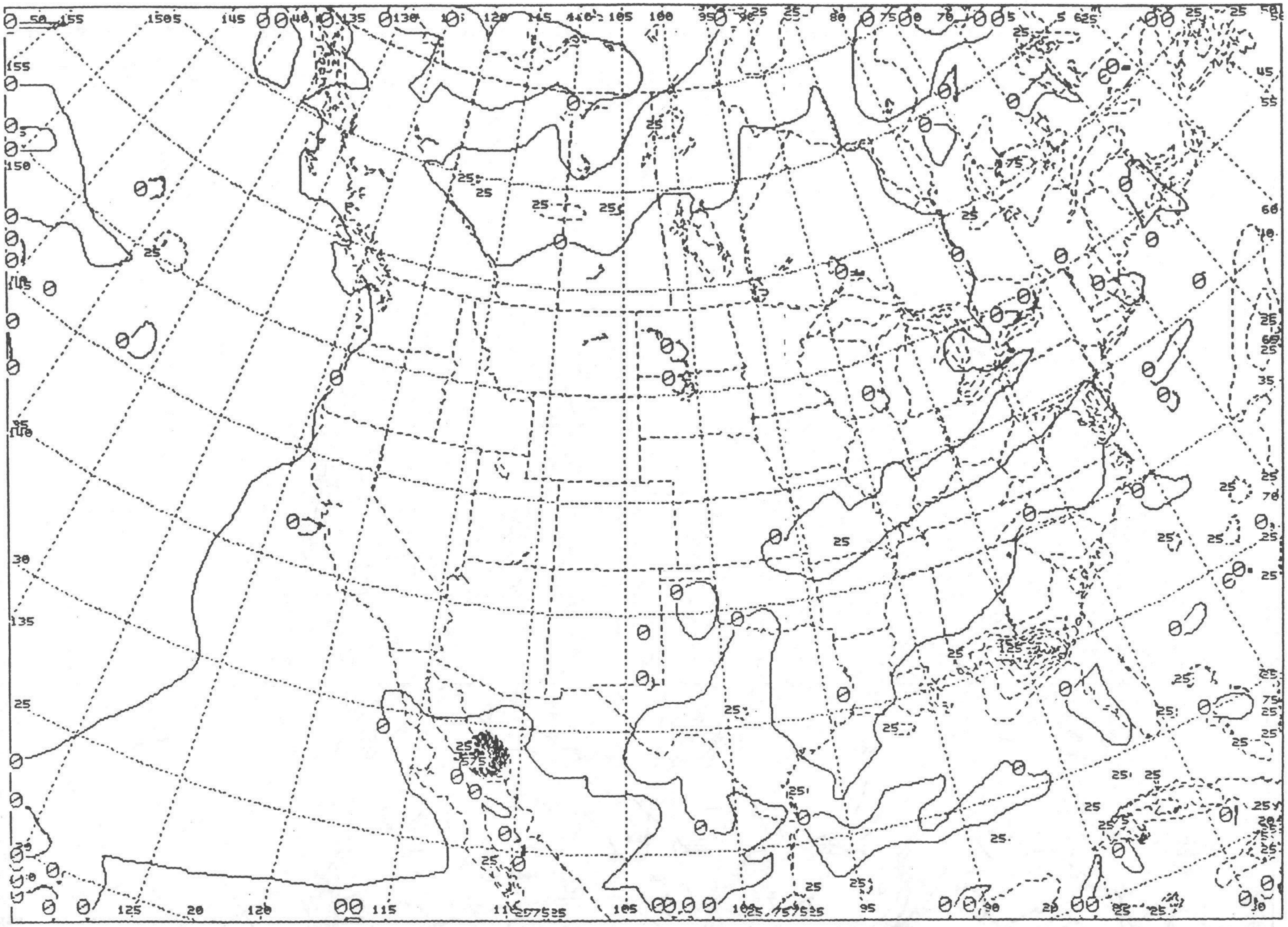

Figure 7. Same as Fig. 4 except for accumulated precipitation. Units are hundredths of inches. Contours are dashed except for the zero contour which is solid and has larger line labels.

APPENDIX

Algorithm for Determining Groups for Second-Order Minima Removal

This Appendix contains the FORTRAN 77 listing of the subroutine to determine groups for Methods 4, 5, and 8. For Method 4, INC would be input equal to MINPK. The statements with a "D" in column 1 are diagnostic and can be compiled only if needed to monitor the subroutine's operation. The size of the last group is constrained to be not less than MINPK/2 in size. Inputs and outputs are explained in the comments in the code. For Methods 4 and 5, IC( ) would hold the gridpoint values with the overall minimum removed, and NDP would be NX*NY. For Method 8, IC( ) would hold the second-order differences, and NDP would be NX*NY-2.

```
      SUBROUTINE VRBLPK(KFIL10,KFIL12,IC,NDP,MINPK,INC,
     1                  JMIN,JMAX,LBIT,NOV,NDQ,LX)
C
C        AUGUST 1992   GLAHN   TDL   MICROVAX
C
C        PURPOSE
C            TO DETERMINE GROUPS OF VARIABLE SIZE, BUT AT LEAST OF
C            SIZE MINPK, AND THE ASSOCIATED MAX AND MIN OF EACH GROUP,
C            THE NUMBER OF BITS NECESSARY TO PACK EACH GROUP, AND THE
C            NUMBER OF VALUES IN EACH GROUP.  THE ROUTINE IS DESIGNED
C            TO DETERMINE THE GROUPS SUCH THAT A SMALL NUMBER OF BITS
C            IS NECESSARY TO PACK THE DATA WITHOUT EXCESSIVE
C            COMPUTATIONS.  IF ALL VALUES IN THE GROUP ARE ZERO, THE
C            NUMBER OF BITS TO USE IN PACKING IS DEFINED AS ZERO.
C            ALL VARIABLES ARE INTEGER.
C
C        DATA SET USE
C           KFIL10 - UNIT NUMBER FOR CURRENT CONSOLE. (OUTPUT)
C           KFIL12 - UNIT NUMBER FOR OUTPUT (PRINT) FILE. (OUTPUT)
C
C        VARIABLES IN CALL SEQUENCE
C              KFIL10 = UNIT NUMBER FOR CURRENT CONSOLE.  (INPUT)
C              KFIL12 = UNIT NUMBER FOR OUTPUT (PRINT) FILE.  (INPUT)
C                IC( ) = ARRAY TO HOLD DATA FOR PACKING.  THE VALUES
C                        DO NOT HAVE TO BE POSITIVE AT THIS POINT, BUT
C                        MUST BE IN THE RANGE -9999999 TO +9999999.
C                        THESE INTEGER VALUES WILL BE RETAINED EXACTLY
C                        THROUGH PACKING AND UNPACKING.  (INPUT)
C                  NDP = NUMBER OF VALUES IN IC( ).  ALSO TREATED
C                        AS ITS DIMENSION.  (INPUT)
C                MINPK = THE MINIMUM SIZE OF EACH GROUP, EXCEPT POSSIBLY
C                        THE LAST ONE.  (INPUT)
C                  INC = THE NUMBER OF VALUES TO ADD TO AN ALREADY
C                        EXISTING GROUP IN DETERMINING WHETHER OR NOT
C                        TO START A NEW GROUP.  IDEALLY, THIS WOULD BE
C                        1, BUT EACH TIME INC VALUES ARE ATTEMPTED, THE
C                        MAX AND MIN OF THE NEXT MINPK VALUES MUST BE
C                        FOUND.  THIS IS "A LOOP WITHIN A LOOP," AND
C                        A SLIGHTLY LARGER VALUE MAY GIVE ABOUT AS GOOD
C                        RESULTS WITH SLIGHTLY LESS COMPUTATIONAL TIME.
C                        IF INC IS LE 0, 1 IS USED, AND A DIAGNOSTIC IS
```

```
C                       OUTPUT.  (INPUT)
C             JMIN(J) = THE MINIMUM OF EACH GROUP (J=1,LX).  (OUTPUT)
C             JMAX(J) = THE MAXIMUM OF EACH GROUP (J=1,LX).  (OUTPUT)
C             LBIT(J) = THE NUMBER OF BITS NECESSARY TO PACK EACH GROUP
C                       (J=1,LX).  IT IS ASSUMED THE MINIMUM OF EACH
C                       GROUP WILL BE REMOVED BEFORE PACKING, AND THE
C                       VALUES TO PACK WILL, THEREFORE, ALL BE POSITIVE.
C                       HOWEVER, IC( ) DOES NOT NECESSARILY CONTAIN
C                       ALL POSITIVE VALUES.  IF THE OVERALL MINIMUM
C                       HAS BEEN REMOVED, THEN IC( ) WILL CONTAIN
C                       ONLY POSITIVE VALUES.  (OUTPUT)
C              NOV(J) = THE NUMBER OF VALUES IN EACH GROUP (J=1,LX).
C                       (OUTPUT)
C                 NDQ = THE DIMENSION OF JMIN( ), JMAX( ), LBIT( ), AND
C                       NOV( ).  (INPUT)
C                  LX = THE NUMBER OF GROUPS DETERMINED.  (OUTPUT)
C
C        INTERNAL VARIABLES
C                KINC = WORKING COPY OF INC.  MAY BE MODIFIED.
C                MINA = MINUMUM VALUE IN GROUP A.
C                MAXA = MAXIMUM VALUE IN GROUP A.
C               IBITA = NUMBER OF BITS NEEDED TO HOLD VALUES IN GROUP A.
C                MINB = MINUMUM VALUE IN GROUP B.
C                MAXB = MAXIMUM VALUE IN GROUP B.
C               IBITB = NUMBER OF BITS NEEDED TO HOLD VALUES IN GROUP B.
C                MINC = MINUMUM VALUE IN GROUP C.
C                MAXC = MAXIMUM VALUE IN GROUP C.
C              KTOTAL = COUNT OF NUMBER OF VALUES IN IC( ) PROCESSED.
C               NOUNT = NUMBER OF VALUES ADDED TO GROUP A TO MAKE GROUP .
C
C        NON SYSTEM SUBROUTINES CALLED
C           NONE
C
      DIMENSION IC(NDP)
      DIMENSION JMIN(NDQ),JMAX(NDQ),LBIT(NDQ),NOV(NDQ)
C
      IF(INC.LE.0)WRITE(KFIL12,100)INC
 100  FORMAT('0INC ='I8,' NOT CORRECT.  1 IS USED.')
      KINC=MAX(INC,1)
      KSTART=1
      KTOTAL=0
      LX=0
 110  KOUNTA=0
      IBITA=0
      MINA=9999999
      MAXA=-9999999
C
C        FIND THE MIN AND MAX OF GROUP A.  THIS WILL INITALLY BE OF
C        SIZE MINPK (IF THERE ARE STILL MINPK VALUES IN IC( )), BUT
C        WILL INCREASE IN SIZE IN INCREMENTS OF INC UNTIL A NEW
C        GROUP IS STARTED.
C
      NEND=MIN(KSTART+MINPK-1,NDP)
      IF(NDP-NEND.LE.MINPK/2)NEND=NDP
C
```

```
      DO 120 K=KSTART,NEND
      MINA=MIN(MINA,IC(K))
      MAXA=MAX(MAXA,IC(K))
      KOUNTA=KOUNTA+1
 120  CONTINUE
C
C        INCREMENT KTOTAL AND FIND THE BITS NEEDED TO PACK THE A GROUP.
C
      KTOTAL=KTOTAL+KOUNTA
 125  IF(MAXA-MINA.LT.2**IBITA)GO TO 130
      IBITA=IBITA+1
      GO TO 125
C
 130  CONTINUE
D     WRITE(KFIL12,131)KOUNTA,KTOTAL,MINA,MAXA,IBITA
D131  FORMAT(' AT 130'10I6)
      IF(KTOTAL.GE.NDP)GO TO 200
C
C        MORE VALUES LEFT IN IC( ).  TRY TO ADD INC VALUES TO GROUP A.
C        THIS AUGMENTED GROUP IS CALLED GROUP C.
C
      MINC=MINA
      MAXC=MAXA
 135  NOUNT=0
      IF(NDP-(KTOTAL+INC).LE.MINPK/2)KINC=NDP-KTOTAL
C        ABOVE STATEMENT CONSTRAINS THE LAST GROUP TO BE NOT LESS THAN
C        MINPK/2 IN SIZE.  IF A PROVISION LIKE THIS IS NOT INCLUDED,
C        THERE WILL ALMOST ALWAYS BE A VERY SMALL GROUP AT THE END.
C
      DO 140 K=KTOTAL+1,MIN(KTOTAL+KINC,NDP)
      MINC=MIN(MINC,IC(K))
      MAXC=MAX(MAXC,IC(K))
      NOUNT=NOUNT+1
 140  CONTINUE
C
D     WRITE(KFIL12,141)KOUNTA,KTOTAL,MINA,MAXA,IBITA,MINC,MAXC,NOUNT
D141  FORMAT(' AT 141'10I6)
C        IF THE NUMBER OF BITS NEEDED FOR GROUP C IS GT IBITA,
C        THEN THIS GROUP A IS A GROUP TO PACK.
      IF(MAXC-MINC.GE.2**IBITA) GO TO 200
C
C        THE BITS NECESSARY FOR GROUP C HAS NOT INCREASED FROM THE
C        BITS NECESSARY FOR GROUP A.  FIND PACKING BITS OVER MINPK VALUES
C        FOLLOWING GROUP A, THAT IS GROUP B.
C
      MINB=9999999
      MAXB=-9999999
      IBITB=0
      JOUNT=0
C
      DO 160 K=KTOTAL+1,MIN(KTOTAL+MINPK,NDP)
      MINB=MIN(MINB,IC(K))
      MAXB=MAX(MAXB,IC(K))
      JOUNT=JOUNT+1
 160  CONTINUE
```

```
C
 165  IF(MAXB-MINB.LT.2**IBITB)GO TO 170
      IBITB=IBITB+1
      GO TO 165
C
C        DETERMINE WHETHER THE NEXT MINPK VALUES CAN BE PACKED IN
C        LESS BITS THAN GROUP A.  IF SO, PACK GROUP A AND START
C        ANOTHER GROUP.
C
 170  CONTINUE
D     WRITE(KFIL12,171)KOUNTA,KTOTAL,MINA,MAXA,IBITA,MINB,MAXB,IBITB,
D    1                 JOUNT
D171  FORMAT(' AT 171'10I6)
      IF(IBITB.LT.IBITA)GO TO 200
C
C        IBITB GE IBITA.  THEREFORE, ADD THIS INCREMENT TO A.
C
      KTOTAL=KTOTAL+NOUNT
      KOUNTA=KOUNTA+NOUNT
      MINA=MINC
      MAXA=MAXC
C        KOUNTA IS THE NUMBER OF VALUES IN GROUP A.  THIS GROUP WILL
C        NEVER BE SPLIT.
      IF(KTOTAL.LT.NDP)GO TO 135
C
C        GROUP A IS TO BE PACKED.  STORE VALUES IN JMIN( ), JMAX( ),
C        LBIT( ), AND NOV( ).
C
 200  LX=LX+1
      IF(LX.LE.NDQ)GO TO 205
      WRITE(KFIL12,201)
 201  FORMAT('0LX NOT LARGE ENOUGH.  STOP IN VRBLPK AT 201')
      STOP 201
C
 205  JMIN(LX)=MINA
      JMAX(LX)=MAXA
      LBIT(LX)=IBITA
      NOV(LX)=KOUNTA
      KSTART=KTOTAL+1
D     WRITE(KFIL12,209)KOUNTA,NDP,KTOTAL,LX,JMIN(LX),JMAX(LX),LBIT(LX),
D    1                 NOV(LX),KSTART
D209  FORMAT(' AT 209'10I6)
      IF(KTOTAL.LT.NDP)GO TO 110
C        WITH THE ABOVE TRANSFER, A NEW GROUP A OF SIZE MINPK WILL
C        BE DEFINED.
C
      RETURN
      END
```