Surface Water Statistics - Webinar

USGS Office of Surface Water

March 7, 2011

USGS
science for a changing world

Agenda

- What's new in USGS SW statistics
  - OSW technical memorandum 2011.03
  - SW Statistics website
  - Status of software updates: PeakFQ, SWSTAT, GLSNET/WREG
  - GetNWISQ, SREF, QSTATS, …
- WREG program
  - Overview of program capabilities
  - WREG Shell: Greg Granato
- Future webinars
  - If useful, how frequent?
  - Suggested topics?

OSW Technical Memorandum 2011.03

Technical resources for regionalization studies

Just released Friday

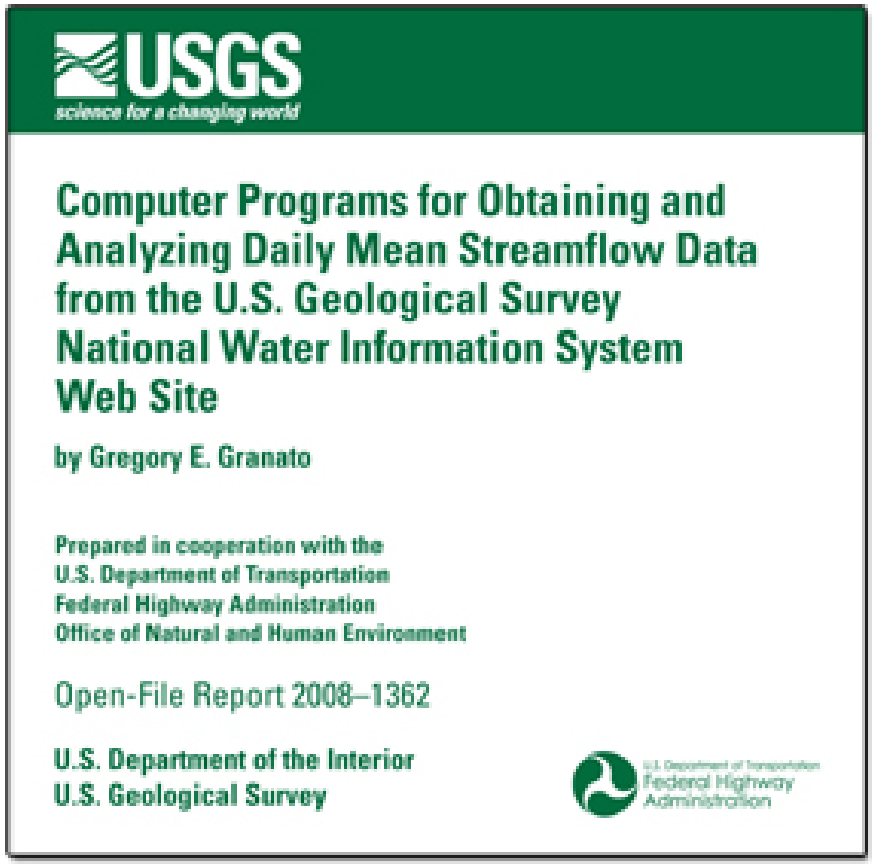
1. New website with resources for SW statistics
2. Joint studies with other WSCs to define study areas based on hydrology and to pool expertise
3. Training – SW1523
4. Proposals will be forwarded to OSW

   Status of studies posted on website. Send me updates.
5. Mentoring available
6. Workplans suggested
7. Report outline, data archiving

Website with technical resources

http://water.usgs.gov/usgs/osw/swstats

Newish software: GNWISQ, QSTAT, SREF, …


USGS
science for a changing world

Computer Programs for Obtaining and Analyzing Daily Mean Streamflow Data from the U.S. Geological Survey National Water Information System Web Site

by Gregory E. Granato

Prepared in cooperation with the
U.S. Department of Transportation
Federal Highway Administration
Office of Natural and Human Environment

Open-File Report 2008–1362

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of Transportation Federal Highway Administration

http://pubs.usgs.gov/of/2008/1362/

Suite of programs by Greg Granato, 2009

GNWISQ: get NWIS daily flow time series in batch jobs

QSTAT: calculate basic statistics for daily time series downloaded by GNWISQ

SREF: extend or augment records using MOVE.1 or MOVE.3

Status of software: PeakFQ

- Peak flow frequency analysis program, Windows GUI version (2006)

- Working on new version with new capabilities:
  - Expected Moments Algorithm (EMA) for fitting frequency curve
  - Kendall's Tau statistics (trend analysis of peak flow time series)
  - Multiple Grubbs Beck test for low outliers

- New export format

  flood estimates
  annual exceedance probabilities
  Variance of Estimate (EMA)
  Confidence Intervals (EMA) / Limits (traditional B17B calc)
  K values

- Serious testers welcome (program available soon)

- Expected release in 2011

Input window in new PeakFQ

PKFQWin

File Help

Use File menu to Open PeakFQ data or PKFQWin spec file.
Update Station, Threshold and Output specifications as desired.
Click Run PeakFQ button to generate results.

PeakFQ Data File: Z:\kmflynn\Presentations\Data\PEAKS.INP

PKFQWin Spec File:

Station Specifications | Input/View | Output Options | Results

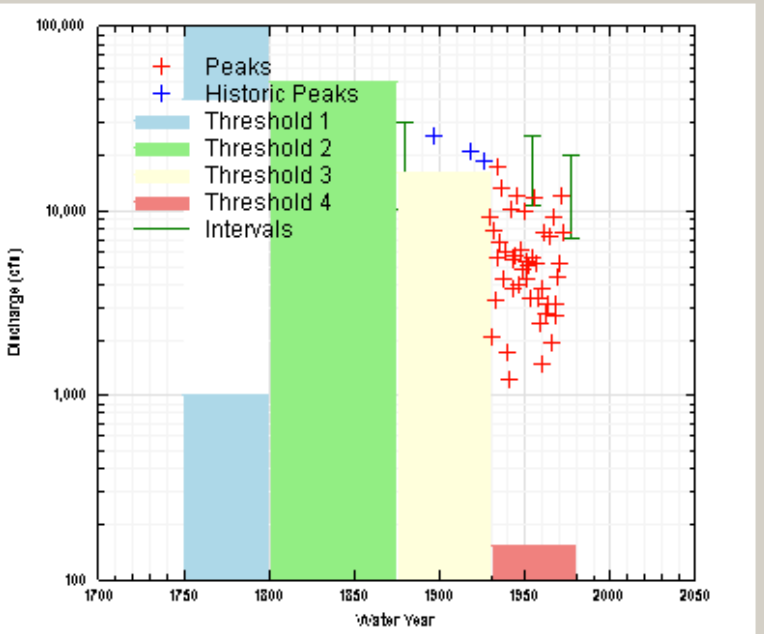
Station: 03606500

Perception Thresholds & Periods

Start Year	End Year	Low Threshold	High Threshold	Comments
1750	1800	1000	40000	Paleo flood record WRIR 87-111
1801	1875	50000	1.0e20	Journal accounts, WRIR 56-012
1876	1930	16000	1.0e20	High Water marks, WRIR 56-012
1931	1980	150	1.0e20	Observed Record
			1.0e20	

Peak and Interval Data

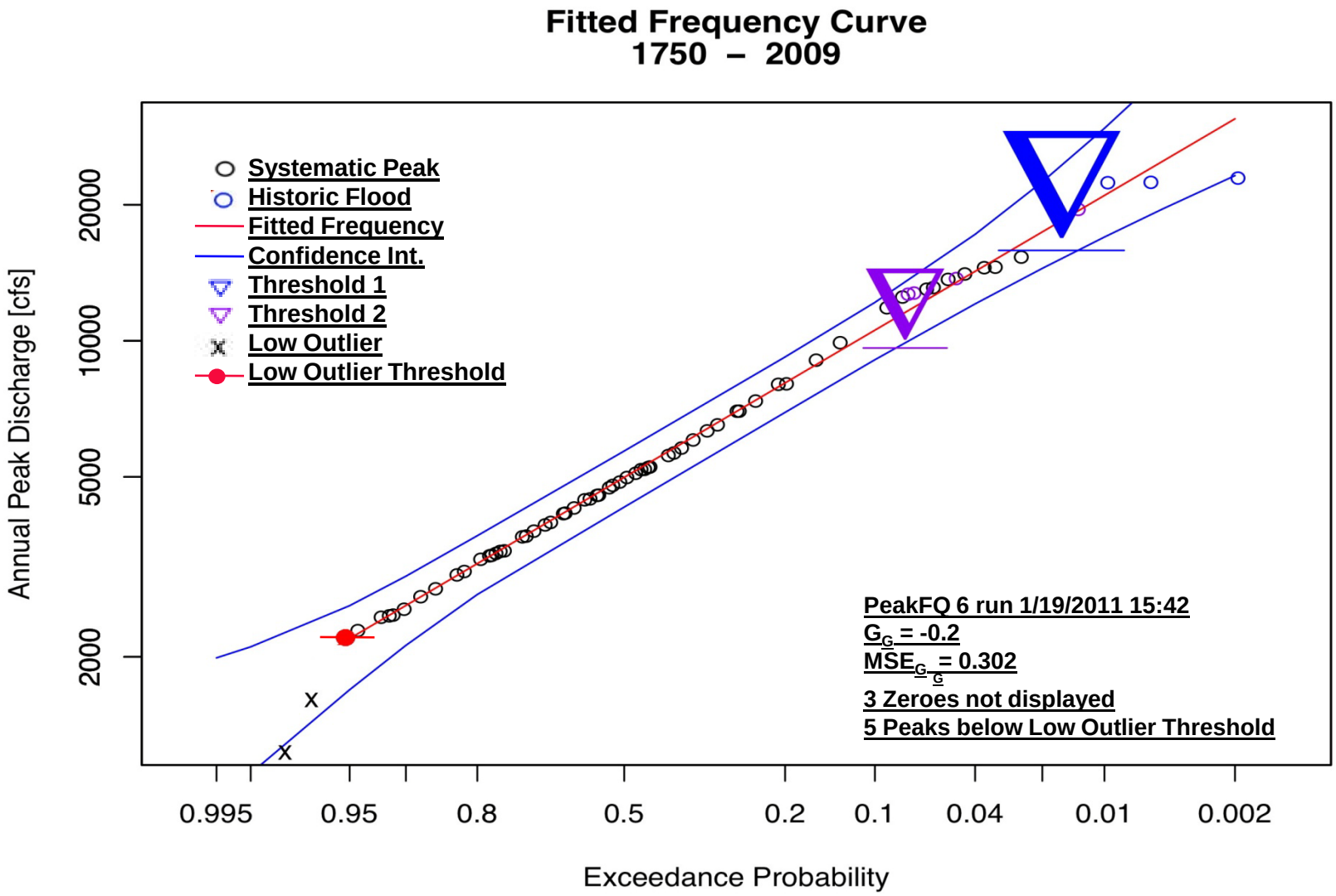
Water Year	Peak Value	Remark Codes	Low Interval	High Interval	Comment
1972	12000				
1973	7640		6112	9168	20% +- review of slope-area
1880			10000	30000	WRIR 56-012
1955			10500	25000	WRIR 56-012
1978			7000	20000	OFR 80-1245



Run PEAKFQ | Save Specs | Exit

USGS science for a changing world

Plotted frequency curve in new PeakFQ



Status of software: SWSTAT

- Surface water statistics program
  - Basic statistics (mean, min, max, etc.)
  - Calculation of n-day flows from daily values
  - Frequency analysis of annual time series
  - Flow duration (and comparison of flow durations at 2 stations)
  - Duration hydrograph?
  - Trend analysis – Kendall Tau

- New version SWStat 4.1 under development:
  - integrated into EPA BASINS environment for better input data handling (direct import of data from NWIS-Web into WDMs)

- Improved outputs, similar to PeakFQ

- Serious testers welcome now

- Expected release in 2011

Status of software: GLSNET/WREG

Multiple linear regression for regionalization of surface water statistics

-GLSNET

- being phased out, but still ok for GLS regressions
- Requires WDM files for input

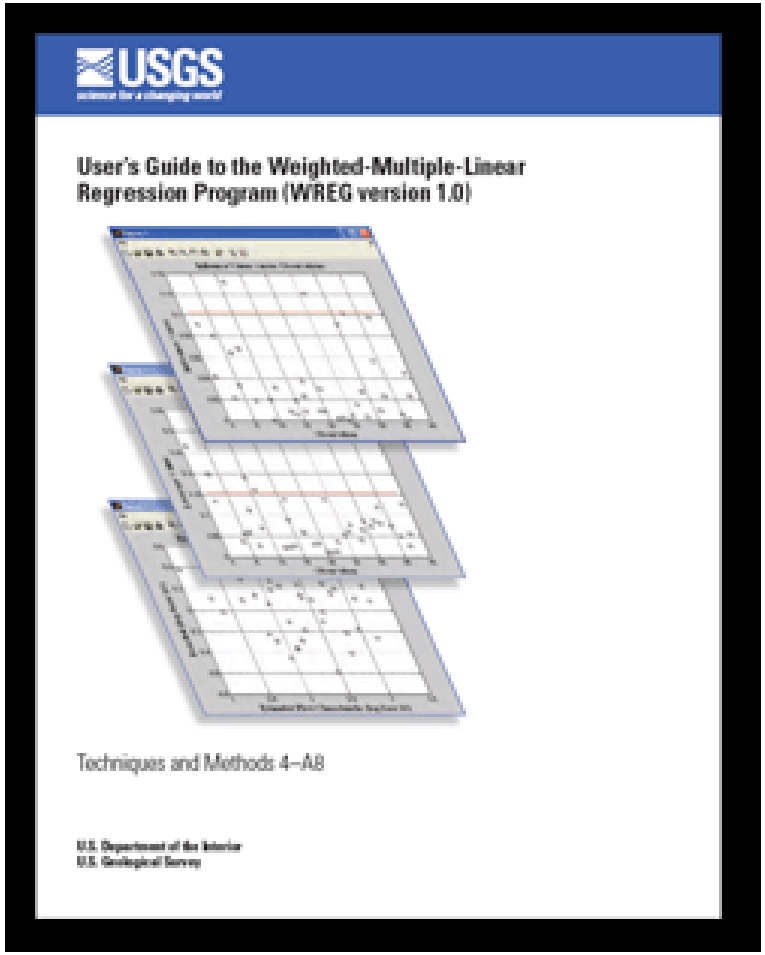
-WREG – developed by Ken Eng and others, released 2010

- GUI interface, text inputs
- OLS, WLS, or GLS
- Additional updates to GLS – considers uncertainty in skew
- RoI models

- WREG shell – written by Greg Granato

- Enhances original program's input/output handling

Weighted Multiple Linear Regression Program (WREG)



http://pubs.usgs.gov/tm/tm4a8/

Least squares regressions:

-Ordinary least squares (OLS)

-Weighted least squares (WLS)

- Standard (Tasker, 1980 for frequency stats) and user-defined

-Generalized least squares (GLS)

- Stedinger and Tasker series of papers, for peak flow frequency stats.
- Same as GLSNET, or with additional option for uncertainty in skew

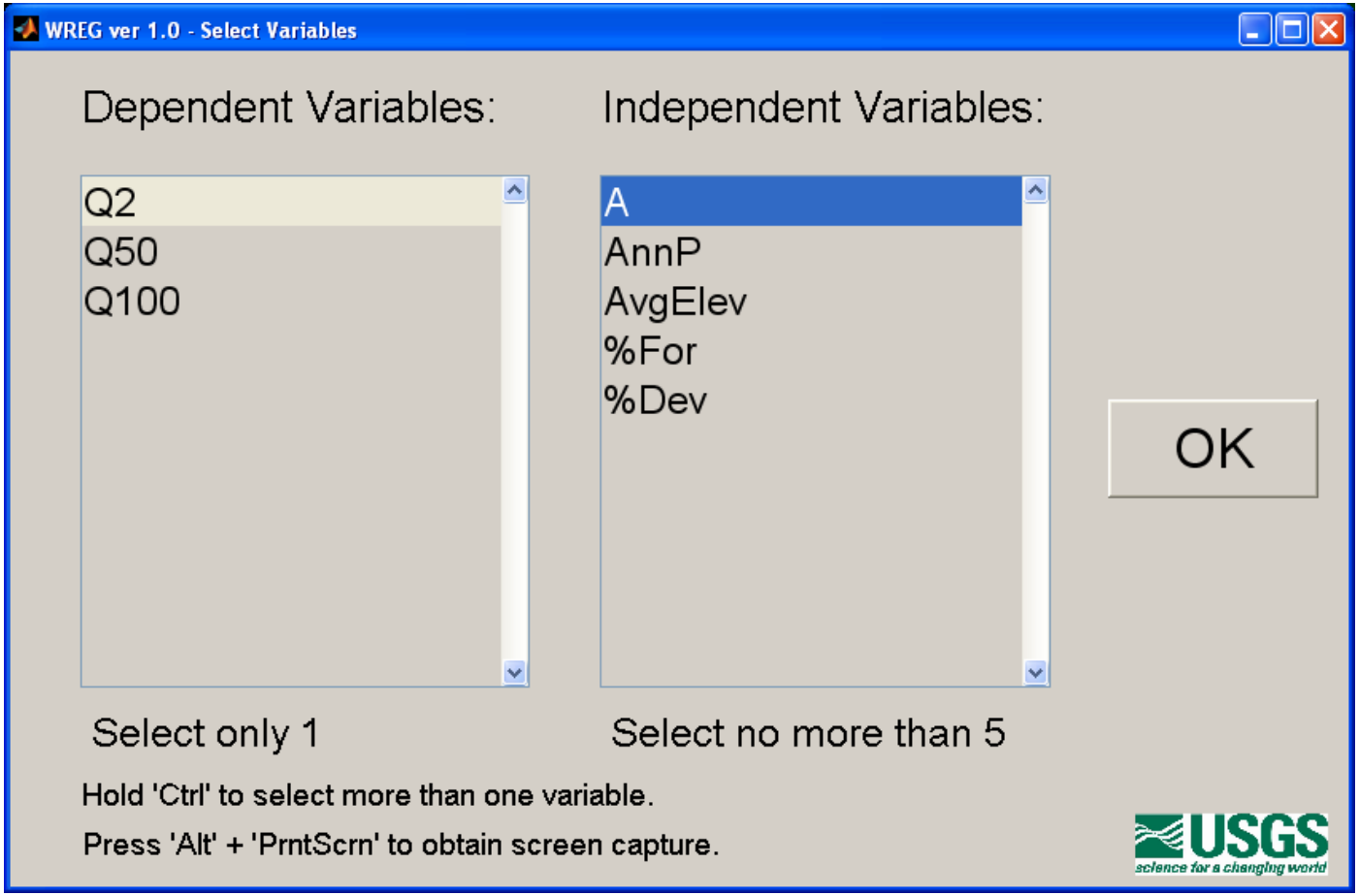
Region of influence:

-Regions defined using geographic space, predictor variable space, or a hybrid space.

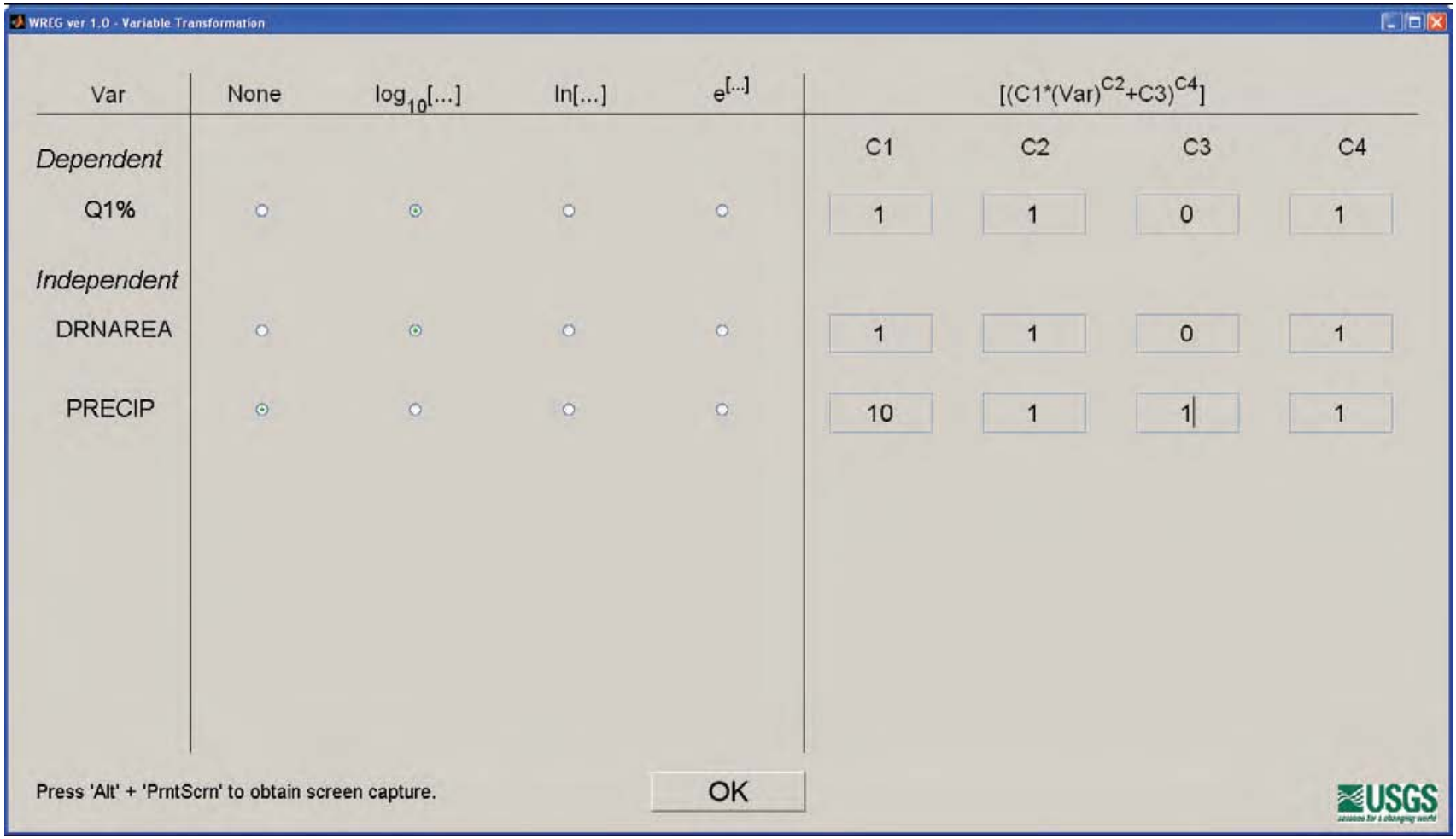
-Set up and test model, but cannot be used for ungaged sites

Program Features

- Point and click GUI interface
- Text input and output files
- Plots of major diagnostics



Select variable transformations

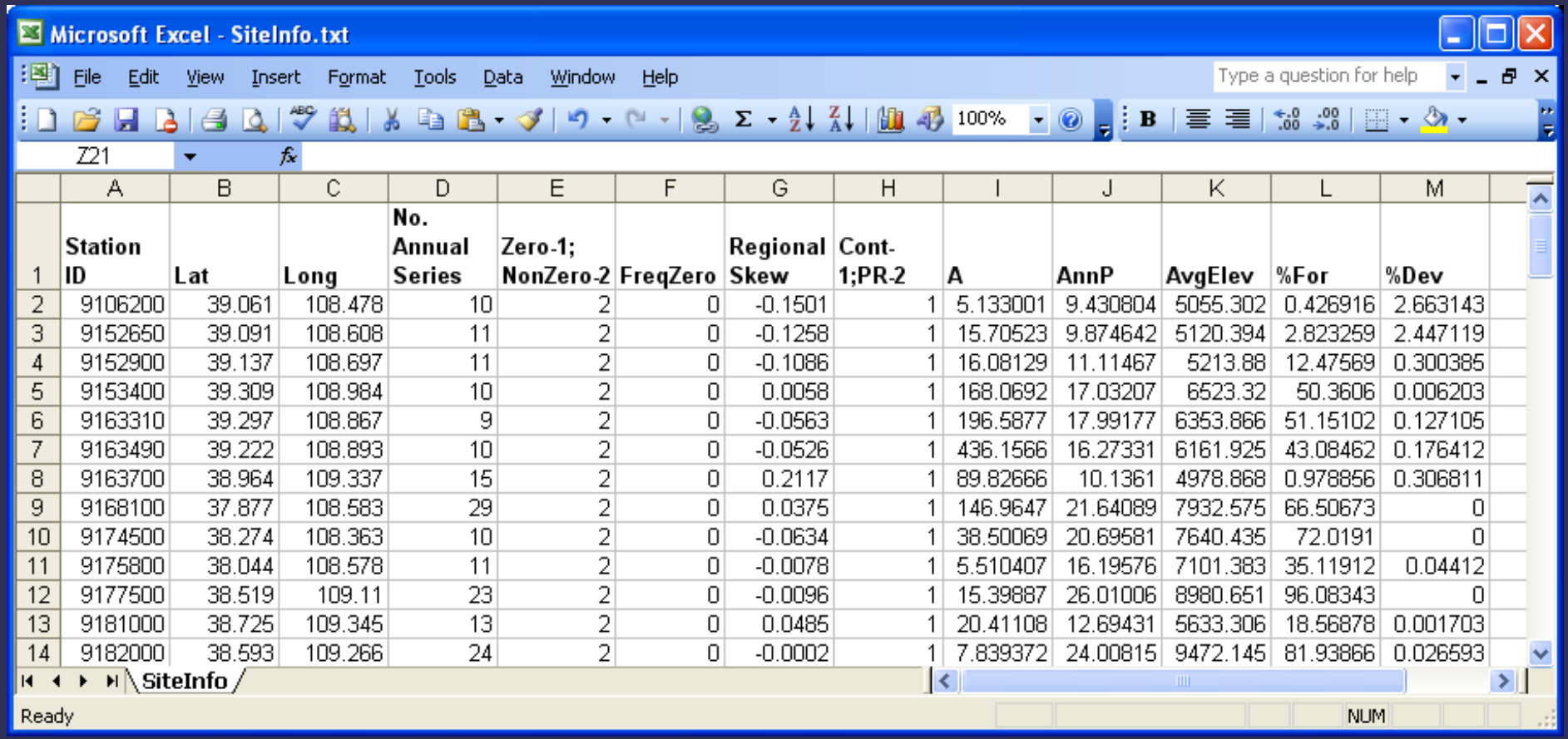


Choose model

Input Files: ASCII text, tab delimited

File name	Description	WREG requirements
SiteInfo.txt	Site information and basin characteristics to be used in the regression (the independent vars).	Always required.
FlowChar.txt	Flow characteristics to be used in the regression (the dependent vars).	Always required.
LP3G.txt	Skew for Log-Pearson Type III distribution .	Always required.
LP3K.txt	K for Log-Pearson Type III Distribution.	Always required.
LP3s.txt	Standard deviation for Log-Pearson Type III distribution.	Always required.
UserWLS.txt	User specified weighting matrix.	Required only if the user-defined WLS option is selected.
USGS########.txt	Annual time series of flow at streamflow-gaging stations.	One file per station, required only when using either the GLS option.

SiteInfo.txt



Microsoft Excel - SiteInfo.txt

File Edit View Insert Format Tools Data Window Help

Type a question for help

100%

Z21

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Station ID	Lat	Long	No. Annual Series	Zero-1; NonZero-2	FreqZero	Regional Skew	Cont-1;PR-2	A	AnnP	AvgElev	%For	%Dev
2	9106200	39.061	108.478	10	2	0	-0.1501	1	5.133001	9.430804	5055.302	0.426916	2.663143
3	9152650	39.091	108.608	11	2	0	-0.1258	1	15.70523	9.874642	5120.394	2.823259	2.447119
4	9152900	39.137	108.697	11	2	0	-0.1086	1	16.08129	11.11467	5213.88	12.47569	0.300385
5	9153400	39.309	108.984	10	2	0	0.0058	1	168.0692	17.03207	6523.32	50.3606	0.006203
6	9163310	39.297	108.867	9	2	0	-0.0563	1	196.5877	17.99177	6353.866	51.15102	0.127105
7	9163490	39.222	108.893	10	2	0	-0.0526	1	436.1566	16.27331	6161.925	43.08462	0.176412
8	9163700	38.964	109.337	15	2	0	0.2117	1	89.82666	10.1361	4978.868	0.978856	0.306811
9	9168100	37.877	108.583	29	2	0	0.0375	1	146.9647	21.64089	7932.575	66.50673	0
10	9174500	38.274	108.363	10	2	0	-0.0634	1	38.50069	20.69581	7640.435	72.0191	0
11	9175800	38.044	108.578	11	2	0	-0.0078	1	5.510407	16.19576	7101.383	35.11912	0.04412
12	9177500	38.519	109.11	23	2	0	-0.0096	1	15.39887	26.01006	8980.651	96.08343	0
13	9181000	38.725	109.345	13	2	0	0.0485	1	20.41108	12.69431	5633.306	18.56878	0.001703
14	9182000	38.593	109.266	24	2	0	-0.0002	1	7.839372	24.00815	9472.145	81.93866	0.026593

SiteInfo

Ready NUM

First 8 columns are required.

Remaining columns contain basin attributes

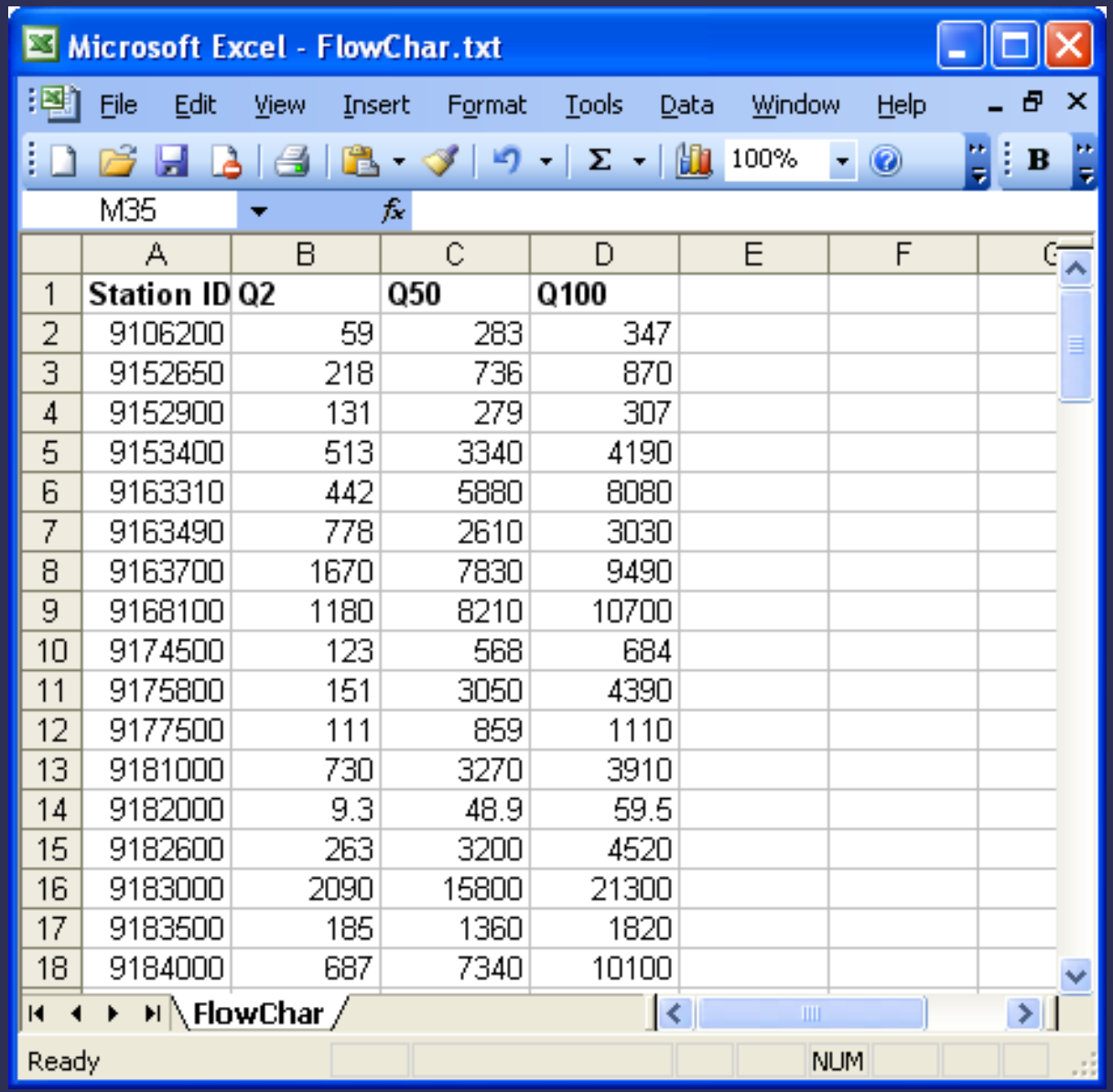
- Stations MUST be listed in ascending order.

- Columns for Zero-1;NonZero-2 and FreqZero are placeholders for logistic regression, not yet implemented by WREG. Values must be provided, but they can be dummy values.

FlowChar.txt

Stations MUST be:

- **listed in ascending numerical order**
- **listed in the same order in all files**

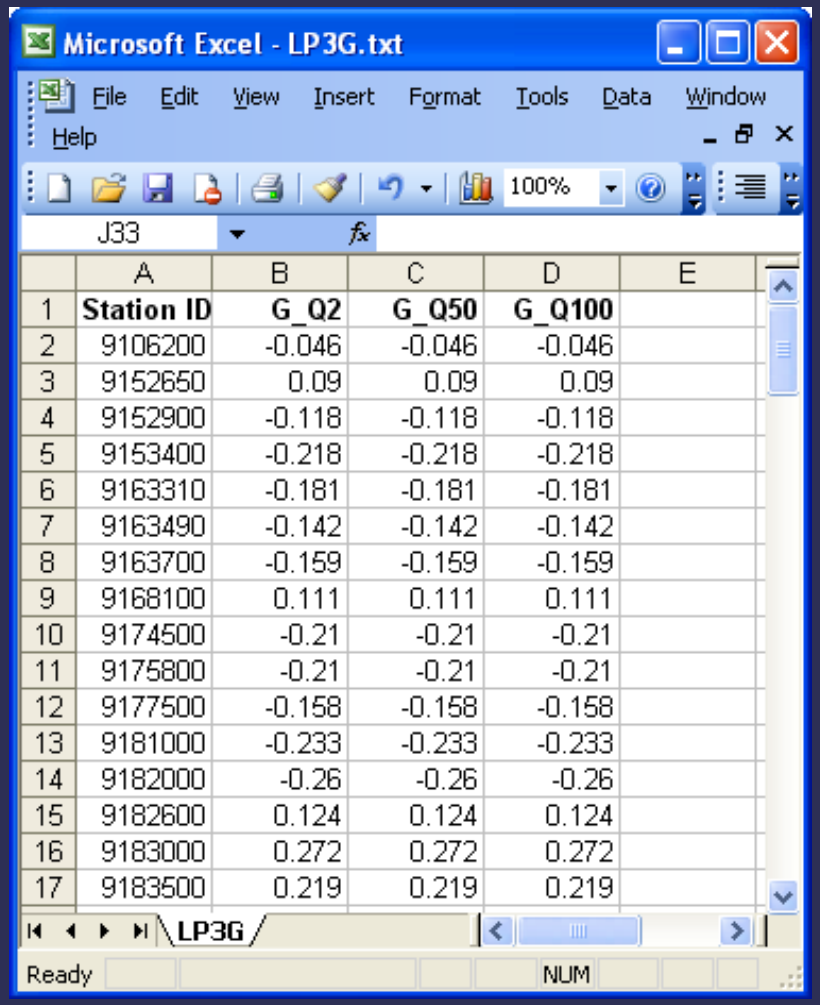


	A	B	C	D
1	Station ID	Q2	Q50	Q100
2	9106200	59	283	347
3	9152650	218	736	870
4	9152900	131	279	307
5	9153400	513	3340	4190
6	9163310	442	5880	8080
7	9163490	778	2610	3030
8	9163700	1670	7830	9490
9	9168100	1180	8210	10700
10	9174500	123	568	684
11	9175800	151	3050	4390
12	9177500	111	859	1110
13	9181000	730	3270	3910
14	9182000	9.3	48.9	59.5
15	9182600	263	3200	4520
16	9183000	2090	15800	21300
17	9183500	185	1360	1820
18	9184000	687	7340	10100

LP3G, LP3K, LP3s.txt

LP3G: Skew for LP3 distribution
LP3K: K-factor for LP3 distribution
LP3s: Std dev for LP3 distribution

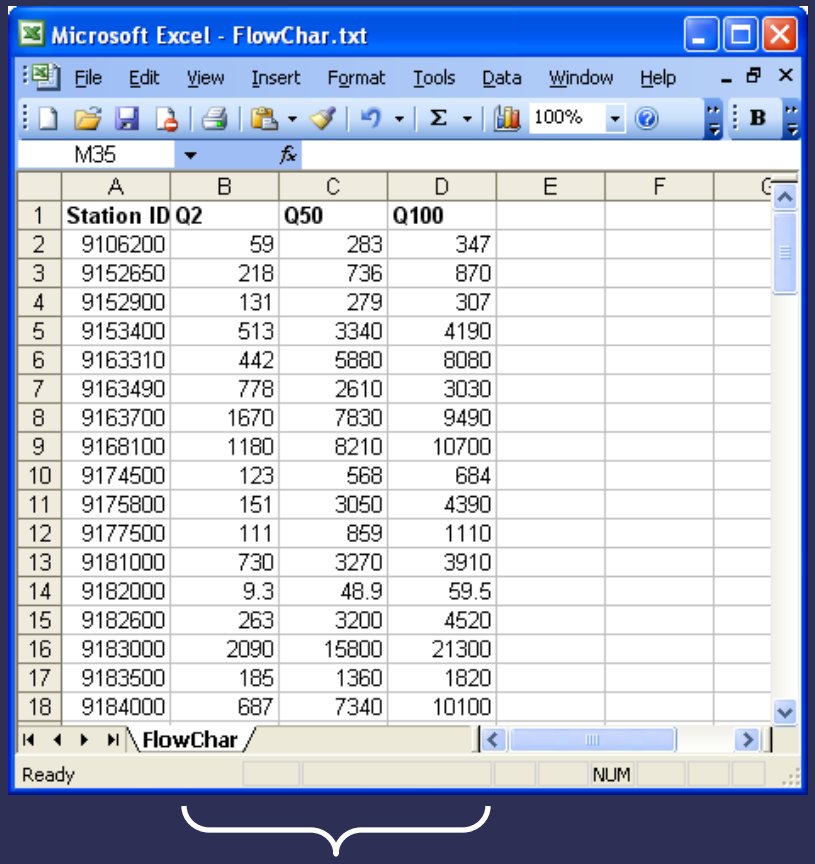
If frequency stats are not used, a dummy value must be supplied (for example, -99.99)



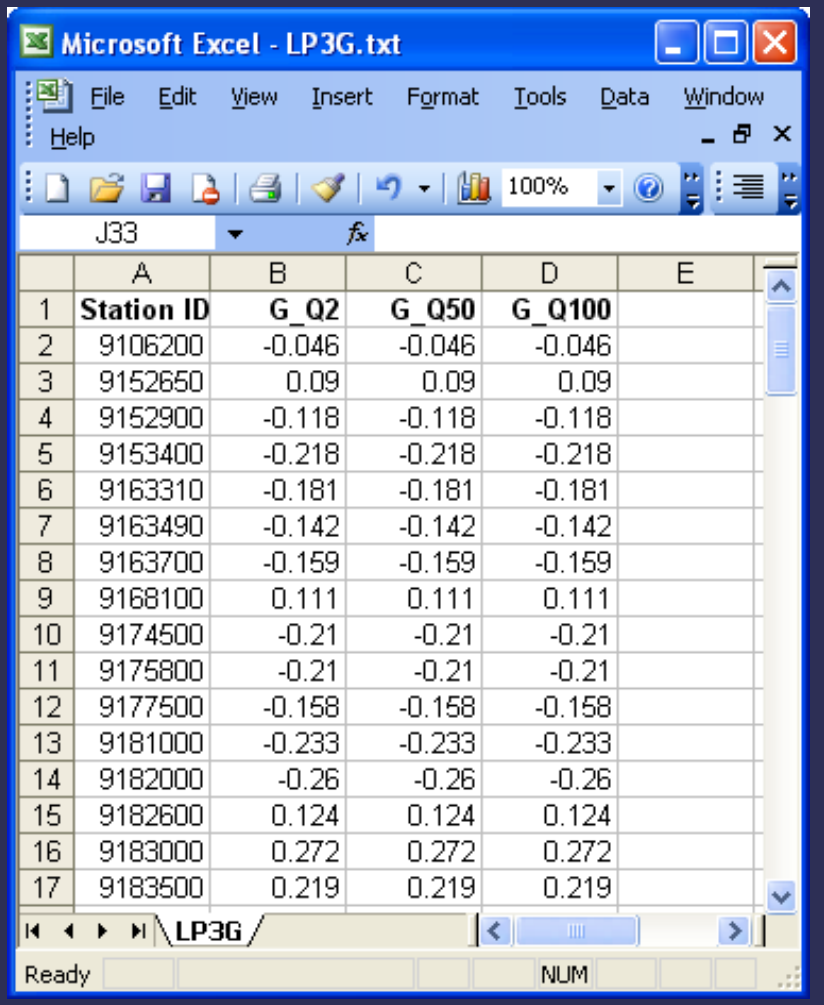
	A	B	C	D	E
1	Station ID	G_Q2	G_Q50	G_Q100	
2	9106200	-0.046	-0.046	-0.046	
3	9152650	0.09	0.09	0.09	
4	9152900	-0.118	-0.118	-0.118	
5	9153400	-0.218	-0.218	-0.218	
6	9163310	-0.181	-0.181	-0.181	
7	9163490	-0.142	-0.142	-0.142	
8	9163700	-0.159	-0.159	-0.159	
9	9168100	0.111	0.111	0.111	
10	9174500	-0.21	-0.21	-0.21	
11	9175800	-0.21	-0.21	-0.21	
12	9177500	-0.158	-0.158	-0.158	
13	9181000	-0.233	-0.233	-0.233	
14	9182000	-0.26	-0.26	-0.26	
15	9182600	0.124	0.124	0.124	
16	9183000	0.272	0.272	0.272	
17	9183500	0.219	0.219	0.219	

Order must correspond to that used in FlowChar.txt

LP3G, LP3K, LP3s.txt



	A	B	C	D	E	F
1	Station ID	Q2	Q50	Q100		
2	9106200	59	283	347		
3	9152650	218	736	870		
4	9152900	131	279	307		
5	9153400	513	3340	4190		
6	9163310	442	5880	8080		
7	9163490	778	2610	3030		
8	9163700	1670	7830	9490		
9	9168100	1180	8210	10700		
10	9174500	123	568	684		
11	9175800	151	3050	4390		
12	9177500	111	859	1110		
13	9181000	730	3270	3910		
14	9182000	9.3	48.9	59.5		
15	9182600	263	3200	4520		
16	9183000	2090	15800	21300		
17	9183500	185	1360	1820		
18	9184000	687	7340	10100		

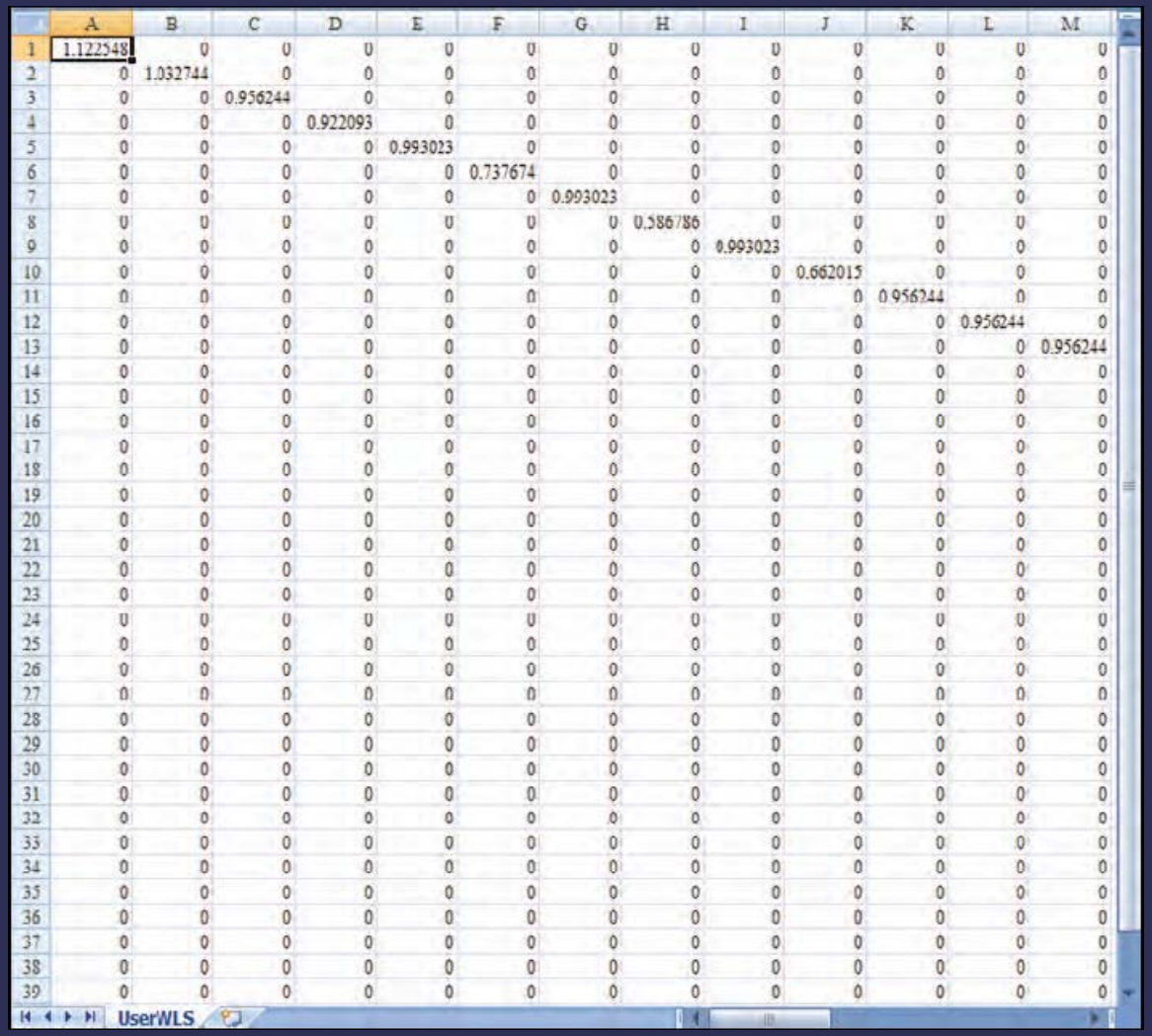


	A	B	C	D	E
1	Station ID	G_Q2	G_Q50	G_Q100	
2	9106200	-0.046	-0.046	-0.046	
3	9152650	0.09	0.09	0.09	
4	9152900	-0.118	-0.118	-0.118	
5	9153400	-0.218	-0.218	-0.218	
6	9163310	-0.181	-0.181	-0.181	
7	9163490	-0.142	-0.142	-0.142	
8	9163700	-0.159	-0.159	-0.159	
9	9168100	0.111	0.111	0.111	
10	9174500	-0.21	-0.21	-0.21	
11	9175800	-0.21	-0.21	-0.21	
12	9177500	-0.158	-0.158	-0.158	
13	9181000	-0.233	-0.233	-0.233	
14	9182000	-0.26	-0.26	-0.26	
15	9182600	0.124	0.124	0.124	
16	9183000	0.272	0.272	0.272	
17	9183500	0.219	0.219	0.219	

Order must correspond to that used in FlowChar.txt

UserWLS.txt

Contains matrix of weights for a user-defined WLS regression.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	1.122548	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1.032744	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0.956244	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0.922093	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0.993023	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0.737674	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0.993023	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0.586786	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0.993023	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0.662015	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0.956244	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0.956244	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0.956244
14	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0

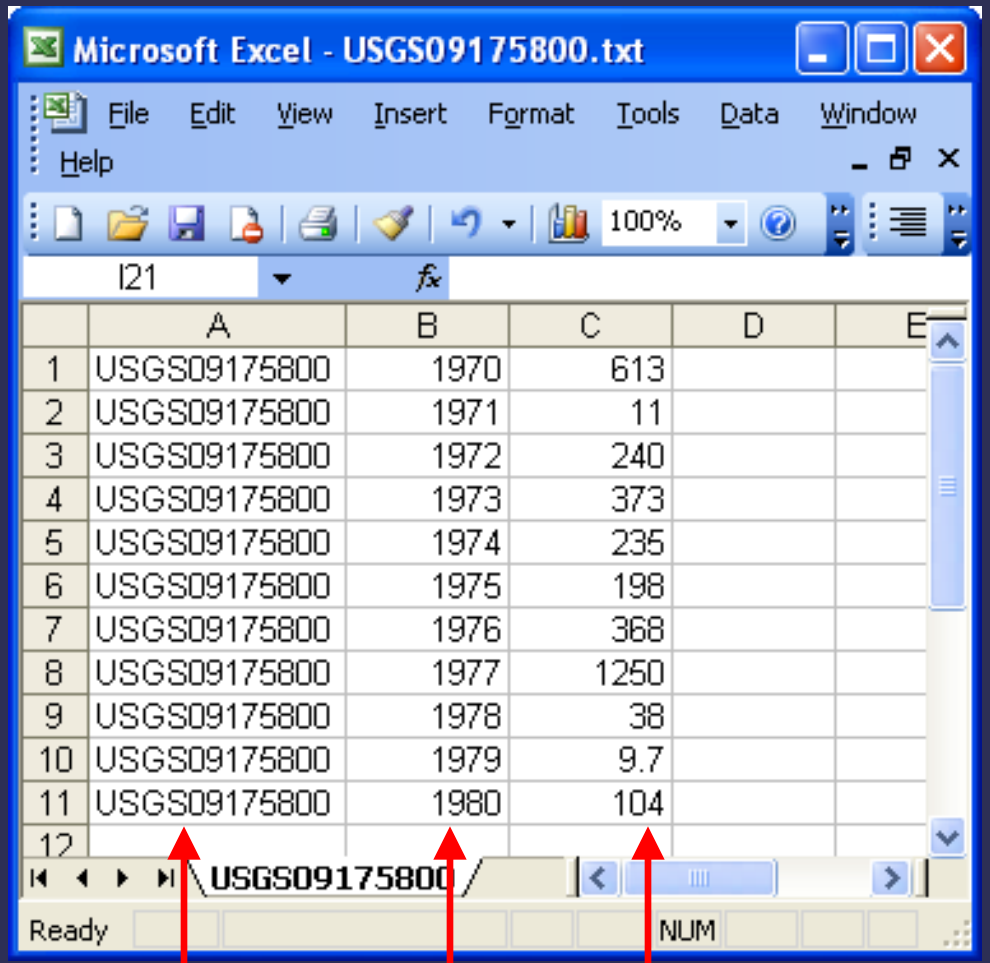
UserWLS

USGS#########.txt

Time series files: one file per station.

Naming convention: WREG looks for a file with "USGS" in its name. Then it uses these files in ascending numerical order.

Time series should correspond to that used to develop frequency statistic.



	A	B	C
1	USGS09175800	1970	613
2	USGS09175800	1971	11
3	USGS09175800	1972	240
4	USGS09175800	1973	373
5	USGS09175800	1974	235
6	USGS09175800	1975	198
7	USGS09175800	1976	368
8	USGS09175800	1977	1250
9	USGS09175800	1978	38
10	USGS09175800	1979	9.7
11	USGS09175800	1980	104
12			

Output Files

File name	Description
ConventionalOLS.txt	Results from a OLS regression.
ConventionalWLS.txt	Results from an WLS regression.
ConventionalGLS.txt	Results from a GLS regression.
RegionofInfluenceOLS.txt	Results from an OLS regression using RoI.
RegionofInfluenceWLS.txt	Results from a WLS regression using RoI.
RegionofInfluenceGLS.txt	Results from a GLS regression using RoI.
RegressionModel.txt	The regression equation (including transformations) calculated by WREG. Not output for ROI.
InvXLX.txt	Covariance of the regression parameters. Output only for conventional regression.
SSres.txt	Sum of squared residuals of the regression. Output only for conventional regression.
EventLog.txt	Record of the program's execution. Helpful for diagnosing runtime errors.

Inputs/Output limitations

- Executable file must reside in same directory as input files and output files

- Output files easily overwritten

- Input files somewhat cumbersome to create

Add-ons for WREG

WREG shell:

(Greg Granato, MA WSC)

- input/output handling
- error checking of input files
- run WREG from within shell program

Scripts for input file development :

(Charles Berenbrock, CA WSC)

-Takes output from PeakFQ and writes FlowChar.txt, LP3s.txt, LP3G.txt, and LP3K.txt files (batch jobs)

-Still reviewing, will post to website soon

WREG Shell Demo

USGS
science for a changing world

WREG versions

WREGv1.01:

Outputs regression coefficients to 3 decimal places (instead of 2)

WREGv1.02:

If it exists, reads in one large file of annual time series, rather than many individual files –WREG shell rewrites individual annual time series files (USGS########.txt) into one large file (USGSAnnualTimeSeries.txt).

New version is distributed with WREGshell program, but MCR_Installer.exe must first be run to install MATLAB runtime libraries.

DISCUSSION

Would additional webinars be useful?

- If so, how frequently?

- Suggested topics?

USGS
science for a changing world

Future webinar topics