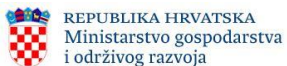


Assessing uncertainty in spatial data analysis

Zagreb 1 / 4 / 2026



REPUBLIKA HRVATSKA
Ministarstvo gospodarstva
i održivog razvoja



AGENCIJA ZA
PLAĆANJA U
POLJOPRIVREDI,
RIBARSTVU I
RURALNOM
RAZVOJU



HRVATSKE
ŠUME



EKONERG

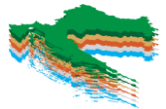


REPUBLIKA HRVATSKA
Državna geodetska uprava



FOND ZA ZAŠTITU OKOLIŠA I
ENERGETSKU UČINKOVITOST





Uncertainty in spatial data analysis

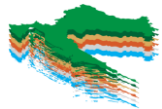


CROLIS Sub-Action 6 : Results & Analysis

What is uncertainty?

- **Precision** is the repeatability of an outcome (values of an estimator over different samples);
- In the CROLIS historic LULUCF model:
 - **Estimator** is the unknown area of a certain LC change (within a certain time period);
 - **Outcome** for a given sample is the estimated area of a certain LC change;
- Precision - *closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions* (JCGM, 2012);
- Our analysis does not include *replicate measurements* - instead we use Monte Carlo sampling methods - choosing random subsets of our one sample.
- **Uncertainty** is the opposite of precision - low precision (low repeatability) means high uncertainty of the estimates.





Uncertainty in spatial data analysis



CROLIS Sub-Action 6 : Results & Analysis

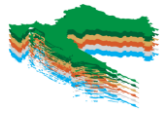
Uncertainty of sample-based estimates

- Characterized by the variance (dispersion) of the estimates - high variance corresponds to low precision and high uncertainty;
- **Variance** - measure of uncertainty/precision - calculated as the mean (over all possible samples) of the squared differences between the estimate for a sample and the population parameter - here the population parameter is the true area of a certain LC change;
- **Exact variance cannot be observed!** But it can be estimated from the one sample obtained.
- Say we have 100 different samples:
 - each sample produces an estimate of the considered population parameter;
 - all samples combined produce a mean estimate of the same population parameter;
- **Sample variance** = calculated as the mean (over all 100 samples) of the squared differences between the estimate for a sample and the mean estimate.

Standard error - the square root of the variance of an estimator;

- The standard error characterizes variability on the same scale as the estimate;
- If we are estimating LC change area in hectares, the standard error is also in hectares, whereas the variance has units of hectares².





Uncertainty in spatial data analysis



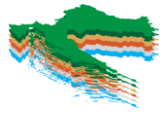
CROLIS Sub-Action 6 : Results & Analysis

Confidence intervals

- Estimates are usually reported together with corresponding 95% (99%) confidence intervals;
- **Confidence intervals** = estimates of lower and upper bounds (confidence limits) for a parameter, forming a range in which plausible values of that parameter lie;
 - The percentage (95%, 99%) indicates the level of plausability.
- Unlike (simple) estimates, intervals take into account variability due to sampling - important since an estimate produced from a sample is unlikely to match the population parameter.
- Assuming normal distribution of parameter estimates from different sample realizations (this has to be checked!), the 95% confidence interval is usually computed as

Point estimate \pm 1.96 x standard error.





Uncertainty in LC maps



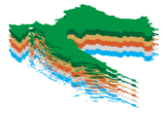
Map accuracy assessment

Uncertainty in land-cover products is fundamentally about map errors (accuracy).

Why does map accuracy matter for uncertainty assessment?

- **Determining whether the map can be trusted at all** - LC maps contain errors, and those errors must be assessed against an independent reference sample; otherwise the map is just a prototype or untested hypothesis.
- **What kind of uncertainty is present** - Overall accuracy gives one summary number, class-specific metrics are more informative than a single uncertainty score - different classes can have very different reliability.
- **Quantifying uncertainty statistically** - using random sub-samples we can report **standard errors and confidence intervals** for accuracy and area estimates.
 - Not only a single estimate, but a reliable measure of how uncertain that estimate is.
 - **Without a rigorous accuracy assessment based on independent reference data, uncertainty remains poorly defined and area estimates derived from the map may be misleading.**





Uncertainty in LC maps

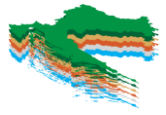


Map accuracy assessment

Confusion matrix and overall accuracy

- **Confusion matrix (error matrix):** a cross-tabulation comparing **map labels** (predicted classes) with **reference labels** (actual classes) for a sample of map units.
 - **Main diagonal:** correctly classified cases.
 - **Off-diagonal entries:** misclassifications, showing where classes are confused.
 - The confusion matrix is the basic tool from which standard map accuracy measures are derived.
- **Overall accuracy:** the proportion or percentage of total area classified correctly.
 - It is obtained by summing the diagonal elements of the confusion matrix when these are expressed as **area proportions**.
 - Overall accuracy is a **single summary measure** for the full set of land-cover classes.
 - It can be computed for an entire map or for subregions, such as countries, biomes, or other domains.





Uncertainty in LC maps



Map accuracy assessment

Class-specific accuracy metrics

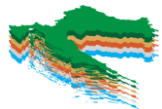
- **Class-specific accuracy metrics** describe performance for an individual land-cover class.
- **Omission error:** a case truly belonging to class X is mapped as another class; the class has been wrongly **left out**.
- **Producer's accuracy:** how well the map captures a real class on the ground.
 - Producer's accuracy = **100% – omission error**
 - In binary classification, this is also called **sensitivity** or **recall** for the target class.
- **Commission error:** a case is wrongly **included** in a mapped class.
- **User's accuracy:** how often a mapped class is actually correct on the ground.
 - User's accuracy = **100% – commission error**
 - In binary classification, this is also called **precision**.
- **F1 score:** the harmonic mean of **user's accuracy** and **producer's accuracy** - a balanced score (measure).



Weight selection in CROLIS historic LULUCF model:

- Weight selection can be informed by uncertainty/precision;
 - Choosing weight between two extreme weight models which produce least uncertainty.
 - Question: Overall uncertainty, uncertainty per LC change class, ...
- **Model uncertainty** = refers to the repeatability of results between different model runs;
 - Here different model runs mean using different subsets of the training dataset (our sample).
- Produce 95% confidence intervals for LC change areas (for each class).





University of Nottingham
LE 10EN | 801058



Committee on Earth Observation Satellites
Working Group on Calibration and Validation
Land Product Validation Subgroup
Land Cover Focus Area



**Land Cover and Change Map Accuracy Assessment and
Area Estimation Good Practices Protocol**

Version 1.0 - 2025

