

Remote Sensing - Sentinel 2

Spectral Indices, Classification, and Spatio-Temporal Analysis

Chris Reudenbach

Introduction

Remote sensing is the only measurement technology that enables continuous observation of the Earth's surface across full spatial extents — from landscapes to entire countries and global regions. To use remote sensing effectively, it is essential not only to apply established analytical methods but also to adapt and extend them as new environmental questions and large datasets emerge.

Sentinel-2 is currently the most important platform for Earth observation in all areas, but especially for climate change, land use and ecological issues at all levels, from the upper micro to the global scale.

There are two operational Sentinel-2 satellites: Sentinel-2A and Sentinel-2B, both in sun-synchronous polar orbits and 180 degrees out of phase. This arrangement allows them to cover the mid-latitudes with an orbital period of about 5 days.

The Sentinel-2 data are therefore predestined to record spatial and temporal changes on the Earth's surface (the forest was green in early summer, it has disappeared by late summer). They are ideal for timely studies before and after natural disasters, or for biomass balancing, although complementary systems such as Landsat-8/9 (long-term continuity), PlanetScope (daily repeat), or Sentinel-1 SAR (cloud-penetrating structural information) are also operationally important; nevertheless, Sentinel-2 remains the most widely used because it uniquely combines global free access, high radiometric quality, a consistent 10–20 m multispectral design, and a reliably maintained 5-day revisit cycle.

Change detection and time-series analysis are central in environmental informatics because most Earth-surface processes — drought-induced forest dieback, land-use conversion, vegetation dynamics, bark-beetle outbreaks, flooding, and climate impacts — unfold simultaneously across space and time.

A robust spatial-temporal workflow therefore requires:

1. Acquiring Sentinel satellite data,

2. Extracting spectral and structural information, and
3. Applying classification and temporal analysis methods to quantify and interpret changes.

The following sections outline the conceptual foundations of image-based change detection and demonstrate how classification, spectral indices and Sentinel-2 data together reveal spatial-temporal patterns. The Burgwald serve as examples of vegetation change between 2018 and 2023. For a broader conceptual framework of GIScience and remote sensing as components of system-based spatial information science, see Tolpekin and Stein (2012).

Change Detection and Spatial–Temporal Analysis

Unprocessed Sentinel-2 imagery is rich in information, but not immediately meaningful. While humans can interpret true-color scenes intuitively, scientific interpretation requires quantitative methods. The strength of remote-sensing analysis lies in deriving information that is invisible to the naked eye: vegetation health, stress, canopy density, moisture content, structural changes, and land-cover transitions.

Vegetation indices

Spectral vegetation indices are dimensionless, mathematically defined combinations of Sentinel-2 bands that compress multi-band reflectance into a few variables targeted at specific biophysical properties of the land surface. By using ratios, differences or more complex transforms of visible, red-edge, NIR and SWIR bands, they enhance signals related to chlorophyll content, canopy density and structure, leaf water status, or exposed soil background while suppressing illumination and atmospheric effects. Because no single raw band can reliably isolate these processes, spectral indices are a crucial intermediate layer between radiance/reflectance data and higher-level products such as land-cover classes, crop types, or degradation maps. In practice, remote-sensing workflows therefore rely heavily on vegetation and moisture indices (e.g. NDVI, EVI, SAVI, red-edge-based indices, NDWI/NDMI) as input features for time-series analysis, change detection and machine-learning classification, particularly when exploiting the rich spectral sampling of Sentinel-2’s red-edge and SWIR bands (Bannari et al. 1995; Xue and Su 2017; Huang et al. 2020; Montero et al. 2023).

The Normalized Difference Vegetation Index (NDVI) measures chlorophyll absorption in red vs. near-infrared reflectance and is a fundamental indicator of greenness, photosynthetic activity, and canopy density. NDVI is the most widely used vegetation index, but also frequently misinterpreted or over-generalised (Huang et al. 2020). The Modified Soil Adjusted Vegetation Index (MSAVI2) reduces soil-brightness influence in sparse or partially damaged canopies, making it useful in drought-affected stands and early dieback where soil background is exposed. The Moisture Stress Index (MSI), derived from the SWIR/NIR ratio, is sensitive to water loss and canopy desiccation and often serves as an early-warning indicator before

visual decline or NDVI reductions become apparent. The Soil Adjusted Vegetation Index (SAVI) explicitly accounts for soil reflectance in mixed pixels and is highly effective in clearcuts, forest edges, young regrowth and partially disturbed sites. The Enhanced Vegetation Index (EVI) reduces atmospheric scattering effects and mitigates NDVI saturation in dense canopies, which is particularly important in dense conifer forests and for monitoring spruce dieback with Sentinel-2 (Bannari et al. 1995; Xue and Su 2017).

These indices serve as continuous measures of canopy condition, moisture content, and structural state. When combined with classification, they provide the physical basis required to interpret spatial-temporal change patterns.

Structural and textural descriptors

Structural and textural descriptors capture how reflectance values are arranged in space, rather than their absolute value in a single pixel. In remote sensing, they typically quantify local neighbourhood patterns (for example via grey-level co-occurrence matrices, local binary patterns, Gabor filters or wavelets) or higher-level structure such as object shape, size, orientation and canopy architecture (Fekri-Ershad 2018). These descriptors become particularly important in high-resolution imagery and heterogeneous landscapes, where many land-cover types may be spectrally similar but differ in spatial pattern — for example forest vs. shrubland, tiled roofs vs. asphalt, or dense vs. sparse crowns. Modern classification workflows therefore often combine spectral indices with structural/textural features to form joint spatial–spectral feature sets, which have been shown to improve land-cover mapping, change detection and fine-grained object recognition in multispectral and hyperspectral data (Datta et al. 2022; Mehmood et al. 2022; Li et al. 2024; Wang, Yan, and Liu 2025).

Classification logic: MLC, k-means, Random Forest

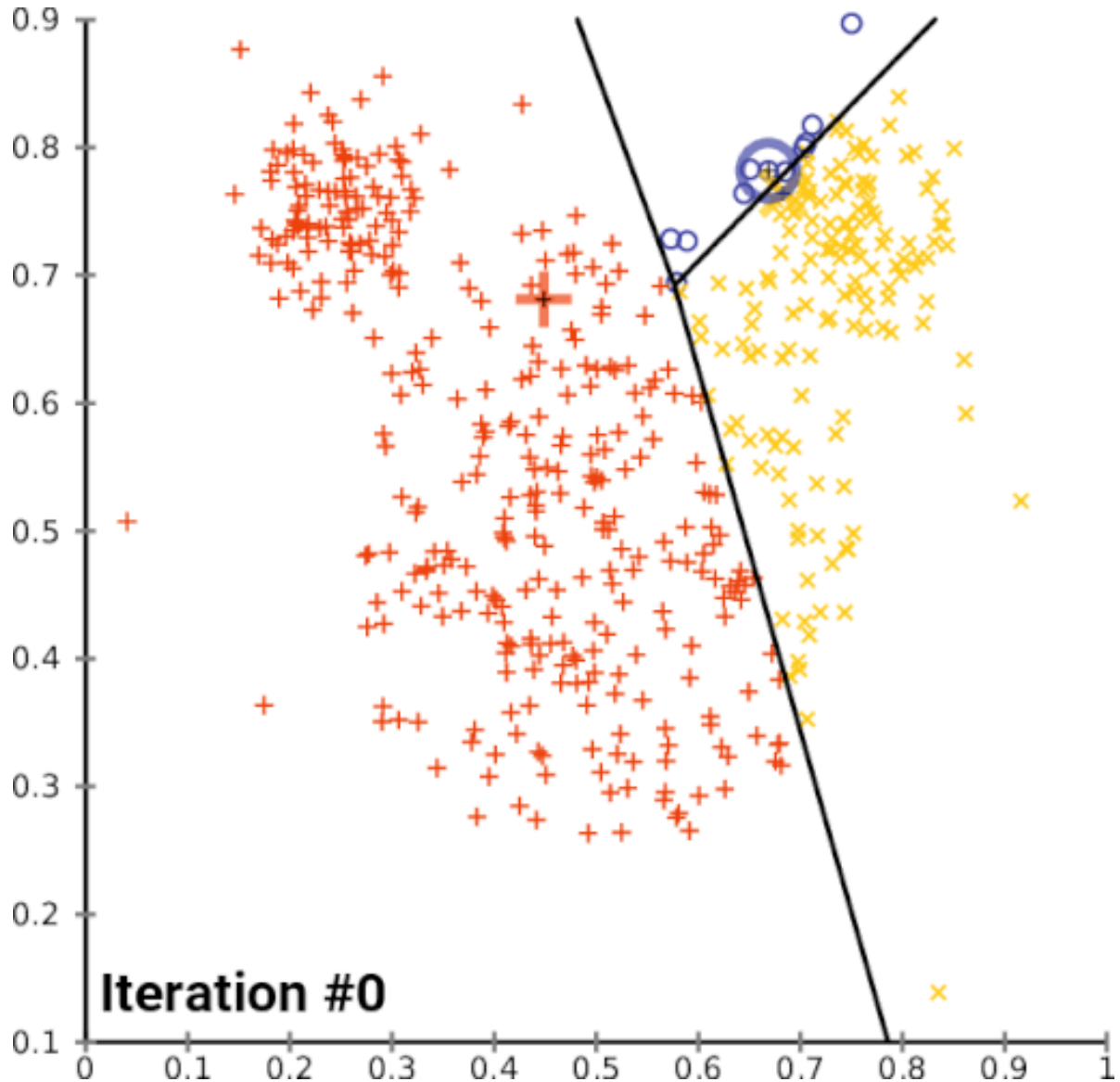
From a methodological perspective, three families of classifiers illustrate the logic from classical statistics through unsupervised clustering to modern ensemble learning: Maximum Likelihood Classification (MLC), k-means clustering, and Random Forest (RF). Together, these three methods illustrate the progression from distribution-based decision rules (MLC), via geometry-based partitioning (k-means), to data-driven ensemble learning (RF) that is now standard in remote-sensing workflows and frequently used as a benchmark in comparative studies (Lu and Weng 2007; Hogland, Billor, and Anderson 2013; Paradis et al. 2022; Belgiu and Drăguț 2016).

1. K-means (Unsupervised Clustering)

K-means is an unsupervised clustering algorithm that partitions pixels into k clusters by iteratively minimising within-cluster variance. It is simple, fast and requires no labelled training data, but has no notion of semantic classes, is sensitive to initialisation and scale, and ignores

spatial context. In land-cover workflows, it is therefore mainly used for exploratory grouping or as a pre-step before manual labelling rather than as a final map (Paradis et al. 2022).

K-means is often considered the most accessible unsupervised classification algorithm. It groups pixels solely by spectral similarity, without requiring training data, making it ideal for exploratory analysis. Each pixel is an observation in multidimensional feature space. The algorithm assigns each pixel to the nearest cluster centroid, iteratively adjusting centroid positions until convergence.



Convergence of k-means clustering from an unfavourable starting position (two initial cluster centres are fairly close).

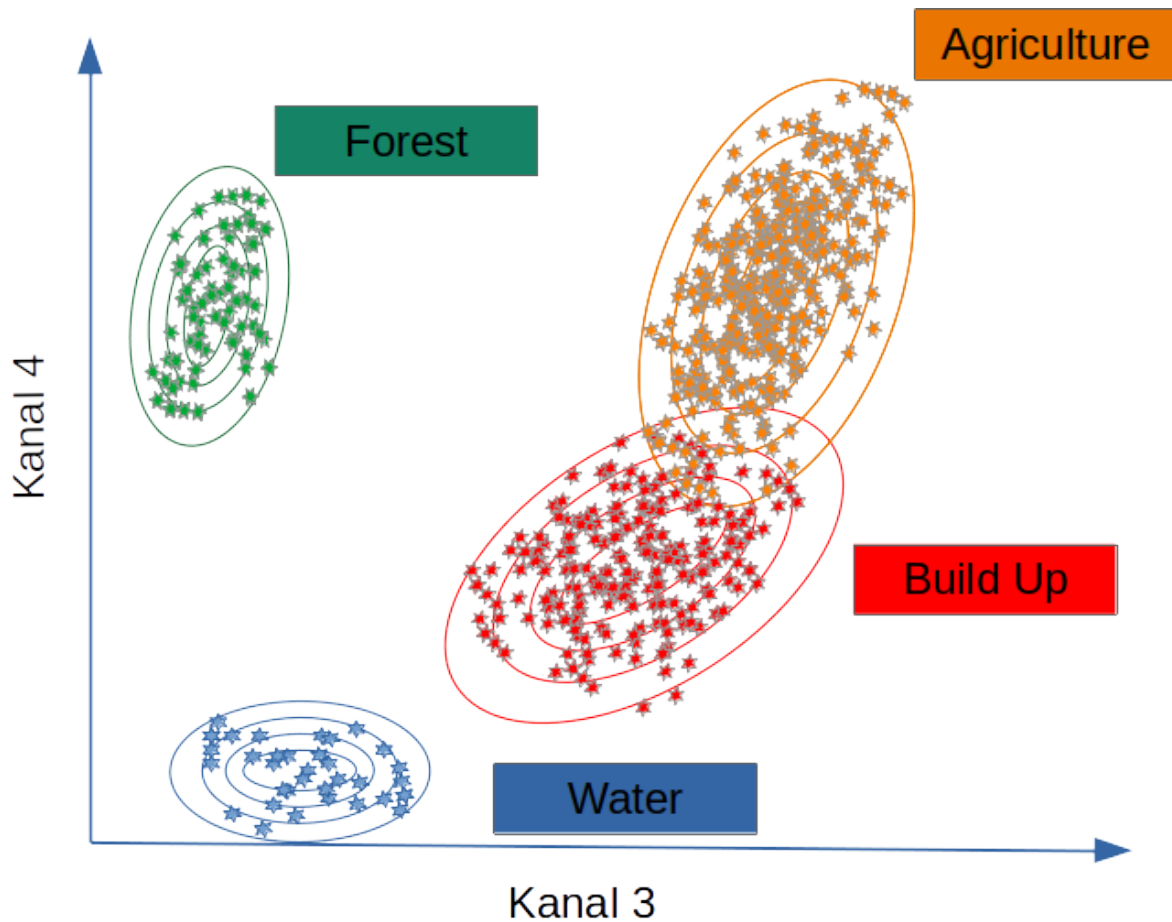
In vegetation-change studies, k-means is used to check whether classes like forest, dead wood, and clearcuts are spectrally distinct — an essential requirement for stable supervised classification.

2. Maximum Likelihood Classification (MLC)

Maximum Likelihood Classification (MLC) is a parametric supervised classifier that assumes each class follows a multivariate normal distribution in feature space. It assigns each pixel to the class with the highest posterior probability given class means, covariance matrices and, optionally, class priors. MLC works well when classes are spectrally compact and approximately Gaussian, but degrades for heterogeneous or non-Gaussian classes (Hogland, Billor, and Anderson 2013).

MLC assumes that each class follows a multivariate normal distribution in spectral (or feature) space. It computes the probability that a pixel belongs to each class and assigns it to the most likely one. Performance is strongly influenced by the covariance structure of the classes and by how well the training data represent within-class variability (Hogland, Billor, and Anderson 2013).

MLC performs well when spectral differences between classes (for example, forest vs. clearcut) are large and the Gaussian assumption is approximately valid.



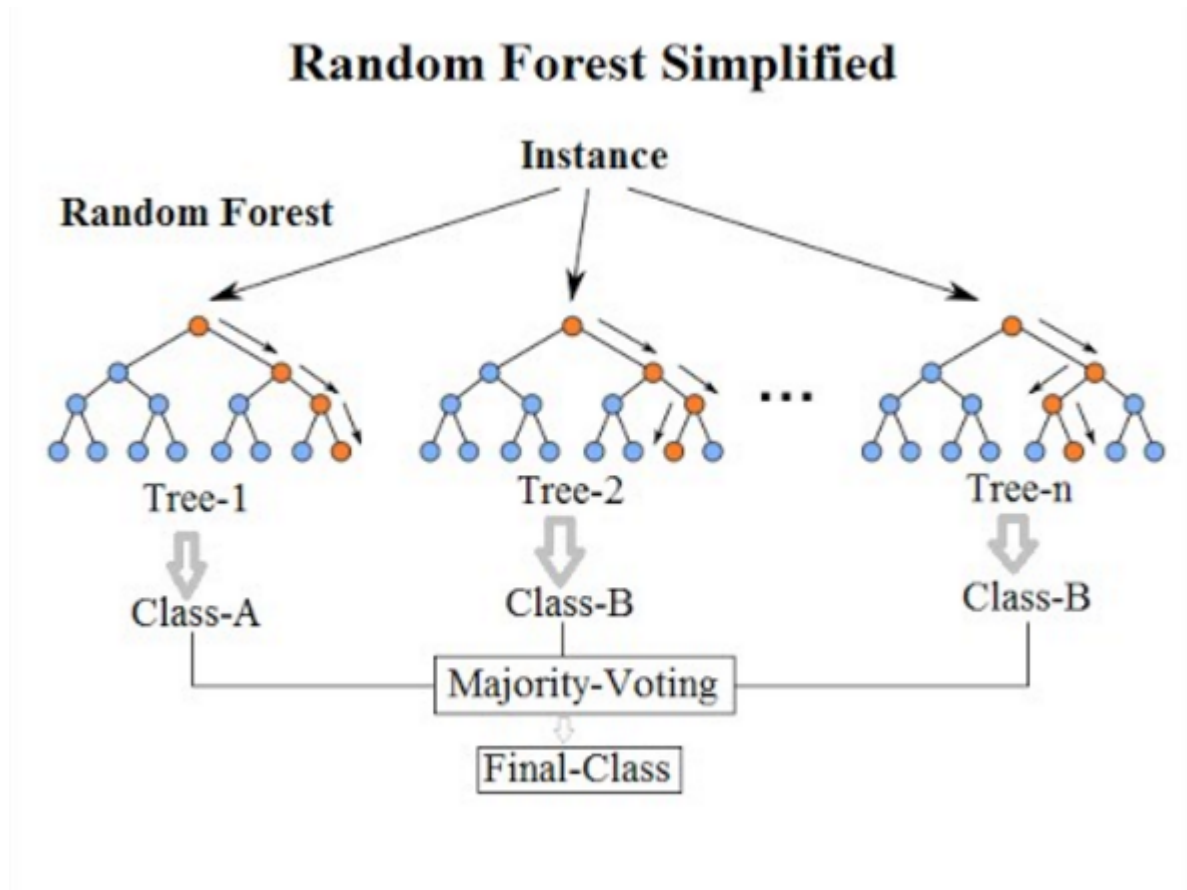
Classifiers such as MLC use training data to determine descriptive models representing statistical signatures. Within the limits of the quality and representativeness of the training data, such models can be applied to all pixels for which the predictor variables are available (Lu and Weng 2007; Belgiu and Drăguț 2016).

3. Random Forest Classification (RF)

Random Forest (RF) is a non-parametric supervised ensemble approach. Many decision trees are trained on bootstrapped samples and random subsets of features, and predictions are aggregated by majority vote (classification) or averaging (regression). RF can model complex, non-linear class boundaries, handles mixed feature sets (spectral indices, structural/textural descriptors, ancillary GIS layers), and is relatively robust to noise and unbalanced data (Belgiu and Drăguț 2016).

Random Forest builds an ensemble of decision trees, each learning different feature partitions. A pixel is classified via majority vote across all trees. RF excels in complex, heterogeneous

landscapes because it handles nonlinear relationships, high-dimensional predictors (bands plus indices plus structural/textural descriptors), correlated variables, and spectral noise and outliers.



Random Forest is now considered a baseline method in environmental machine learning, alongside Gradient Boosting and Support Vector Machines, particularly for land-cover mapping and change detection (Belgiu and Drăguț 2016; Mehmood et al. 2022).

In this tutorial, Random Forest is used to predict the spatial distribution of clear-felling/no forest alongside MLC, combined with standard methods of random validation and model quality assessment. Both methods are widely used in forestry, agriculture, urban studies, ecosystem monitoring, and climate-change research.

Combining Spectral, Structural, and Classification Features

In practice, many workflows **combine** three types of information:

- Spectral indices (NDVI, MSAVI2, MSI, SAVI, EVI, kNDVI),

- Structural or textural descriptors, and
- Classification models (unsupervised and supervised).

However, this is a *design choice*, not a universal requirement. Brute-force experiments and feature-selection studies show that classification performance depends strongly on the **correlation structure and composition of the predictor set** rather than on simply “using everything” (Lu and Weng 2007; Mehmood et al. 2022). Highly collinear indices can dilute model signal, and some feature types become redundant once others are included.

A common strategy is therefore to derive structural or textural descriptors not on every individual band, but on a **reduced spectral subspace**, for example on the first principal component of an energy-rich band stack (e.g. red–NIR–SWIR). This preserves most spatial structure while limiting dimensionality and redundancy. In many Random Forest applications, the most important splits still come from raw spectral bands or a few key indices, with texture contributing mainly in cases where **spatial pattern** (patchiness, edge density, canopy roughness) matters more than pure reflectance (Belgiu and Drăguț 2016; Datta et al. 2022). For RGB-only data, by contrast, spatial/structural features and RGB-based indices often carry a larger fraction of the discriminative information.

Equally important is that model training and feature selection are **spatially explicit**. Spatial autocorrelation and clustered sampling can cause over-optimistic accuracies and misleading variable importance. Target-oriented spatial cross-validation and forward feature selection, as implemented for example in the CAST framework, explicitly account for prediction geometry and help to remove predictors that only exploit spatial structure instead of genuine environmental relationships (Meyer et al. 2018, 2019, 2024; Meyer and Pebesma 2021, 2022). In other words, the “best” combination of spectral, structural and ancillary predictors is **task- and data-specific** and must be derived under an appropriate spatial validation scheme.

Additionally, long-term studies increasingly rely on cloud-native geospatial techniques, such as STAC catalogues, COGs, and distributed computing via gdalcubes (Appel and Pebesma 2019), rstac (Brazil Data Cube Team 2021), or openeo (Lahn 2024). These approaches are essential when analyses scale to hundreds of scenes or multiple years.

Assessing Classification Quality

Supervised classification requires validation, usually by splitting data into training and test subsets or by using independent reference data. Performance metrics include:

- Sensitivity (true positive rate for clearcuts),
- Specificity (true negative rate for non-clearcuts),
- Predictive values, adjusting for actual class frequencies, and
- Global measures such as overall accuracy and Cohen’s kappa.

Even high accuracy may mask structural misclassification in subtle transitions, such as thinning canopies or partially damaged stands.

Continuous spectral indices such as kNDVI — derived using the gdalcubes framework — help diagnose such errors by revealing sub-pixel gradients that discrete classes cannot capture.

Advanced time-series methods like BFAST (Verbesselt, Zeileis, and Herold 2012; Wu et al. 2020) breakpoints in vegetation dynamics, including timing and magnitude of structural change. BFAST is particularly valuable in drought–forest interactions and post-disturbance recovery monitoring.

Digitising Training Data

Training data must represent all land-cover types to be classified. Digitisation can be done interactively in R (for example with **mapedit**) or externally in QGIS:

https://docs.qgis.org/3.34/en/docs/training_manual/create_vector_data/create_new_vector.html#basic-ty-digitizing-polygons

After digitisation, raster values for all predictors (bands, indices, structural descriptors) are extracted for each polygon and used to train the classifier.

Best practices include:

- Ensuring balanced class representation,
- Avoiding polygons that span multiple land-cover types, and
- Using high-resolution aerial imagery to refine training locations.

Further Reading

Technical tutorials & blogs

These resources are **practical, implementation-focused** guides. They are useful to understand concrete R workflows, but they do **not** replace primary literature or your own methodological reasoning.

- **Goldstein, S. (2019).** Classifying satellite imagery in R. *Urban Spatial Blog*. https://urbanspatial.github.io/classifying_satellite_imagery_in_R/ (Goldstein 2019)
- **Hijmans, R. J. (2023).** Supervised classification. *rspatial.org*. <https://rspatial.org/raster/rs/5-supclassification.html> (Hijmans 2023)
- **Lizarazo, I. (2021).** A tutorial on pixel-based land cover classification using random forests in R. *RPubs*. https://rpubs.com/ials2un/rf_landcover (Lizarazo 2021)

- **Gonçalves, J. (2018).** Advanced techniques with raster data: Supervised classification. R-exercises. <https://www.r-exercises.com/2018/03/07/advanced-techniques-with-raster-data-part-2-supervised-classification/> (Gonçalves 2018)
- **Stefan, V. (2020).** Satellite image classification in R. <https://valentinitnelav.github.io/satellite-image-classification-r/> (Stefan 2020)

💡 About blog posts and tutorials...

Use tutorials and blog posts as practical support only. They are extremely helpful for code and workflow patterns, but they cannot replace independent thinking, methodological scrutiny, or engagement with peer-reviewed research.

Final Remarks

This tutorial integrates both conventional and cloud-based processing paradigms. Traditional workflows download individual Sentinel scenes and process them locally. Cloud-native methods use STAC catalogues, COGs, and distributed computation (for example gdalcubes, rstac, openeo) to build scalable workflows for long time series and large spatial extents.

For cloud-native methods see:

- <https://gdalcubes.github.io/source/tutorials/>
- <https://openeo.org/>
- <https://open-eo.github.io/openeo-r-client/>

Both paradigms remain relevant. Local processing supports transparency, teaching, and technical fundamentals. Cloud-native processing is essential for scaling to continental or multi-year analyses.

References

- Appel, Marius, and Edzer Pebesma. 2019. “On-Demand Processing of Data Cubes from Satellite Image Collections with the Gdalcubes Library.” *Data* 4 (3). <https://www.mdpi.com/2306-5729/4/3/92>.
- Bannari, A., D. Morin, F. Bonn, and A. R. Huete. 1995. “A Review of Vegetation Indices.” *Remote Sensing Reviews* 13 (1-2): 95–120. <https://doi.org/10.1080/02757259509532298>.
- Belgiu, M., and L. Drăguț. 2016. “Random Forest in Remote Sensing: A Review of Applications and Future Directions.” *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Brazil Data Cube Team. 2021. *Rstac: Client Library for SpatioTemporal Asset Catalog*. <https://github.com/brazil-data-cube/rstac>.

- Datta, D., P. K. Mallick, A. K. Bhoi, M. F. Ijaz, J. Shafi, and J. Choi. 2022. “Hyperspectral Image Classification: Potentials, Challenges, and Future Directions.” *Computational Intelligence and Neuroscience* 2022: 3854635. <https://doi.org/10.1155/2022/3854635>.
- Fekri-Ershad, S. 2018. “A Review on Image Texture Analysis Methods.” *International Online Journal of Image Processing and Pattern Recognition* 1 (1): 1–63. <https://doi.org/10.48550/arXiv.1804.00494>.
- Goldstein, Sydney. 2019. “Classifying Satellite Imagery in r.” https://urbanspatial.github.io/classifying_satellite_imagery_in_R/.
- Gonçalves, João. 2018. “Advanced Techniques with Raster Data: Supervised Classification.” <https://www.r-exercises.com/2018/03/07/advanced-techniques-with-raster-data-part-2-supervised-classification/>.
- Hijmans, Robert J. 2023. “Supervised Classification.” <https://rspatial.org/raster/rs/5-supclassification.html>.
- Hogland, J., N. Billor, and N. Anderson. 2013. “Comparison of Standard Maximum Likelihood Classification and Polytomous Logistic Regression Used in Remote Sensing.” *European Journal of Remote Sensing* 46: 623–40.
- Huang, S., L. Tang, J. P. Hupy, Y. Wang, and G. Shao. 2020. “A Commentary Review on the Use of Normalized Difference Vegetation Index (NDVI) in the Era of Popular Remote Sensing.” *Journal of Forestry Research* 32 (1): 1–6. <https://doi.org/10.1007/s11676-020-01155-1>.
- Lahn, Florian. 2024. *OpenEO: Client Interface for 'openEO' Servers*. <https://CRAN.R-project.org/package=openEO>.
- Li, J. et al. 2024. “A Review of Remote Sensing Image Segmentation by Deep Learning Methods.” *International Journal of Digital Earth* 17 (1). <https://doi.org/10.1080/17538947.2024.2328827>.
- Lizarazo, Ivan. 2021. “A Tutorial on Pixel-Based Land Cover Classification Using Random Forests in r.” https://rpubs.com/ials2un/rf_landcover.
- Lu, D., and Q. Weng. 2007. “A Survey of Image Classification Methods and Techniques for Improving Classification Performance.” *International Journal of Remote Sensing* 28 (5): 823–70. <https://doi.org/10.1080/01431160600746456>.
- Mehmood, M., A. Shahzad, B. Zafar, A. Shabbir, and N. Ali. 2022. “Remote Sensing Image Classification: A Comprehensive Review and Applications.” *Mathematical Problems in Engineering* 2022: 5880959. <https://doi.org/10.1155/2022/5880959>.
- Meyer, Hanna, Carles Milà, Marvin Ludwig, Jan Linnenbrink, and Fabian Schumacher. 2024. *CAST: 'Caret' Applications for Spatial-Temporal Models*. <https://github.com/HannaMeyer/CAST>.
- Meyer, Hanna, and Edzer Pebesma. 2021. “Predicting into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models.” *Methods in Ecology and Evolution* 12 (9): 1620–33. <https://doi.org/10.1111/2041-210X.13650>.
- . 2022. “Machine Learning-Based Global Maps of Ecological Variables and the Challenge of Assessing Them.” *Nature Communications* 13 (1): 2208. <https://doi.org/10.1038/s41467-022-29838-9>.
- Meyer, Hanna, Christoph Reudenbach, Tomislav Hengl, Marwan Katurji, and Thomas Nauss.

2018. “Improving Performance of Spatio-Temporal Machine Learning Models Using Forward Feature Selection and Target-Oriented Validation.” *Environmental Modelling and Software* 101: 1–9. <https://doi.org/10.1016/j.envsoft.2017.12.001>.
- Meyer, Hanna, Christoph Reudenbach, Stephan Wöllauer, and Thomas Nauss. 2019. “Importance of Spatial Predictor Variable Selection in Machine Learning Applications – Moving from Data Reproduction to Spatial Prediction.” *Ecological Modelling* 411: 108815. <https://doi.org/10.1016/j.ecolmodel.2019.108815>.
- Montero, D., C. Aybar, M. D. Mahecha, F. Martinuzzi, M. Söchting, and S. Wieneke. 2023. “A Standardized Catalogue of Spectral Indices to Advance the Use of Remote Sensing in Earth System Research.” *Scientific Data* 10 (1): 197. <https://doi.org/10.1038/s41597-023-02096-0>.
- Paradis, E. et al. 2022. “Probabilistic Unsupervised Classification for Large-Scale Analysis of Spectral Imaging Data.” *Computers & Geosciences* 159: 104967.
- Stefan, Valentin. 2020. “Satellite Image Classification in r.” <https://valentinitnelav.github.io/satellite-image-classification-r/>.
- Tolpekin, V., and A. Stein. 2012. *The Core of GIScience: A System-Based Approach*. ITC, University of Twente. <http://www.charim.net/sites/default/files/handbook/datamanagement/3/3.3/The%20core%20of%20GIScience%2C%20a%20system%20-based%20approach.pdf>.
- Verbesselt, Jan, Achim Zeileis, and Martin Herold. 2012. “Near Real-Time Disturbance Detection Using Satellite Image Time Series.” *Remote Sensing of Environment* 123: 98–108. <https://doi.org/10.1016/j.rse.2012.02.022>.
- Wang, D., Z. Yan, and P. Liu. 2025. “Fine-Grained Interpretation of Remote Sensing Image: A Review.” *Remote Sensing* 17 (23): 3887. <https://doi.org/10.3390/rs17233887>.
- Wu, Ling, Zhaoliang Li, Xiangnan Liu, Lihong Zhu, Yibo Tang, Biyao Zhang, Boliang Xu, Meiling Liu, Yuanyuan Meng, and Boyuan Liu. 2020. “Multi-Type Forest Change Detection Using BFAST and Monthly Landsat Time Series for Monitoring Spatiotemporal Dynamics of Forests in Subtropical Wetland.” *Remote Sensing* 12 (2). <https://doi.org/10.3390/rs12020341>.
- Xue, J., and B. Su. 2017. “Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications.” *Journal of Sensors* 2017: 1353691. <https://doi.org/10.1155/2017/1353691>.