Formats of Data Files

Data files downloadable from the CESMD are compressed (zipped) files containing one or more data files in one of several standard simple text (Ascii) formats. The files are multichannel, containing all the files recorded at the station (3 for a triaxial ground station, 6 to 100 or more for geotechnical arrays and structures). The channels are generally downloadable as individual channel files for structures. The data files in general have three sections for each channel: a text header, an integer-values header, a real-values header and the data values. The files are of three types:

1) Volume 1, raw acceleration time series, in units of acceleration. No instrument correction or filtering has been applied to Volume 1 data, so it has all the noise in the original recording still present. These files are most often used by researchers who want the compete recording, signal plus noise, to do research into short or long period noise characteristics, or to do their own specialized processing.

   Accelerogram files obtained via digitization of film record will contain the digitized acceleration-time pairs, with unequal spacing. The format for the first film records digitized in the 1970s Bluebook project at Caltech is given in Hudson et al. (1969). For records subsequently digitized and processed by the California Strong Motion Program the format differs only slightly, and is given in [Shakal and Huang (1985)](#). For records processed by USGS National Strong Motion Project the [SMC format](#) is used (Converse and Brady, 1992).

   Accelerograms recorded by digital accelerographs have equal time steps, and the format is slightly simpler. For CSMIP records, the format is described in Shakal and Huang (2004), for USGS records, it is described in Converse and Brady (1992). Volume 1 files are generally named with a ".v1" suffix.

2) Volume 2, processed time series files, contain the processed acceleration, velocity and displacement time series obtained by integrating and filtering the raw acceleration. The bandwidth is generally given in the file header for contemporarily produced files. The files have a format similar to Volume 1, and the format is described in the above reports. The files are generally named with a ".v2" suffix.

3) Volume 3, or response spectrum files, contain the response spectra at a set of periods, listed in the file, that typically range from 0.04 to 15 seconds (the Caltech files had 91 periods as a standard). Several response spectra are usually given, including Sa, Sv, Sd and FS.

An effort is underway to standardize on a single [COSMOS format](#) to improve user convenience. That format preserves many of the important features of the individual formats. The process of converting software to this new standard is underway. Some tools are available for converting to the new format, available at COSMOS and the CESMD.