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

TDL OFFICE NOTE 98-3

PACKING OF INDIVIDUAL RADAR PRODUCTS

Harry R. Glahn

November 1998

PACKING OF INDIVIDUAL RADAR PRODUCTS

Harry R. Glahn

1. INTRODUCTION

This office note is one in a series dealing with the efficiency of data packing in the World Meteorological Organization (WMO) GRIB (GRIdded Binary) (WMO 1988; Dey 1996) format. The Advanced Weather Interactive Processing System (AWIPS) will soon be distributing selected radial and raster radar products[1] from Weather Forecast Offices (WFOs). This set of products will undoubtedly grow considerably in the coming years; therefore, since transmission bandwidth, especially on the terrestrial Wide Area Network (WAN)[2], is limited, it is desirable to have an efficient data packing[3] method for the distribution of the products.[4] Since raster radar data are essentially "gridpoint" and a radial data set is productized in a way that can be processed like gridpoint data,[5] it seems appropriate that GRIB be used if it is cost effective in terms of packing, unpacking, and compression ratios, as compared to other schemes and specifically the run-length encoding currently being used for many radar products.

The current form of GRIB (i.e., GRIB Edition 1) has a number of deficiencies, discussed by Glahn (1993). A new GRIB, Edition 2, is being designed and has recently been approved by the WMO for experimental use. This new version will be more flexible and will explicitly accommodate radar and satellite data--options not considered in Edition 1 of GRIB. It is expected that Edition 2 will include as an option essentially the packing capabilities of TDLPACK developed for data archival used by the Techniques Development Laboratory for guidance product development (e.g., MOS). TDLPACK makes use of (1) second-order packing of values with a more efficient method of defining groups than the current "Complex" GRIB method of using a (secondary) bit map, (2) a more efficient method of identifying missing values than with the current (primary) bit map, (3) variable widths for group sizes rather than the 8 bits specified in GRIB Edition 1, (4) boustrophedonic ordering of values, and (5) packing second-order spatial differences when appropriate.

---

[1]"Radial" products contain data arranged by azimuth and range; "raster" products contain data on a cartesian grid.

[2]Current plans are for each WFO's radar data to be transmitted to the Network Control Facility (NCF) over the WAN, and then distributed to the other WFOs over the Satellite Broadcast Network (SBN).

[3]No distinction is made here between "packing" and "compression." Both terms just mean sending the same information in fewer bits.

[4]The storage and archival of the products might also use the same compression algorithm.

[5]A product with a constant number of values per radial can be treated as an X by Y array, just as a raster product.

The three packing methods that are referenced in this office note are described briefly in the next section. Section 3 presents results of packing several types of radar products, and Section 4 summarizes for comparison some previous results.

2. DESCRIPTION OF PACKING METHODS

The description here will be very brief; other references can be consulted for more detail (WMO 1988; Dey 1996; Glahn 1997, 1992, 1993, 1994, 1995).

A. Simple GRIB

GRIB, in this simple form, does basically two things: (1) the overall minimum value is subtracted from the field (the values at the gridpoints), making all values non-negative and having no larger magnitude than necessary, and (2) only as many bits per value are used as the largest value in the field (after subtraction of the overall minimum) requires, both after the desired decimal and/or binary scaling is done.[6] Missing values, if there are any, can be designated by a "0" in a bit map,[7] and the packed grid will then not include a value for that point.

B. Complex GRIB

Complex GRIB, in addition to doing what simple GRIB does, packs "groups" of "adjacent" values with only the number of bits required for that group after the minimum for the group (as well as the overall minimum) is subtracted. The definition of the groups is left to the originating organization. The groups can be of constant size (such as defined by grid row or column, which for fields covering a large area is not very useful) or the group size can vary. Unfortunately, the way GRIB carries variable group sizes is quite inefficient[8] (see Glahn 1992, 1993), but not enough to offset the advantage of "grouping." However, the groups must be picked with care (see Glahn 1994).

C. TDLPACK

TDLPACK takes advantage of the second-order spatial difference scheme defined by the Office of the Federal Coordinator for Meteorology (OFCM 1990), and a more efficient packing algorithm. Groups are defined and packed in much the same manner as in complex GRIB, but the way the groups are identified in the product is more efficient. Missing values can be designated by reserving the largest value in a group for a missing indicator. This is (almost) guaranteed to be more efficient than the GRIB schemes--usually considerably so.

---

[6]Decimal (binary) scaling means that the data values are multiplied by a power of 10 (2) before packing them as integer values.

[7]A bit map is itself a grid of the same size as the grid of values being packed, in which each value occupies one bit and is either a 1 or 0 indicating, respectively, whether or not a data value is actually present for that corresponding gridpoint.

[8]Complex GRIB actually requires <u>another</u> bit map that indicates where each group starts.

3. RESULTS

Several radial and raster radar products were decoded from their original run-length scheme (or used directly if not packed), packed with TDLPACK, and unpacked. The results are shown in Tables 1 through 7. In each table, the products are ordered by original size. All products are from the WFO at Sterling, Virginia, except as noted. In the tables, the Date Time is given as month and day, followed by hour and minute; Elev means beam elevation; and Ratio means the Output Size divided by the Original Size.

Consistent with the earlier reports by Glahn (e.g., 1998), the adjustable parameters MINPK and INC for TDLPACK were used as 14 and 1, respectively (see Glahn 1994). For determining processing times, a Hewlett Packard (HP) 755[9] was used that was otherwise only performing "housekeeping" chores with about 1 to 2 percent of its cp cycles. The timing software available was precise, but gives clock time, not actual cp time. Timing results were quite consistent, as shown by replication, but differences of only 1 or 2 percent between values are in the noise level.

Table 1 shows base reflectivity, Product ID 19, a 16-data level, run-length encoded, 0.54 n mi X 1 deg, radial product. The maximum number of values is 84,410. The original run-length encoded product size ranged from about 8,000

Table 1. Base reflectivity, Product ID 19. All data were from Sterling, Virginia, except for the second in the list, which was from Wakefield, Virginia. Data were from 1998 except for the third entry, which was from 1996.

Product Date Time	Elev (deg)	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)
0517 1205	3.4	8,238	136	0.02	0.258	0.021
0409 1723	0.5	15,470	7,560	0.49	0.272	0.030
0624 1916	0.5	17,110	10,192	0.60	0.249	0.036
0518 1744	1.5	18,066	8,664	0.48	0.257	0.029
0517 0235	0.5	21,818	13,304	0.61	0.251	0.039
0508 1310	0.5	26,358	17,800	0.68	0.240	0.043
0507 1102	0.5	32,226	19,696	0.61	0.258	0.039

to 32,000 bytes. Except for the first, which consisted of only a few values, the output product was 48 to 68 percent of the original. Packing times were about a quarter of a second per product and did not depend on original or final product size. Unpacking was about 0.02 to 0.04 seconds per product and did depend on product size. Figures 1, 2, and 3 show three base reflectivity products, those for 0409 1723, 0508 1310, and 0507 1102 (see Table 1). The latter two are rather "busy" products, being about double the size of the first. Base reflectivity at this resolution is one of the products proposed

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

to be distributed initially by AWIPS. Note that if largely "blank" products are to be sent, entry No. 1 may be representative of those products.

Table 2 shows base velocity, Product IDs 26 and 27. These are radial, 16 data level, 0.27 n mi X 1 deg and 0.54 n mi X 1 deg products, respectively. The sizes of these run-length encoded products could be reduced but not by as great a factor as base reflectivity. Packing times were similar to those for base reflectivity; unpacking times were 0.03 to 0.04 seconds per product. Base velocity at 0.54 n mi X 1 deg resolution is also proposed for initial AWIPS distribution.

Table 2. Base velocity, Product IDs 26 (the first three) and 27 (the last two). Products are ordered by original size within those two groups. Data are from 1998.

Product Date Time	Elev (deg)	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)
0517 1205	3.4	13,224	8,448	0.64	0.261	0.030
0508 1310	0.5	17,330	15,504	0.89	0.238	0.035
0506 2229	0.5	23,592	18,160	0.77	0.252	0.038
0518 1744	1.5	13,942	7,240	0.52	0.253	0.029
0518 1912	1.5	15,200	7,752	0.51	0.250	0.028

Table 3 contains 16 data level, 2.2 X 2.2 n mi, raster composite reflectivity, Product ID 38. These run-length encoded products of about 5,500 to 9,000 bytes were reduced to 31 to 60 percent of their original sizes. Packing times were about 0.15 seconds per product; unpacking was about 0.02 seconds. Composite reflectivity is also proposed to be distributed initially by AWIPS, but the 0.54 X 0.54 n mi product rather than the 2.2 X 2.2 n mi product. This higher resolution product could be approximately 8 times as large as the ones shown in Table 3.

Table 3. Composite reflectivity, Product ID 38. Data are from 1998.

Product Date Time	Elev (deg)	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)
0417 2044	--	5,496	1,696	0.31	0.152	0.017
0518 1733	--	5,596	1,776	0.32	0.154	0.016
0521 1040	--	6,128	2,392	0.39	0.148	0.020
0517 0229	--	7,202	3,856	0.54	0.149	0.019
0508 1304	--	8,208	4,456	0.54	0.160	0.020
0508 1704	--	9,060	5,472	0.60	0.151	0.020

Table 4 contains 16 data level, 2.2 X 2.2 n mi raster Vertically Integrated Liquid (VIL), Product ID 57. This product is also proposed for initial distribution on AWIPS. These are quite small products of less than 2,000 bytes. The packed output is 30 to 40 percent of the run-length encoded originals, except for the first entry in which all data values were zero. As with the first entry in Table 1, this may not be an unusual product size. Packing times are less than 0.04 seconds per product; unpacking times are about 0.004 sec. Figure 4 shows a VIL product, the one in Table 4 for 0517 0229.

Table 4. Vertically Integrated Liquid, Product ID 57. The data for 0503 2358 are from 1957; 0624 1910 are from 1996; the rest are from 1998.

Product Date Time	Elev (deg)	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)
0518 1733	--	1,318	96	0.07	0.038	0.004
0503 2358	--	1,576	456	0.29	0.038	0.002
0417 2044	--	1,586	432	0.27	0.037	0.005
0517 0229	--	1,722	672	0.39	0.038	0.003
0624 1910	--	1,844	752	0.41	0.037	0.004

Table 5 contains 8 data level, 0.54 n mi X 1 deg radial, base spectrum width, Product ID 30. These results are similar those in Table 2.

Table 5. Base spectrum width, Product ID 30. Data are from 1998.

Product Date Time	Elev (deg)	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)
0527 2018	0.5	21,966	12,176	0.55	0.251	0.035
0527 1918	0.5	23,974	13,968	0.58	0.267	0.037
0514 1433	0.5	21,464	9,120	0.42	0.250	0.033

Table 6 contains 16 data level 1.1 X 1.1 n mi raster, 1-h precipitation accumulation. These are relatively small products, undoubtedly containing many zeros. Packing ratios of 0.29 to 0.53 were achieved. Packing times were also relatively small--smallest of any set of products except for VIL. Storm total rainfall accumulation at this same resolution (not 1-h accumulation) is planned to be sent over AWIPS.

Table 7 contains digital hybrid scan reflectivity, Product ID 32. These original products, containing 85,110 values, were not encoded. They reduce through packing to 31 to 43 percent of their original values. Packing and unpacking times were similar to those in Tables 1 and 2.

Table 6. One-hour precipitation accumulation, Product ID 78. Data are from 1998.

Product Date Time	Elev (deg)	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)
0504 2332	--	6,488	1,888	0.29	0.123	0.011
0517 0229	--	6,822	2,672	0.39	0.114	0.012
0508 1058	--	7,790	4,104	0.53	0.120	0.015
0504 1310	--	8,310	3,656	0.44	0.111	0.015
0409 1954	--	10,738	5,336	0.50	0.116	0.017

Table 7. Digital Hybrid Scan, Product ID 32. The products are ordered according to the output size. Data are from 1997.

Product Date Time	Elev (deg)	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)
0613 0927	--	85,110	26,256	0.31	0.246	0.035
0613 0727	--	"	29,280	0.34	0.284	0.036
0613 0827	--	"	29,992	0.35	0.247	0.038
0613 1027	--	"	36,560	0.43	0.247	0.039

4. PREVIOUS RESULTS

For comparison, Table 8 gives results abstracted from Glahn (1997 and 1998). These results pertain to gridded data produced by the National Centers for Environmental Prediction (NCEP), satellite data produced by the National Environmental Satellite, Data, and Information Service (NESDIS), and a composite (mosaic) radar product being prepared for AWIPS SBN distribution. The NCEP gridded products are normally sent in Simple GRIB (see Section 2A); they could be reduced in size to about 40 to 45 percent of their original size, based on this sample. As shown by Glahn (1997), using TDLPACK would increase packing time of 40-km eta grids by 0.041 seconds per product, but with an <u>improvement</u>[10] in unpacking time by 0.004 sec per product. The corresponding times for 20-km eta grids were 0.177 sec per product for packing and 0.016 <u>improvement</u> for unpacking. Satellite data, both 1-km visible and 4-km water vapor, originally at 8 bits per point, could be reduced to 60 percent of their original size. The radar mosaic could be reduced to 15 percent of a Simple GRIB product.

[10]The improvement comes about because there are missing data at some gridpoints (the eta grid does not cover the complete AWIPS grid on which the data are being sent), and this necessitates a bit map. A bit map is really another grid of the same size as the one on which the data are being sent of 1 bit per point; this bit map must also be unpacked.

Table 8. Previous results by Glahn (1997; 1998) for eta model gridpoint data, satellite data, and composite radar data. In all cases the output was in TDLPACK.

Product	Original Size (byte)	Output Size (byte)	Ratio	Packing Time (sec)	Unpacking Time (sec)	Refer
40-km eta[1]	29,328	12,491	0.43	0.087	0.020	1997
20-km eta[2]	116,045	48,621	0.42	0.363	0.084	1997
Satellite Water Vap[3]	243,561	144,980	0.60	0.830	0.190	1997
Satellite Visible[4]	5,120	3,033	0.59	0.012	0.003	1997
Radar Mosaic[5]	53,207	8,190	0.15	0.510	0.048	1998

[1]An average of 1908 fields, which included missing points, initially packed in Simple GRIB (see Section 2A).

[2]An average of 372 fields, which included missing points, initially packed in Simple GRIB.

[3]Original product was 1 byte per point; values up to 256 were accommodated.

[4]Visible images are sent "horizontal" line by line as the data are received. There are 5,120 such lines, each 5,120 bytes (values) in length. They need to be sent by line, and therefore were packed in that manner. The total volume of data, however, is over 26 megabytes, one byte per value, and about 15 megabytes packed in TDLPACK.

[5]This is a non-existing product, but one that is planned. It is assumed reasonable options are Simple GRIB (which is currently being used for NCEP products) and TDLPACK. The best for Simple GRIB [Table 1 "missing" in Glahn (1998)] and for TDLPACK [Table 1, "no missing" in Glahn (1998)] were used.

5. CONCLUSIONS

It should be of considerable benefit to the National Weather Service (NWS), and to other users of meteorological data, if all "gridded" products could be transmitted, archived, and possibly stored for real-time use in a common data format. At the present, there are at least three formats in general use: (1) Simple GRIB for model data; (2) a run-length encoded format for radar data; and (3) essentially no particular code for satellite data--just 8 bits per value.

Each compression form has been derived for its specific use, with no regard for other kinds of gridded data. GRIB Edition 1 did not consider anything except model or model-related data. It was, in effect, a "first effort" that has served a very useful purpose, but harbors several problems and inefficiencies. The run-length encoding used for radar data was derived with radar data specifically in mind, and concentrates on efficiency for 3- and 4-bit data. For instance, a string of only up to 15 identical values can be accommodated in a "run." Satellite data are sent over AWIPS in a manner married to an 8-bit byte. It is doubtful existing software could be used for transmitting 10-bit data, as some meteorologists deem important, and a first cut at doing this might be to use two bytes for the 10 bits--a great waste.

The algorithms in TDLPACK--the all-important one for determining groups, and for actual compression--are targeted for GRIB Edition 2 under a name something like "second-order packing." Header information has been formulated for single-product and mosaiced radar data, as well as for model data. Satellite data will fit into the same framework.

Other compression schemes could be used. Many are geared to character-based values, and are not good for model data. They would be appropriate for data, such as radar and satellite, where the absolute range of values could be specified, especially if they are 8-bit or less. However, many times commercial packages require license fees and are not maintained indefinitely. Source code in a perpetual language (e.g., Fortran) for software to be used for communication, archival, and storage purposes must be available to the NWS so that changes can be made which invariably are needed.

It is our intent to provide documented Application Programming Interfaces (APIs) for both packing and unpacking for this new GRIB Edition 2 option. This software will balance computer efficiency and maintainability. It is hoped that GRIB Edition 2 and the simultaneous movement to provide software to implement that code form will draw users to this common format.

Computer efficiency is always a factor. TDLPACK as well as its progenitor Complex GRIB Edition 1, must determine "groups" of similar values; this is relatively expensive--somewhat like determining "runs" for run-length encoding. Unpacking may require slightly more time for TDLPACK than Simple GRIB, but only when there are no missing values; in that case, TDLPACK requires less time for unpacking.

ACKNOWLEDGEMENTS

Thanks go to Ronald Reap for reviewing a draft of this office note and especially to David Kitzmiller for review, providing the run-length decoder, the radar data, and the graphics.

REFERENCES

Dey, C. H., 1996: 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 Centers for Environmental Prediction Office Note 388, National Weather Service, NOAA, U.S. Department of Commerce, 5 sections and an appendix.

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.

_____, 1995: GRIB product sizes for AWIPS. TDL Office Note 95-3, National Weather Service, NOAA, U.S. Department of Commerce, 29 pp.

_____, 1997: Satellite, gridpoint, and vector data packing. TDL Office Note 97-1, National Weather Service, NOAA, U.S. Department of Commerce, 15 pp.

_____, 1998: Packing of Radar Mosaics. TDL Office Note 98-1, National Weather Service, NOAA, U.S. Department of Commerce, 8 pp.

OFCM, 1990: Standard Formats for Weather Data Exchange among Automated Weather Information Systems FCM-S2-1990. Office of the Federal Coordinator, NOAA, U.S. Department of Commerce, 10 chapters and appendices.

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

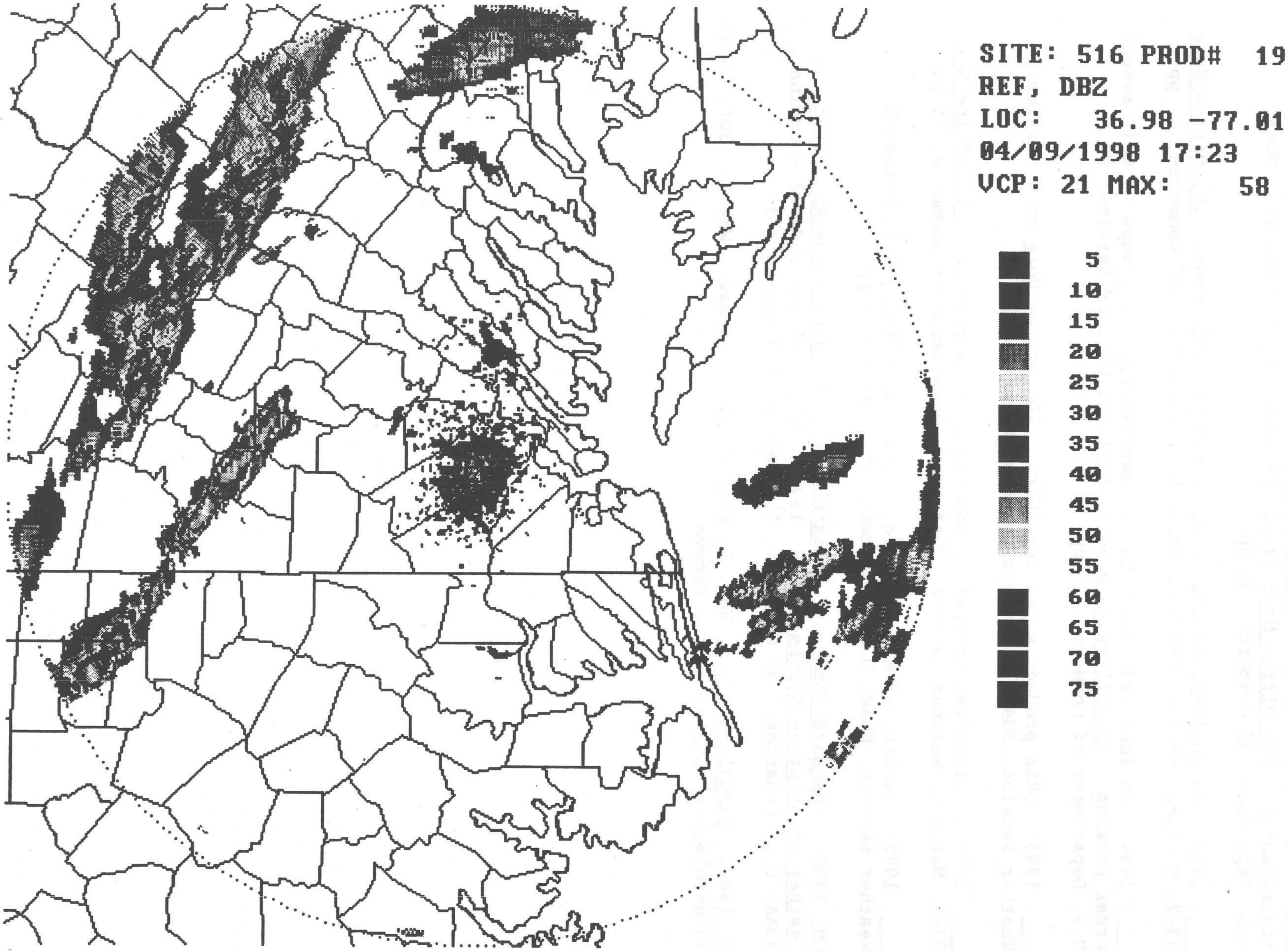

Figure 1. Base reflectivity for Wakefield, Virginia, for April 9, 1998, at 1723 UTC.

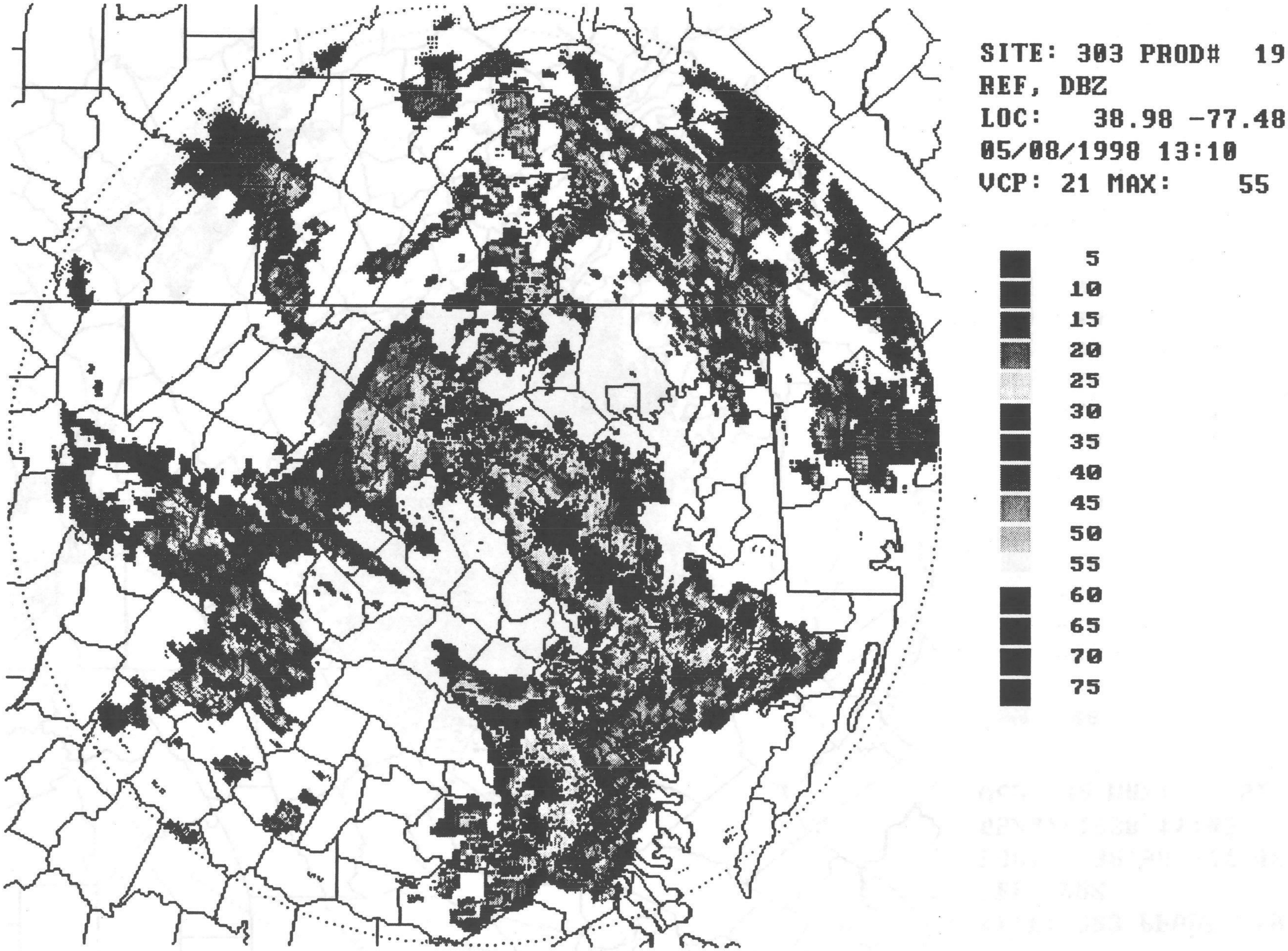

Figure 2. Base reflectivity for Sterling, Virginia, for May 8, 1998, at 1310 UTC.

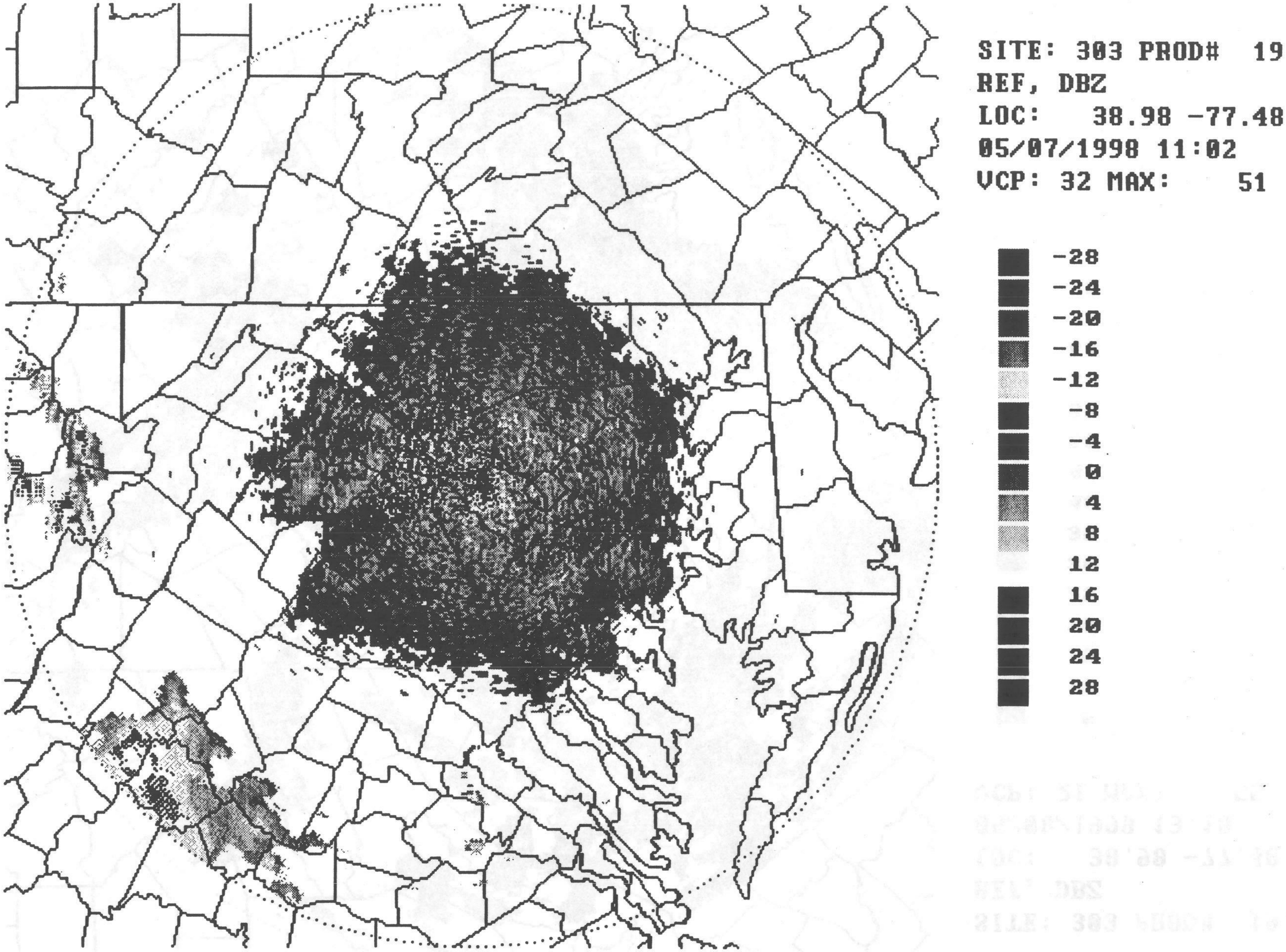

Figure 3. Base reflectivity for Sterling, Virginia, for May 7, 1998, at 1102 UTC.

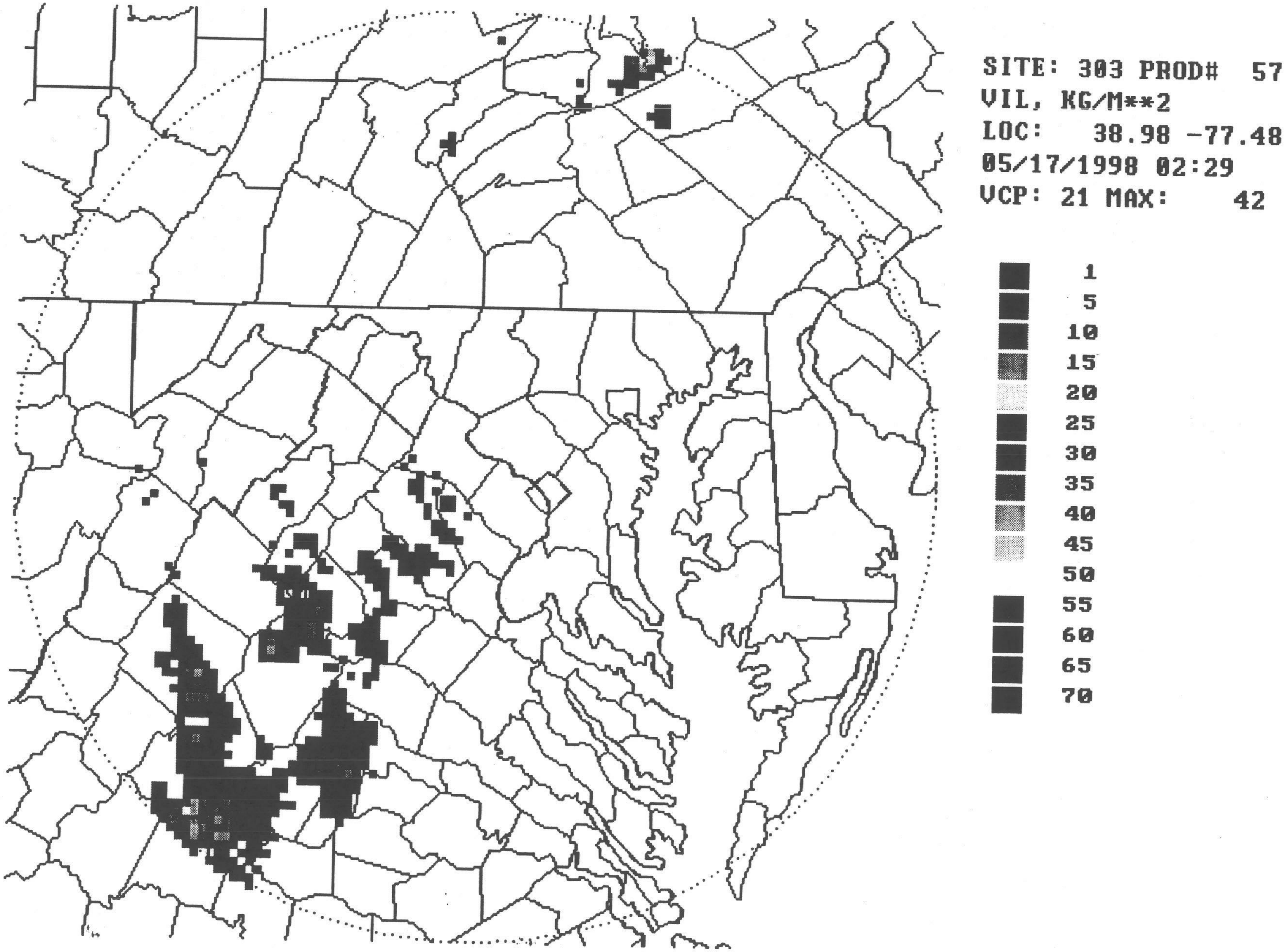

Figure 4. VIL for Sterling, Virginia, for May 17, 1998, at 0229 UTC.