NOTES AND CORRESPONDENCE

Quality Control and Archiving of Digital Data Gathered by the Bureau of Reclamation's Weather Radar

GERARD E. KLAZURA[†] AND RA AMAN[**]

Bureau of Reclamation, Denver, Colorado

5 December 1988, and 13 July 1989

ABSTRACT

This paper describes the logic and procedures used to process high volumes of digital radar data recorded on magnetic tape to a compressed, quality-checked archival format. The basic philosophy in the processing is to retain all echo data that exceed the background noise threshold, and to convert fixed-length raw data records to a space-saving variable-length format, having about 70 data quality conditions checked and flagged in binary masks. Edit flags provide information to analysts regarding errors or potential errors within the header fields (e.g. azimuth, elevation, date, time, range delay, range interval, and samples per averaged return) and data fields (e.g. suspicious reflectivity gradients, anomalously high reflectivities, and reflectivities significantly below the noise threshold).

1. Introduction

This paper describes the logic and procedures used to process high volumes of digital radar data recorded on magnetic tape to a compressed, quality-checked archival format. Computer programs to perform these functions have been developed and tested. The basic philosophy in the processing is to retain all echo data that exceed the background noise threshold, and to convert fixed-length raw data records to a space-saving variable-length format, having about 70 data quality conditions checked and flagged in binary masks. Edit flags provide information to analysts regarding errors or potential errors within the header fields (e.g., azimuth, elevation, date, time, range delay, range interval, and samples per averaged return) and data fields (e.g., suspicious reflectivity gradients, anomalously high reflectivities, and reflectivities significantly below the noise threshold).

The radar is an incoherent, reflectivity-only system with a 5-cm wavelength and a 1° beamwidth. The radar characteristics and earlier processing procedures are described in Schroeder and Klazura (1978). This paper updates the changes in processing that have occurred since then.

[†] Present Affiliation: NEXRAD Operational Support Facility
[**] Contractor computer programmer.

Corresponding author address: Gerard Klazura, NEXRAD Operational Support Facility, W/JSx1, 1200 Westheimer Drive, Norman, OK 73069.

2. Data acquisition and flow

a. Data acquisition

Data are generally recorded in a volume scan mode. A typical volume scan is a sequence of 360° azimuthal sweeps starting at 1° in elevation and ending at 12° to 20° in elevation, with elevation steps of 1°. The time interval for a complete volume scan is about 2.5 to 5 minutes depending on the total number of 360° azimuthal sweeps and the rotational speed of the antenna. The antenna azimuthal sweep rate is controlled by the pulse repetition frequency (PRF) setting, the number of pulse samples per averaged return (SA), and the azimuth increment (AI). The PRF is typically 414 s^{-1}; AI is 1.0°; and the SA is 16 or 8, which translates to a sweep duration (one antenna rotation) of 19 s and 9.5 s, respectively. Echo power returned from targets is digitized and averaged with other returned signals at the same space interval in a digital video integrator and processor (DVIP). A DVIP number (0 to 255), which is the returned power in counts, is assigned to each of 250 range-bins, with the largest numbers representing the strongest returned signals. Actually, the lowest DVIP values assigned, which represent returned signals not detectable in the receiver (i.e., below threshold), are not equal to zero, but rather are usually set in the range of 20 to 45.

A "blue sky" elimination feature is included in the radar's microprocessor/hardware system so that only records containing at least one radial with at least one of the 250 range bin DVIP values above a preset threshold level are recorded. The first and last records of a constant elevation sweep are always recorded to maintain an accurate antenna sweep history.

b. Description of magnetic tape record

Each magnetic-tape record (Fig. 1) consists of 22 frames of header ("housekeeping") information (date, time, record count, and parameter and status settings) and four sets of the following: (i) Four frames of azimuth and elevation data; (ii) 250 frames of averaged-echo range-bin data; (iii) three frames of project aircraft positions; and (iv) three frames of averaged-echo range bin data. Thus each record is a fixed length of 1062 frames where each frame is an 8-binary-bit byte. The data are recorded onto reels of tape of density 1600 characters per inch (cpi) and a length of approximately 730 m. Each record is nearly 1.7 cm long with inter-record gaps of approximately 1.6 cm separating each record. A full tape can hold approximately 22 000 records, representing from 1 to 6 hours of data acquisition time, depending on percent of surrounding echo coverage and data acquisition-rate settings.

Housekeeping information includes a range delay setting (frame 7 in Fig. 1) which is selectable from 00 to 99 km (two Binary Coded Decimal digits) and is used to specify the range of the first data bin. The blue sky setting (frame 9), which is also selectable from 00 to 99 DVIP units, is set by the radar operator following a perusal of the background noise information for data collected when the antenna is pointed to yield minimum-return echo signals. For instance, the operator may see that the background noise values are generally 34 to 40 DVIP units and might set a 43 DVIP unit value for the blue sky setting. A change of 3 DVIP units is equivalent to a change in power returned of 1 dBm. The ID Code (frame 12) is selectable from 00 to 99 and is a number used to identify a particular project site. The first bit in frame 8 and all eight bits in frame 11 define important status conditions. The various bits configurations are defined in the six "Status Format" tables in Fig. 1.

The digitized power-returned data are usually recorded in the logarithmic receiver mode. A typical bit configuration for status frame 11 is 11000010 (hexadecimal C2). This indicates PPI mode, which is a normal 360° azimuthal sweep (1), an azimuth increment of 1° (1), a PRF of 414 Hz (0), 8 pulse samples per averaged return (00), and a range-bin interval of 0.5 km (010).

The spatial bounds of each range bin are defined by the range delay, the range interval, and the azimuth increment. For instance, if these three parameters are set to 30 km, 0.5 km, and 1°, respectively, then the first range bin (frames 27, 287, 547, 807) is 30.0 to 30.5 km by 1° in width. The second range bin (frames 28, 288, 548, 808) is 30.5 to 31.0 km by 1° in width, etc. The DVIP values assigned to each range bin contain the binary equivalent of a number up to 255, which is directly related to the average power received for the corresponding range and azimuth intervals.

SWR - 86 RADAR RAW DATA TAPE FORMAT

LOGICAL TRACK NUMBER

0	1	2	3	4	5	6	7	FRAME (BYTE) COUNT	FIELD
0	0	200	100	80	40	20	10	1	DAY OF YEAR
8	4	2	1	Z	E	R	O	2	DAY OF YEAR
0	0	20	10	8	4	2	1	3	HOURS (C.U.T)
0	40	20	10	8	4	2	1	4	MINUTES
* RES	40	20	10	8	4	2	1	5	SECONDS
80	40	20	10	8	4	2	1	6	YEAR
80	40	20	10	8	4	2	1	7	RANGE DELAY
L/L		Z	E	R	O			8	LIN/LOG RECVR
80	40	20	10	8	4	2	1	9	BLUE SKY
		Z	E	R	O			10	0
RHI	AI	PRF	SA1	SA0	RI2	RI1	RI0	11	STATUS
80	40	20	10	8	4	2	1	12	ID CODE
Z	E	R	O	80K	40K	20K	10K	13	RECORD COUNT
8K	4K	2K	1K	800	400	200	100	14	RECORD COUNT
80	40	20	10	8	4	2	1	15	RECORD COUNT
ZERO								16	
ZERO								—	BYTES (16 -
ZERO								21	22) - '0'
ZERO								22	
0	0	200	100	80	40	20	10	23	AZIMUTH
8	4	2	1	.8	.4	.2	.1	24	AZIMUTH
0	0	0	100	80	40	20	10	25	ELEVATION
8	4	2	1	.8	.4	.2	.1	26	ELEVATION
128	64	32	16	8	4	2	1	27	DVIP, BIN 1
BYTES 27 - 276 RADAR ECHO									(DVIP'S,
								•	BINS 2-250)
128	64	32	16	8	4	2	1	277	AIRCRAFT #1
128	64	32	16	8	4	2	1	278	AIRCRAFT #2
128	64	32	16	8	4	2	1	279	AIRCRAFT #3
								280	(DVIPS,BINS
								281	254-256
								282	UNUSED)
0	0	200	100	80	40	20	10	283	AZIMUTH
8	4	2	1	.8	.4	.2	.1	284	AZIMUTH
0	0	0	100	80	40	20	10	285	ELEVATION
8	4	2	1	.8	.4	.2	.1	286	ELEVATION
128	64	32	16	8	4	2	1	287	DVIP, BIN 1
BYTES 287-536 RADAR ECHO								288	(DVIP'S,
								•	BINS 2-250)
128	64	32	16	8	4	2	1	537	AIRCRAFT #1
128	64	32	16	8	4	2	1	538	AIRCRAFT #2
128	64	32	16	8	4	2	1	539	AIRCRAFT #3
								540	(DVIPS,BINS
								541	254-256
								542	UNUSED)
0	0	200	100	80	40	20	10	543	AZIMUTH
8	4	2	1	.8	.4	.2	.1	544	AZIMUTH
0	0	0	100	80	40	20	10	545	ELEVATION
8	4	2	1	.8	.4	.2	.1	546	ELEVATION
128	64	32	16	8	4	2	1	547	DVIP, BIN 1
BYTES 547 - 796 RADAR ECHO								548	(DVIP'S,
								•	BINS 2-250)
128	64	32	10	8	4	2	1	797	AIRCRAFT #1
128	64	32	10	8	4	2	1	798	AIRCRAFT #2
128	64	32	10	8	4	2	1	799	AIRCRAFT #3
								800	(DVIPS,BINS
								801	254-256
								802	UNUSED)
0	0	200	100	80	40	20	10	803	AZIMUTH
8	4	2	1	.8	.4	.2	1	804	AZIMUTH
0	0	0	100	80	40	20	10	805	ELEVATION
8	4	2	1	.8	.4	2	1	806	ELEVATION
128	64	32	16	8	4	2	1	807	DVIP, BIN 1
BYTES 807-1056 RADAR ECHO								808	(DVIP'S,
								•	BINS 2-250)
128	64	32	16	8	4	2	1	1057	AIRCRAFT #1
128	64	32	16	8	4	2	1	1058	AIRCRAFT #2
128	64	32	16	8	4	2	1	1059	AIRCRAFT #3
								1060	(DVIP'S,BINS
								1061	254-256
								1062	UNUSED)
END OF RECORD GAP									

STATUS FORMAT

L/L	RECEIVER
0	LOG
1	LINEAR

RHI	
0	ON (RHI)
1	OFF (PPI)

AI	AZIMUTH INCREMENT
0	0.5
1	1.0 DEG.

PRF	PULSE REPETITION FREQUENCY
0	414
1	207 HZ.

SA	SAMPLES/ AVERAGE
0 0	8
0 1	16
1 0	32
1 1	64

RI	RANGE INTERVAL DISTANCE
0 0 0	.125
0 0 1	.25
0 1 0	.5
0 1 1	1
1 0 0	2 km

* - NOTE, RES = RESERVED EXCLUSIVELY FOR USE BY ARCHIVING PROGRAM.

FIG. 1. Magnetic tape format of raw data records on SWR-86 radar.

Bin positions 251 to 253 are used to indicate the bin positions of up to three project aircraft. A zero value in these positions indicates that no aircraft was detected. The presence of a number 1 to 250 indicates the range at which a project aircraft was located. For instance, if bin position 252 (frames 278, 538, 798, and/or 1058) contains the number 37, this means the project aircraft designated as number 2 has been detected by a separate aircraft detection antenna to be located at bin position 37. Range-bin 37 would still contain a DVIP number which indicates an echo signal returned. The reflectivity value may be higher due to the existence of the aircraft at that location.

3. Computer processing software

Most of the computer routines that make up the processing software are written in FORTRAN 77. Briefly, the processing procedure consists of 1) copying each raw record to disk, except those with bad record lengths (not 1062 bytes); 2) deleting nonecho (noise) data bins from this (raw-record) disk copy; 3) performing numerous quality and accuracy tests on the housekeeping information and a few tests on the data bins; 4) optimally compressing the raw data files; 5) comparing the compressed data files with the raw data files; and 6) producing a "first look" analysis product. Reports of these processing steps are printed. These include a listing of each error condition flagged (header edit report), a noise analysis report, and a comprehensive antenna-sweep summary report. The latter two reports along with the first-look analysis products (maximum reflectivity and echo top PPI displays) are also copied to microfiche.

a. Raw data edit

Original data can contain an assortment of errors. For example, date and/or time may be set or incremented incorrectly; azimuth, elevation, and record count may be erroneous; or preset variables such as pulse repetition frequency and azimuth increment may have been reset by accident. When questionable conditions or errors are detected, they are flagged and noted in two locations. First, they are displayed in hard copy form in a header (housekeeping) edit report. Second, the flags are recorded into the permanent compressed archival file giving analysts the ability to filter, or isolate for special treatment, radials with any combination of flagged problems.

A condensed summary of the tests performed on the data is listed in Table 1. The error or alert conditions tested are illustrated in seven broad categories to give the reader a general idea of the approach used. Most of the conditions tested in "Change Alert" are to alert the analyst to parameter changes that were expected to be stable. For instance, range delay would be expected to stay constant through the day.

Other flags not specifically shown in Table 1 also occur. For instance, if a parity error occurs while the raw tape is being read, a flag is appropriately inserted. A flag is inserted when a beginning of an azimuthal sweep is detected, or if a value is corrected. The only items that are allowed to be corrected in going from

TABLE 1. Summary of data quality and consistency tests performed. Digits 1–3 indicate the number of distinct conditions of a given type.

	Change alert	Possibly inconsistent	Improbable change	Discontinuity spike	Unlikely value	Illegal or too big/small	Illegal BCD digit
I.D. code	1						1
Year	1				1		1
Day			2			1	1
Hour		1	1			1	1
Minute			1			1	1
Second			2			1	1
Azimuth	2	3	1	1		1	1
Azimuth increment	1				1		
Elevation	1	2		1		2	1
Elevation increment	1	1					
Range increment	1				1	1	
Range delay	1						1
DVIP values	1			1		2	
Blue sky DVIP	1					1	1
LIN/LOG receiver	1						
Pulse repetition freq.	1				1		
Samples averaged	1				1		
Raw record count			1				1
Range/height mode	1						
Required 0, S not all 0					3		
Parity error count	1						1

the raw data to the archival file are the date and ID code, which are changed to the correct value if they were improperly set originally.

The great majority (approximately 95 percent) of each raw record is comprised of the data bin information. Frequently, only a few of the data bins actually contain detectable echo data information. A very important part of the processing of the data from raw records to archival file records is the filtering of DVIP data below the noise threshold while retaining valid echo data. The DVIP noise threshold information is automatically computed and frequently recomputed through a statistical analysis of the DVIP values in the range bins. The approach used is to first find 101 consecutive range bins in a relatively echo-free region, then find a DVIP modal value that repeats at least 30 times. It is highly improbable that a signal (echo) region could be involved in the computation of the modal value. At the same time, a nonspurious minimum DVIP value is found. The noise threshold DVIP value is computed as follows:

$$\text{NOISE} = \text{MODE} + (\text{MODE} - \text{MINIMUM}) + 3.$$

For example, assume the 101 consecutive range bins for a given azimuth had the following distribution of DVIP values:

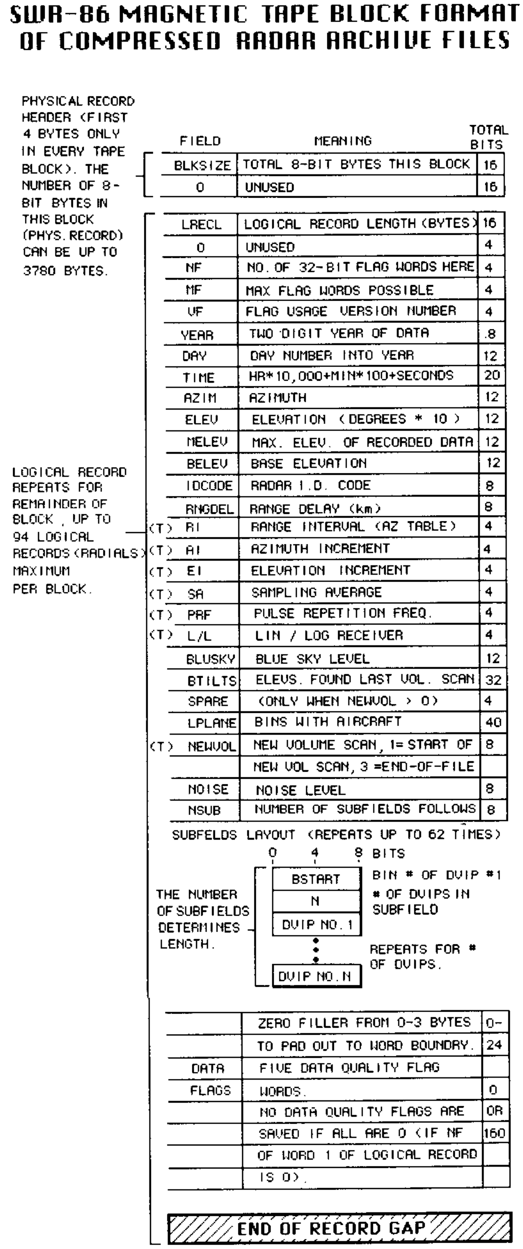
19(1), 24(1), 25(18), 26(31), 27(19), 28(7), 30(4), 31(1), 32(3), 34(1), 35(3), 36(2), 38(2), 44(1), 51(1), 52(1), 54(3), 55(2), 56(1).

The minimum DVIP value of 19 is rejected as spurious since there is a gap of at least one missing DVIP value before the next one present. The minimum used is 24. Mode 26 is accepted for noise level calculation since it occurred >29 times. The noise value becomes $31 = 26 + (26 - 24) + 3$. DVIP values exceeding 31 would be considered as valid echo data in this example.

Addition of the three DVIP units accounts for small noise fluctuations of the radar system and enhances the chances that system noise is not apt to be identified as returned echo data. The automated DVIP noise value computation provides a check on the accuracy of the "blue sky" threshold set by the radar operator, serves as a useful comparison for noise value thresholds computed during calibrations, and verifies that the processor operation is stable.

b. Compressed archival file

The compressed file created during the processing primarily contains data for bins that exceed the computed noise level. Up to two noise bins are allowed to be enclosed between non-noise bins. This results in a significant size reduction which is most convenient and cost effective for long-term storage. In fact, when archived on 1600-cpi magnetic tape, the compressed data have an average space reduction of 3.3:1 (winter storm



(T) TRANSLATION TABLE

	RI	AI	EI	* +8	PRF	SA	L/L	NEWVOL
0	0 125	.5	0.25	-.25	~414	8	LOG	NO
1	0 25	1.0	0 5	-.5	~207	16	LIN'R	YES
2	1.00		0.75	-.75		32		
3	0.50		1.0	-1.0		64		E-O-DAT
4	2 0		1.5	-1.5				
5			2.0	-2.0				
6			2.5	-2.5				
7			3.0	?				

* EI IS COMPUTED BY COMPARING CURRENT ELEVATION WITH NEXT EXPECTED STEP. IF THE NEXT STEP IS A DROP, THE 8'S (LEFT) BIT IS SET TO INDICATE NEGATIVE.

FIG. 2. Magnetic tape block format of compressed archival file.

data) to 5.3:1 (summer convective storm data) when compared with the raw data.

The contents of the compressed archival file are listed in Fig. 2. The frames are not all 8-bit bytes as they were in Fig. 1. Whereas the raw radar record has one header (housekeeping) field and four radials of range-bin data, the basic archival file logical record structure has a header for each radial, and each radial primarily has range bins with above-noise threshold DVIP values. The logical records are concatenated to create a large blocked physical record. The error flag information (if at least one condition has been flagged) is appended to each logical record as a 160-bit flag array. Each bit set in the flag array has a specific meaning that pertains to the radial (logical record) of which it is a part.

All of the header information from the raw records except RHI/PPI (which indicates whether antenna scan was at constant elevation or azimuth angle) and record count is carried over to the archival records. In addition, certain parameters computed during the processing become part of the archival records. Examples are elevation increment, beginning of new volume scans, DVIP noise threshold, and a bit map that describes the elevation steps found in each volume scan.

Each subfield contains one or more (consecutive) range bins with above-noise threshold DVIP values, and, at most, pairs of noise bins enclosed between non-noise signals. If a radial (250 bins) has an echo signal in at least every third range bin, there would be one subfield for that logical record. If range bins 10, 12, 13, 16, 19, 30 to 35, and 164 of a particular radial have echo signals, there would be three subfields in that logical record: 10 bins starting at bin 10; 6 bins starting at 30; and 1 bin starting at 164.

DVIP values for bins 254 to 256 in the three range bins following aircraft location information in each of four radials in the raw records (see Fig. 1) are not processed over to the archival file.

The archival files are used by scientists and engineers for meteorological data analyses. These analysts generally convert the DVIP units to meteorological reflectivity units using a modified version of the Probert-Jones (1962) radar equation. Schroeder and Klazura (1978) discuss pertinent calibration and transfer function information that is relevant to this conversion.

4. Concluding remarks

This paper describes a digital radar data-processing scheme that has evolved during a 12-year period from a time-consuming computer/human procedure to a fully automated system. The edit checking flags, which until recently could only be scanned visually, have now been incorporated into the data records. This allows analysts to use marginal data records much less tediously than in the past by allowing them to automate special treatment of selected problems as opposed to conducting visual searches for problems from microfiche reports.

This automated processing system, which includes flagging into the archival format, was used for data collected during a full season. Raw radar tapes for any given day were processed to a final archival version within a couple of days. Previously, human inspection and decisions to correct or discard records slowed this process immensely.

Before implementing a scheme such as this, data managers need to consider carefully the risks along with the advantages. One obvious risk that needs to be considered, especially when raw data are not kept permanently, is the possibility of irretrievable data loss due to software errors in processing raw data to a new format. It is safer to retain the raw data as the archival form. However, the advantages of significantly reduced data volumes and data that are quality checked are great motivators to assume some of the risks. In addition, risks are significantly reduced by employing a processing step that compares information contained within the archival file with original information in the raw file. (Note: This is performed on each radar tape.) As breakthroughs in data storage technology continue to occur, it may become feasible not to destroy raw data. Even now, the optical disk and related WORM (Write Once Read Many) technologies may make this option viable for some. However, the raw data should be considered only a safety backup source.

The archival format has other advantages over the raw format in addition to saving space. Analysts are relieved of the complexity of comparing and detecting suspicious values, determining data noise thresholds, or deciding when a new sweep or volume scan starts.

Acknowledgments. The authors would like to thank Ron Miller for his assistance in computer-generating Figs. 1 and 2.

REFERENCES

Probert-Jones, J. R., 1962: The radar equation in meteorology. *Quart. J. Roy. Meteor. Soc.*, **88**, 485–495.

Schroeder, M. J., and G. E. Klazura, 1978: Computer processing of digital radar data gathered during HIPLEX. *J. Appl. Meteor.*, **17**, 498–507.