GEOSTATISTICS

MSc in Petroleum Geoengineering

First semester 2018/2019

COURSE COMMUNICATION DOCUMENT

University of Miskolc
Faculty of Earth Science and Engineering
Institute of Geophysics and Geoinformatics

Course datasheet

Course Title: Geostatistics	**Credits: 3**
Type (lec. / sem. / lab. / consult.) and Number of Contact Hours per Week: **lec. 2, sem. 1**	
Neptun code: MFGFT710007	
Type of Assessment (exam. / pr. mark. / other): **exam (oral)** Attendance at lectures is regulated by the university code of education and examination. Writing two tests during the semester and preparing one powerpoint presentation on an assigned topic (condition of signature). **Grading limits:** >86%: excellent, 71-85 %: good, 61-70 %: medium, 46-60 %: satisfactory, <45 %: unsatisfactory.	
Position in Curriculum (which semester): **first**	
Pre-requisites *(if any)*:	
Course Description:	
Acquired store of learning: Study goals: The course provides an introduction to the principles and hydrocarbon applications of mathematical statistical methods, which equips the students with necessary skills to apply statistical methods in building both deterministic and stochastic reservoir models. Course content: The probability density function (pdf) and cumulative distribution function (cdf). Gaussian and non-Gaussian data distributions. The most frequent value method as robust statistical estimator. The linear and rank correlation coefficient. Covariance and correlation matrices. Linear and non-linear regression analysis. Spatial correlation of petrophysical parameters, variogram models and kriging. Multidimensional scaling, modeling and data analysis. Hierarchical and non-hierarchical cluster analysis, the K-means clustering method. Principal component analysis, factor analysis and their applications in petroleum geosciences. Linearized and global optimization methods and their statistical aspects. Discrete inverse theory and its application to geophysical datasets. Evolutionary computation algorithms. The calculation of the estimation error of model parameters. Characterization of the accuracy and reliability of the inversion result. Theory of neural networks. Education method: Lectures with projected MS-PowerPoint presentation. Demonstration of statistical methods using own developed MATLAB codes (recipes) and the MATLAB Statistical Toolbox. **Competencies to evolve:** T1, T4, T5, T7, T8, T10, T11, T12, K1, K2, K3, K5, K7, K8, K10, A1, A9, F2	
The 3-5 most important compulsory, or recommended **literature** (textbook, book) **resources**:	
• Edward H. Isaacs, R. Mohan Srivastava, 1989. An introduction to applied geostatistics. Oxford University Press. • Troyan V., Kiselev J., 2010. Statistical methods of geophysical data processing. World Scientific Publishing Co. • Reyment R. A., Jöreskog K. G., 1996. Applied factor analysis in the natural sciences. Cambridge Univ. Press.	

- Csernyák L., Hajagos B., Hursán G., Steiner F., Szűcs P., Turai E., 1997. Optimum methods in statistics. Akadémiai Kiadó, Budapest.
- Clark I., 1979: Practical geostatistics. Elsevier Applied Science.
- Szabó N. P., 2017. Geostatistics. Electronic textbook. http://www.uni-miskolc.hu/~geofiz/education.html

Responsible Instructor (*name, position, scientific degree):*

Norbert Péter Szabó Dr., associate professor, PhD, dr. habil.

Course schedule

Date	Lecture
12 September	Data distributions. Datasets, histograms, pdf and cdf types. Determination of the mode. Characterization of uncertainty. The most frequent value method as robust statistical estimator.
Holiday	-
26 September	The Maximum Likelihood method. Confidence intervals. Skewness and kurtosis. Propagation of error. The linear and rank correlation coefficient. Multivariate linear functional relationships. Covariance and correlation matrices.
3 October	Linear and non-linear regression analysis. Robust regression methods. Petrophysical examples. Spatial correlation of petrophysical parameters, variogram models and kriging.
10 October	Writing test 1. Introduction to multivariate statistical methods. Multi-dimensional samples and scaling. Multidimensional modeling and data analysis.
17 October	Hierarchical and non-hierarchical cluster analysis, the K-means clustering method. Petrophysical examples.
24 October	Reduction of dimensionality. Principal component analysis, factor analysis and their applications in petroleum geosciences. Shale volume and permeability estimation.
31 October	Discrete inverse theory. Linearized and global optimization methods. Geophysical inverse problems.
7 November	Linear regression using inversion tools. The Gaussian Least Squares method. Weighted norms to be minimized. Well-logging applications.
14 November	The quality check of inversion results. The relation between the data and model covariance matrices.
21 November	Simulated Annealing methods. Classical and float-encoded genetic algorithm. Artificial neural networks.

28 November	Simulated conference I.
5 December	Writing test 2. Simulated conference II.
12 December	Repeating the writing tests.

Date	**Seminar**
12 September	Computer practice using MATLAB recipes.
Holiday	Computer practice using MATLAB recipes.
26 September	Computer practice using MATLAB recipes.
3 October	Computer practice using MATLAB recipes.
10 October	Writing test 1.
17 October	Computer practice using MATLAB recipes.
24 October	Computer practice using MATLAB recipes.
31 October	Computer practice using MATLAB recipes.
7 November	Computer practice using MATLAB recipes.
14 November	Computer practice using MATLAB recipes.
21 November	Computer practice using MATLAB recipes.
28 November	Simulated conference I. Students deliver the powerpoint presentations on the assigned topics. Evaluation of presentations.

5 December	Writing test 2. Simulated conference II. Students deliver the powerpoint presentations on the assigned topics. Evaluation of presentations.
12 December	Repeating the writing tests.

Sample of writing test II

1. What kind of multidimensional scaling methods do you know?
2. How do we extract the principal components from a multivariate data set?
3. How does the hierarchical cluster analysis work?
4. What are the operators of the float-encoded genetic algorithm?
5. Please characterize the training phase of ANN models. What is called a perceptron?

Solution

The answers can be found in the course material "Geostatistics" (and the recommended literature) uploaded to the site of the Department of Geophysics:
http://www.uni-miskolc.hu/~geofiz/education.html

1. See the slide titled "Scaling of observations" in the above course material.
2. See the slide titled "Estimation of principal components" in the above course material.
3. See the slides titled "Hierarchical clustering", "Dendrogram" and "Linkage criteria" in the above course material.
4. See the slide titled "Genetic operators" in the above course material.
5. See the slides titled "Training of ANNs" and "Perceptron" in the above course material.

Exam questions

1. Data distributions. Datasets, histograms, pdf and cdf types. Determination of the mode.
2. Characterization of uncertainty. The most frequent value method as robust statistical estimator.
3. The Maximum Likelihood method. Confidence intervals. Skewness and kurtosis. Propagation of error. The linear and rank correlation coefficient.
4. Multivariate linear functional relationships. Covariance and correlation matrices.
5. Linear and non-linear regression analysis. Robust regression methods. Petrophysical examples.
6. Spatial correlation of petrophysical parameters, variogram models and kriging.
7. Hierarchical and non-hierarchical cluster analysis, the K-means clustering method. Petrophysical examples.
8. Principal component analysis, factor analysis and their applications in petroleum geosciences. Shale volume and permeability estimation.
9. Discrete inverse theory. Linearized and global optimization methods. Geophysical inverse problems.

10. Linear regression using inversion tools. The Gaussian Least Squares method. Weighted norms to be minimized. Well-logging applications.
11. The quality check of inversion results. The relation between the data and model covariance matrices.
12. Simulated Annealing methods. Classical and float-encoded genetic algorithm. Artificial neural networks.