Course Title: Statistical applications to experimental and technical design and data analysis

Modality: CFT- Transversal Training Course

Orientation:

- ☐ Ocean Observation and Global Change
- ☐ Sustainable use of Marine Resources
- ☐ Integral Management of the Sea
- ☐ Technological progress. Engineering and Business Management

Dates:
13 to 16 and 20 to 23 January 2020
Timetable:
Monday, Tuesday and Wednesday from 10h to 14h. Thursday from 10h to 13h
Duration:
30h
Location:
DOMAR Videoconference Room, Torre-CACTI building, Campus As Lagoas-Marcosende, Universidade de Vigo
Language:
English

Academic coordinators:

Name	Institution	e-mail
Jacobo de Uña Álvarez	Universidade de Vigo	jacobo@uvigo.es
Ángeles Saavedra González	Universidade de Vigo	saavedra@uvigo.es

Lecturers:

Name	Institution	e-mail
Jacobo de Uña Álvarez	Universidade de Vigo	jacobo@uvigo.es
Ángeles Saavedra González	Universidade de Vigo	saavedra@uvigo.es

General description:
The aim of this course is to provide theoretical background and statistical tools commonly used for solve problems in marine science, technology and management

Contents:
Multiple linear regression:

Covariance and correlation. Simple correlation, partial correlation, multiple correlation. The multiple linear regression model. Interpretation of the model. Estimation of the model and hypothesis testing. Explained and residual variance. ANOVA table. Multiple R squared. Adjusted R squared. Residuals and residual analysis. Variable selection: extra sum of squares. Log-linear regression. Polynomial regression. Model with interaction. Prediction.
Practical exercises with R (r-project.org).

Advanced regression and multivariate analysis:

Regression with categorical response. Binary and multinomial logistic regression. Nonparametric regression. Principal components analysis. Cluster analysis: K-means and hierarchical clustering. Discriminant analysis: Fisher linear discriminant rule, logistic discrimination. ROC analysis in classification problems.

Practical exercises with R (r-project.org).

Time series:

Time series data: Time series. Plots, trends and seasonal variation. Decomposition of series. Correlation: Stationary time series. Autocorrelation. The correlogram. Forecasting strategies. Stationary models: The ARMA process. Non stationary models: ARIMA and SARIMA processes.

Practical exercises with R (r-project.org).

Spatial statistics:

Types of spatial dependence. Spatial correlation: Fitting a variogram model. Spatial prediction: The kriging methods.

Practical exercises with R (r-project.org).

Teaching methodologies:

Theoretical lectures and practical exercises with R

Evaluation system:

Based on attendance and test exam

Brief CV of the lecturers:

Jacobo de Uña Álvarez:

Department of Statistics and Operations Research / SiDOR research group (sidor.uvigo.es).
Chair of the Department of Statistics and OR (2004-2007, 2017-2019).
Chair of SiDOR research group (1998 -)
Degree in Mathematical Sciences. University of Santiago de Compostela, 1995
PhD in Statistics. University of Santiago de Compostela, 1998
Identification number of the researcher:

- WoS Researcher ID K-5667-2014
- SCOPUS Author ID 6603553828
- Open Researcher and Contributor ID (ORCID) 0000-0002-4686-8417

General quality indicators of scientific production:

- Number of research six-year periods: 3
- Number of PhD conducted in the last 10 years: 7
- Publications included in Scopus: 88
- Total citations: 953 in 649 documents
- H-Index: 15

Lines of research: nonparametric curve estimation, survival analysis and multi-state models, goodness-of-fit tests, multiple comparison procedures, application of statistics to other areas
I have published papers in prestigious statistical journals like Journal of Multivariate Analysis, Biometrics, Statistics in Medicine, Statistical Methods in Medical Research, Computational Statistics and Data Analysis, or Journal of Statistical Software; I have also publications in journals of other areas such as Medicine, Biology, Bioinformatics, Computer Sciences, Nanosciences or Economics. I have contributed a number of software packages in R. I am Associate Editor of several statistical

journals like Biometrics, Test or Annals of the Institute of Statistical Mathematics. I have been Principal Investigator of research projects in the area of Statistics (MTM) funded by the Spanish Ministry of Science continuously since 2005. My teaching activity (1996 -) includes many courses for PhD students, MSc students, companies and research institutes, on topics like Time Series Analysis, Regression Models, Reliability Engineering, Multi-state models, Multivariate Statistics and Experimental Design.

Ángeles Saavedra González:

Department of Statistics and Operations Research / School of Mining and Energy Engineering.
Degree in Mathematical Sciences. University of Santiago de Compostela, 1989
PhD in Mathematics. University of Oviedo, 1997
Identification number of the researcher:

- WoS Researcher ID R-9077-2018
- SCOPUS Author ID 7006526997
- Open Researcher and Contributor ID (ORCID) 0000-0002-4845-2271

General quality indicators of scientific production:

- Number of research six-year periods: 3
- Number of PhD conducted in the last 10 years: 5
- Publications included in Scopus: 51
- Total citations: 463 in 381 documents
- H-Index: 13

Lines of research: Applied statistics, where much of my current publications are concentrated. Most of my publications are encompassed within various engineering problems. The statistical techniques developed in these publications are also diverse: Multivariable statistical methods, Fuzzy statistics, Geostatistics, Gray theory, Time series, and Quality control. Currently my research is focused on the development and study of Bayesian networks.

Relevant references:

Cowpertwait, PSP and Metcalfe, AV (2009) Introductory Time Series with R, Springer New York.

Diggle, PJ and Ribeiro, PJ (2006) Model-based geostatistics, Springer New York.

Fox, J and Weisberg, S (2011) An R companion to applied regression. SAGE Publications.

Hosmer, DW and Lemeshow, S (2000) Applied Logistic Regression. Wiley

Lattin J, Carroll JD, Green PE (2003) Analyzing Multivariate Data. Thomson.

Peña, D (2002) Análisis de Datos Multivariantes. McGraw-Hill (in Spanish)

Peña, D (1999) Estadística. Modelos y métodos. Vol. 2: Modelos lineales y series temporalis. Alianza Universidad Textos (in Spanish)

Sheather, SJ (2009) A Modern Approach to Regression with R. Springer.

Zar, JH (1996) Biostatistical Analysis. Prentice Hall.

Zuur, AF; Ieno, EN and Smith, GM (2007) Analysing ecological data. Springer New York.