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

TDL OFFICE NOTE 95-3

GRIB PRODUCT SIZES FOR AWIPS

Harry R. Glahn

December 1995

GRIB PRODUCT SIZES FOR AWIPS

Harry R. Glahn

1. INTRODUCTION

With the transmission of large volumes of gridded products to National Weather Service (NWS) field stations via the Advanced Weather Interactive Processing System (AWIPS) becoming a reality, we need to consider carefully the efficiencies, and lack thereof, of the GRIB code form. GRIB stands for the World Meteorological Organization (WMO) GRIdded Binary code for exchanging meteorological gridpoint data (WMO 1988; Stackpole 1994). It is clear that GRIB will be used to provide the large quantities of gridpoint data to the AWIPS contractor, it is expected that the same code will be used over the AWIPS communications networks, and a GRIB decoder is being provided for use at field stations. However, GRIB, in an attempt to be all things to all people, has a large number of options that can be used. Most of the options for routine use in AWIPS have been agreed upon (e.g., the order of the gridpoints, single values at each gridpoint, primary bit map to be part of the message[1] if one is necessary, and the grid description section to be included in the product).[2] However, whether simple packing[3] or the so called "complex" method (see Section 2) will be used has not been prescribed.

The method of packing, simple or complex, affects (1) bandwidth needed between any devices where the message is transmitted, (2) storage requirements on any device where the message is to be stored, (3) cpu time required for processing, and (4) the clock time on the computer where processing is done.

Early in the process of establishing requirements for AWIPS, the capacity needed for point-to-multipoint transmission of information was estimated in Appendix K to the System Requirements Specification (U.S. Government 1989). The sizes of the (rectangular) grids, in terms of number of gridpoints, were specified, and an estimate made of the number of bytes that would be needed to transmit a product containing those grids. The early estimate was for one byte per gridpoint, but later consideration led to increasing that estimate by 25%. These estimates were made with little or no actual GRIB experience, but were based on experience with the "2-pack" scheme that was used on the IBM 360-195[4] at the National Meteorological Center at the time (2 gridpoints per

---

[1] In this paper, "message" is used to refer to the complete GRIB code form, Sections 0 through 5, but does not include any communications header information. While GRIB can also be used for a storage format, in which case "file size" could be used to describe the size, use of "message" is consistent with usage in WMO (1988).

[2] This does not imply that the decoder will be limited to those options.

[3] In this paper, "gribbing" and "packing" are used interchangeably.

[4] No endorsement of specific equipment or companies is expressed or implied in this document.

32-bit computer word) and even earlier experience with the "5-pack" used on the CDC 6600 (5 gridpoints per 60-bit word). As experience with GRIB was acquired, it was thought that the original estimate of one byte per gridpoint was sufficient with complex packing. Therefore, the current version of Appendix K gives that estimate.

The data reported in this office note were generated primarily to show the advantages of using the complex method of packing. However, the opportunity presented itself to make some timing runs on the decoder being readied for AWIPS. Both product size and processing time are important in reaching a decision concerning the method of packing.

2. SIMPLE AND COMPLEX PACKING

Simple Packing is just that--quite simple. It involves the following steps:

1. Decide the degree of precision that is needed for the packed data. That is, one could decide that providing relative humidity to the nearest percent is good enough. Each value in the grid would be rounded to whole percent, and specified as an integer. Factors of 10 and/or 2 can be used in GRIB to get more or less precision.

2. Subtract the minimum value (called the "reference" value) in the grid from every value. This will give a range of numbers from zero to some positive value. This serves two purposes--negative numbers do not have to be dealt with, and the bits necessary to hold the new set of numbers may be less, depending on the field being packed.[5]

3. Determine the number of bits necessary to portray the largest number in the grid (after the subtraction of the minimum).

4. Use that number of bits to transmit each value.

Of course, there are many housekeeping chores associated with this packing, including the encoding of the minimum value itself,[6] the definition of the

---

[5] When both negative and positive values are 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, but 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. Not having to deal with a sign for each packed value is very advantageous.

[6] One of the weaknesses of GRIB, one that makes portability more difficult than it needs to be, is that the minimum value is in a particular floating point (REAL) form. Since this form is not standard across platforms, difficulties can arise. See Glahn (1993), footnote No. 4.

product, the grid used, etc. But the actual packing of the data is quite simple.[7]

Complex packing is not all that complex in theory, but the housekeeping of implementing it requires more steps. Briefly, Steps 1 and 2 for simple packing above are taken, followed by:

3. Determine groups of adjacent points (considering the grid to be a one-dimensional array of numbers) such that when the minimum of the group is subtracted the numbers in the group will be small.

4. Subtract the group minimum from each value in the group and determine the number of bits necessary to portray the largest number in the group (after subtraction of the group minimum).

5. Use that number of bits to transmit each value in the group.

The housekeeping chores are "more complex," because the minimum of each group must also be transmitted, as well as the number of bits needed to represent the values in the group and where the groups start and stop.

Former studies (Glahn 1992, 1993, 1994) have shown more efficient ways of packing data, still in the general GRIB complex framework, but have not been adopted by the WMO. GRIB uses a full byte to transmit the number of bits required to pack the groups, whereas about half that number is needed, and can be determined "on the fly." Also, the size of the groups (where they start and stop) is represented by a "secondary" bit map, again an inefficient method that is <u>quite</u> costly to the complex method. However, even with these drawbacks, the complex method can save considerable storage or transmission space.

3. 40-KM ETA GRID MESSAGE SIZES

Grids from the Eta model (Mesinger et al. 1990) now running at the National Centers for Environmental Prediction (NCEP) are gribbed and sent to the National Weather Service Telecommunications Gateway (NWSTG), where they are stored on a "server." A number of grids are packaged in a file, which is available to certain users, and the grids are also sent individually over communication lines, which are a precursor to the AWIPS communication network. These 40-km AWIPS grids correspond to the grid NMC has defined as No. 212 (Stackpole 1994), which has a gridspacing of exactly 40 km at 35° North Latitude. These grids are in reference to the Lambert Conformal map projection oriented at 95° West Longitude and tangent to the earth at 25° North Latitude. The lower left corner point is at 12.190° West Longitude and 133.459° North Latitude. These 40-km grids are 185 by 129, for a total of 23,865 points.

---

[7] Actually, if the numbers range about, or include, zero, there is little to be gained by subtracting the minimum value, except for the convenience of not dealing with negative numbers. However, for a field where each value is quite large (e.g., 500-mb height), the "efficiency" or removing the minimum is quite good.

The files for 0000 UTC, October 19, 1995, were degribbed, gribbed with both the simple and complex methods, and again degribbed.[8] Various statistics were kept, and are presented in Table 1 and discussed below. All of the original gribbed fields were simply packed, and with varying degrees of decimal (and occasionally, binary) resolution. For packing, the same decimal resolution was used, the only difference being in the method of packing and the occasional neglect of the other than zero binary scaling factor.

The algorithm used for determining the groups for complex packing is that in Glahn (1994); it requires two adjustable parameters, the minimum number of values to put into a group (MINPK) and the number of values to try to add (or subtract with the lookback feature) at a time to a group (INC). For all results shown in Tables 1 through 4, MINPK = 14 and INC = 1 were used. (See the Appendix I for a discussion of MINPK and INC.)

It was noted that the grid is not completely filled with data; 1332 points on the lower left are missing. This required a (primary) bit map to be sent, which consumes 1 bit per gridpoint.

As shown in Table 1, the improvement in message size (also bits/point) for complex versus simple packing ranged from 32% for the 324 geopotential height fields to 4% for the 264 relative humidity fields (top rows; decimal scale = 1), for an overall improvement of 20% for the 1884 fields (i.e., the simple-packed fields were, on the average, 25% larger than the complex-packed fields).

Some of the scaling of the NCEP packed fields was surprising. For instance, the vertical velocity fields were generally X10^2 for the heights from 975 mb through 250 mb inclusive (24 levels and 12 projections--hours 0 through 33 at 3-h intervals).[9] However, the scaling was X10^3 for 100 and 150 mb for all projections, for 200 mb for projections 12 through 33 hours, and for 1000 mb for all projections except hour 0, where it was X10^2.

Also of note is that relative humidity was scaled X10^1 (first two lines for relative humidity in Table 1), which means that the relative humidity is being transmitted to tenths of percent. That is undoubtedly an overkill, since relative humidity can hardly be measured to whole percent, let alone be forecast by a model that accurately. Because of the large degree of precision used, and because the range of numbers is relatively small, little is gained by complex packing. However, with precision only to whole percent (see last two lines for relative humidity in Table 1), considerably more is gained

---

[8] The files on the server had names like "us008_gf085_95101903_YxRAx".

[9] Note that in the GRIB code, a positive decimal scaling factor "D" means more significance, while a positive binary scaling factor "E" means less significance according to the formula

$$Y \text{ X } 10^D = R + (X_i + X_j) \text{ X } 2^E$$

where Y is the original value, R is the reference value and X_i and X_j are the group minima and deviations, respectively, for complex packing. $X_i + X_j = X$ is just the deviation from R for simple packing (WMO 1988, p. I-B1-6).

Table 1. Statistics associated with packing 40-km Eta model fields by both the simple and complex GRIB methods. The two numbers in parentheses in the field definition are from the WMO Tables 2 and 3 respectively (WMO 1988).

Field Definition	No. Fields	Message Length (byte)	Bits/ Point	Improvement (%)	Points/ Group	Degrib Time (sec)	Dec. Scale D
Geopotential Height (7, 100)							
Simple	324	30323.8	10.17	--	--	17	0
Complex	324	20717.2	6.94	32	17.3	29	0
Temperature (11, 100)							
Simple	324	29234.8	9.80	--	--	20	1
Complex	324	21614.2	7.25	26	17.4	29	1
Relative Humidity (52, 100)							
Simple	264	32474.0	10.89	--	--	12	1
Complex	264	31071.3	10.42	4	17.7	24	1
Simple	264	23656.0	7.93	--	--	20	0
Complex	264	20967.5	7.03	11	17.8	29	0
U-, V-Wind Components (33 & 34, 100)							
Simple	648	30927.0	10.37	--	--	42	1
Complex	648	25494.0	8.55	11	17.6	68	1
Vertical Velocity 39, 100)							
Simple	324	29035.4	9.73	--	--	16	2
Complex	324	22868.2	7.67	21	18.3	25	2
Overall							
Simple	1884	30423.7	10.20	--	--	104	-
Complex	1884	24335.2	8.16	20	17.6	160	-

(11% versus 4%), and the message is much shorter[10].

---

[10] The overall totals at the bottom of Table 1 do not include these less-precise fields; the scaling was held at 10^1.

The decimal scaling for wind components was usually X10^{1}, but certain exceptions were noted for v-wind where it was X10^{0}: 300 mb at 3 hours, 300 and 250 mb at hours 6, 9, 12, 15, 18, and 21; 350, 300, and 250 mb at hours 24 and 27; and 400, 350, 300, and 250 mb at hours 30 and 33.

Of the 324 geopotential height fields, 267 were packed X10^{0} and 57 were packed X10^{-1}. The levels at which accuracy was retained to only tens of meters were generally 150, 200, 250, 300, and 350 mb, but not above 150 mb (in elevation at 100 mb, or below 350 mb.

Generally, the binary scaling was zero (i.e., 2^{0}); differences to this general rule were for geopotential height at 200 mb at hour 0, 775 mb at hour 6, and 850 mb at hour 24; and for u-wind at 600 mb at hour 3 and 925 mb at hour 33. In these few cases the binary scaling was 2^{1}. When these fields were packed for comparison, and the values in Table 1 determined, the binary scaling was used as zero. This resulted, for those very few grids, in (one bit) greater precision being used than was present in the degribbed fields.

4. 20-KM ETA GRID MESSAGE SIZES

Other fields are available from the same NWSTG server on the AWIPS 20-km grid. All characteristics of the grid are the same except for the gridspacing. The grid is 369 by 257 for a total of 94,833 points. Table 2 gives the results for the 20-km grids similar to Table 1 for the 40-km grids for 0000 UTC, November 11, 1995.

The fields on the 20-km resolution were generally for the surface and low sigma levels, whereas the fields at 40-km resolution were for isobaric surfaces throughout the atmosphere. As with the 40-km fields, the lower left of the grid was missing, and a bit map was necessary.

The improvement in message size for complex versus simple packing ranged from 13% for relative humidity to 70% for convective and total precipitation fields,[11] for an overall improvement of 29% for the 366 fields in the sample, 279 of which are separately shown (i.e., the simple-packed fields were, on the average, 42% larger than the complex-packed fields).

Although Table 2 shows the decimal scaling generally used, there were a few exceptions. For instance, for vorticity, the decimal scaling was X10^{6} except for 1000 and 850 mb at 9 and 12 hours, where it was X10^{5}. Also, for precipitation, two of the 0-hour fields has a scaling of X10^{2} instead of X10^{1}.

The relative humidity was scaled X10^{1} (tenths of percent) which resulted in a somewhat larger product size with simple packing than any other field and considerably larger for complex packing--9.40 bits/point versus the next

---

[11] The "zero" values in precipitation fields actually have the value -0.25 (when scaled to 10^{2} and -0.3 when scaled to 10^{1}). This is retained from past practice to aid in placing the zero contour when contouring the grid (G. DiMego, personal communication). Since the unit of measurement used (kg/m^{2}) equates to 1 millimeter of precipitation (Stackpole 1994, p. 3 of Section 0, last paragraph), this value is a negative quarter of a millimeter of precipitation.

Table 2. Statistics associated with packing 20-km Eta Model fields by both the simple and complex GRIB methods. The two numbers in parentheses in the field definition are from the WMO Tables 2 and 3 respectively (WMO 1988).

Field Definition	No. Fields	Message Length (byte)	Bits/ Point	Improvement (%)	Points/ Group	Degrib Time (sec)	Dec. Scale D
Vorticity (41, 100)							
Simple	45	121560	10.25	--	--	10	6
Complex	45	93665	7.90	23	17.8	16	6
Temperature (11, 105 & 108)							
Simple	54	125585	10.59	--	--	10	1
Complex	54	85886	7.25	32	17.2	19	1
Relative Humidity (52, 105 & 108)							
Simple	54	128832	10.87	--	--	12	1
Complex	54	111465	9.40	13	17.6	18	1
U-, V-Wind Components (33 & 34, 105 & 108)							
Simple	108	123529	10.42	--	--	26	1
Complex	108	93774	7.91	24	17.3	41	1
Precipitation (61 & 63, 1)							
Simple	18	107404	9.06	--	--	4	1
Complex	18	32720	2.76	70	75.8	4	1
Overall							
Simple	366	122605	10.34	--	--	80	-
Complex	366	86560	7.30	29	18.9	126	-

largest of 7.91 bits/point for wind components. As indicated in the above section for 40-km grids, this would undoubtedly be below 7 bits/point when packed to whole percent.

Generally, the binary scaling was 2^0, the only difference being noted was for temperature at 30 hours at one level, where it was 2^1.[12]

---

[12] Another sample of 447 fields for 0000 UTC, November 12, 1995, also had only one field with binary scaling of 2^1, but it was for the u-wind component at one level at 12 hours. Referring back to vorticity fields, all had decimal scaling of 10^6.

5. TIMES FOR GRIBBING AND DEGRIBBING

While the primary purpose of this study was not to document gribbing or degribbing times, the opportunity afforded itself. Unfortunately, the timing function available only recorded to whole seconds, so a considerable sample was necessary to get reasonably reliable times--the more fields involved, the better the times. The times presented for gribbing and degribbing are the times during which the respective subroutines had control. That is, the times are not true cpu times. If the subroutine were not using the cpu (i.e., because of other processes running) the times would include this anyway. All runs were made from a Hewlett Packard (HP) 755 (Blizzard) in Room 10201 of SMCC2, logged onto an HP 755 (Thunder) in Room 7386 of SSMC2. While each machine is involved in numerous housekeeping chores, the runs were minimally impacted by such activity, the runs being made on weekends or very early in the morning. It is believed the times for the gribbing and degribbing on the 1884 40-km fields and 366 20-km fields are reliable, having been replicated.

It should be emphasized that normal (buffered) FORTRAN READs and WRITEs were used, and that the files being dealt with were on the same machine (Thunder) as where the processing was being done. Large volumes of data were not being transported across a LAN. For the gribbing, the data were being read in binary, 4-bytes per word, so the volume was quite large.

A. 40-km Fields

With the buffering on input and output going on, Thunder showed the cpu usage to be 40-50% for simple packing and 50-60% for complex packing. The average gribbing times per field were 0.113 and 0.188 seconds for simple and complex packing, respectively (8.8 grids per second versus 5.3). The clock run times were about 9.4 minutes for simple packing and 11.3 minutes for complex for the 1884 fields.

For unpacking, the cpu usage shown was above 95%. These runs were made with no grid output; only the statistics were kept. The average times per grid were 0.055 and 0.085 seconds for simple and complex methods, respectively (18.1 grids per second versus 11.8). Because of the high cpu usage, the clock run times were less than 10% higher than the actual degribbing times.

Therefore, the gribbing of a complex-packed field took about 0.075 seconds longer than a simple-packed field and the degribbing took 0.030 seconds longer.

B. 20-km Fields

Thunder showed the cpu usage to be about 50% for complex packing and somewhat less for simple packing. The average gribbing times per field were 0.481 and 0.710 seconds for simple and complex packing, respectively[13] (2.1 fields per second versus 1.4). The clock run times were about 8.1 minutes for simple packing and 9.5 minutes for complex for the 366 fields.

---

[13] The sample of 447 fields for November 12 gave 0.497 and 0.705 seconds for simple and complex packing, respectively--quite consistent with the earlier sample.

For unpacking, the cpu usage shown was above 85%. These runs were made with no grid output; only the statistics were kept. The average times per field were 0.219 and 0.344 seconds for simple and complex methods, respectively[14] (4.6 fields per second versus 2.9). The clock run times were 15 to 20% higher than the actual degribbing times.

Therefore, the gribbing of a complex-packed field took about 0.229 seconds longer than a simple-packed field and the degribbing took 0.125 seconds longer.

6. DISCUSSION

A. Message Sizes

"Why are the sizes sooo much bigger....grib?" came scrolling across my PC screen a few weeks ago. This followed on the heels of an earlier message "Wouldn't you say that Appendix K is a bit off?" That is what instigated this investigation. That is, why are the actual Eta model fields larger than the Appendix K estimates? The reasons the Appendix K estimates are smaller than those in the current 40- and 20-km Eta messages are basically:

1. The current estimates were based on complex packing. Verbal agreement had been reached about 2 years ago that complex packing would be used--not any of the non-standard features that could further enhance the size efficiency, but standard, complex packing.[15] As has been shown on a sample in this document, the 40-km sizes would be reduced by 20% and the 20-km sizes by 29%. While these sizes are for a particular sample, they should be fairly stable.[16] Note that overall, for both resolutions together, and the number of fields used in the comparison, only 7.79[17] bits per point would have been required with the decimal resolution used. This is <u>below</u> the Appendix K estimate of 8.00 bits/point. Complex packing does pay off size-wise.

2. A primary bit map is required for these Eta grids, because the model does not cover the complete area designated to be sent for AWIPS. This possibility was never considered in making the estimates; it was assumed any model data would be available for the complete grid. While accounting for missing data with a bit map may seem trivial, it is not. Suppose that 8 bits were required for each point (the Appendix K estimate) and the grid were full. For the 1332 points on the 40-km grid

---

[14] The November 12 sample gave 0.210 and 0.340 seconds, respectively. The two samples were quite consistent in this respect.

[15] J. Stackpole, then Chief, Production Management Branch, Automation Division, National Meteorological Center, personal communication.

[16] The November 12 sample also gave a reduction of 29% for the 447 20-km fields. For this sample, the number of bits used per gridpoint was 7.29--very consistent with Table 2.

[17] $[(1884 \times 23865 \times 8.16) + (366 \times 94833 \times 7.30)]/(1884 \times 23865 + 366 \times 94833) = 7.79$.

that are missing, 1332 bytes would have been required to send the data. However, to not send the data (23865 X 1)/8 = 2983 bytes are required! That is, one bit is required for each gridpoint for the bit map; for a relatively small number of missing points, the message is increased by upwards of 10% (in this case, 7%). Note that the 7.79 bits/point value for this sample includes the use of the bit map, even though it was not originally considered.

Appendix K estimates were made without knowing at what resolution the data would be sent. The resolutions of the data (e.g., tenths of m/sec for wind components) were decided somewhere along the line, and those are what are in the files used in this study. Presumably, the resolutions are what were thought to be needed. However, as a result of this current study, a proposal has been made[18] by NCEP to modify considerably the scaling used in packing. The packing proposed is discussed in Appendix II, along with the implications of changing the accuracy retained as a function of grid spacing.

B. Processing Time

Processing time, especially at the field sites, is important. The gribbing code used in this study is not what is used operationally at NCEP. However, the basic algorithms must be similar; the relative efficiency is not known. Besides the extra processing necessary for the actual packing in the complex method, the "groups" must be determined, thereby increasing the time substantially. It is doubtful that on the supercomputers used at NCEP this difference is important.

Taking the appropriate numbers from Tables 1 and 2, we see that the degribbing times per point for the 40- and 20-km fields for simple packing are, respectively,

$$104 \text{ sec}/(23865 \text{ points} \times 1884 \text{ fields}) = 2.313 \times 10^{-6} \text{ sec/point, and}$$

$$80 \text{ sec}/(94833 \text{ points} \times 366 \text{ fields}) = 2.305 \times 10^{-6} \text{ sec/point.}$$

We also see from Appendix K, Table K.2.7.B, that the Eta model data for the Regional areas over the continental United States totals 226,691 kilobytes for the two primary runs or 113,345 kilobytes per run. In addition, Appendix K states (p. SRSI-K-28) that:

> "...it is assumed that the scheduled outputs of the NMC Mesoscale Model data, the RAFs data and the Aviation Model data do not overlap in time and therefore, only the largest of the three, the NMC Mesoscale Model output data, is used to compute peak bit rate. The output from the NMC Mesoscale model will be provided to the AWIPS contractor within four, 50 minute time slots. Model output will be available to the contractor by time projections before the model run is complete."

Appendix K also states (p. SRSI-K-5) that these data (being produced 4 or less times per day):

---

[18] R. Petersen, personal communication.

"...shall be available for use at each AWIPS site needing that product within 15 minutes of its delivery from the Government to the AWIPS contractor."

With this background, 2.31 being the "average" of 2.313 and 2.305, and knowing that the Appendix K estimate is 1 byte per point, we can calculate

$$(1.133 \text{ X } 10^{8} \text{ points X } 2.31 \text{ X } 10^{-6} \text{ sec/point})/(60 \text{ sec/min}) = 4.36 \text{ min}$$

This 4.36 minutes is an estimate of the time out of the 50-minute slot that the cpu of an HP 755 class computer would be busy decoding simply-packed data if all the data were to be unpacked upon receipt.

From the numbers in Tables 1 and 2 relating to complex packed data, the estimates are 3.559 (3.630) X 10^{-6} sec/point for 40-km (20-km) data. Finally, with an average of 3.59 X 10^{-6} per point, the estimate for the number of minutes an HP 755 class cpu would be busy during a 50-minute period when the Eta model data would be arriving would be 6.78.[19]

The current AWIPS design is for the GRIB data to be stored in that form and decoded "on-the-fly." In this case, the relevant estimates are 0.055 and 0.085 seconds for simple and complex methods, respectively, for each 40-km grid, and 0.219 and 0.334 seconds, respectively, for each 20-km grid.

One possible complication with second order packing is that it would be more difficult to "clip" out an area "on-the-fly" while unpacking. That is, if one wanted to save a sub-area, it would be possible to skip, with some dexterity, the points not needed and save some unpacking time. Second order packing complicates this process and makes it less efficient, but does not rule it out. The actual saving that could be achieved is unknown. The process of not unpacking the whole message is complicated, for either simple or complex packing, by the presence of a primary bit map (the non-presence of some data values) and especially by the possibility that the order of the gridpoints in the message do not conform to the order needed for the clipped grid. Clipping in this way is not planned for AWIPS implementation, at least not initially; only if processing time becomes an issue will effort be devoted to development and testing of this capability.

C. Storage and Transmission Considerations

The use of simple packing over complex packing increases the storage necessary by about 30% (32% for this sample) in every device and over all transmission paths. In some cases of storage, where the storage time is short (e.g., a grid is held only long enough to transmit it), this is undoubtedly unimportant (provided, of course, transmission can occur without significant delay). However, when 24 hours of data, say, are to be held in a server, 30% extra may be important. Probably of more importance is the transmission over all paths necessary for the messages to reach their destination where they will be unpacked. At a local AWIPS site, this includes the transfer between

---

[19] Since the cpu was busy a high percentage of the time, this is probably only a slight overestimate for actual degribbing time. Not only that, other processes associated with degribbing will also be required.

communication devices over a LAN, and storage in the database. It also may include, depending on the final implementation, the transfer over that same LAN to the processor where the data are to be decoded. No matter were the degribbing is done, the data must be transferred to the decoder for processing.

The estimated peak rate of other than satellite data for point-to-multipoint transmission includes 302 kilobits/sec over each of four 50-minute periods for the Eta model data (Appendix K, p. SRSI-K-30). Based on data in the sample reported on here, this estimate would increase to about 387 kilobits/sec with simple packing.[20] This is an increase in estimated bit rate for other than satellite data of 14%. Other models' data would also be increased.[21]

7. CONCLUSIONS

For message size, the complex method of packing gets more efficient, relative to the simple method, the smaller the gridlength (see Appendix II for a discussion of this topic).[22] In addition, the impact on the communications and storage will be greater just due to shear volume. On the other hand, the impact of more required cpu cycles to unpack with the complex method will become more important as data volumes grow.

What to do? Since it seems that the maximum cpu cycles necessary for field site processing are likely available for the presently prescribed set of products (based on these results on an HP 755, a maximum of 6.8 minutes would be required out of a 50-minute period, four times per day; only 2.4 minutes more each time than for simply packed data), and since the transfer of data within the system may actually become of considerable importance, it is probably wise to plan to use the complex method at this time. A 30% increase in storage for simple packing may be important in some circumstances, its importance at a WFO depending partially on the final AWIPS detailed design. In communication outages, large volumes of data may have to be stored for later transmission. Possible transmission delays should be minimized with complex packing. It seems unlikely the difference in packing time on a supercomputer would be noticeable.[23] It would always be possible to

---

[20] The overall bits/point for this sample is
[(1884 X 23865 X 10.20) + (366 X 94833 X 10.34)/
(1884 X 23865 + 366 X 94833) = 10.26. This is 10.26/8 = 1.28 times higher than the estimate of 8 bits/point on which the Appendix K estimate was based. 302 X 1.28 = 387.

[21] Of course, the satellite data requirements, again as stated in Appendix K (p. SRSI-K-28), are considerably larger, having a peak rate of 2,150 kilobits/sec.

[22] It should be noted that all comparisons given here are based on the "grouping" algorithm for complex packing given by Glahn (1994). Similar results on substitutes are not guaranteed.

[23] If there are multiple users of a GRIBbed product (i.e., other than NWS use over AWIPS), any user should be able to handle the unpacking of complex-packed data, since it is a standard GRIB option.

(partially) go to simple packing if the various system processes demanded it (e.g., the cpu cycles available at the field site became more critical than the transfer of data from place to place).

ACKNOWLEDGMENTS

I appreciate the review of an early version of this document by Wendy Wolf and Ralph Petersen; Appendix II is a result of the latter review. I also am indebted to David Kitzmiller who prepared the figures for me.

REFERENCES

Glahn, H. R., 1992: On the packing of gridpoint data for efficient transmission. TDL Office Note 92-11, National Weather Service, NOAA, U.S. Department of Commerce, 32 pp.

_____, 1993: An analysis of some features of the GRIB code. TDL Office Note 93-2, National Weather Service, NOAA, U.S. Department of Commerce, 42 pp.

_____, 1994: An improved algorithm for determining the groups in GRIB second-order packing. TDL Office Note 94-1, National Weather Service, NOAA, U.S. Department of Commerce, 25 pp.

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.

Stackpole, J. D., 1994: 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.

U.S. Government, 1989: AWIPS-90 System Requirements Specification, Vol. I, Appendix K.

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

APPENDIX I

Effect of the Values of MINPK and INC on Message Size and Packing Time

The 324 40-km geopotential height fields for which data are shown in Table 1 were gribbed with varying values of MINPK; the results are shown in Figs. 1 and 2.

Figure 1 indicates that for values of MINPK between 1 and 186 inclusive, the resulting packed message size was a minimum at about 10, but there was no practical difference for MINPK values from about 8 through 14. This generally agrees with earlier results by Glahn (1994).

For many of the values of MINPK, repeated runs were made to get more stable values of gribbing times. Due to the natural variation in the timing of processes running on a multiprocessing machine and the poor resolution of the timing function available, small variations in the gribbing times in Fig. 1 are not significant. However, it is reasonable that the minimum packing time is for a value of MINPK the same or slightly higher than the value for minimum message size. The horizontal lines show message size and packing time for simple packing for comparison.

A value of MINPK = 1 would be, of course, ridiculous for use; it is included for interest. MINPK = 1 forces each group to contain only points of exactly the same value.

The aberration around MINPK = 92 is real; 92.5 is half the row length of the field. As indicated by Glahn (1994), MINPK should most assuredly be less than half the row length for best results. In the case of large fields, this is of little importance, because half the row length is much larger than a reasonable value for MINPK.

The final three values to the right in Fig. 1 are for the row length = 185 ± 1. Because INC = 1, rather than some large value, the gribbing time is quite high.

Figure 2 shows the average number of points put into each group for the various values of MINPK. Again the jitter about a relatively straight line is apparent at about MINPK = 92; the transition indicated in Figs. 1 and 2 would probably be more abrupt is the row length were <u>evenly</u> divisible by 2. A MINPK near the row length creates very large groups. Below MINPK = 92, the average number of points per group was about 1.25 X MINPK.

A few runs were made with varying values of MINPK and INC = 3, Results were as expected (not shown). For each value of MINPK from 2 through 14, the message size was slightly larger and the gribbing time somewhat smaller.

While these results are for one cycle on one day and for only one type of field, the general nature of the results should hold for many large scale (in relation to grid spacing) fields.

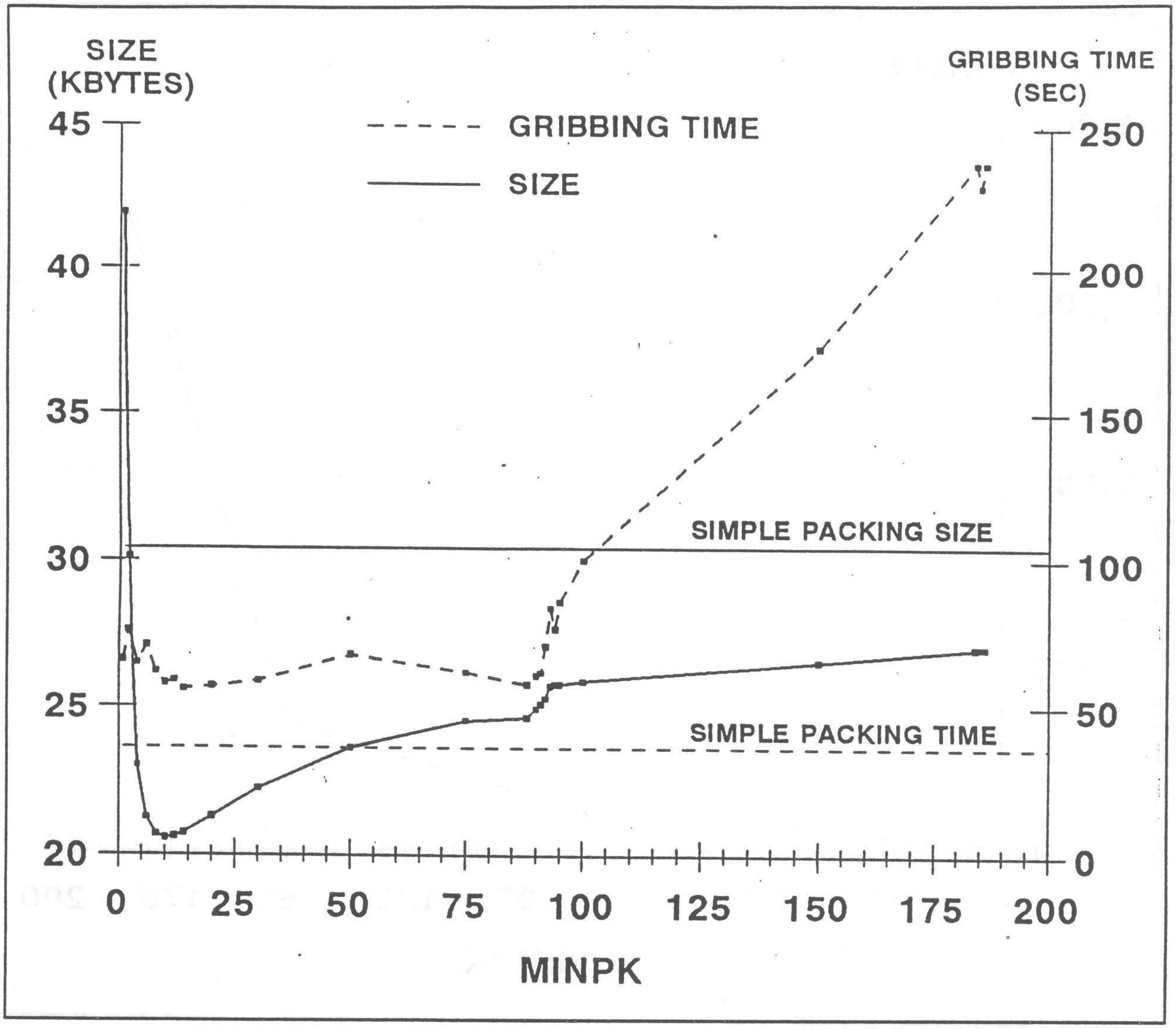

Figure 1. Message sizes and gribbing times for complex packing as a function of MINPK for INC = 1. Straight lines are shown for simple packing for reference.

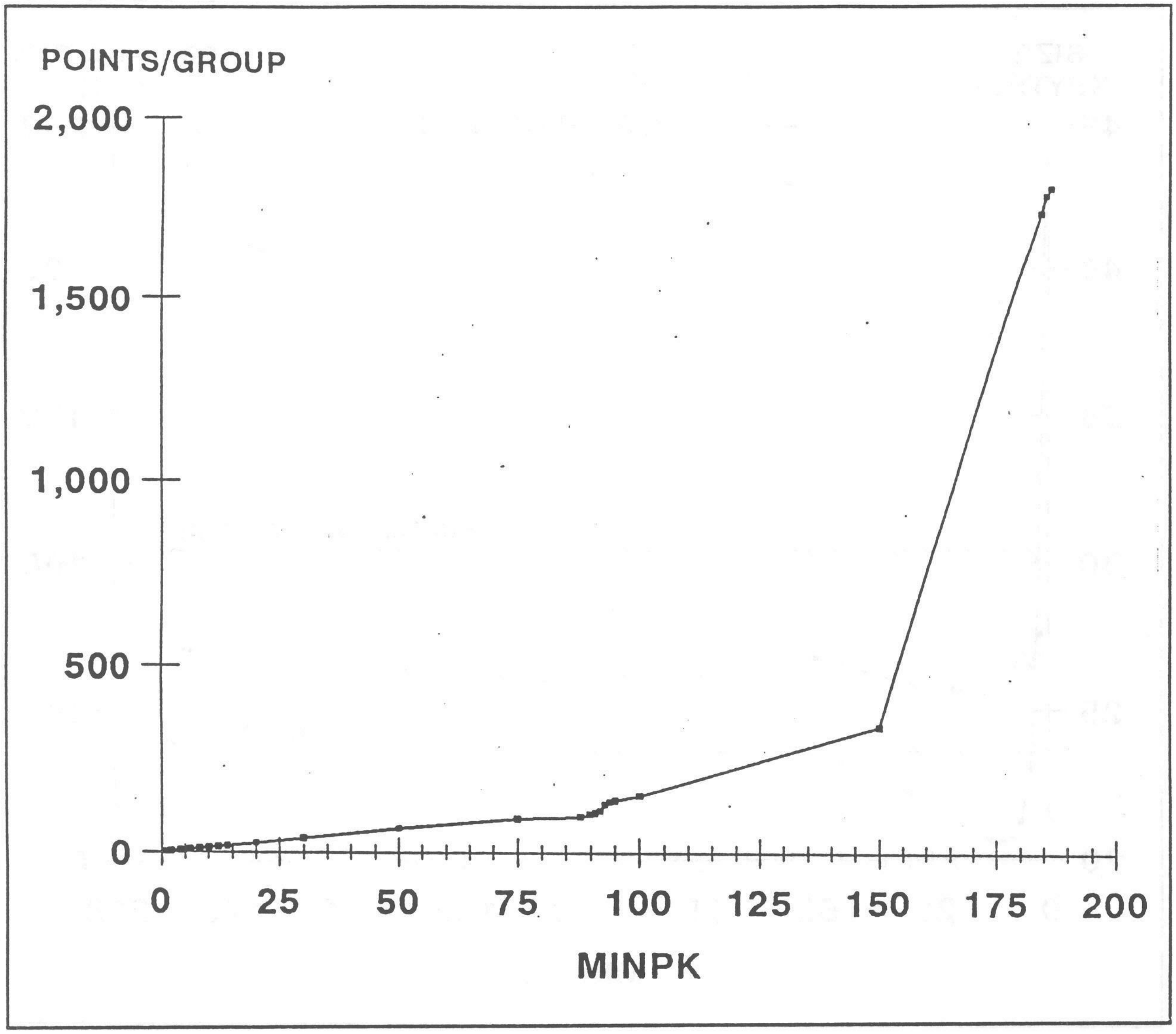

Figure 2. Points per group for complex packing as a function of MINPK for INC = 1.

APPENDIX II

Proposed Scaling for Eta Model Grids

A study is underway at NCEP designed to lead to the scaling of data in GRIB messages more appropriate than current practice. The thrust is to provide the various types of data to field stations with scaling appropriate not only for viewing but also sufficient for certain computations to be made with the necessary accuracy to be most useful. To be rather precise in doing this, more use would be made of binary, rather than decimal, scaling. For instance, the accuracy for geopotential heights on an 80-km grid necessary to compute geostrophic wind would have to be increased by a factor of 2 (4) for a 40-km (20-km) grid to get the same geostrophic wind accuracy on those grids. Tables 3 and 4, respectively, show for the same 40-km and 20-km grids used in the body of this office note a comparison of message sizes for simple and complex packing (similar to Tables 1 and 2, respectively) with scaling factors tentatively proposed by NCEP.

In Tables 3 and 4, D and E are the decimal and binary scaling factors, respectively, as defined in the GRIB document. As noted earlier, a positive D and a negative E mean more accuracy. Use of these new scaling factors would not change the conclusions reached in the body of this office note. For instance, the overall improvement for the 1884 40-km fields is still 20% for complex over simple packing (simple-packed fields 25% larger). The overall improvement for 279 20-km fields is 25% (simple-packed fields 33% larger). (In Table 2, all 366 fields in the files were used for the overall total, whereas in Table 4 only those included in the individual lines in the table were included because those are the only fields dealt with in the proposal as of this date.) Note that the numbers for vorticity in Tables 2 and 4 do not match because not all scaling was D = 6 for Table 2 whereas it was for Table 4.

For some of the fields, the original accuracy as packed did not equal the accuracy implied in Tables 3 and 4. The worst discrepancy is for geopotential height where the units were meters but the values were packed and the results shown in Table 3 to 0.125 m. In order to see what effect this might have, a random component was added to the heights such that each gridpoint value when rounded would give the original value. That is, each gridpoint could be different from the original by 0.5 m. While a random process inserted here may not be readily accepted, it is likely the low order digit in a height field in tenths of meters is reasonably random. Also, the random component will give the worst results that could be expected. The difference in results was small, the improvement of complex over simple packing being decreased by only a few tenths of percent. It is likely the added accuracy before packing would seldom affect the bits/point for simple packing (although it could actually increase or decrease a whopping 1 bit/point), and would have somewhat more, but small, effect on complex packing.

The efficiency in message size of complex over simple packing as a function of grid spacing deserves consideration. Complex packing is relatively most efficient for fields that have a large scale pattern (high spatial redundancy) and/or that have a large gradient from one "side" of the grid to the other, for example, large north-south gradient of temperature or geopotential height; this relative efficiency is illustrated in Tables 1 through 4. Since most of

Table 3. Statistics associated with packing 40-km Eta model fields by both the simple and complex GRIB methods, similar to Table 1 but with scalings tentatively proposed by NCEP.

Field Definition	No. Fields	Message Length (byte)	Bits/ Point	Improvement (%)	Points/ Group	Dec. Scale D	Bin. Scale E
Geopotential Height (7, 100)							
Simple	324	41172.4	13.80	--	--	0	-3
Complex	324	31273.0	10.48	24	17.3	0	-3
Temperature (11, 100)							
Simple	324	28999.0	9.72	--	--	0	-3
Complex	324	20695.6	6.94	29	17.5	0	-3
Relative Humidity (52, 100)							
Simple	264	23656.0	7.93	--	--	0	0
Complex	264	20967.4	7.03	9	17.8	0	0
U-, V-Wind Components (33 & 34, 100)							
Simple	648	30300.8	10.16	--	--	0	-3
Complex	648	24962.6	8.37	18	17.6	0	-3
Vertical Velocity 39, 100)							
Simple	324	33634.3	11.27	--	--	2	-2
Complex	324	27713.8	9.29	18	21.0	2	-2
Overall							
Simple	1884	31588.7	10.59	--	--	-	-
Complex	1884	25227.3	8.46	20	17.7	-	-

the variation in those fields (from the minimum value, say) is in the large-scale features, it is intuitive that for a given accuracy (i.e., given decimal and/or binary scaling), the complex method gets more efficient (more improvement over simple packing) for reduced grid spacing. To illustrate this point, the 54 temperature fields shown in Table 4 were packed at various binary scaling factors E (and decimal scaling D = 0) for not only the original 20-km gridpoint data, but also for 40-km and 80-km grids covering the exact same

Table 4. Statistics associated with packing 20-km Eta model fields by both the simple and complex GRIB methods, similar to Table 2 but with scalings tentatively proposed by NCEP.

Field Definition	No. Fields	Message Length (byte)	Bits/ Point	Improvement (%)	Points/ Group	Dec. Scale D	Bin. Scale E
Vorticity (41,100)							
Simple	45	125715.2	10.61	--	--	6	0
Complex	45	96910.0	8.18	23	17.8	6	0
Temperature (11, 105 & 108)							
Simple	54	128832.0	10.87	--	--	0	-4
Complex	54	93914.6	7.92	27	17.2	0	-4
Relative Humidity (52, 105 & 108)							
Simple	54	93770.0	7.91	--	--	0	0
Complex	54	71792.3	6.06	23	18.0	0	0
U-, V-Wind Components (33 & 34, 105 & 108)							
Simple	108	130888.2	11.04	--	--	0	-4
Complex	108	102055.2	8.61	22	17.3	0	-4
Precipitation 61 & 63, 1)							
Simple	18	102858.9	8.68	--	--	1	0
Complex	18	32590.8	2.75	68	75.8	1	0
Overall							
Simple	279	119658.0	10.09	--	--	-	-
Complex	279	89184.6	7.52	25	18.4	-	-

area and using every second and fourth point, respectively, in each direction.[24]

The original 20-km temperature data were available to tenths of degrees, which is slightly greater accuracy than E = -3. Therefore, the original data

---

[24] AWIPS grids have been defined so that such is possible. That is, the 369 X 257 20-km grids can become 185 X 129 40-km grids, 93 X 65 80-km grids, 47 X 33 160-km grids, or 24 X 17 320 km grids covering exactly the same area (i.e., the outer rows and columns of the grids at those gridspacings coincide).

carried the necessary accuracy for the comparisons to be legitimate, and perhaps, even believable. Figure 3 shows that for a given accuracy, the smaller the gridlength the more improvement of complex over simple packing. It also indicates that the improvement at a given gridlength is less the more accuracy that is retained. Generally, there is an additional 10% improvement when the grid spacing is halved and the accuracy retained is held constant.[25]

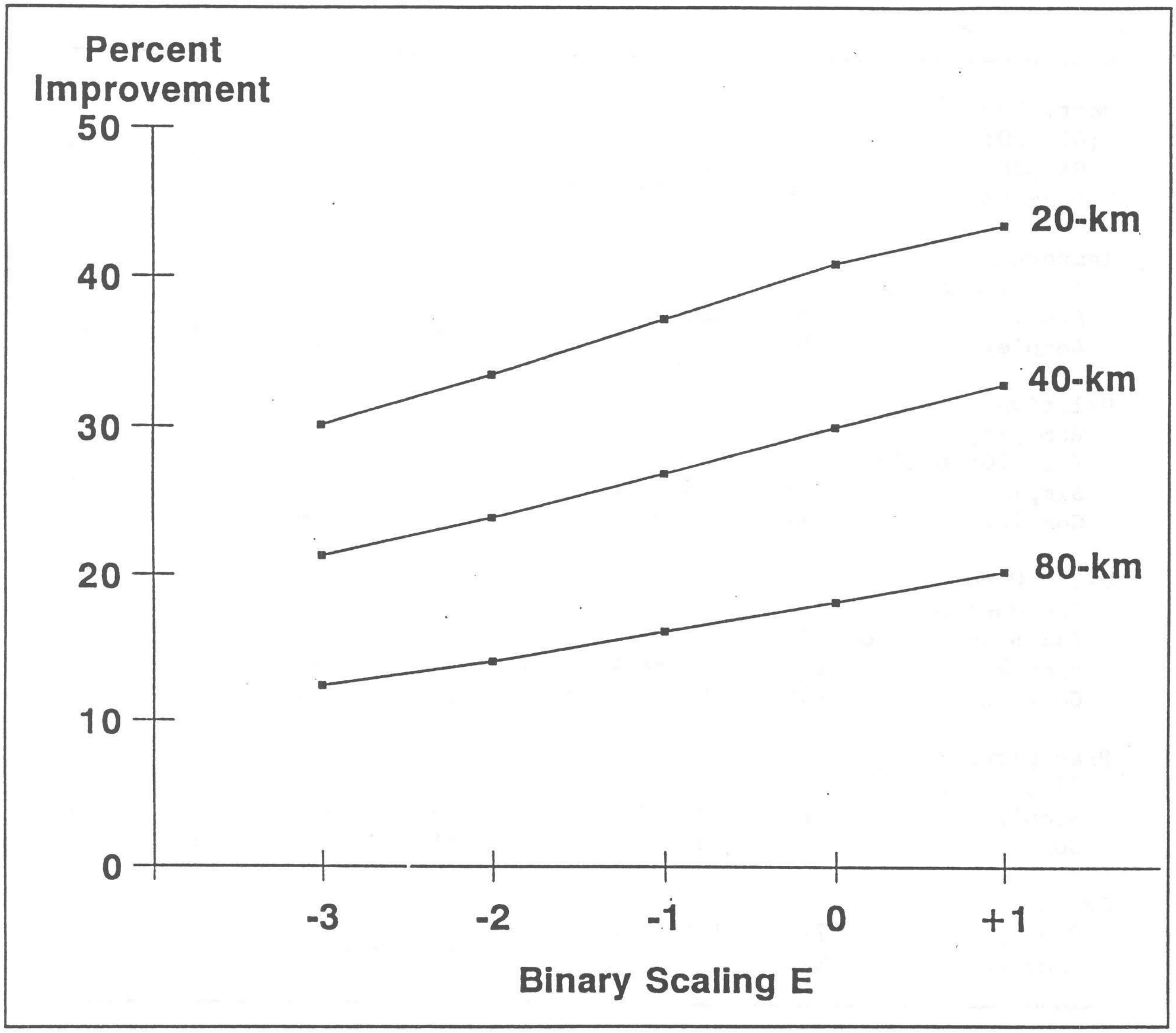

Figure 3. Percent improvement of complex packing over simple packing for 54 temperature fields as a function of binary scaling for each of 20-km, 40-km, and 80-km grids.

However, what happens when there needs to be more accuracy retained as the grid length is reduced? Figures 4 through 7 give different views of the results from the 54 temperature grids. Figure 4 indicates that when the

---

[25] The accuracy is 2, 1, 0.5, 0.25, and 0.125 degrees Celsius for E = +1, 0, -1, -2, and -3, respectively, for D = 0. There is no suggestion here that 2-degree accuracy is sufficient for any purpose.

gridlength is decreased from 80 to 40 to 20 km, and the accuracy is increased, respectively, from E = -1 to E = -2 to E = -3, the efficiency increases. This is what would be expected for temperature if the NCEP proposal is implemented.

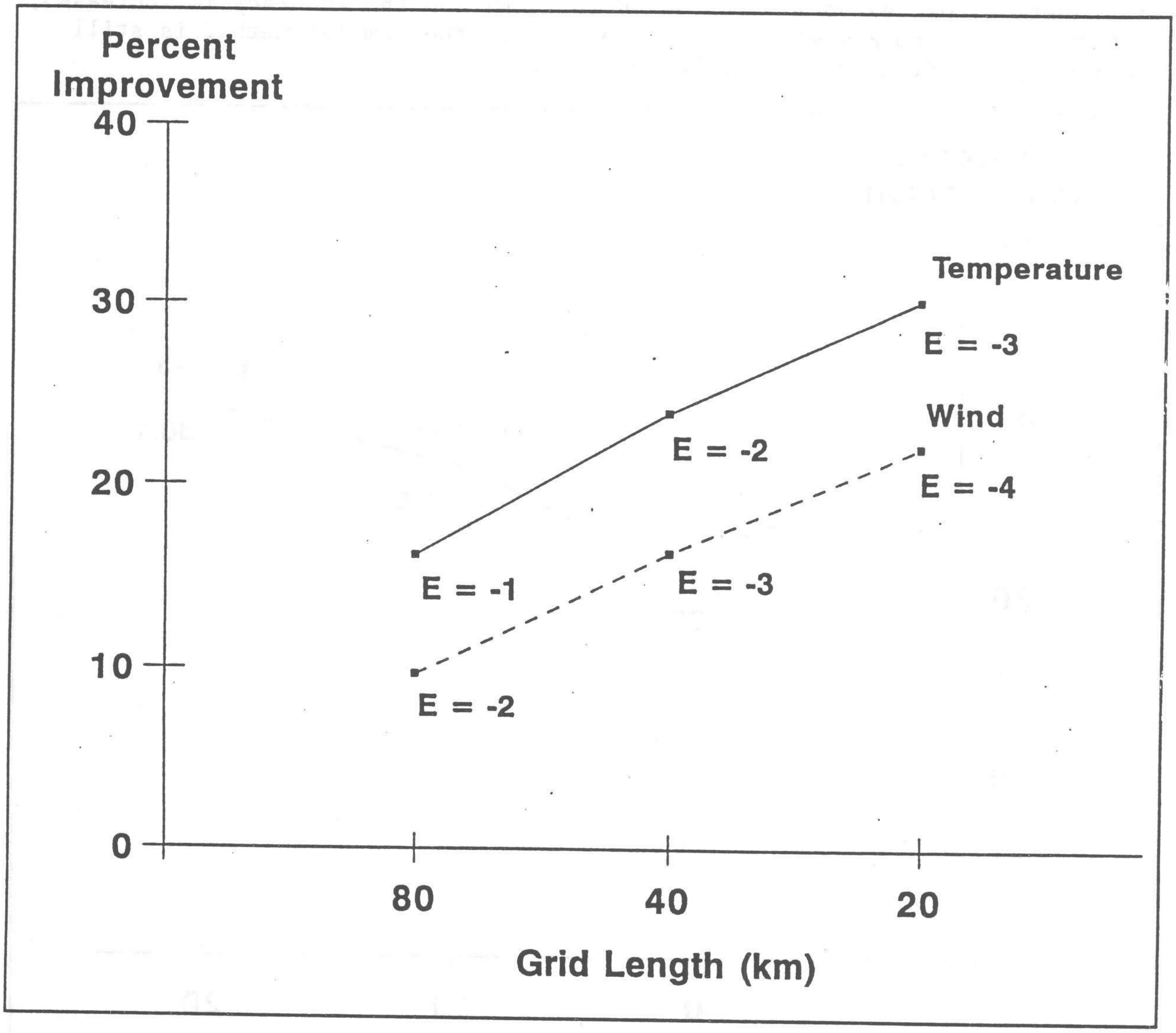

Figure 4. Percent improvement of complex packing over simple packing as a function of gridlength where the binary accuracy is increased by a factor of 2 for each halving of the gridlength.

As an added test, the 108 wind component fields shown in Table 4 were packed with E = -2, E = -3, and E = -4 as the gridlength was decreased from 80 to 40 to 20 km. The original 20-km data were scaled to tenths of m/sec, so the accuracy nearly supported E = -4, and generally, the low order digit does not control the message size to any great extent. The wind component results, also plotted in Fig. 4, show an increase in (message size) efficiency of complex over simple packing similar to temperature as the gridlength is decreased.

For geopotential height, it may be necessary to increase accuracy by a factor of 4 for each halving of the grid spacing. While the geopotential height data were not packed in the sample available to an accuracy that would make experimentation believable, the temperature data--being much like height data in overall characteristics--will suffice. Figure 5 shows that when the gridlength is decreased from 80 to 40 to 20 km and the accuracy in increased, respectively, from E = +1 to E = -1 to E = -3, the complex method is still increasingly efficient for smaller gridlengths.

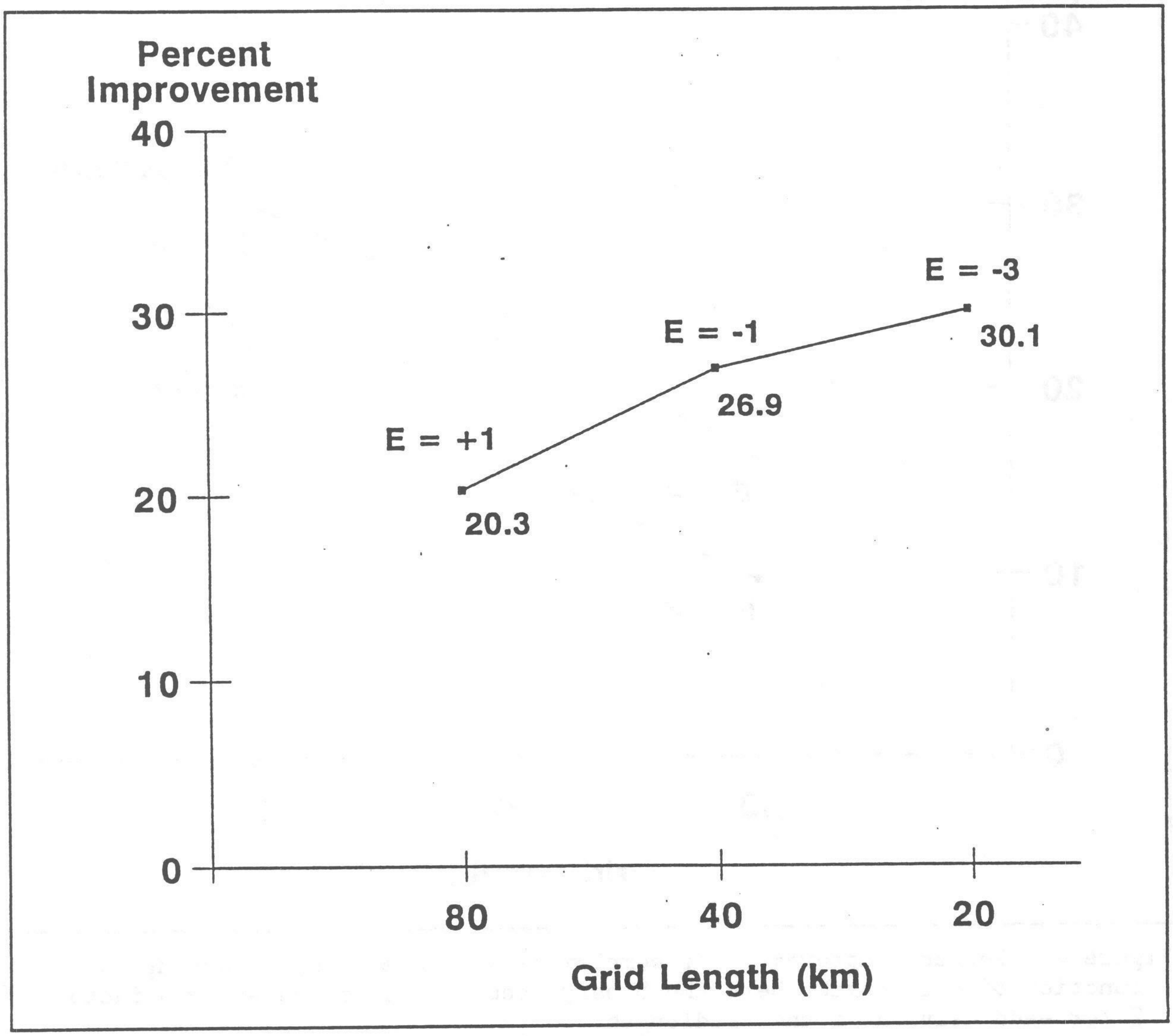

Figure 5. Percent improvement of complex packing over simple packing as a function of gridlength where the binary accuracy is increased by a factor of 4 for each halving of the gridlength.

Because the difference between the highest and lowest values on the grid determine the number of bits needed per point for the simple method and because this difference is essentially constant regardless of grid spacing, the bits per gridpoint will increase by approximately 1 for each binary unit increase in accuracy. (In the event that the "omitted" gridpoints in the field with the larger grid spacing determine the bits required, the bits/point could increase by 2 for a unit binary accuracy increase which is also associated with a decrease in grid spacing.) On the other hand for the complex method, the bits required per point may remain relatively constant, as indicated in Fig. 6.

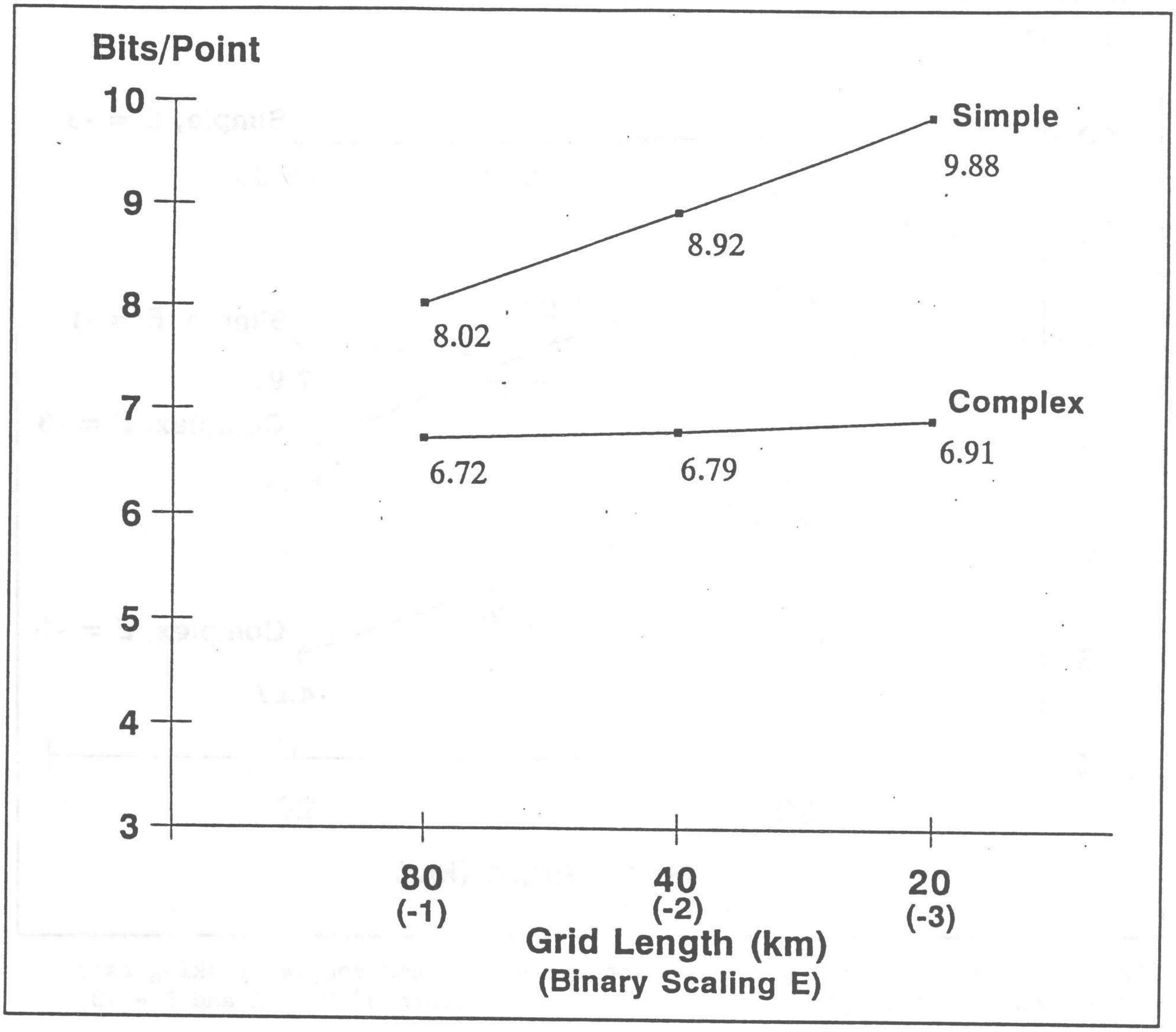

Figure 6. Bits required per gridpoint for complex packing over simple packing as a function of gridlength where the binary accuracy is increased by a factor of 2 for each halving of the gridlength.

Figure 7 shows for simple and complex packing and for binary scaling E = -1 and E = -3, the bits required per point as a function of gridlength. The constancy of values, and the approximate separation by 2 bits/point, for the simple method is obvious, while the bits/point for the complex method decreases with smaller gridspacing.

In summary, Fig. 4 gives an idea of the relative efficiencies that might be obtained for temperature and wind components when the accuracy is doubled for a halving of the grid spacing, and Fig. 5 gives a similar picture of what might be obtained for geopotential heights when the accuracy is quadrupled for a halving of the grid spacing. For many fields, the accuracy would not need to be increased as the grid spacing is decreased. For instance, this would

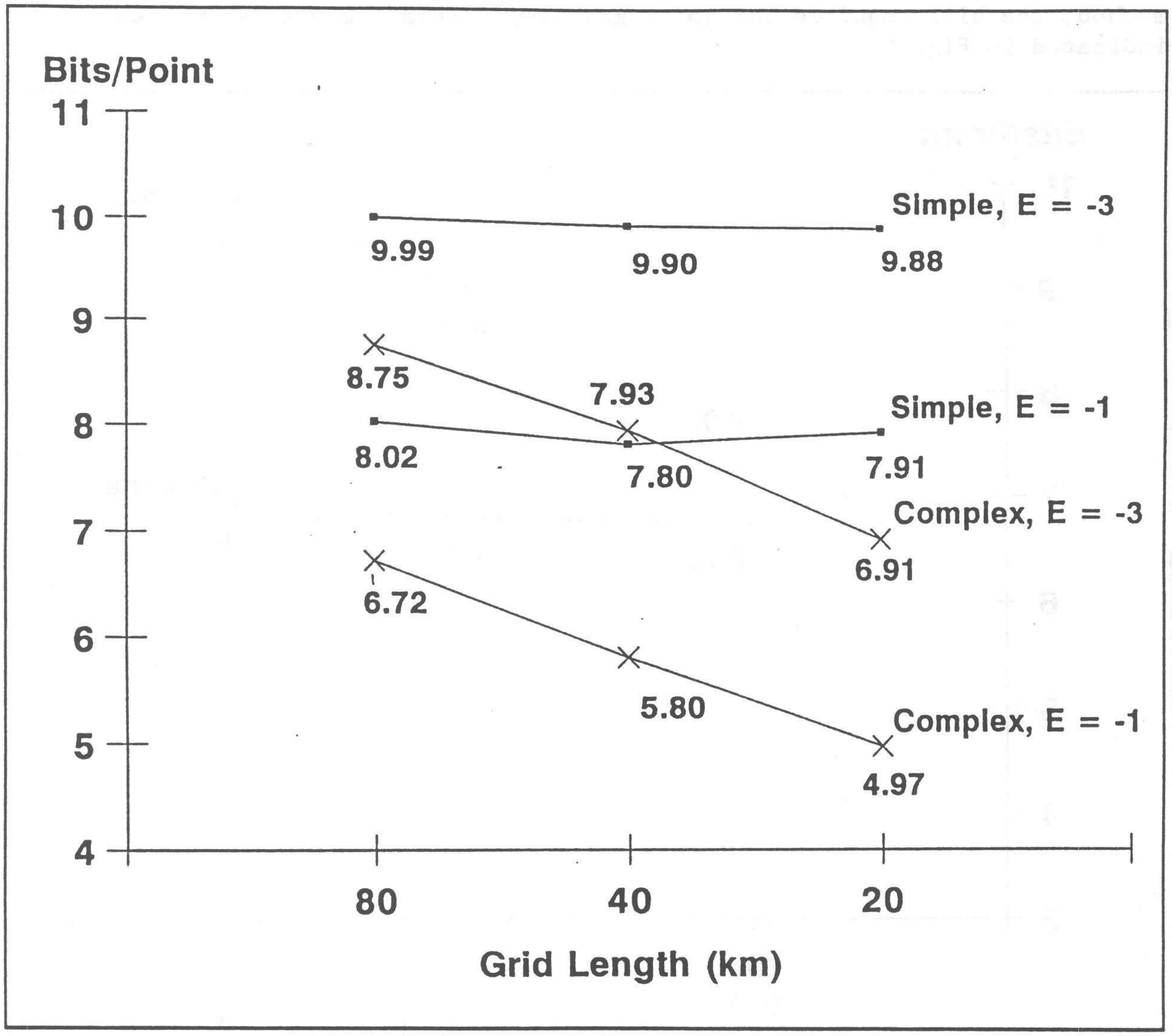

Figure 7. Bits required per gridpoint for simple and complex packing as a function of gridlength at constant scaling factors of E = -1 and E = -3.

likely apply to vertical velocity, precipitation amount, and relative humidity. For these fields, the increase in relative efficiency of complex over simple packing is obvious as the grid spacing is reduced (see Fig. 3).

Are these results representative? What would make them inappropriate? First, of course, the sample is from only one day in late November. However, the area is large enough for the pattern to be "representative" of what might be expected on a day that is neither summer nor deep winter; the results by season would be expected to vary about those shown here. Secondly, some fields do not benefit from complex packing as much as temperature. However, these are generally the ones for which the retained accuracy does not need to depend on grid spacing, and as the grid spacing becomes less, there is no question that the relative efficiency goes up. Thirdly, as models become more bonded with high resolution terrain features, the lower level fields may have considerable small scale variation that will militate against the relative efficiency of the complex packing. That is, if the very small scale variation is a large part of the total variation, complex packing will not be as advantageous as for fields where most of the variation is in patterns on a scale of greater than MINPK times the gridlength. In the case of 20-km fields, this would be a scale of about 280 km. For these efficiency comparisons, all data were from the 20-km eta fields and were degraded to lower resolutions as needed. The eta data may not have the fine scale detail that future models will have, especially at the very low levels.

Every situation, with a given numerical model, grid spacing used for the message, possible smoothing before packing to remove "noise," accuracy to be retained, time of year, meteorological variable, and areal extent of the grid, etc., will give specific results that will not duplicate what are contained here. However, it is believed that these results, results that support complex packing, are relevant and should help determine the packing methods to be used by the NWS for AWIPS.

The question still remains, and is a matter of judgment, as to how much resolution is needed for the various fields. To shed some light on this, an experiment was done in which geopotential heights were used for geostrophic wind computations. It was assumed the heights were rounded to meters on a 40-km grid, and that the part rounded off was a random component. As indicated earlier, this would be the worst condition that would occur, and is probably realistic because the low order digits at alternate gridpoints would likely not have much spatial pattern. The distribution of errors in a geostrophic wind component that would occur because of rounding is somewhat "normal," because the errors are computed as (a function of) the differences between two random numbers (taken from a rectangular distribution). From simulation, the numbers in Table 5 show characteristics of errors at 35° latitude for an individual wind component and for the total wind.

Table 5. Characteristics of geostrophic wind errors in m/sec at 35° latitude caused by rounding heights to meters on a 40-km grid.

Error Characteristic	Single Wind Component	Single Wind Component With Smoothing	Total Wind
Maximum Error	1.464	1.016	2.053
Average Absolute Error	0.488	0.192	0.764
Standard Deviation of Error	0.598	0.239	0.363

So, for an individual wind component, the maximum error is 1.46 m/sec but the average (absolute) error is only 0.49 m/sec. If a simple 5-point smoother were used on these errors, the reduction in absolute error (due to rounding) would be from 0.49 to 0.19 m/sec. In Fig. 8, the dotted curve shows the relative frequency of errors of an individual wind component. For comparison, and with the same mean and standard deviation, the normal curve is shown. Also shown is the distribution of errors after the application of a 5-point smoother (the values of σ on the abscissa apply to the errors and the normal curve, but not to the smoothed errors).[26] So, even though the maximum error is 1.46 m/sec, the average error is much smaller, and 90 percent of the errors are below 1 m/sec in magnitude.

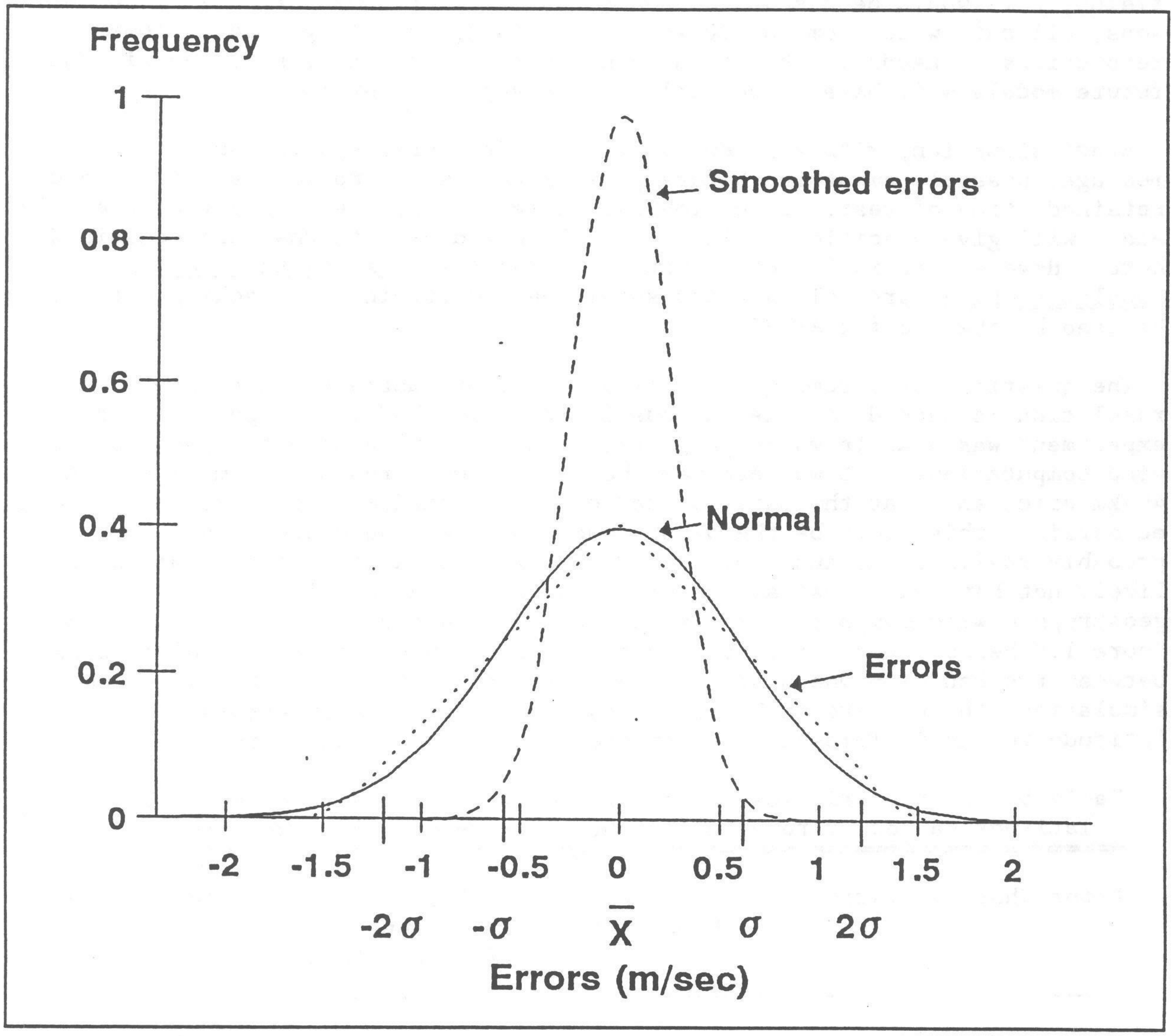

Figure 8. Distribution of errors in m/sec in a geostrophic wind component when heights are rounded to meters on a 40-km grid. Smoothed errors and the normal curve are shown for comparison.

For the total wind component, the maximum error is 2.05 m/sec and the average error is 0.76 m/sec. Note that the maximum error for the total wind is approximately the maximum error for a single component multiplied by the

---

[26] It is not suggested, necessarily, that a smoother be used. However, certain calculations made on model fields can be expected to have some "noise" irrespective of rounding of the transmitted data. It is not unusual to smooth such fields before viewing them so that they will have better "patterns." In these cases, the highly random errors due to rounding of the data largely disappear when viewed against the real meteorological information content.

square root of 2; this is not true for the average absolute error. The distribution of total errors is indicated in Fig. 9,[27] and there compared to the normal distribution. So, while the maximum error can be as high as 2.05 m/sec, such an error rarely occurs, and 90 percent of them are below 1.25 m/sec.

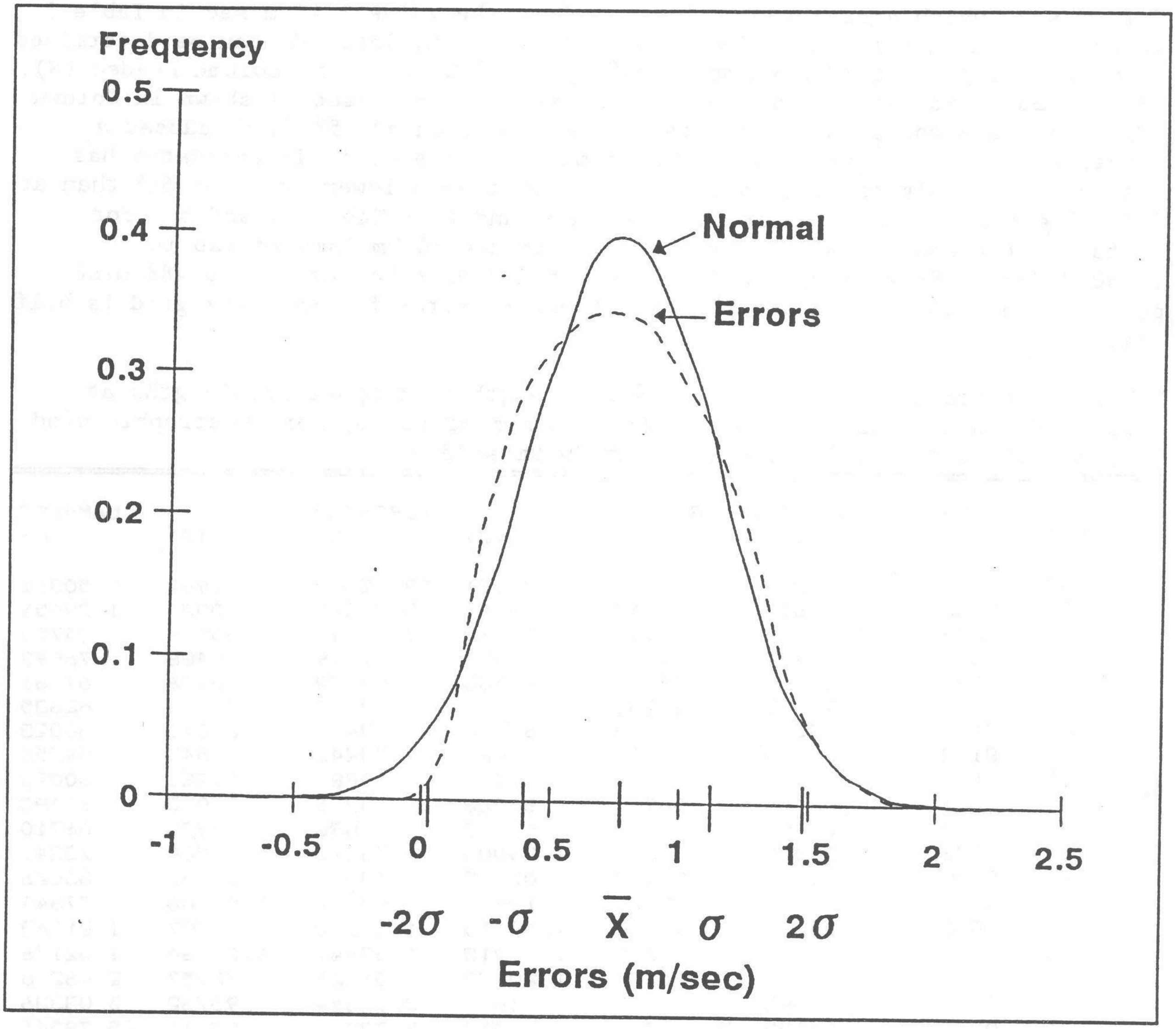

Figure 9. Distribution of errors in m/sec of total geostrophic wind speed as a result of rounding heights to meters on a 40-km grid. The normal curve is shown for comparison.

The values in Table 5 hold for reductions in grid spacing by factors of 2^n which is accompanied with increased packing resolution of 2^{-n}. That is, the error produced by rounding heights to meters on a 20-km grid would be double those in Table 5; however, if the heights were rounded to 0.5 m instead of whole meters, the errors would be the same as shown in Table 5.

---

[27] In Figs. 8 and 9, each curve is scaled such that the area under the curve is unity where the abscissa is in units of σ as shown.

It is noted that the values in Table 5 apply to latitude 35°. The latitude enters in two different ways into the calculation--the coriolis parameter and the map factor. The variation of error with latitude can be determined from Tables 6, 7, and/or 8. For instance, Table 6 is most appropriate for the AWIPS Lambert Regional maps which are tangent at 25° and have a scale (gridlength) of 80 km at 35°. At 35°, the maximum error, shown in Column (G), is 0.73 m/sec, which agrees with Table 5, since the value 1.46 m/sec in Table 5 is for a 40-km gridlength. The grid spacing at any latitude can be determined by multiplying the grid spacing at 35° by the factor in the column headed (S). The maximum error at 60° for the Lambert map is 0.60 m/sec as shown in Column (G). Although the grid spacing is less at 60° than at 35° which causes a larger error in the geostrophic wind component, the coriolis parameter has less effect at higher latitudes and the result is a lower error at 60° than at 35°. The mean absolute error can also be found from Tables 5 and 6. For instance, the mean absolute error at 60° for the 80-km Lambert map is (0.60/0.73)(0.488/2) = 0.20 m/sec, the 2 being used because the 0.488 m/sec pertains to a 40-km grid and the corresponding error for an 80-km grid is half that.

Table 6. Factors (S) to multiply by gridlength to compute gridlengths at latitudes other than 35°. Also, the maximum error (G) for geostrophic wind calculations on a grid of gridlength 80 km at 35°.

LATITUDE DEG	SIN LAT	POLAR STEREOGRAPHIC (S)	(G)	(S)	MERCATOR (G)	(S)	LAMBERT (G)
89	.9998	1.27089	.33060	.02131	19.72080	.11992	3.50361
85	.9962	1.26857	.33242	.10640	3.96345	.30327	1.39051
80	.9848	1.26134	.33819	.21199	2.01229	.45044	.94703
75	.9659	1.24934	.34812	.31596	1.37649	.56488	.76993
70	.9397	1.23267	.36268	.41753	1.07072	.65976	.67761
65	.9063	1.21145	.38262	.51592	.89844	.74004	.62635
60	.8660	1.18585	.40906	.61039	.79472	.80816	.60023
55	.8192	1.15606	.44361	.70021	.73242	.86547	.59256
50	.7660	1.12231	.48863	.78470	.69886	.91281	.60078
45	.7071	1.08486	.54763	.86322	.68825	.95075	.62488
40	.6428	1.04398	.62602	.93517	.69886	.97970	.66710
35	.5736	1.00000	.73241	1.00000	.73242	1.00000	.73241
30	.5000	.95324	.88140	1.05722	.79472	1.01196	.83026
25	.4226	.90407	1.09951	1.10640	.89844	1.01588	.97849
20	.3420	.85285	1.44021	1.14715	1.07072	1.01207	1.21363
15	.2588	.79997	2.02898	1.17918	1.37649	1.00084	1.62176
10	.1736	.74585	3.24361	1.20223	2.01229	.98257	2.46216
5	.0872	.69088	6.97668	1.21613	3.96344	.95762	5.03336
1	.0175	.64659	37.22776	1.22059	19.72077	.93315	25.79541

In a similar fashion, Tables 7 and 8 can be used easily for the AWIPS polar stereographic and mercator maps, respectively. The former have the scale quoted at 60° latitude and the latter at 20°; for the standard Regional Scale, the gridlengths are 95.25 and 80 km.

These tables indicate, of course, that geostrophic wind should not be computed near the equator because of the coriolis parameter. This is exacerbated by the map factor for the polar stereographic projection. The tables also show, as expected, that the map factor creates large errors near the pole for the mercator and to a considerably lesser extent for the Lambert.

Table 7. Factors (S) to multiply by gridlength to compute gridlengths at latitudes other than 60°. Also, the maximum error (G) for geostrophic wind calculations on a grid of gridlength 95.25 km at 60°.

LATITUDE DEG	SIN LAT	POLAR STEREOGRAPHIC		MERCATOR		LAMBERT	
		(S)	(G)	(S)	(G)	(S)	(G)
89	.9998	1.07172	.32928	.03490	10.11009	.14839	2.37815
85	.9962	1.06976	.33109	.17431	2.03191	.37526	.94384
80	.9848	1.06366	.33684	.34730	1.03163	.55736	.64282
75	.9659	1.05354	.34672	.51764	.70567	.69897	.52261
70	.9397	1.03948	.36122	.68404	.54892	.81637	.45994
65	.9063	1.02159	.38109	.84524	.46060	.91571	.42515
60	.8660	1.00000	.40742	1.00000	.40742	1.00000	.40742
55	.8192	.97488	.44183	1.14715	.37548	1.07091	.40221
50	.7660	.94642	.48667	1.28558	.35828	1.12949	.40779
45	.7071	.91484	.54544	1.41421	.35284	1.17643	.42415
40	.6428	.88037	.62351	1.53209	.35828	1.21225	.45281
35	.5736	.84328	.72948	1.63830	.37548	1.23737	.49714
30	.5000	.80385	.87787	1.73205	.40742	1.25218	.56356
25	.4226	.76238	1.09510	1.81262	.46060	1.25703	.66417
20	.3420	.71919	1.43443	1.87939	.54892	1.25231	.82378
15	.2588	.67460	2.02084	1.93185	.70567	1.23842	1.10080
10	.1736	.62896	3.23060	1.96962	1.03163	1.21580	1.67124
5	.0872	.58261	6.94870	1.99239	2.03190	1.18494	3.41650
1	.0175	.54525	37.07846	1.99970	10.11007	1.15465	17.50920

Table 8. Factors (S) to multiply by gridlength to compute gridlengths at latitudes other than 20°. Also, the maximum error (G) for geostrophic wind calculations on a grid of gridlength 80 km at 20°.

LATITUDE DEG	SIN LAT	POLAR STEREOGRAPHIC		MERCATOR		LAMBERT	
		(S)	(G)	(S)	(G)	(S)	(G)
89	.9998	1.49018	.28195	.01857	22.62277	.11849	3.54588
85	.9962	1.48746	.28350	.09275	4.54668	.29966	1.40729
80	.9848	1.47897	.28843	.18479	2.30841	.44507	.95846
75	.9659	1.46490	.29689	.27543	1.57905	.55814	.77922
70	.9397	1.44535	.30931	.36397	1.22828	.65189	.68578
65	.9063	1.42048	.32632	.44974	1.03065	.73122	.63391
60	.8660	1.39046	.34887	.53209	.91166	.79853	.60747
55	.8192	1.35553	.37833	.61039	.84019	.85515	.59971
50	.7660	1.31596	.41673	.68404	.80170	.90193	.60802
45	.7071	1.27204	.46705	.75249	.78952	.93941	.63242
40	.6428	1.22412	.53390	.81521	.80170	.96802	.67515
35	.5736	1.17254	.62464	.87172	.84019	.98808	.74125
30	.5000	1.11772	.75170	.92160	.91166	.99990	.84028
25	.4226	1.06006	.93772	.96447	1.03065	1.00377	.99030
20	.3420	1.00000	1.22828	1.00000	1.22828	1.00000	1.22828
15	.2588	.93800	1.73041	1.02792	1.57904	.98891	1.64133
10	.1736	.87454	2.76630	1.04801	2.30841	.97085	2.49187
5	.0872	.81009	5.95004	1.06013	4.54668	.94621	5.09409
1	.0175	.75815	31.74959	1.06402	22.62274	.92202	26.10668

Other simulation studies such as these can be made for other variables that will need to be calculated at NWS field stations. It is emphasized that while the maximum error can many times be readily calculated, a more meaningful number may be the average (absolute) error, which in the case of geostrophic wind is about 1/3 of the maximum error.