Technical Report

GenStats User's Guide

Marshall L. Buhl Jr.

F

GenStats User's Guide

Marshall L. Buhl Jr.

NREL

National Renewable Energy Laboratory

NOTICE

Available electronically at http://wind.nrel.gov/designcodes/genstats/genstats.pdf

Document Revision Record

Revision	Date	Description
1.00	29-Apr-2002	First release.

Introduction

GenStats is a fairly simple utility that generates statistics for one or more files. Although they may be of different lengths, all files read during a single run of GenStats must be the same format.

GenStats was designed to process wind-turbine test or simulation data, but it may be useful for most any kind of tabular data.

GenStats is the predecessor to Crunch and is a subset of that program. However, it can do one thing that Crunch cannot do—it can handle files with different lengths.

Retrieving Files from the Archive

You should download the GenStats archive from our web server page http://wind.nrel.gov/designcodes/genstats/. The file should have a name such as *GenStats_v172.exe*. Create a GenStats folder somewhere on your file system and put this file there. You can double-click on it from Windows Explorer or enter *GenStats_v172* at a command prompt with the GenStats folder as the current directory. This will create some files and folders. Please see *Installing NWTC Design Codes on PCs Running Windows NT®*[1] for details on how to set up your system to run GenStats.

Distributed Files

The files included in the archive of GenStats are as follows:

14u001.sts	The sample output file
14u001.tim	The sample input data file
ArcFiles.txt	The list of files that are written to the archive
Archive.bat	The batch file that creates the archive
ChangeLog.txt	The list of changes to the GenStats Perl script
GenStats.doc	This user's guide in Word format
GenStats.exe	The GenStats executable for Windows®
GenStats.inp	The sample input parameter file
GenStats.pdf	This user's guide in PDF format
Source*.f90	The GenStats source code

Running GenStats

Before using GenStats for production work, you should run it with the sample input file. To do so, rename the sample output file (*14u001.sts*) to anything else, double click on GenStats.exe in an Explorer win-

dow, and GenStats will produce a new *14u001.sts*. Check the new file against the old one.

If you followed the instructions found in *Installing NWTC Design Codes on PCs Running Windows NT®*,[1] you will be able to run GenStats from any folder on your system. Open a command prompt in the folder in which you wish to work.

To determine what to do, GenStats reads a parameter input file. By default, it reads a file called *GenStats.inp* if you enter only "GenStats" at the command prompt. You can override the default name by specifying it on the command line. Enter `GenStats /h` for command-line help.

Processing Input

GenStats can reorder and rename your channels. It will process only those channels and data rows you want. It can apply scales and offsets as it reads in your data. Channel names and units are each limited to 10-character strings.

You can skip as much of the beginning and/or end of the file as you wish. You specify the number of lines to skip before storing data and the total number of records to store in the data array. If you set the total number of records to 0, GenStats will store everything it finds after skipping the beginning. GenStats will allocate the storage array at run time, so there is virtually no limit on the number of channels or records. You are limited only by the available virtual memory in your computer.

After specifying the amount of data to read in, you can also tell GenStats to ignore any data in rows that occur before time reaches a certain value. These records will be deducted from the total number of records.

Statistics

The statistics calculated are:

- Minimum
- Mean
- Maximum
- Maximum range (maximum-minimum)
- Standard deviation
- Skewness
- Kurtosis
- Mean crossing frequency (how often the signal crosses its mean in a positive-going direction)

Output

GenStats can generate three different types of output. You can specify one, two, or all three to be generated during a single run.

One is a table of the statistics for all the channels in a file. These tables are written to files with the same root name as their data file but with an *.sts* extension. One *.sts* file is generated for each data file.

The second type of output is the summary file. For all input data files in a single run, GenStats generates one table of statistics for each requested channel. These tables are written to files called *<ChannelName>.sum*.

The third type of output is an echo of the processed data. It will contain only the data used for statistical analyses. If you asked GenStats to use only part of the files, you will see only that part. If you applied scales and offsets, the output will be modified accordingly. These tables are written to files with the same root name as their data file but with a *.mod* extension. One *.mod* file is generated for each data file.

You can tell GenStats to output data in fixed columns or delimited by tabs. The former is best for viewing with an editor or for printing. The latter is best for importing into spreadsheets. You can also specify how to format the numbers.

Azimuth Averages

GenStats can generate azimuth averages (AA) of selected channels. It will then subtract the AA from the original signal to create a new detrended channel. GenStats calculates the statistics for both the AA and detrended channels. The AA channels have the same channel name as their original channels but have an "aa" added to the AA column in the statistics table. The detrended channels will have a "-aa" in the AA column of that table. For the summary tables, the file names will be *<ChannelName>_aa.sum* and *<ChannelName>-aa.sum*, which represent the AA channel and the detrended channel.

Peak Finding

You can ask GenStats to find peaks and valleys in the signals and fit them with parabolas. It will then replace the peak and valley values with the maxima and minima of the parabolas. The time values for these peaks are not adjusted. The statistics calculations will use these new values. If you generate .mod files, they will contain the modified data.

Compiling GenStats

You should not need to compile GenStats unless you want to make changes to the code. The archive contains code primarily for the Compaq Visual Fortran compiler. All of the compiler-specific code should reside in files called *Sys_PCD.f90* and *Mod_PCD.f90*. All source code resides in GenStats' *Source* folder.

Running GenStats

To run GenStats, open a command prompt in the directory that contains the parameter input file. The syntax for the command is:

```
GenStats [/h] [<parameter file>]
```

where the optional switch "/h" generates syntax help and the optional "<parameter file>" tells GenStats what calculations to make. The default name of the parameter file is *GenStats.inp*.

Input-Parameter File Format

Use the sample *GenStats.inp* file as a template. Except for the list of input channels and the list of data files, no lines should be added or removed from the sample input file. For the input channels, there should be one line for each channel. For data files, list one file name per line. One blank line separates each section of the parameter-input file from other sections.

Although my sample input file is written using a sentence-like structure, you do not need to do it that way. For example, one of the first lines in my sample input file says:

```
1 header record in each file.
```

If you prefer, you can rewrite that line as:

```
1        Number of header record in each file.
```

As long as the parameter being read is the first "word" on the line, and it is separated from the comment portion of the line with any sort of white space, GenStats will not mind. The amount of spacing is not important—use whatever looks good to you. For all lines in the input file, the non-comment information must occur within the first 100 characters.

For lines listing the channels used for azimuth averaging, peak finding, and summary files, enter the list of columns with the numbers separated by any combination of white space and an optional comma.

Known Bugs

- None.

Possible Future Enhancements

- None are planned.

Caveats

The National Renewable Energy Laboratory (NREL) makes no promises about the usability or accuracy of GenStats, which is essentially a beta code. NREL does not have the resources to provide full support for this program. *You may use GenStats for evaluation purposes only*.

Acknowledgments

Marshall Buhl of the National Wind Technology Center (NWTC) wrote the majority of GenStats. Some algorithms were written or designed by Norm Weaver of InterWeaver Consulting for other programs and were used in GenStats. Paul Veers of Sandia National Laboratories (SNL) was the inspiration for the peak finding algorithm. Larry Schluter of SNL coded Veers' algorithm in LIFE2, and we started with his code. Funding for GenStats development came from the U.S. Department of Energy under contract No. DE-AC36-83CH10093 to the National Renewable Energy Laboratory (NREL). The original work was performed under task WE80.4030, which was managed by C.P. "Sandy" Butterfield of the NWTC.

Maintenance and improvements were funded by the U.S. Department of Energy under contract No. DE-AC36-98-GO10337 to NREL. This work was per-

formed under tasks WE90.1210 (managed by Alan Wright) and WER2.1250 (managed by Sandy Butterfield).

Feedback

If you have problems with GenStats, please contact Marshall Buhl. If he has time to respond to your needs, he will do so, but please do not expect an immediate response. Please send your comments or bug reports to:

Marshall L. Buhl, Jr.
NWTC/3811
National Renewable Energy Laboratory
1617 Cole Blvd.
Golden, CO 80401-3393
United States of America

Web: http://wind.nrel.gov/designcodes/
Email: marshall_buhl@nrel.gov
Voice: +1 (303) 384-6914
Fax: +1 (303) 384-6901

References

[1] Buhl Jr., M.L. *Installing NWTC Design Codes on PCs Running Windows NT®*. NREL/EL-500-29384. Available at http://wind.nrel.gov/designcodes/papers/setup.pdf. Golden, Colorado: National Renewable Energy Laboratory, December 2000.