[Updated 2022-11-22]

MODEL AND INPUT DATA

We developed a random forest model, which runs on Google's Earth Engine platform, that classifies fused monthly composite images to differentiate between mined and unmined areas. The imagery that the model classifies is a fused data product that we create composed of monthly Sentinel-1 median composites and monthly, cloud-free Sentinel-2 median composite images. The model was trained using 8,000 sample points distributed across the Madre de Dios region of Peru, using fused Sentinel-1/Sentinel-2 imagery from both the wet and dry season. These points were randomly distributed across areas delineated by remote sensing experts that were identified as either mining, agricultural, or non-mining; the agricultural areas were delineated to improve the model's accuracy during months in the dry season, when many fields change from green to brown.

POST-PROCESSING

To improve the results of our detection model, we also employ a number of post-processing steps. We use data from the European Commission's Global Surface Water dataset to remove surface water bodies that are erroneously classified as mining activity. In addition, we implement a simple count filter to declassify any pixel which is not classified as mining a minimum of three times (we do not apply this count filter to the last two months of imagery available).

ACCURACY ASSESSMENT

To assess the accuracy of our model, we conducted a manual classification of randomly selected points across twelve monthly composite images from 2019–2021, with four months selected per year, as follows:

Year	Selected months
2019	Feb, Apr, Jul, Dec
2020	Feb, Jun, Sep, Oct
2021	Mar, Jun, Aug, Nov

A set of 10 sample areas were selected (these remained consistent across all months), and 5,000 random points were distributed across each sample area for each month classified (these

were generated uniquely for each month) and classified as either non-mine or mine. After the classification of points was finished, a random set was chosen for each month. For each sample area, 50 points classified as mine (if 50 were not available, all mine classified points were used) and 150 points classified as non-mine were randomly selected. In instances where no mine points were present for a sample area, 200 non-mine points were used.

The key metric for determining the accuracy of our model is the User's Accuracy for Mines, which indicates the percent of pixels labeled as mine that actually represent mines on the ground; we find that for the 12 months classified, we have a minimum user accuracy score of 86% and a median user accuracy score of 94%.