

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. Two 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. In the fit, the primary cause of this attenuation is map error. If a strong slope bias is present, attention should be focused on the quality of the data.

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 MAP FOR THE CAL

This is the final output from Step 2 – the map of fitted density of deforestation in the CAL. It is fitted in the sense that it is estimated from the relative frequency of deforestation experienced in the HRP within modeling regions.

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. Areas that are to be excluded from consideration (such as planned forestry concessions) should be marked with 0's. 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.

ACTUAL DEFORESTATION DURING THE CAL

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

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.

NAME FOR THE PLOT

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”.

NAME FOR THE CHART DATA

The chart data will be saved in a Common-Separated-Values file. Specify the name desired with a “.csv” file extension.

NAME FOR THE ASSESSMENT POLYGONS

The assessment polygons will be saved as a Shape file. Specify the name desired with a “.shp” file extension.