CHAPTER III

Digital GEOS Data from Garni, Armenia
as Archived on Optical Disk

G. Glassmoyer

(U. S. Geological Survey)

PROCESSING

As of this writing 245 GEOS tape cartridges from the Garni Array, representing the time period 26 June 1990 through 10 August 1992, have been played back at the National Strong Motion Data Center.

The playback procedure involves demultiplexing of the data on tape cartridge using the computer programs RDGEOS on a DEC VAX computer or PCGEOS on an IBM-compatible PC. These programs write separate time-series files to disk for each channel recorded on GEOS, plus a time-stamp file that contains the start times of each (0.42 second) block of data recorded on the GEOS.

Format

The format used for the digital data is a block-binary file consisting of two header blocks (512 bytes each), one 2-byte integer, the other 4-byte real, followed by the data as 2-byte integers. Integer and real values are stored in the format used by DEC PDP and VAX computers. 2-byte integer values are stored as low-byte-first, high-byte-second signed integers in two's-complement form. 4-byte real values are stored as high-word-first, low-word-second. The integers are in the same format as used by PCs, however the real values are not the same format as PC real values. Programs RECSEC or VECTOR, written by E. Cranswick and R. Banfill, can be used to read the data on a PC.

Filenames on the VAX are constructed from the station name and the start-of-record time. Characters 1-3 are the day-of-year (001-366), characters 4-5 the hour (00-23), characters 6-7 the minute (00-59), and character 8 represents the second (A-T, where A=0-2, B=3-5, ..., T=57-59). Character 9 is the component (4 is vertical, 5 and 6 are horizontal). The suffix is the station name. On the PC and the optical disk one less character is available for the filename so the hour is represented by a letter from A to X.

Fixes

Normally the GEOS can be relied upon to write to tape cartridge the complete recording sufficiently well to be read and accurately interpreted by the demultiplexing programs. It has been a recognized phenomenon that the GEOS instruments may occasionally fail in part or whole to write an accurate tape. These types of failures occurred in several forms in this study. In the case of data that can still be read by the programs, these failures can be grouped into two types: bit errors and block errors.

An exact correction can usually be made for bit errors. These are what are often referred to as "glitches." The symptom is an occasional data value that is a large power of two different from the preceding and subsequent data values. The cause typically is memory board failures, but can also be in the analog-to-digital circuit. The computer program BITFIX is used to find and correct bit errors whenever the bit errors are sufficiently large compared to the true rate of change between data values.

Block errors usually cannot be fully corrected. These are instances where the data block to be written to GEOS tape cartridge either is not written, or that which is written is corrupt. During playback the demultiplexing program (RDGEOS or PCGEOS) attempts to identify missing data blocks by examining the data-block time stamps. When these are identified, the null data value (integer header offset 3 of Table 3-1a) is written to the data file in place of the missing data. The G5 recordings have several instances of these data gaps.

If the first data block of an event is corrupt, then the files created for that event will initially be given a name that does not indicate the true time of that event. Such files may also contain initial meaningless data. Such files are easily recognized by the program CHKTIME which compares the start-of-block time stamps, as written to the time-stamp files, to their expected time values. Start-of-record times have been corrected for these files using the program CHGTIME.

In addition to allowing files to be given the correct start times, the program CHKTIME, reading the time-stamp files, also verifies the integrity of timing within the recordings. By this means several recordings have been found to have faulty internal timing. The cause of these failings is not always certain, but often is the result of reading a corrupt data block in the midst of an event. Recordings that contain internal timing errors have been corrected to the degree possible by using null values to mark those samples where no data are available, or by removing extra null values that the demultiplexing programs may have mistakenly added. These corrections are made using the editing capability of the programs RAWRI2 and RAWWI2.

Recording parameters are stored in the two header blocks described in Table 3-1. The most critical values are shown in boldface. Many of these such as amplifier gain (see log-files in Figure 3-1) are derived from the GEOS cartridge. Other parameters such as station location or sensor orientation are added from the parameter list file GARNI.AHC. While usually accurate, there are instances where the polarity may be opposite from that which is indicated for vertical components. A careful evaluation of each clock correction in the manner described in the log-file section later should be made if these are important to a study.

PLOTS OF SELECTED EVENTS

Plots are included for those seismic events that were recorded at five or more three-component stations. These events are listed in Table 3-2, where the second-code of the filename is used to indicate those stations that recorded an event. The plots have been prepared using E. Cranswick's RECSEC program. They show each trace at equal amplitude, bundled by three-component station, for a duration proportional to the number of stations recording the event (5 seconds per station). The time (nominally UTC) of the first sample in each plot is indicated in large numbers on the right-hand side of the plot page. This is given as year (less 1900), day-of-year, hour, minute, second, and millisecond. The time scale at the top of the plot page is in seconds. Indicators on the right-hand side identify the individual trace. For example, G5B.V2 is station G5B's 2nd velocity channel - the first channel is up, the second is nominally north, and the third is nominally east. See Table 2-1 for actual horizontal orientations indicated for components 5 (=.V2) and 6 (=.V3). Numbers on the left-hand side of each trace indicate the maximum absolute value in digital counts through the end of the plotted portion of each trace (32767 is full-scale).

ORGANIZATION OF OPTICAL DISK

Two IBM-3363 write-once-read-many (WORM) optical disks have been prepared containing calibration records and seismic triggers as listed in Table 3-3. One disk contains the 1990 data and the other contains the data from 1991 through 10 August 1992. Data are organized into subdirectories by day-of-year of event. These subdirectories are within parent directories by year. Log-files created by the demultiplexing programs from each tape are included in the LOG directory on the 1991-1992 disk. Also included are ORDARR files for use by the Cranswick and Banfill (WFPS) processing programs for each year. The WHERE list of Table 3-3 is also provided as text files for each year. Seismic triggers were selected by visual inspection and may still include such events as aircraft triggers or other signals that appear as possibly of seismic origin.

LOG-FILES

Four examples of log-files are provided in Figures 3-1a through 3-1d. These show most of the useful types of information that can be read from a log-file.

LOG28.G3 (Figure 3-1a) is a typical log-file from the program RDGEOS. The first line indicates the date and time the program was run and the program name and version. The log-file number (28) is added after the RDGEOS program has completed its operation, so it does not appear on this line which shows simply LOG.G3. Following is the channel/component naming description. The default refers to the log file itself as well as to the time-stamp files. Channel 1 data is component 4 of data for G3A, *etc*. The next lines indicate the tape/experiment number, location number, GEOS unit serial number, and the year. If any of these parameters change during the tape, such as the year, a new indication is made.

In the next line file numbers include all files be they data or clock files as counted by the demultiplexing program (RDGEOS). The event numbers (Evt) are data files as counted by the GEOS. The type can be a trigger, continuous record, manual record, sensor calibration, amplifier calibration, or a clock file. Clock files can be master-clock synchronization (synch), manual synch, WWVB synch, WWVB synch failures (VB Synch TMO), or WWVB corrections. The time indicated for data files is of the first sample, as measured by the internal GEOS clock and is shown as day-of-year, hour, minute, second, and millisecond. The time standard can be none, manual, WWVB, or external. For the Garni data the time standard shown in this column is probably not reliable. An indication of none means that the clock has not been set and shows the elapsed time since the GEOS was turned on (starting day 001). Manual (MAN.) would indicate an operator set the clock by hand, such as by reading the time off of a watch. WWVB would indicate the clock was set by interpreting a WWVB time code. External (EXT.) is the only time standard that properly should have been used in this data set, and indicates that the internal GEOS clock was set by an external clock. The clock correction (Corr) as it is stored in the data files is shown in seconds that should be subtracted from the internal GEOS time to get the corrected time. See the later clock-correction example for a more complete discussion. DR100 name is a generic representation of the the names of the data files as they would be stored on a VAX. Duration is an estimate of the length of the recording in minutes and seconds. Srate is the sampling rate in samples per second. Pre-event is the length of time in seconds of GEOS blocks that were filled before the active block of the tape turn-on decision. Volts is the recorder battery voltage at the time of a trigger.

The last four columns refer to parameters of the trigger algorithm. Trg ch is the trigger channel. STA is the short-term-average interval in seconds. LTA is the effective length of the long-term-average interval in seconds. Ratio is the threshold triggering level by which the short-term-average must exceed the long-term-average in order to trigger. Ratio is expressed in Fortran notation of 2 raised to a power. Motion type, amplifier gain, and anti-aliasing filter frequency are shown for each channel.

Examples

File 1 of G3 tape number 28 is a sensor calibration. It was recorded on GEOS unit number 5 beginning in the year 1991, day-of-year 121, hour 14, minute 47, second 27.382 according to the internal GEOS clock that was set by a master-clock. The clock correction stored for this record is 0.000 seconds. Channel 1 is ground-motion velocity recorded with 54 “decibels” (more accurately 2**(54/6) = 512 x) gain using a 7-pole, 50-hertz anti-aliasing filter, and given the filename 1211447J4.G3A (121O47J4.G3A on the PC optical disk). It was sampled at 1200 samples per second and is a little less than 37 seconds long. Battery voltage is not reported for calibrations.

File 2 is an amplifier calibration. The notation about 506 null values refers to this file; in ampliflier calibrations, however, it is due to a minor bug in the GEOS, and the data is complete without the nulls which have since been removed. In other data files each 506 null values represent a lost data block with the data value of -32768 being used in the place of the lost samples.

File 7 is a triggered event. The GEOS was set to trigger by comparing for channel 1 the short-term-average of 0.4 seconds data to 8 (= 2**3) times the long-term-average representing 20 seconds of data. The batteries providing power to the GEOS were at 26.66 volts. Optimum battery voltage is about 24 to 28 volts.

File 3 is a failure to read a WWVB time code. It should not have been set to try since no WWVB-encoded time-code was available.

File 4 is the all-important clock correction! Eight attempts are made by the GEOS to read the one-second pulse of an external signal. This need not be a WWVB signal. As measured by the internal GEOS clock these pulses were detected 1, 0, 0, 0, 1, 0, 0, and 0 milliseconds after the second so the internal and external clocks are in near-perfect agreement.

File 13 is also a clock correction. At its time the internal GEOS clock is 4 milliseconds fast relative to the external clock. For data files that follow the clock correction is recorded in the header as 0.004 seconds.

Clock-Correction Example

File 9 is a triggered recording of a magnitude 5.1 aftershock of the Ratchi, Georgia Earthquake. To get the best estimate of a clock correction for this event it is necessary to interpolate between two clock corrections. Using 0.25 milliseconds (the mean of the eight clock-correction measurements) from file 4 (121:14:56:44) and 4.00 milliseconds from file 13 (122:02:57:52) one can get by linear interpolation 3.5 milliseconds as a clock correction at 122:01:26:00. This 3.5 milliseconds would be subtracted from the 122:01:26:00.618 start-of-record time to get the external clock referenced time of 122:01:26:00.6145.

When similar calculations are made for units G1, G2 and G4 their clock corrections are found to be -0.7, -7.0, and 16.2 milliseconds, respectively, and their externally referenced start-of-record times are ...:01.8837, ...:01.0980, and ...:01.5068. All units are using the same external clock signal, a master clock, so relative times have thus been determined.

The absolute time could then also be determined since there is data in Table 3-4 indicating when the master clock was set to Omega-time. The master clock appears to have been synched to Omega sometime between 116:09:13 and 116:21:14. Later it was found to need adjusting by -5 milliseconds sometime between 125:23:47 and 126:08:05. The minimum estimate of drift for the master clock may be made by assuming the master clock was synched at 116:21:14 and was found to need an adjustment of -4.5 milliseconds at 126:08:05. The maximum estimate of drift for the master clock may be made by assuming the master clock was synched at 116:09:13 and was found to need an adjustment of -5.5 milliseconds at 125:23:47. This yields a range of -2.5 to -3.2 milliseconds, averaging to -2.8 milliseconds. This master-clock correction should be added to the previously calculated relative clock corrections to get absolute clock corrections. Thus respectively for G1, G2, G3 and G4 the best absolute clock corrections would be -0.0035, -0.0098, 0.0007 and 0.0134 seconds to be subtracted from GEOS start-of-record times. These are different from the values -0.002, -0.006, 0.000 and 0.011 seconds that are stored by default in the headers. All clock corrections stored on the optical disks are the inaccurate default values. The destinctions are small in this case but may approach one full second during intervals when clock synchs are infrequent.

Under ideal circumstances, such as in this example, relative clock corrections could at best be accurate to 1 to 2 milliseconds and absolute clock corrections to 2 to 4 milliseconds. During the Spring of 1991 clock corrections were usually well recorded on all GEOSs. At other times, however, at least one recorder may have been misprogrammed, permitting only partial recovery of relative times. Also, the recording of master-clock resynchs (Table 3-4) was inconsistant, thus limiting the opportunities for the recovery of absolute time.

Other Examples

Following file 14 is a page break.

At the end of the log-file is the line indicating the number of files processed.

LOG03.G1 (Figure 3-1b) is a short example from the PCGEOS program, version 04.07.

The first 2 files on this tape are corrupted clock files. The first 6 words (cells) stored in each block are the start time for that block. PCGEOS attempted to search ahead in the files for intact time information, but found none, so these files were skipped.

The first file from this tape was corrupted. PCGEOS used default header information from file G1.GHD in the “grand-parent” directory for the file(s) it encountered before finding the readable header information in file 4 of this tape. The default header information uses serial number 99 and voltage equal to 20.00 as flags that the header information did not come from the tape.

File 5 is a master-clock synch. At 183:08:26 the GEOS internal clock was set to match the external clock. The remaining information on this line is not reliable.

LOG22.G1 (Figure 3-1c) begins with a master-clock synch, shows a faulty serial number and indicates an operator-initiated (continuous) recording.

LOG06.G2 (Figure 3-1d) begins with clock information extracted from the header of the first trigger. PCGEOS will use this information, which is from the last clock correction on the previous tape, as a starting point for the default value of the clock correction. There is a risk that this value can be very out-of-date and inaccurate (as it may be here). The other examples did not have this either because they were from RDGEOS, had a corrupted header, or began with a clock file.

Not shown is an example which ends with Physical End of Tape indicating that the tape was read all the way through instead of just to the end-of-tape mark. This would be the case for 100% full tapes.

KNOWN EVENTS

153 of the included events have been identified in the National Earthquake Information Center (NEIC) catalog. These are listed in Table 3-5 by date, Table 3-6 by distance, and Table 3-7 by magnitude. Body-wave magnitudes are used unless otherwise indicated. A travel-time plot of the larger events is shown in Figure 3-2.