

TESTING STAGE / FITTING PHASE: 3. MODEL FIT ASSESSMENT

The third step undertakes an assessment of the quality of the model fit. The primary purpose is to provide a basis for comparing models to select the best approach. To do this, it divides up the jurisdiction into a coarse grid where the grid cells have an area the same size as typical REDD projects. Cells which intersect the edge of the jurisdiction are removed (loss of data is not a concern – the purpose is simply to yield a representative assessment for model comparison). The suggested size is the median area (in ha) of REDD projects in the jurisdiction. If this is unknown, a size of 100,000 ha is suggested.

The procedure then displays a performance chart with actual deforestation on the X axis and fitted deforestation on the Y axis. Each dot represents one of the coarse grid cells on these axes. Three lines are provided. The red line is the 1:1 line which indicates the expected fitted deforestation if it matched the actual deforestation within a coarse grid cell. A summary statistic called the Median Absolute Error (MedAE) shows the median departure of the blue dots from the 1:1 line. It is measured in hectares.

The 1:1 line is the important reference line. However, the OLS best-fit regression line is also provided as a visual reference. Ideally, the best-fit line will coincide with the 1:1 line. Commonly, it will be found to have a slope that is flatter than the 1:1 line. There are several causes of this phenomenon. In the fit, a key cause of this attenuation is map error. If a strong slope bias is present, attention should be focused on the quality of the data. Note that it's not only errors in the forest cover maps that are a problem, but also errors in covariates used in model development. The third line that's provided is a Theil-Sen regression line. This is a robust regression line that is resistant to the effects of outliers. If the Theil-Sen fit is similar to the 1:1 line, it indicates that outliers are fewer than 29% (approximately) of the samples.

To qualify for use in the final assessment, an alternative modeling procedure must beat the benchmark for both the fit (this result) and the prediction using the MedAE statistic. Otherwise, the benchmark procedure should be used.

INPUTS

WORKING FOLDER

The computer folder where inputs are expected and outputs are written.

FITTED DENSITY FOR THE CAL IMAGE (HA/PIXEL)

This is the final output from Step 2 – the map of fitted density of deforestation in the CAL.

MASK OF THE JURISDICTION

This is a binary map (contains 0's and 1's) where the 1's indicate areas inside the jurisdiction and suitable for consideration.

Important: This mask should cover the entire jurisdiction and not consider excluded areas. Note that maps with 1's and NAN's are not equivalent. All binary maps must be 1's and 0's with this tool.

DEFORESTATION IN THE CAL IMAGE

This is also a binary map (contains 0's and 1's) where the 1's indicate areas that were deforested during the CAL and 0's indicate areas of forest persistence or non-forest at the start of the CAL.

ASSESSMENT GRID CELL AREA OR 100,000 (HA)

Specify the median size of REDD projects in the jurisdiction, in hectares. If this is unknown, use 100,000 ha.

OUTPUTS

TITLE OF PLOT

A descriptive title to be used for the performance chart.

OUTPUT PLOT FILENAME

Specify a name for the file to be used to save the performance chart graphic. Several formats are supported. Select the format by specifying the appropriate file extension: “.png”, “.jpg”, “.pdf”, “.svg”, “.eps”, “.ps” and “.tif”. Note that a CSV file of the data used for the graph will also be created using the same name and a “.csv” extension.

OUTPUT RESIDUAL IMAGE

The residual image is a map of the differences between the fitted deforestation and the actual deforestation in each assessment grid cell.

SET X-AXIS MAX MANUALLY

You may optionally indicate a specific maximum to use rather than the default for the X-axis.

SET Y-AXIS MAX MANUALLY

You may optionally indicate a specific maximum to use rather than the default for the Y-axis.