getUsgs

Retrieve Real Time Gage Data from the United States Geological Survey

Overview

The getUsgs program is used to retrieve real time stream flow and other gaged data from the National Water Information Service (NWIS) of the United States Geological Survey (USGS) and output the data in a variety of formats. Data may be output in any or all of the following manners:

- output as USGS RDB format text (as retrieved from NWIS)
- output as Standard Hydrometeorological Exchange Format (SHEF) messages
- stored to a HEC-DSS file
- stored to a CWMS Oracle database

The data to be retrieved is controlled by a list of locations to retrieve and a specification of the number of hours prior to the current time to retrieve.

Time Zones

Gage data retrieved from the NWIS system is time stamped in one of the following time zones:

- EST – Eastern Standard Time (UTC-5 hours)
- CST – Central Standard Time (UTC-6 hours)
- MST – Mountain Standard Time (UTC-7 hours)
- PST – Pacific Standard Time (UTC-8 hours)
- AKST – Alaska Standard Time (UTC-9 hours)
- HST – Hawaii Standard Time (UTC-10 hours)

Unless specified otherwise, data output as SHEF messages or stored to a HEC-DSS file will be in the time zone of the retrieved data. The program allows outputting SHEF messages and storing to a HEC-DSS in any of the listed time zones, as well as in Coordinated Universal Time, which may be specified as UTC, GMT, or Z.

Input Files

The program requires three input files to specify:

- locations
- parameters
- parameter aliases

Each of these files is required to be in Comma-Separated Variable (CSV) format with a header line specifying the order of fields within the file. This format provides for easy editing using a spreadsheet

program such as Microsoft Excel. Fields in CSV format files are separated by the comma (,) character, and need not be quoted unless the field contains embedded comma characters. The first field in each file is a USGS Identifier, which must be enclosed in square brackets ([ ]) to prevent a spreadsheet program from treating the identifiers as integer numbers and removing leading zeros. Each file has a default filename in the directory from which the program is executed, which may be overridden by a command line option.

Locations Input File

The locations input file contains information necessary to generate output locations from the USGS location identifiers. Field names in the header line that are used by the program are:

- **[USGS_LOC]** – USGS location identifier. This field is always required in the file, must be the first field, and must be enclosed in square brackets. The field may not be empty.
- **SHEF_LOC** – SHEF location identifier. This field must be included in the file if data is to be output as SHEF messages. The field may not be empty.
- **DSS_A-Part** – HEC-DSS "A" pathname part. This field is required in the file if data is to be stored to a HEC-DSS file. The field may be empty.
- **DSS_B-Part** – HEC-DSS "B" pathname part. This field is required in the file if data is to be stored to a HEC-DSS file. The field may not be empty.
- **DSS_F-Part** – HEC-DSS "F" pathname part. This field is required in the file if data is to be stored to a HEC-DSS file. The field may be empty.
- **CWMS_LOC** – CWMS location identifier, including the base location and any sub-location. This field is required in the file if data is to be stored to a CWMS Oracle database. The field may not be empty.
- **CWMS_VER** – CWMS version identifier. This field is required in the file if data is to be stored to a CWMS Oracle database. The field may be empty.
- **PARAMETERS** – Location parameters. This field is always required. It contains a comma-separated list of parameters to be processed for the location. The parameters may be specified as USGS parameter identifiers or as parameter aliases specified in the parameter aliases file. This field may be empty.

The default name of the locations input file is Locations.csv in the directory from which the program is executed. A sample locations input file is shown in Figure 1.

Parameters Input File

The parameters input file contains information necessary to generate output parameters from the USGS parameter identifiers. Field names in the header line that are used by the program are:

- **[USGS_PARAMETER]** – USGS parameter identifier. This field is always required in the file, must be the first field, and must be enclosed in square brackets. This field specifies the USGS parameter identifier the field must not be empty.

- **SHEF_PARAMETER** – SHEF parameter identifier. This field must be included in the file if data is to be output as SHEF messages. The field may not be empty.
- **SHEF_FACTOR** – Parameter conversion factor from USGS to SHEF. This field must be included in the file if data is to be output as SHEF messages. The field may not be empty.
- **SHEF_UNIT** – Unit system for SHEF message. Must be SI or ENGLISH. This field must be included in the file if data is to be output as SHEF messages. The field may not be empty.
- **DSS_PARAMETER** - HEC-DSS "C" pathname part. This field is required in the file if data is to be stored to a HEC-DSS file. The field may be empty.
- **DSS_FACTOR** – Parameter conversion factor from USGS to HEC-DSS. This field must be included in the file if data is to be stored to a HEC-DSS file. The field may not be empty.
- **DSS_UNIT** – Data stored to a HEC-DSS will be marked as having this unit. This field must be included in the file if data is to be stored to a HEC-DSS file. The field may not be empty.
- **DSS_TYPE** – Data stored to a HEC-DSS will be marked as having this type. Must be one of INST-VAL, INST-CUM, PER-AVER, PER-CUM. This field must be included in the file if data is to be stored to a HEC-DSS file. The field may not be empty.
- **CWMS_PARAMETER** – CWMS parameter identifier, including base parameter and any sub-parameter. This field is required in the file if data is to be stored to a CWMS Oracle database. The field may be empty.
- **CWMS_FACTOR** – Parameter conversion factor from USGS to CWMS. This field must be included in the file if data is to be stored to a CWMS Oracle database. The field may not be empty.
- **CWMS_UNIT** – Data stored to a CWMS Oracle database will be marked as having this unit. This field must be included in the file if data is to be stored to a CWMS Oracle database. The field may not be empty.
- **CWMS_TYPE** – CWMS parameter type identifier. This field must be included in the file if data is to be stored to a CWMS Oracle database. The field may not be empty.

The default name of the parameters input file is Parameters.csv in the directory from which the program is executed. A sample parameters input file is shown in Figure 2.

Parameter Aliases Input File

The parameter aliases input file contains text aliases for USGS parameter identifiers. Field names in the header line that are used by the program are:

- **[USGS_PARAMETER]** – USGS parameter identifier. This field is always required in the file, must be the first field, and must be enclosed in square brackets. This field specifies the USGS parameter identifier the field must not be empty.
- **ALIAS** – Parameter alias to be used in PARAMETERS field of the location input file. This field must be included in the file and may not be empty.

The default name of the parameter aliases input file is Parameter_Aliases.csv in the directory from which the program is executed. A sample parameter aliases input file is shown in Figure 3.

Usage

The getUsgs program is a Jython script named getUsgs.py which can be executed in any environment in which hec.jar and heclib.jar are available on the classpath. In order to store data to a CWMS Oracle database, dbiClient.jar and cwmsdb.jar are also required.

The program includes a "shebang" for automatic interpreter loading on UNIX-like environments, allowing execution using the "getUsgs.py" command or, if renamed, "getUsgs". On Windows the program can be executed using the command "jython getUsgs.py", which also works on UNIX-like environments. On Windows client installations of HEC-DSSVue, the program can also be executed using the HEC-DSSVue.cmd file as "HEC-DSSVue getUsgs.py". In the discussion below, "getUsgs" is used for simplicity.

The program writes to the standard output device (stdout), as well as the standard error device (stderr). Any output of USGS RDB format text or SHEF messages is written to stdout. All other output, including program status and error messages are written to stderr. In Windows and on most shells in UNIX-like environments, these two output streams can be separated and redirected to different files by using the following redirection command at the end of the command line:

- `> filename` or `1> filename` – redirect stdout to file, overwriting existing content
- `>> filename` or `1>> filename` – redirect stdout to file, appending to existing content
- `2> filename` – redirect stderr to file, overwriting existing content
- `2>> filename` – redirect stderr to file, appending to existing content
- `2> &1` – redirect stderr to stdout, used to combine stdout and stderr to a single output stream, can be combined with stdout redirection to redirect both streams to a file (e.g., `> filename 2> &1`)

The C-Shell (csh) and its work-alikes (tcsh, etc…) on UNIX-like environments provide the following output redirection commands, which do not support separation of stdout and stderr:

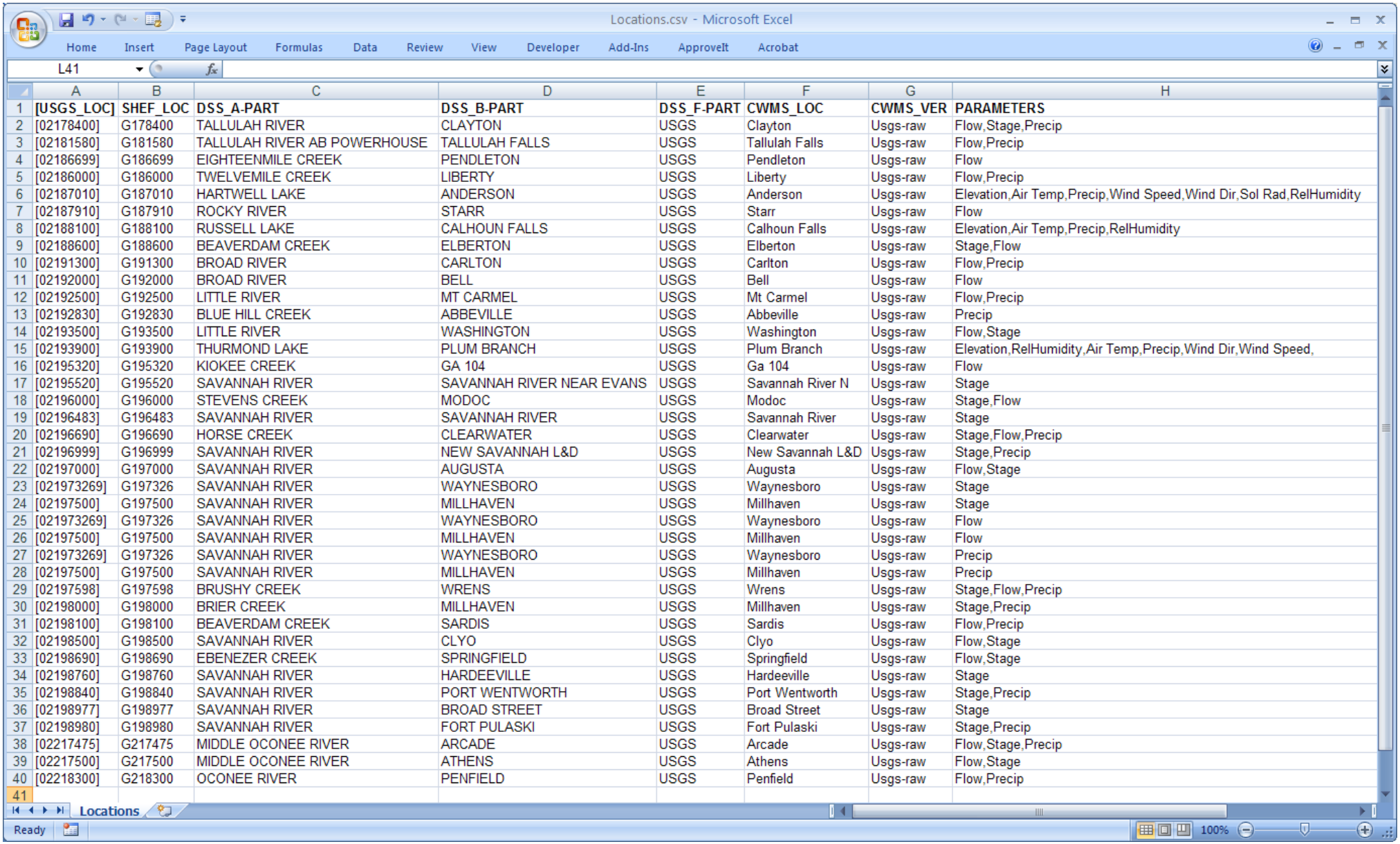
- `> filename` – redirect stdout to file, overwriting existing content
- `>& filename` – redirect stdout to file, appending to existing content
- `>> filename` – redirect stdout and stderr to file, overwriting existing content
- `>>& filename` – redirect stdout and stderr to file, appending to existing content

Since the CWMS execution environment uses the tcsh shell be default, it is necessary to explicitly execute the program from a different shell in order to capture stdout and stderr to different files in the CWMS environment.

Command Line

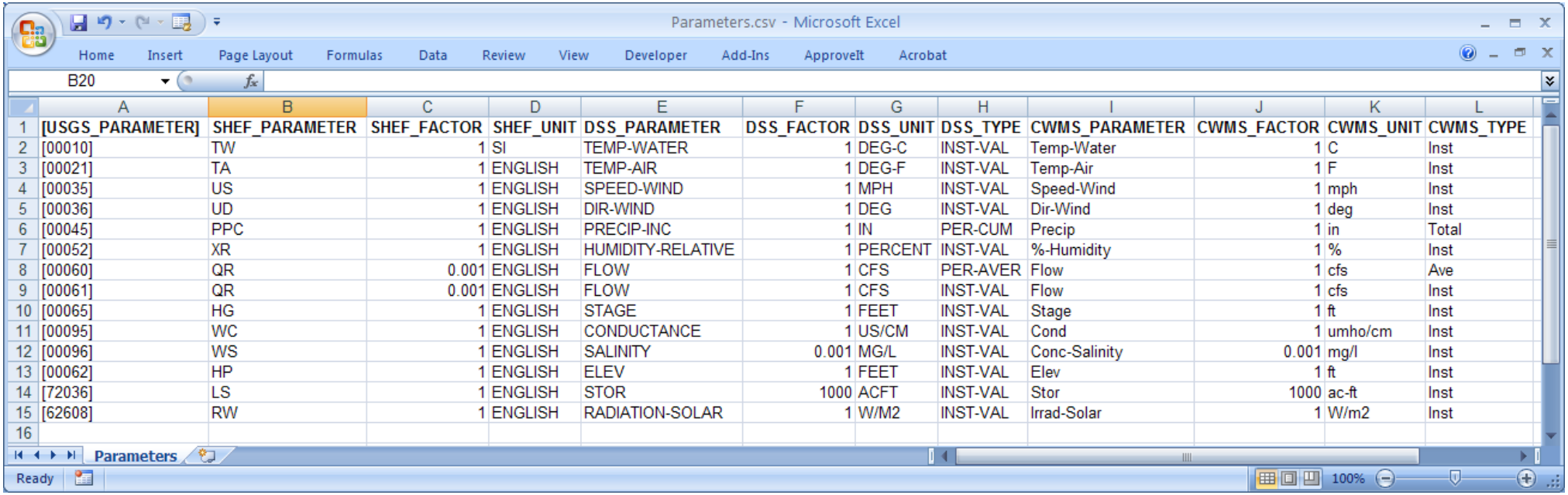
Program execution via the command line has the form "getUsgs *program_options redirection_options"*, where *redirection_options* are discussed above and *program_options* is comprised of the follwing

- `-l` *`locations_filename`* (or `--locations` *`locations_filename`*) – specifies locations input file. Defaults to Locations.csv in the directory the program is executed from.
- `-p` *`parameters_filename`* (or `--parameters` *`parameters_filename`*) – specifies parameters input file. Defaults to Parameters.csv in the directory the program is executed from.
- `-a` *`parameter_aliases_filename`* (or `--aliases` *`parameter_aliases_filename`*) – specifies parameter aliases input file. Defaults to Locations.csv in the directory the program is executed from.
- `-u` (or `--usgs`) – specifies outputting data as USGS RDB format text.
- `-s` (or `--shef`) – specifies storing data to a HEC_DSS file.
- `--tzshef` *`time_zone`* –specifies time zone to use for SHEF messages. See Time Zone section above for list of valid time zones. If not specified, the SHEF messages will be in the time zone specified in the USGS text.
- `-d` *`dss_filename`* (`or --dss` *`dss_filename`*) – specifies storing data to a HEC-DSS file. dss_filename specifies the HEC-DSS file to use.
- `--tzdss` *`time_zone`* –specifies time zone to use for data stored to a HEC-DSS file. See Time Zone section above for list of valid time zones. If not specified, the data will be stored to a HEC-DSS file will be in the time zone specified in the USGS text.
- `-c` (or `--cwms`) – specifies storing data to a CWMS Oracle database
- `--rule` *`store_rule`* –specifies CWMS store rule to use for data stored to the CWMS Oracle database. If not specified, DELETE INSERT will be used. Valid values for *`store_rule`* are:
  - `REPLACE ALL`
  - `DO NOT REPLACE`
  - `REPLACE WITH MISSING VALUES ONLY`
  - `REPLACE WITH NON MISSING`
  - `DELETE INSERT`
- `-h` *`hours_to_retrieve`* (or `--hours` *`hours_to_retreieve`*) – specifies the number of hours of data to retrieve from the USGS. Defaults to 24 hours
- `-o` *`output_level`* (or `--output` *`output_level`*) – specifies the level of output generated by the program. Defaults to NORMAL. Valid values for *`output_level`* are:
  - `NONE`
  - `NORMAL`
  - `VERBOSE`



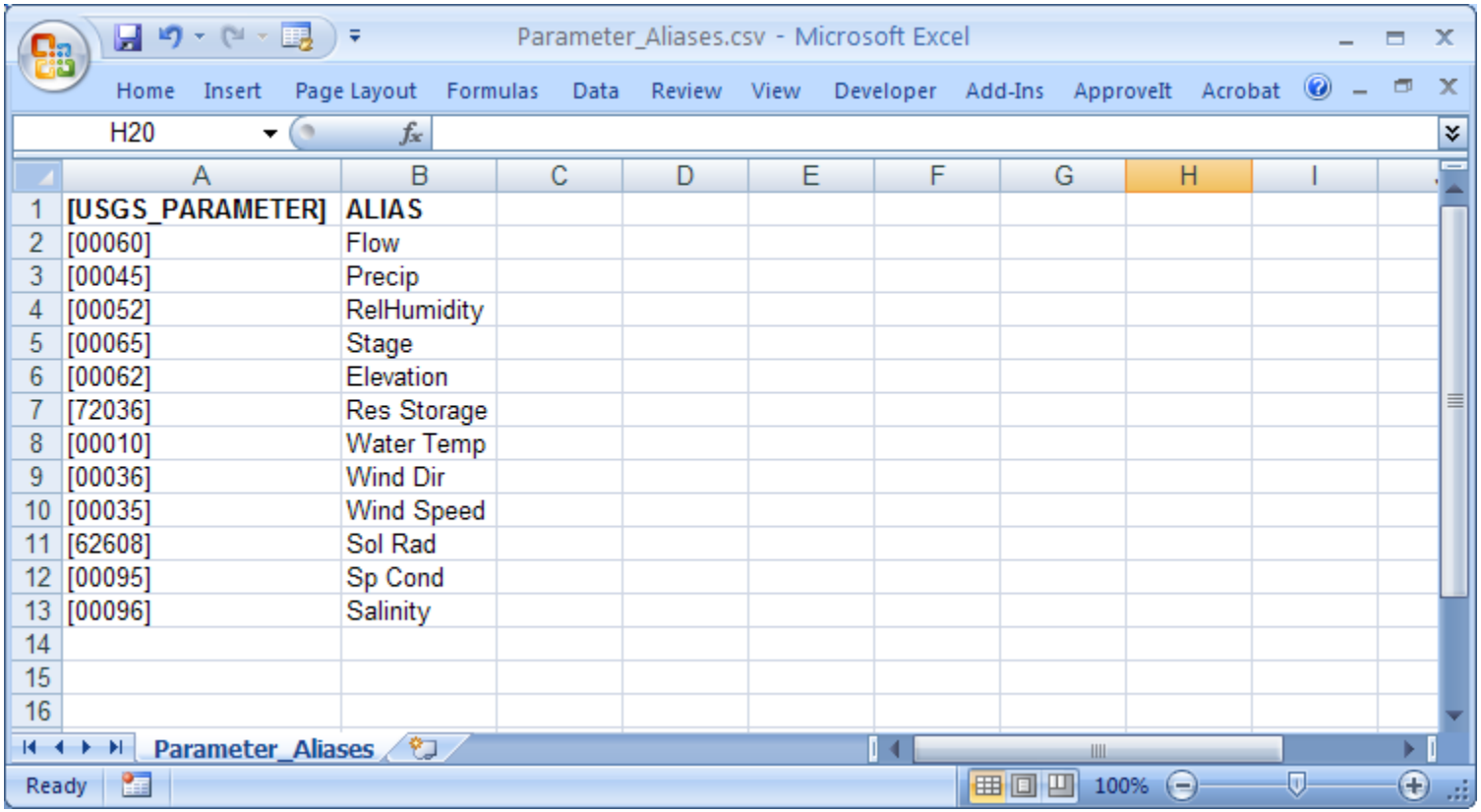
[USGS_LOC]	SHEF_LOC	DSS_A-PART	DSS_B-PART	DSS_F-PART	CWMS_LOC	CWMS_VER	PARAMETERS
[02178400]	G178400	TALLULAH RIVER	CLAYTON	USGS	Clayton	Usgs-raw	Flow,Stage,Precip
[02181580]	G181580	TALLULAH RIVER AB POWERHOUSE	TALLULAH FALLS	USGS	Tallulah Falls	Usgs-raw	Flow,Precip
[02186699]	G186699	EIGHTEENMILE CREEK	PENDLETON	USGS	Pendleton	Usgs-raw	Flow
[02186000]	G186000	TWELVEMILE CREEK	LIBERTY	USGS	Liberty	Usgs-raw	Flow,Precip
[02187010]	G187010	HARTWELL LAKE	ANDERSON	USGS	Anderson	Usgs-raw	Elevation,Air Temp,Precip,Wind Speed,Wind Dir,Sol Rad,RelHumidity
[02187910]	G187910	ROCKY RIVER	STARR	USGS	Starr	Usgs-raw	Flow
[02188100]	G188100	RUSSELL LAKE	CALHOUN FALLS	USGS	Calhoun Falls	Usgs-raw	Elevation,Air Temp,Precip,RelHumidity
[02188600]	G188600	BEAVERDAM CREEK	ELBERTON	USGS	Elberton	Usgs-raw	Stage,Flow
[02191300]	G191300	BROAD RIVER	CARLTON	USGS	Carlton	Usgs-raw	Flow,Precip
[02192000]	G192000	BROAD RIVER	BELL	USGS	Bell	Usgs-raw	Flow
[02192500]	G192500	LITTLE RIVER	MT CARMEL	USGS	Mt Carmel	Usgs-raw	Flow,Precip
[02192830]	G192830	BLUE HILL CREEK	ABBEVILLE	USGS	Abbeville	Usgs-raw	Precip
[02193500]	G193500	LITTLE RIVER	WASHINGTON	USGS	Washington	Usgs-raw	Flow,Stage
[02193900]	G193900	THURMOND LAKE	PLUM BRANCH	USGS	Plum Branch	Usgs-raw	Elevation,RelHumidity,Air Temp,Precip,Wind Dir,Wind Speed,
[02195320]	G195320	KIOKEE CREEK	GA 104	USGS	Ga 104	Usgs-raw	Flow
[02195520]	G195520	SAVANNAH RIVER	SAVANNAH RIVER NEAR EVANS	USGS	Savannah River N	Usgs-raw	Stage
[02196000]	G196000	STEVENS CREEK	MODOC	USGS	Modoc	Usgs-raw	Stage,Flow
[02196483]	G196483	SAVANNAH RIVER	SAVANNAH RIVER	USGS	Savannah River	Usgs-raw	Stage
[02196690]	G196690	HORSE CREEK	CLEARWATER	USGS	Clearwater	Usgs-raw	Stage,Flow,Precip
[02196999]	G196999	SAVANNAH RIVER	NEW SAVANNAH L&D	USGS	New Savannah L&D	Usgs-raw	Stage,Precip
[02197000]	G197000	SAVANNAH RIVER	AUGUSTA	USGS	Augusta	Usgs-raw	Flow,Stage
[021973269]	G197326	SAVANNAH RIVER	WAYNESBORO	USGS	Waynesboro	Usgs-raw	Stage
[02197500]	G197500	SAVANNAH RIVER	MILLHAVEN	USGS	Millhaven	Usgs-raw	Stage
[021973269]	G197326	SAVANNAH RIVER	WAYNESBORO	USGS	Waynesboro	Usgs-raw	Flow
[02197500]	G197500	SAVANNAH RIVER	MILLHAVEN	USGS	Millhaven	Usgs-raw	Flow
[021973269]	G197326	SAVANNAH RIVER	WAYNESBORO	USGS	Waynesboro	Usgs-raw	Precip
[02197500]	G197500	SAVANNAH RIVER	MILLHAVEN	USGS	Millhaven	Usgs-raw	Precip
[02197598]	G197598	BRUSHY CREEK	WRENS	USGS	Wrens	Usgs-raw	Stage,Flow,Precip
[02198000]	G198000	BRIER CREEK	MILLHAVEN	USGS	Millhaven	Usgs-raw	Stage,Precip
[02198100]	G198100	BEAVERDAM CREEK	SARDIS	USGS	Sardis	Usgs-raw	Flow,Precip
[02198500]	G198500	SAVANNAH RIVER	CLYO	USGS	Clyo	Usgs-raw	Flow,Stage
[02198690]	G198690	EBENEZER CREEK	SPRINGFIELD	USGS	Springfield	Usgs-raw	Flow,Stage
[02198760]	G198760	SAVANNAH RIVER	HARDEEVILLE	USGS	Hardeeville	Usgs-raw	Stage
[02198840]	G198840	SAVANNAH RIVER	PORT WENTWORTH	USGS	Port Wentworth	Usgs-raw	Stage,Precip
[02198977]	G198977	SAVANNAH RIVER	BROAD STREET	USGS	Broad Street	Usgs-raw	Stage
[02198980]	G198980	SAVANNAH RIVER	FORT PULASKI	USGS	Fort Pulaski	Usgs-raw	Stage,Precip
[02217475]	G217475	MIDDLE OCONEE RIVER	ARCADE	USGS	Arcade	Usgs-raw	Flow,Stage,Precip
[02217500]	G217500	MIDDLE OCONEE RIVER	ATHENS	USGS	Athens	Usgs-raw	Flow,Stage
[02218300]	G218300	OCONEE RIVER	PENFIELD	USGS	Penfield	Usgs-raw	Flow,Precip

Figure 1. Sample Locations Input File



	A	B	C	D	E	F	G	H	I	J	K	L
1	[USGS_PARAMETER]	SHEF_PARAMETER	SHEF_FACTOR	SHEF_UNIT	DSS_PARAMETER	DSS_FACTOR	DSS_UNIT	DSS_TYPE	CWMS_PARAMETER	CWMS_FACTOR	CWMS_UNIT	CWMS_TYPE
2	[00010]	TW	1	SI	TEMP-WATER	1	DEG-C	INST-VAL	Temp-Water	1	C	Inst
3	[00021]	TA	1	ENGLISH	TEMP-AIR	1	DEG-F	INST-VAL	Temp-Air	1	F	Inst
4	[00035]	US	1	ENGLISH	SPEED-WIND	1	MPH	INST-VAL	Speed-Wind	1	mph	Inst
5	[00036]	UD	1	ENGLISH	DIR-WIND	1	DEG	INST-VAL	Dir-Wind	1	deg	Inst
6	[00045]	PPC	1	ENGLISH	PRECIP-INC	1	IN	PER-CUM	Precip	1	in	Total
7	[00052]	XR	1	ENGLISH	HUMIDITY-RELATIVE	1	PERCENT	INST-VAL	%-Humidity	1	%	Inst
8	[00060]	QR	0.001	ENGLISH	FLOW	1	CFS	PER-AVER	Flow	1	cfs	Ave
9	[00061]	QR	0.001	ENGLISH	FLOW	1	CFS	INST-VAL	Flow	1	cfs	Inst
10	[00065]	HG	1	ENGLISH	STAGE	1	FEET	INST-VAL	Stage	1	ft	Inst
11	[00095]	WC	1	ENGLISH	CONDUCTANCE	1	US/CM	INST-VAL	Cond	1	umho/cm	Inst
12	[00096]	WS	1	ENGLISH	SALINITY	0.001	MG/L	INST-VAL	Conc-Salinity	0.001	mg/l	Inst
13	[00062]	HP	1	ENGLISH	ELEV	1	FEET	INST-VAL	Elev	1	ft	Inst
14	[72036]	LS	1	ENGLISH	STOR	1000	ACFT	INST-VAL	Stor	1000	ac-ft	Inst
15	[62608]	RW	1	ENGLISH	RADIATION-SOLAR	1	W/M2	INST-VAL	Irrad-Solar	1	W/m2	Inst
16												

Figure 2. Sample Parameters Input File



	A	B
1	[USGS_PARAMETER]	ALIAS
2	[00060]	Flow
3	[00045]	Precip
4	[00052]	RelHumidity
5	[00065]	Stage
6	[00062]	Elevation
7	[72036]	Res Storage
8	[00010]	Water Temp
9	[00036]	Wind Dir
10	[00035]	Wind Speed
11	[62608]	Sol Rad
12	[00095]	Sp Cond
13	[00096]	Salinity
14		
15		
16		

Figure 3. Sample Parameter Alias Input File