Hydrologic Engineering Center

Training Course on

STATISTICAL ANALYSIS IN WATER RESOURCES

8 – 12 April 2019
Davis, California

Course Objectives

The objectives of this course are to provide the participant with the theoretical background and practice with analysis tools required for practical application of techniques for univariate and multivariate frequency analysis, regional analysis of hydrometeorological data and regression analysis. Data utilized in this course come from a variety of sources relevant to water resources management such as streamflow, precipitation, snow measurement and more. HEC computer programs and commercial software packages will be used in workshop application problems.

HEC Instructors
Mike Bartles
Beth Faber
Greg Karlovits

Guest Instructors
Chuck Parrett, USGS

SUMMARY

Day		Topic
Monday	am	Introduction to Probability, Statistic and Distributions (Workshop exercises during intro lecture)
	pm	Exploratory Data Analysis; and Use of HEC-SSP
Tuesday	am	Parametric Modeling; Extreme Value theory
	pm	Extreme Value workshop, Bulletin 17C and EMA
Wednesday	am	Monte Carlo Analysis theory and exercises, Uncertainty in frequency analysis, Regional Skew
	pm	Regionalization – Index Flood; Precipitation Frequency Analysis
Thursday	am	Multivariate Analysis; Coincident Frequency Analysis,
	pm	Linear Regression & workshop and Multiple Linear Regression; Applications of Regression
Friday	am	Multiple Linear Regression Application Workshop

STATISTICAL METHODS IN HYDROLOGY

Monday, April 8

8:00 – 8:15 am	Introductions
8:15 – 8:30 am	**Administrative Tasks**
8:30 – 11:30 am	1.1 Lecture: **Introduction to Probability, Statistics and Distributions** (Faber) Objectives of statistical analyses; definitions; types of variables; histograms; cumulative distributions; concepts of parent population; parameters used to define distributions - mean, median, mode, standard deviation, skew coefficient. Use of binomial equation to determine risk. 1.1 Workshop: **Within-Lecture exercises** Exercises are interspersed within Lecture 1.1 Exploration of using relative frequency to estimate probability, and fitting probability distributions to data.
11:30 – 12:30 pm	Lunch
12:30 – 1:15 pm	1.2 Lecture: **Exploratory Data Analysis** (Karlovits) How to examine data for probability studies. Suggest hypotheses about the causes of observed phenomena; Assess assumptions on which statistical inference will be based; Support the selection of appropriate statistical tools and techniques; Provide a basis for further data collection through surveys or experiments
1:15 – 2:30 pm	1.3 Workshop: **Visualize your Data** (Karlovits) Students will explore an unfamiliar multivariate dataset to visualize the data, form hypotheses about the data, and create recommendations regarding the modeling of the data.
2:30 – 2:45 pm	Break
2:45 – 3:45 pm	1.4 Lecture: **Introduction to HEC-SSP** (Bartles) Overview and demonstration of the Statistical Software Package (HEC-SSP) and its various uses.
3:45 – 5:00 pm	1.5 Workshop **HEC-SSP for time-series analysis** (Bartles) Introduction to HEC-SSP with multiple examples of importing, viewing, and manipulating data. Students will then create, compute, and view results of

various analyses.

Tuesday, April 9

8:00 – 9:00 pm 2.1 Lecture: **Parametric Modeling** (Bartles)

Introduction to parametric modeling and fitting analytical probability distributions. Discuss advantages, disadvantages, and data requirements when performing parametric modeling. Provide examples of commonly used distributions and fitting methods.

9:00 – 10:45 pm 2.2 Workshop: **Parametric Modeling Workshop** (Bartles)

Students will utilize HEC-SSP to analyze snow depth, snow water equivalent, and precipitation data. Exercises will focus on data analysis, parametric modeling, and results visualization.

10:45 – 11:00 am Break

11:00 – 12:00 pm 2.3 Lecture: **Extreme Value Theory** (Karlovits)

Role and importance of extremes in hydrologic analysis. Models, data and assumptions when modeling extremal data. Development of extremal datasets using block maxima and peaks-over-threshold. Using the First and Second Extreme Value Theorems to choose the appropriate model.

12:00 – 1:00 pm Lunch

1:00 – 2:15 pm 2.4 Workshop: **Extreme Value Theory Workshop** (Karlovits)

Developing an extremal dataset and estimating extreme-value models for hydrometeorological data

2:15 – 3:15 pm 2.5 Lecture: **Flood Frequency Guidance, Bulletins 17B and 17C** (Faber)

Bulletin 17B for Flood Frequency Analysis has been updated to Bulletin 17C. Topics include parameter estimation methods (EMA), treatment of outliers and historical information, confidence intervals and new regional skew studies performed by USGS.

3:15 – 3:30 pm Break

3:30 – 5:00 pm 2.6 Workshop: **Bulletin 17C in HEC-SSP** (Bartles)

Students will utilize HEC-SSP to estimate peak flow-frequency using a data set that contains low outliers, historical information, and paleoflood data.

Wednesday, April 10

8:00 – 10:00 pm	3.1 Lecture:	**Monte Carlo Analysis** (Faber)
	Replacing distribution with sample, integration, example applications, Risk-based uncertainty analysis.	
	3.1 Workshop: **Within-Lecture exercises**	
10:00 – 10:15 am	break	
10:15 – 11:45 am	3.2 Lecture:	**Uncertainty in Frequency Estimates** (Faber)
	Understanding uncertainty in estimation of flood frequency curves; sampling errors of mean, standard deviation, and skew; confidence limits.	
11:45 – 12:45 pm	Lunch	
12:45 – 1:45 pm	3.3 Lecture:	**Regionalization: Skew Coefficient** (Faber)
	Need for Regionalization: uncertainty in regional frequency curves, regional skew. Worth of regionalized results, sensitivity of process to regional values, mapping techniques, applications to ungaged areas,	
1:45 – 3:00 pm	3.4 Lecture:	**Regionalization: Index Flood** (Karlovits)
	Application of the index flood and related methods for regionalization of hydrologic data. Discussion of L-moments and their role in frequency analysis.	
3:00 – 3:15 pm	break	
3:15 – 4:30 pm	3.5 Lecture:	**Precipitation Frequency Analysis** (Karlovits)
	Analysis of precipitation extremes using regional frequency analysis with L-moments. Additional variants for regionalizing data. Choosing a regional probability distribution. Insights into NOAA Atlas 14. Recent advances in precipitation-frequency analysis.	

Thursday, April 11

8:00 – 9:00 am 4.1 Lecture: **Multivariate Analysis** (Karlovits)

Review of marginal, conditional and joint probability. Exploration of covariance and correlation. Joint probability distributions and modelling joint probability with copulas.

9:00 – 10:00 am 4.2 Lecture: **Coincident Frequency Analysis** (Faber)

Situations in which coincident frequency analysis is required, assumptions and technique to compute coincident frequency by total probability theorem. Use of Monte Carlo simulation when correlation cannot ignored

10:00 – 10:15 am break

10:15 – 11:45 am 4.3 Workshop: **Coincident Frequency in HEC-SSP** (Faber)

Use of SSP for Coincident Frequency Analysis with independence. Correlated analysis in Excel.

11:45 – 12:45 pm Lunch

12:45 – 1:45 pm 4.4 Lecture: **Simple Linear Regression** (Karlovits)

Development of ordinary least squares regression from the generalized linear model. Evaluating model results, checking assumptions and fixing issues.

1:45 – 2:45 pm 4.5 Workshop: **Simple Linear Regression** (Karlovits)

Development of a simple linear regression model in Excel; demonstration of model development in R.

2:45 – 3:00 pm Break

3:00 – 4:00 pm 4.6 Lecture: **Multiple Linear Regression** (Karlovits)

Extension of the simple linear regression model. Stepwise regression, the art and science of choosing predictors, evaluating model results, checking assumptions and fixing issues.

4:00 – 5:00 pm 4.7 Lecture: **Applications of Multiple Linear Regression** (Parrett)

Application of regression in hydrologic analyses. Streamflow record extension; formulation of regression models, effects of non-linearities, transformations, interpreting results.

Friday, April 12

8:00 – 10:15 am	5.1 Workshop: **Multiple Linear Regression** (Karlovits) Development of a multiple linear regression model in StatTools.
10:15 – 10:30 am	break
10:30 – 11:30 am	**Critique and Completion Activities**