Conceptual support guide and reference repository for Teaching Land Cover and Terrain Characteristics: A Guide for Educators and Teachers.

Introduction:

This guide provides educators with additional resources to strengthen their teaching of land cover and terrain characteristics. It offers a complementary approach that includes a glossary of key terms, links to scholarly articles and specialized publications, and videos that allow for continuous feedback of information. In an environment where access to multiple sources is essential, this guide aims to consolidate knowledge and enrich didactical practices.

The content facilitates the integration of updated and easily accessible teaching materials, promoting a deeper and more applied understanding of the concepts. In this way, teachers can keep students immersed in learning, connecting theory with current research and audiovisual resources.

The objective of the guide to not only to provide theoretical knowledge but also act as a bridge to dynamic and interactive learning. This will allow educators to apply these additional resources to enhance the teaching of land cover and terrain features.

Content

Module 1

Visual material

Introduction to Remote Sensing:
https://www.youtube.com/watch?v=N49PzLDUIFQ&ab_channel=GeospatialWorld
Multispectral Imaging in Remote Sensing:
https://youtu.be/ayp7hP0Xr8Q?si=IM2Y9V6jmaGeVqOt
Synthetic Aperture Radar (SAR) Remote Sensing:
https://youtu.be/sGCJZtJVTWQ?si=20Z3bE83MGQQZ0OR

Module 2

The objective of this activity is to teach students how to make a land cover classification map using R Studio and QGIS software. This approach combines quantitative analysis and geospatial data processing using open-source tools, facilitating hands-on learning of key concepts such as the Normalized Difference Vegetation Index (NDVI) and the use of supervised classification algorithms such as Random Forest.

Students will work with raster data from the area of interest, in this case Fogo Island, and learn how to handle specific raster formats and spatial coordinates. Detailed guidance is provided on how to load and manipulate this data using functions in R Studio, and how to create and evaluate a classification model using remote sensing techniques.

It is essential that students have the latest versions of R Studio and QGIS installed to perform the necessary operations. Throughout this guide, students will acquire skills to combine multispectral data, apply classification models, and evaluate the accuracy of the results, ending in the creation of a land cover classification map of Fogo Island.

Rast function: Function used to load raster images or geospatial data layers to be processed. It is part of the “terra” or “raster” package and allows the manipulation of data in R.

NDVI (Normalized Difference Vegetation Index): NDVI is a vegetation index that uses the red and near infrared bands to measure vegetation density and health. It is calculated using the formula: NDVI = (NIR - Red) / (NIR + Red).

plotRGB function: allows generating true color or false color compositions of a satellite image using different combinations of spectral bands.

Concepts

- Supervised land cover classification: Supervised land cover classification is a remote sensing technique in which labeled training data (training polygons) are used to identify different land cover classes in satellite imagery. A classification algorithm assigns classes based on the spectral characteristics of the trained areas.
- R Studio: is an integrated development environment (IDE) for the R programming language, widely used for statistical analysis and processing of geospatial data, including applications in remote sensing.

- Bands (2, 3, 4, 8): These bands represent different wavelengths of the electromagnetic spectrum: Band 2 (blue), Band 3 (green), Band 4 (red), and Band 8 (Near Infrared, NIR). They are commonly used for multispectral analysis, including land cover classification and NDVI calculation.
- WriteRaster function: Used to save a raster image to a file after processing or analysis. This function is part of the "raster" or "terra" package.
- Training polygons: Training polygons are predefined areas with known classes that are used as a reference for training supervised image classification models. These polygons contain the spectral information used to identify different land cover types.
- QGIS: is an open-source software for geospatial analysis that allows visualization, editing, and analysis of spatial data. It is widely used for map creation, GIS data processing, and land cover classification.
- Google Satellite: is a source of high-resolution satellite imagery available on Google Earth and Google Maps, used as a reference for geospatial analysis and to identify different types of land cover.
- Random Forest classification: is a machine learning model based on the combination of multiple decision trees. It is widely used in land cover classification due to its ability to manage large amounts of data and variables.
- Extract, seq, merged functions: These functions in R allow extracting, sequencing, and merging raster and vector image data for more detailed analysis and classification in geospatial studies.
- Predict function: Function used to apply a previously trained classification model on a satellite image or dataset and generate predictions about land cover classes.
- Accuracy of the model: Accuracy of the model refers to the ability of the classification model to correctly predict land cover classes, as measured by comparison with validation or test data.
- CreateDataPartition function: Function in R used to split the data into training and test sets, allowing the performance of the supervised classification model to be evaluated.
- Confusion matrix: The confusion matrix is a table that compares the model predictions with the actual data, showing how well the model classified each land cover category.
- TrainControl function: Function in R used to set up cross-validation in classification models, helping to evaluate model performance through multiple training and validation iterations.
- Caret::train: Function in R that allows training a predictive model using different algorithms and cross-validations to optimize the classification.
- Land cover classification map: A land cover classification map is a visual output generated from a supervised classification process that shows the different land use and land cover classes in each region.

Reference material module 2

Congalton, R. G., & Green, K. (2019). Assessing the accuracy of remotely sensed data: Principles and practices. CRC Press. Link: Assessing the Accuracy of Remotely Sensed Data | Principles and Practi (taylorfrancis.com)

Foody, G. M. (2002). Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80(1), 185-201. Link: PII: S0034-4257(01)00295-4 (psu.edu)

Graser, A. (2019). Learning QGIS. Packt Publishing Ltd. Link: Learn QGIS: Your step-by-step guide to the fundamental of QGIS 3.4, 4th Edition - Andrew Cutts, Anita Graser - Google Libros

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2(3), 18-22. Link: RN-2002-022.pdf (r-project.org)

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing, 28(5), 823-870. Link: Full article: A survey of image classification methods and techniques for improving classification performance (tandfonline.com)

Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jędrzejewska, B., Lima, M., & Kausrud, K. (2011). The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. *Climate research*, *46*(1), 15-27. Link: Climate Research 46:15 (int-res.com)

Potere, D. (2008). Horizontal positional accuracy of Google Earth's high-resolution imagery archive. Sensors, 8(12), 7973-7981. Link: Horizontal Positional Accuracy of Google Earth's High-Resolution Imagery Archive (mdpi.com)

Lovelace, R., Nowosad, J., & Muenchow, J. (2019). Geocomputation with R. CRC Press. Link: Welcome | Geocomputation with R (geocompx.org)

Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. Remote Sensing of Environment, 62(1), 77-89. Link: (PDF) Selecting and interpreting measures of thematic classification accuracy (researchgate.net)

Kuhn, M. (2008). Building predictive models in R using the caret package. Journal of Statistical Software, 28(5), 1-26. Link: Building Predictive Models in R Using the caret Package | Journal of Statistical Software (jstatsoft.org)

Lu, D., et al. (2014). A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. International Journal of Digital Earth, 7(2), 99-117. Link: Full article: A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems (tandfonline.com)

Visual material

Image Analysis using NDVI to Assess Vegetation Greenness + mini quiz:
https://youtu.be/sPnFkp-25fM?si=SSQ7I_PcH1zyTncK

An introduction to RStudio (for beginners):
https://youtu.be/K9q5wRl4wYw?si=_lqJCtkWhkGg_xPC

What is Random Forest?: https://youtu.be/gkXX4h3qYm4?si=CqGQAMFik50YSwrl

Installing and loading packages in R studio 4.3.3 (using quantmod and tseries as examples): https://youtu.be/ZOlwu5tNoXU?si=rCXu9qkICc564ezF

Confusion Matrix | How to Implement Confusion Matrix In R
https://youtu.be/CuJc1MFY23k?si=3hsfXqYKu2oB7r_y

Module 3

The objective of this module is to guide students through the process of obtaining terrain features, such as elevation, slope and orientation, from a Digital Elevation Model (DEM) using QGIS software. These features are fundamental to understanding the topography of a specific area, in this case Fogo Island, and are essential for applications in geospatial studies, land use planning, and environmental analysis.

The material in this module provides students with a DEM of Fogo Island as a basis for analysis. Through the open-source software QGIS, students will learn how to load and manipulate raster data, using terrain analysis tools such as slope and aspect calculation, as well as verifying that the results conform to expected ranges such as slope in percent and orientation in degrees.

In addition, supplemental resources are provided to help reinforce the theoretical and practical understanding of these processes. The use of DEMs and slope and orientation analysis are crucial tools in disciplines ranging from geomorphology to environmental planning.

This module ensures that students develop practical skills in the use of GIS tools applied to surveying and can analyze terrain features using advanced techniques and specialized software.

Concepts.

- Raster layer: is a matrix data structure used in geographic information systems (GIS) that represents spatial data, such as satellite imagery or elevation layers. Raster layers contain values representing attributes such as elevation, slope or land cover.
- Elevation: refers to the height of a point above sea level, and in GIS, is commonly represented through digital elevation models (DEM), which allow the analysis of terrain features.
- Slope: is a measure of the slope of the terrain and is calculated as the rate of change in elevation over a given area. In GIS, it is often expressed in degrees or as a percentage.
- Aspect: refers to the direction in which a slope is oriented, measured in degrees from north. In GIS, it is crucial for the analysis of sun exposure and wind patterns.
- Fogo Island: is a volcanic island in Cape Verde, which has been the subject of numerous geospatial studies due to its geomorphology and land use changes.
- DEM (Digital Elevation Model): A DEM is a digital terrain representation that shows the elevations of the Earth's surface.
- QGIS: is an open-source software that allows the analysis and visualization of geospatial data, and is widely used in map creation and GIS data processing.
- Slope analysis raster tool: Tool in QGIS that allows to calculation of slope of an area using a DEM as input, generating a raster layer that shows the slope of the terrain.

- Aspect analysis raster tool: Tool in QGIS calculates the orientation of slopes in each area using a DEM, classified in degrees from 0 to 360, corresponding to the cardinal directions.
- Percentage (for slope): Slope expressed as a percentage indicates vertical change relative to horizontal distance and is one of the most common ways of representing slopes in topographic analysis.
- Cardinal direction (for aspect): Cardinal directions (north, south, east, west) are used in aspect analysis to classify slopes according to their orientation.

Reference material module 3

Burrough, P. A., & McDonnell, R. A. (1998). Principles of Geographical Information Systems. Oxford University Press. Link: (PDF) Impact of Seawater on Distribution of Fluoride and Other Ions in Groundwater of Diplo Area, Thar Desert Pakistan (researchgate.net)

Chang, K. T. (2018). Introduction to Geographic Information Systems. McGraw-Hill Education. Link: (PDF) Introduction to Geographic Information Systems (researchgate.net)

González, P. J., Bagnardi, M., Hooper, A. J., Larsen, Y., Marinkovic, P., Samsonov, S. V., & Wright, T. J. (2015). The 2014–2015 eruption of Fogo volcano: Geodetic modeling of Sentinel-1 TOPS interferometry. *Geophysical research letters*, *42*(21), 9239-9246. Link: (PDF) The 2014-2015 eruption of Fogo volcano: Geodetic modeling of Sentinel-1 TOPS interferometry (researchgate.net)

Wilson, J. P., & Gallant, J. C. (2000). Terrain Analysis: Principles and Applications. John Wiley & Sons. Link: (PDF) Impact of Seawater on Distribution of Fluoride and Other Ions in Groundwater of Diplo Area, Thar Desert Pakistan (researchgate.net)

Zhang, J. (2010). Multi-source remote sensing data fusion: status and trends. International Journal of Image and Data Fusion, 1(1), 5-24. Link: Full article: Multi-source remote sensing data fusion: status and trends (tandfonline.com)

Visual material

Slope: https://youtu.be/j5idvvoVlbg?si=QXZqapC7H_V4WcK2
Aspect: https://youtu.be/LfKYEy0w-pc?si=wLEk5KRvkfmGlWWp
How to create slope aspect and slope angle map in QGIS:
https://youtu.be/ccKoG5cRreY?si=FEvfGRHVZ6FTQhGF

Module 4

The objective of this module is to teach students how to perform several types of statistical analysis to establish relationships between land cover classes and terrain characteristics such as elevation, slope and orientation using R Studio software. The analysis of these relationships is fundamental to understanding how the physical properties of the terrain interact with the diverse types of land use in a specific area.

During this module, students must use the results obtained in Modules 2 and 3, where raster layers of land cover and terrain characteristics (elevation, slope and orientation) were generated for Fogo Island. Through R Studio, students will learn how to perform statistical tests such as correspondence analysis, Kruskal-Wallis test and other non-parametric analysis tools, in order to identify significant patterns and correlations between variables.

To perform this analysis, students are encouraged to install the latest versions of R Studio. This module will enable participants to acquire advanced skills in the use of statistical tools applied to geospatial data, strengthening their ability to interpret results and make inferences in environmental studies.

Concepts

- Normality of the data: refers to the extent to which the data conform to a normal or Gaussian distribution. In statistics, many models assume that the data follow this distribution, so it is important to verify it before performing certain analyses. It can be evaluated graphically with histograms and statistically with tests such as Shapiro-Wilk or Kolmogorov-Smirnov.
- Resample function: used to adjust a raster layer to the spatial resolution of another layer, ensuring that both have the same alignment and cell size. This process is important in spatial data analysis when layers of different resolutions are combined.
- As.data.frame function: converts an object, such as a matrix or raster, into a data frame. This is useful when you want to manipulate the data as a table to apply statistical analysis or graphics.
- As.factor function: converts a numeric or character variable into a factor, a categorical data type in R. Factors are commonly used in statistical and graphical models to represent qualitative variables.
- As.numeric function: converts an object into a numeric variable, which is useful when you need to perform mathematical or statistical operations on data originally classified as factors or characters.
- Hist function: used to create histograms, which are bar charts showing the distribution of a continuous variable. They are useful for visualizing the distribution and normality of data.
- Boxplot function: function that generates box plots, which represent the distribution of a continuous variable showing the median, quartiles, and outliers. It is a common tool for comparing data distributions between groups.
- Mutate function: The mutate function from the dplyr package in R is used to create or modify columns within a data frame. It is useful for transforming variables from one form to another in the data analysis process.
- Contingency table: Contingency table is a matrix that shows the frequency of occurrence of categorical variables, allowing to analyze the relationship between

them. It is common in statistical analysis and is used to calculate independence tests such as Chi-square.

- Correspondence analysis (CA): is a multivariate statistical technique used to analyze relationships between categorical variables. It graphically represents the associations between categories in a low-dimensional space.
- Kruskal-Wallis test: is a nonparametric test that assesses whether there is a significant difference between the medians of several groups. It is used as an alternative to ANOVA when the assumption of normality is not met.

Reference material module 4

Freeman, E. A., Frescino, T. S., & Moisen, G. G. (2018). ModelMap: an R package for model creation and map production. *R package version*, *4*, 6-12. Enlace: CMB10 (psu.edu)

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors, and Anderson-Darling tests. Journal of Statistical Modeling and Analytics, 2(1), 21-33. Link: https://www.researchgate.net/publication/267205556_Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirnov_Lilliefors_and_Anderson-Darling_Tests

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: With Applications in R. Springer. Enlace: Driver.dvi (berkeley.edu)

Wickham, H. (2011). The split-apply-combine strategy for data analysis. Journal of Statistical Software, 40(1), 1-29. Enlace: The Split-Apply-Combine Strategy for Data Analysis | Journal of Statistical Software (jstatsoft.org)

McKight, P. E., & Najab, J. (2010). Kruskal-Wallis test. The Corsini Encyclopedia of Psychology, 1-2. Enlace: Kruskal-Wallis Test (researchgate.net)

Visual material

What is Contingency Table?: https://youtu.be/QZjMslZQ2l0?si=5T5LR0W0i5trrsUM

Correspondence Analysis: https://youtu.be/HEPNUgLhwYA?si=ouP0b3LuiMUnGSJj

Kruskal-Wallis-Test (Simply explained): https://youtu.be/l86wEhUzkY4?si=oH7hBO9TcCznECmU