

The third step of the prediction phase undertakes an assessment of the quality of the model prediction for the Confirmation Period (CNF). 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 predicted 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 predicted 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 prediction, the cause of this attenuation is primarily a combination of map error and non-stationarity. Non-stationarity is caused by deforestation activity not behaving in a manner similar to that observed during the time the model was fit. It is expected that the slope bias during the prediction will be stronger than that observed during the fit. This difference is indicative of the effect of non-stationarity. The third line is the robust Theil-Sen regression line. If the Theil-Sen line approximately matches the 1:1 line, it is an indication that outliers (such as from non-stationarity) affect less than 29% of the evaluation grid cells.

To qualify for use in the final assessment, an alternative modeling procedure must beat the benchmark for both the fit and the prediction (this result) 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.

#### ADJUSTED PREDICTED DENSITY IN THE CNF

This is the final output from Step 2 – the map of predicted density of deforestation in the CNF, adjusted so that predicted deforestation matches actual deforestation during the CNF.

#### 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 CNF

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

#### DEFORESTATION IN 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.

---

## FOREST AREAS AT START OF CAL

This is also a binary map (contains 0's and 1's) where the 1's indicate areas that were forest at the start of the CAL and 0's indicate areas of 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". The chart data will also be saved in a Common-Separated-Values file using the same name with a ".csv" file 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.

---

### OUTPUT COMBINED DEFORESTATION REVIEW IMAGE

The Combined Deforestation Review Image is a map that illustrates deforestation that happened at both the CAL and CNF stages.

---

### SET X-AXIS MAX

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

---

### SET Y-AXIS MAX

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

---