OTHA – Omaha Tools for Hydrologic Analysis -Time-Series/Statistical Analysis Programs for Water Resources

William Doan, P.E.¹

Hydraulic Engineer, U.S. Army Corps of Engineers, Omaha District,

Abstract. A suite of programs has been developed by the Omaha District to provide useful tools to aid in the rapid and accurate analysis of hydrologic daily flow time-series data and peak flow data for various water resources/environmental projects. These programs utilize tools not currently found in existing software and all the OTHA programs have: easy-to-use Visual Basic GUI interfaces, FORTRAN engines, connections to WEB data locations, data conversion utilities, and utilize standard data formats: HEC-DSS databases for daily flows and HEC-FFA input formats for instantaneous peak flows. Various water resources related routines within OTHA include: time-series cross-correlation analysis, trend analysis, volume-frequency/balanced-hydrograph generation, extension of records, synthetic daily flow and peak flow generation, mix-population analysis, total probability theorem analysis, risk/binomial distribution analysis, precipitation-frequency analysis using General Extreme Value 1 distribution, and tests and corrects for stationarity or linear trends of statistical parameters.

Model Development. The programs contained in OTHA are predominantly software that had been previously developed by the Omaha District Corps of Engineers in the course of analyzing various water resources projects. The primary purpose of OTHA was to compile all the existing programs into one easy-to-use package and to be able to use these programs in conjunction with traditional statistical analysis software such as HEC-FFA and HEC-STATS. Specific model goals include:

- 1. Provide direct access to data via the Internet.
- 2. Provide means to convert data into a standardized format (DSS for daily data and HEC-FFA input file format for instantaneous peak flow data).
- 3. Minimize amount of required data input.
- 4. Direct display of program output.
- 5. Provide for annual or seasonal analysis when using daily values.
- 6. Provide a common "look or feel" for all routines.
- 7. Provide useful set of tools for working engineers.

Model Use. All routines within OTHA were written with FORTRAN and have Visual Basic graphical user interfaces. The routines within OTHA are called from a main menu or screen and are located within one of four main folders:

¹ Hydraulic Engineer, U.S. Army Corps of Engineers, Omaha District, (402) 221 4582, bill p.doon@useaco.ormy.mil

^{(402) 221-4583,} bill.p.doan@usace.army.mil

NOTHA - Time Series / Statistical Analysis Thomas J. Emerson Foundation	🗮 OTHA - Time Series / Statistical Analysis Thomas J. Emerson Foundation 💷 🗆 🗙
US A Mary Corps of Engineers Omaha District Time Series/Statistical Analysis Programs for Water Resources Data	U.S. Aury Corps of Engineers Dealba District OTHA - Omaha Tools for Hydrologic Analysis Time Series/Statistical Analysis Programs for Water Resources Data
Daily Flow Programs Intl. Peak Flow Programs Miccellaneous Programs Lirks to HEE Programs General Thop an Heamation QWAT Convect: USDS W1510/RE Head to molecular to ito an HECFFA input Flow QGEN Given two pages - will generate missing year peak flow for the 3rdb bared on the first stars (Peak Flow Writeward and Wile Elaboration Builting To Park Flow Builting To Park Flow Given two pages - will generate missing year peak flow for the 3rdb bared on the first stars (Peak Flow Builting To Park Flow Builting	Black Flow Pregnance Inst. Peak Flow Program Miccellaneous Program Links to HEC Program General Program DAIL/DDS Input US65 "Dialy" Data in the finance from thread from threa
Ext Relat	EXT Retail
OTHA - Ime Series / Statistical Analysis Thomas J. Emerson Foundation U.S. Any Corps of Engineers OTHA - Omaha Tools for Hydrologic Analysis U.S. Any Corps of Engineers Time Series/Statistical Analysis Programs for Water Resources Data	 ● OHMA - Time Series / Statistical Analysis ■ ▼ OTHA - Omaha Tools for Hydrologic Analysis I Dis Anay Corps of Engineers O Diaba District Time Series/Statistical Analysis Programs for Water Resources Data
Daily Flow Program Inst. Fisik Flow Program Mitcellaneout Program Links to HEC Programs General Program HECOSSUTL HECOSSUTL If you currently do not have DSSVUE, you can download it from HEC's Web-Site: http://www.bic.urace.amu.nll/cohmen.floc-dis/hecdscode.discue.html HECOFFA DDS - Let DDS - Let If you currently do not have DSSVUE, you can download it from HEC's Web-Site:	Dolp Flow Program: Mit Celleneover Program: Links to HEC Program Greened Program Information MEANPEAK. Computer the nation of instantements and develops mean ratio RISK. Computer total risk, of laske over project life and greene takan to figure over the program of the stantement of the stante over the program processing the stantement of the stante over the stante processing the stantement of the stante over the stante processing the stantement of the stante over the stante processing the stantement over the stantement processing the stantement over the stantement over the stantement processing the stantement over the stantement proces
Edt	ENT
Stop Restart	Stop Restart

Figure 1. OTHA Main Menus

The graphical user interface for each routine is shown below along with a brief description of each routine. For detailed descriptions of all routines, please refer to the OTHA Users Manual, which can be found at:

http://www.nwo.usace.army.mil/hydro/hydrology/otha/index.html

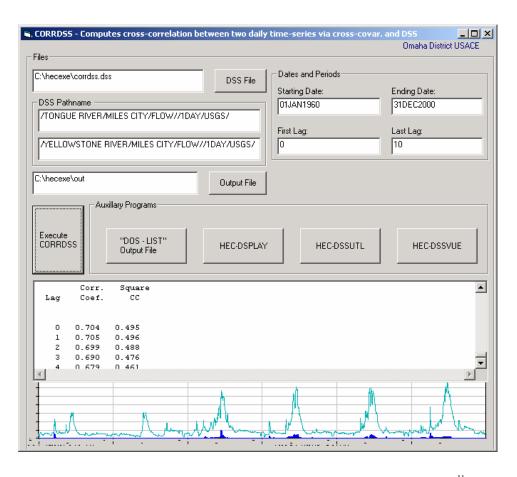
Daily Flow Routines:

DAILYDSS – Retrieves daily flow data from USGS Web-Site and converts the data into HEC-DSS daily time-series. The data from the Web Site needs to be saved as a "Tab-Separated Data File". The name of the DSS file needs to be inputted, as well as the DSS pathname for the data to be placed in are required input.

USGS Web-Site for Daily Data (Save as "Tab-seperated data <u>http://waterdata.usgs.go</u>	file'')	2	19 FEB 04 11:07:29		
C:\hecexe\dailydss.dat Input File (from above		Input File (from above)	/TONGUE RIV CITY/FLOW//1:		
:\hecexe\out		Output File		,,	
:\hecexe\dailydss.dss		DSS File to put data into		Flow	Flow
DSS Pathname to Give Data:			Date	From Text File:	Goes into DSS as:
/TONGUE RIVER/MILES CI	TY/FLOW//1DAY	//USGS/			
			4 2 1938 4 3 1938	150 145	150.00 145.00
			4 3 1938	145	145.00
			4 5 1938	155	155.00
Execute DAILYDSS			4 6 1938	145	145.00
i			4 7 1938	145	145.00
Auxillary Programs			4 8 1938	150	150.00
			4 9 1938	155	155.00
"DOS - LIST"			4 10 1938	165	165.00
Output File	HEC-DS	SSUTL	4 11 1938	175	175.00
oaparrio			4 12 1938	177	177.00
			4 13 1938	166	166.00
			4 14 1938	161	161.00
HEC-DSPLAY	HEC-DS	SSVUE	4 15 1938	171	171.00
			4 16 1938	182	182.00
			4		Þ

DSSSTATS – Quickly and concisely summarizes a long-term period of record of daily flows for any period within a year into: annual minimum flows, annual maximum flows, annual mean flows, annual flow-volumes, annual lag-1 serial or autocorrelation coefficients, and computes graphical frequency-analysis, log-Pearson probability distribution, and trend analysis on annual maximum daily flow values with associated statistical significance tests. Writes out annual maximum flows to an FFA input file for further analysis.

les			an Daily Values				Omaha Di	strict USACE
\hecexe\a	lssstats.dss		DSS	File Dates	and Periods			
					g Date:	E	nding Date:	
				01JAI	V1960	3	31DEC2000	
DSS Pathn	ame ——							
		BEND/FLOW//1D	AV/UCCC/	Startin	g Period:	E	nding Perioc	l:
PLATIE	1./NUN I H	BEND/FLUW//ID	AT/USUS/	01JAI	N	3	BIDEC	
:\hecexe\d	out		Output	File C:\heci	exe\infile.ffa			FFA Input File to Crea
	- Auxilla	ary Programs						
						1		
Execute DSSSTATS		"DOS - LIST" Output File	HEC-I	DSPLAY	HEC-DSSI	лть 📔	HEC-D	SSVUE
Statist:	ical Ana	lysis:						
num		date		date			lag-1	
of	min	of	max	of	accum	mean	serial	
values	value	min value	value	max value	value*	value	corr.	
366	663.00	22DEC1960	67600.00	29MAR1960	3135.66	4326.96	0.8564	
365	530.00	24JUL1961	12400.00	24MAY1961	2389.95	3306.98	0.9137	
365	500.00	11DEC1962	32300.00	26MAR1962	3645.69	5044.55	0.9185	
<u> </u>								
-			1 1					
.			والأستعمان ا			also a		A. Land



$$C_{k}^{ij} = (\frac{1}{N}) \sum_{i=1}^{N-k} (y_{t+k}^{i} - y_{mean}^{i}) (y_{t}^{j} - y_{mean}^{j}) r_{k}^{ij} = \frac{C_{k}^{ij}}{(C_{o}^{ii} C_{o}^{jj})^{1/2}}$$

CORRDSS- Calculates the cross-correlation of two hydrologic daily timeseries using the cross-covariance between the two hydrologic time-series. (Salas, 1992) Essentially, determines the probability that two rivers will flood at the same time – useful for main/tributary confluences and interior drainage analysis.

where:

- y^{1} = one stream gage's time-series of daily flows
- y^J = another stream gage's time-series of daily flows
- c_o = sum of the differences of daily flows from the mean of the daily flows, divided by the total number of observations

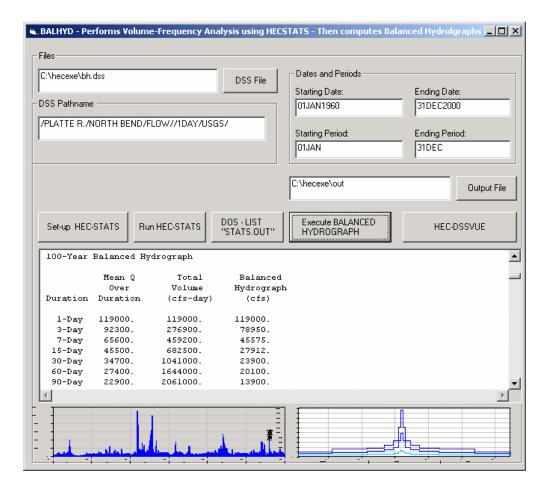
t = time increment

 $r_k = lag-k$ cross-correlation coefficient

c=cross-covariance between the two time-series

k=times-series lag

N=total number of values

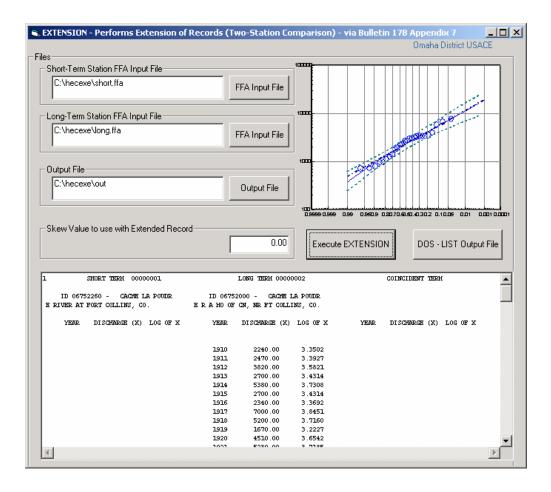


BALHYD – Serves as a graphical interface for HEC-STATS for flow volume-frequency analysis, then computes various return-interval symmetrical hypothetical "balanced" hydrographs and writes the ordinates out to HEC-DSS. Specific periods within the year can be computed, i.e. monthly or seasonal. The generated balanced 10, 25, 50, 100, and 500-year hydrographs are written to the specified DSS file.

Peak-Flow Routines:

QWAT – Retrieves and converts USGS WATSTORE file format from the USGS's Web Site into an HEC-FFA input file format (QR Cards).

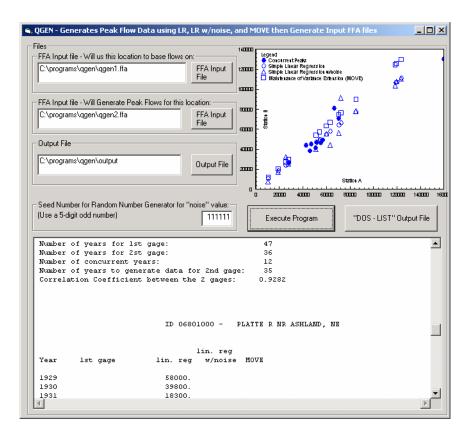
USGS Web-Site for F	eak Flow Data		ΓA	uxillary Programs
(Save as "WATSTO	BE Formatted File")			, ,
	lata.usgs.gov/nwis/peak			
				"DOS - LIST"
:\hecexe\gwat.txt		Input File		Output File
.: \necexe\qwat.txt		(from above)		oupurno
		(nom above)		
:\hecexe\out2ffa.ffa		FFA Input File		
		to Create		
				1150 551
				HEC-FFA
	Execute QWAT			
T 06294900	M F Froze to Dea	th Cr trib 🔺	600 m	
nr Ingomar MT			L L L	USGS 06294900
TT 06294900	Drainage Area =	1.36	L L	M F Froze to Death Cr trib nr Ingomar MT
ID 06294900 -	M F Froze to Dear	th Cr trib	500 -	in the constant of the firmingentarism
nr Ingomar MT			- F	A
QR 716196			400	
			400 -	
QR 621196				
DR 618196			300 -	▲ ·
)R 618196)R 710196				
)R 618196)R 710196)R 715196	6 52		t	
DR 618196 DR 710196 DR 715196 DR 618196	6 52 7 67			
IR 618196 IR 710196 IR 715196 IR 618196 IR 69196	6 52 7 67 8 54		200 -	
IR 618196 IR 710196 IR 715196 IR 618196 IR 69196 IR 77196	6 52 7 67 8 54 9 90		200	
IR 618196 IR 710196 IR 715196 IR 618196 IR 69196 IR 77196 IR 86197	6 52 7 67 8 54 9 90 0 132		Ē	
IR 618196 IR 710196 IR 715196 IR 618196 IR 69196 IR 77196 IR 86197 IR 531197	6 52 7 67 8 54 9 90 0 132 1 45		200 -	. .
618196 JR 710196 JR 715196 JR 618196 JR 619196 JR 6 9196 JR 7 7196 JR 8 6197 JR 531197 JR 814197	6 52 7 67 8 54 9 90 0 132 1 45 2 110		Ē	A AAA ^A AAA
IR 618196 IR 710196 IR 715196 IR 618196 IR 6 9196 IR 7 7196 IR 8 6197 IR 531197	6 52 7 67 9 90 0 132 1 45 2 110 3 60		Ē	



EXTENSION - Extension of records or two-station comparison of peak flows. Utilizes the procedures described in Appendix 7 of Bulletin 17B for extending on gage's record based on the concurrent period of a nearby gage. The program adjusts the log mean flow and standard deviation of the short-term record based on the regression analysis with the long-term record. An inputted skew value can be used with the adjusted parameters to develop the final flow-frequency curve.

(Hydrology Subcommittee, 1982)

 $y_{t} = a + b(x_{t} - \mu_{t})$



QGEN - Will generate individual year's peak flows for a given station by comparison with nearby station. Will generate missing year flows by linear regression, linear regression with "noise", and MOVE statistics. (Salas, 1991) The program puts the original data combined with the synthetic data into three FFA input files.

$$a = \frac{(N_1 + N_2)\mu_y - N_1\mu_{y_1}}{N_2}$$

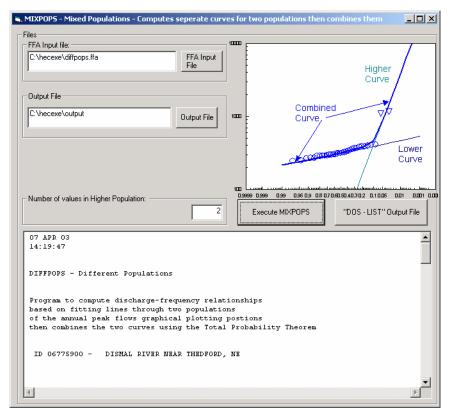
$$\mu_y = \mu_{y_1} + \frac{N_2}{N_1 + N_2}\rho_{xy}\frac{\sigma_{y_1}}{\sigma_{x_1}}(\mu_{x_2} - \mu_{x_1})$$

$$\alpha^2 = \frac{N_2(N_1 - 4)(N_1 - 1)}{(N_2 - 1)(N_1 - 3)(N_1 - 2)}$$

$$b^2 = \frac{[(N_1 + N_2 - 1)\sigma_y^2 - (N_1 - 1)\sigma_{y_1}^2 - N_1(\mu_{y_1} - \mu_y)^2 - N_2(a - \mu_y)^2]}{(N_2 - 1)\sigma_{x_2}^2}$$

$$\sigma_y^2 = \frac{1}{N_1 + N_2 - 1}[(N_1 - 1)\sigma_{y_1}^2 + (N_2 - 1)\rho_{xy}^2\frac{\sigma_{y_1}^2}{\sigma_{x_1}^2}\sigma_{x_2}^2 + \frac{N_1N_2}{N_1 + N_2}\rho_{xy}^2\frac{\sigma_{y_1}^2}{\sigma_{x_1}^2}(\mu_{x_2} - \mu_{x_1})^2$$

X is the set of annual peak flows for the long-term gage Y is the set of annual peak flows for the short-term gage μ_y and σ_y are the mean and standard deviation of the extended sequence y_t is the variable with a missing variable to be generated μ_{y1} is the mean of concurrent observations of Y ρ_{xy} is the cross-correlation between X and Y σ_{y1} is standard deviation of concurrent observations of Y σ_{x1} is standard deviation of concurrent observations of X σ_{x2} is standard deviation of non-concurrent observations of X x_t is the variable with the complete record μ_{x1} is the mean of concurrent observation of X N₁ is the mean of non-concurrent observation of X N₁ is the number of common observations between X and Y N₂ is the number of values of Y to generate and:

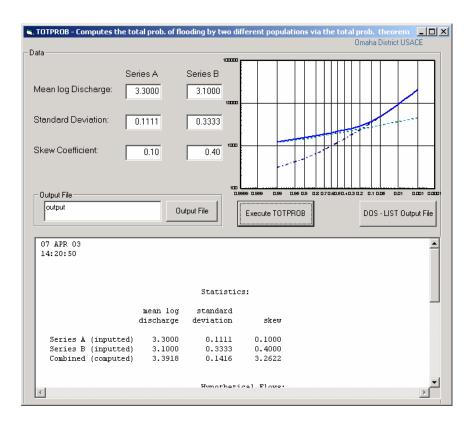


MIXPOPS - For two physically-differentiated populations within one FFA record, the program will compute separate flow-frequency curves for each population then combines the two curves using the Total Probability Theorem. The curves are developed by converting the graphical plotting positions to a linear distance on the probability grid and solves for a best-fit line. (HEC, 1975)

$$P_{Combined} = P_A + P_B - P_A P_B$$

where:

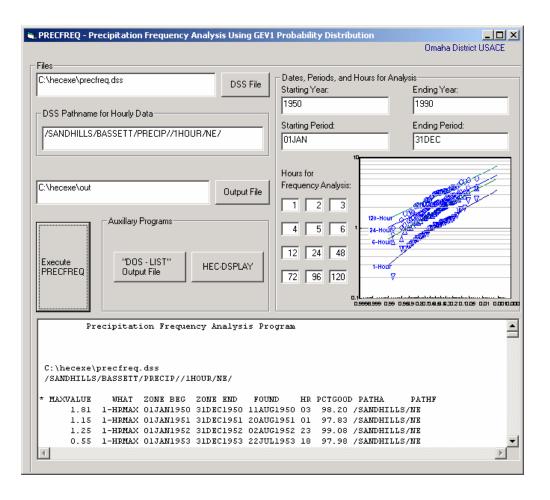
 P_A = Probability of "A" occurring and P_B = Probability of "B" occurring



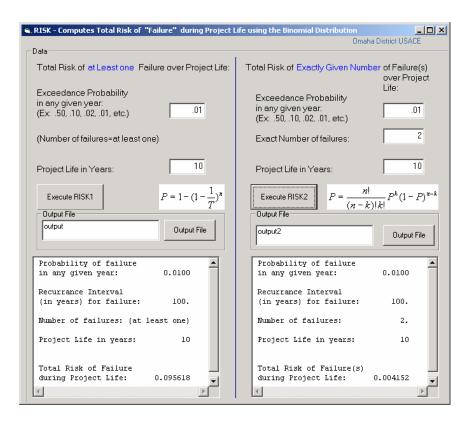
TOTPROB - Given the statistical parameters of mean log discharge, standard deviation, and skew for two different populations, the program will compute the log-Pearson Type III probability distribution for each, then combine the two curves for a smooth curve using the Total Probability Theorem.

$$P_{Combined} = P_A + P_B - P_A P_B$$

<u>Miscellaneous Routines:</u>



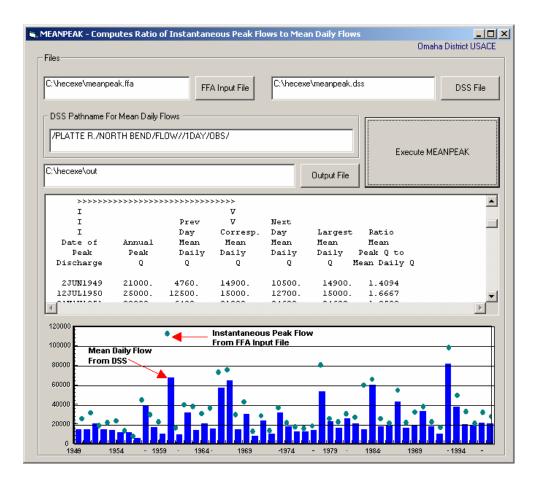
PRECFREQ - Precipitation-frequency analysis of hourly precipitation data using the General Extreme Value 1 probability distribution. The program sorts the hourly DSS records to determine the max "n-hour" for each year then computes graphical frequency analysis using the Weibull plotting positions as well as the GEV-1 analytical solution. Specific periods within the year can be computed, i.e. monthly or seasonal.



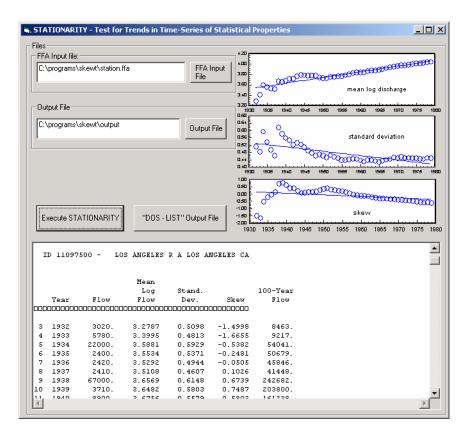
RISK - Computes the total risk of failure over a project life using the Binomial Probability Distribution (Chow, Maidment, Mays, 1988) based on the following equations:

$$R = 1 - (1 - \frac{1}{T})^{n}$$

$$R = \frac{n!}{(n-k)!k!}P^{k}(1-P)^{n-k}$$
where:
R = Risk of occurance
P = Probability in any year
T = Recurrance interval
K = Number of floods
N = Number of years



MEANPEAK - Computes the ratio of mean daily flows to instantaneous peak flows given a FFA input file and mean daily flow DSS file. The programs takes the peak flow from the FFA input file for each year, then searches the mean daily flow DSS file to determine the corresponding mean daily flow for the day the instaneous peak flow occurred and computes the ratio. The program performs linear regression between the ratio of the peak flow to mean daily flow as a function of the mean daily flow or time.



STATIONARITY - Computes time-series of statistical properties of mean log discharge, standard deviation, skews, as well as the analytically-based 100-year flow as a test for basin stationarity. Allows to see how the 100-year flow estimate changes with time and to readily see which statistical parameters influenced the changes. Performs trend-analysis on statistical parameters and tests the statistical significance of the trends using the Students t-test:

Test Statistic:

$$T_c = \frac{b\sigma_x \sqrt{N-1}}{\sigma_\varepsilon}$$

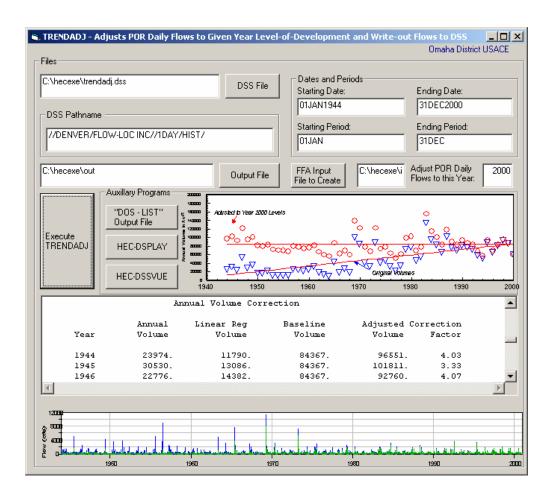
b = slope of regression curve

 σ_x = standard deviation of parameter

 σ_{ε} = standard deviation of residuals between actual values and regression values

N = number of values

Hypothesis test: b=0 is rejected if: $|T_c| > T_{1-\frac{\alpha}{2}}(N-2)$ where α is the significance level.



TRENDADJ - Computes trends in annual flow volume (based on daily flows). If there is a trend, program will adjust the daily flows of the entire period-of-record to a given year level-of-development. The modified or adjusted period-of-record daily flows are written back out to HEC-DSS.

:\tri-lakes\sams\seas_1.txt		Input File	20 FEB 04
		(from SAMS)	13:53:11
:\hecexe\out		Output File	/CHATFIELD RESERVOIR/CHFI/FLOW/31JAN1942/1MON/Trace
			#1/ 25.73
:\hecexe\samsdss.dss		DSS File to	/CHATFIELD
		put data into	RESERVOIR/CHFI/FLOW/28FEB1942/1MON/Trace
	ALC D. N. 1.		#1/ 60.71
DSS Pathname to Give Data:			/CHATFIELD
/CHATFIELD RESERVOIR/	CHFI/FLOW//1MON	177	RESERVOIR/CHFI/FLOW/31MAR1942/1MON/Trace
			#1/ 10.72
			/CHATFIELD
Beginning Year to put Stocha	istic Monthly Data: -		RESERVOIR/CHFI/FLOW/30APR1942/1MON/Trace
19	42		#1/ 45.70 /CHATFIELD
_ · ·	_		RESERVOIR/CHFI/FLOW/31MAY1942/1MON/Trace
			#1/ 91.22
	LIODOO I		/CHATFIELD
Execute SA			RESERVOIR/CHFI/FLOW/30JUN1942/1MON/Trace
1			#1/ 0.00
Auxillary Programs			/CHATFIELD
			RESERVOIR/CHFI/FLOW/31JUL1942/1MON/Trace
"DOS - LIST"			#1/ 92.42
Output File	HEC-DSS	UTL	/CHATFIELD
			RESERVOIR/CHFI/FLOW/31AUG1942/1MON/Trace
			#1/ 193.39
	UEC DOC		/CHATFIELD
HEC-DSPLAY	HEC-DSS		RESERVOIR/CHFI/FLOW/30SEP1942/1MON/Trace
			#1/ 209.16
166			

SAMSDSS – An interface that inputs stochastically generated monthly synthetic data generated by *Stochastic Analysis, Modeling, and Simulation (SAMS)* into HEC-DSS. Any combination of years of simulations and number of simulations will automatically be put into HEC-DSS where it may be directly utilized by other software, such as HEC-DSSVUE, HEC-5, or HEC-RESIM.

HEC Program Links:

A folder was created within OTHA to allow easy access to existing HEC programs. OTHA will make the links to the files if they are already loaded on the PC.

🖷, OTHA - Time Series ,	/ Statistical Analysis	Thomas J. Emerson	Foundation	. 🗆 🗙
U.S. Army Corps of Engineers Omaha District	OTHA - Omaha Tools for Hydr Time Series/Statistical Analy for Water Resources	sis Programs		3
Daily Flow Programs	Inst. Peak Flow Programs Miscellaneous Programs	Links to HEC Programs	General Program Information	
HEC-DSPLAY				
HEC-DSSUTL				
HEC-DSSVUE	If you currently do not have DSSVUE, you ca http://www.hec.usace.army.mil/softwar			
HEC-FFA				
DOS - List				
	EXIT			
Stop	EXII		Be	estart

General Program Information:

A built-in User's Manual, web-link, and mail link reside on the General Information Page.



Conclusion

All routines have example data sets, which are also the default values for the GUI interfaces. The routines within OTHA were developed by the Omaha District for various uses on water resources and environmental projects, but anyone may download and use the program. A link to the Web-Site is found below:

http://www.nwo.usace.army.mil/hydro/hydrology/otha/index.html

References

Chow, Maidment, Mays (1988). Applied Hydrology.

- HEC (1975). Hydrologic Engineering Methods for Water Resources Development, Volume 3 Hydrologic Frequency Analysis.
- Hydrology Subcommittee (1982). Bulletin #17B Guidelines For Determining Flood Flow Frequency.

Salas, Jose D. (1991). Statistical Computer Techniques in Hydrology and Water Resources.

- Salas, Jose D. (1993). Handbook of Hydrology, Chapter 19 Analysis and Modeling of Hydrologic Time Series.
- Salas, J.D., Saada, N., Chung, C.H., Lane, W.L. and Frevert, D.K., 2000, "Stochastic Analysis, Modeling and Simulation (SAMS) Version 2000 - User's Manual", Colorado State University, Water Resources Hydrologic and Environmental Sciences, Technical Report Number 10, Engineering and Research Center, Colorado State University, Fort Collins, Colorado.