OVERVIEW

Space Intelligence HabitatMapper^{TM}

About Space Intelligence

Founded by Dr. Murray Collins and Prof. Edward Mitchard, both former University of Edinburgh academics, Space Intelligence specializes in satellite-based mapping of tropical land cover and carbon storage. Our team of 60 experts, including 13 PhDs, combines scientific and AI expertise to process petabytes of satellite data and produce world-class nature mapping data.

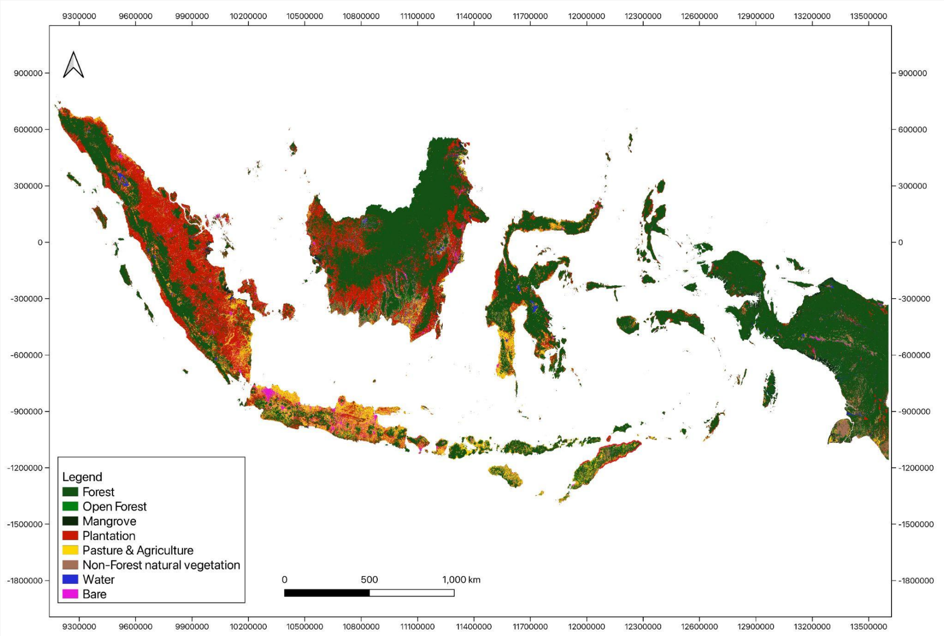
We provide high-quality mapping data to global corporations and governments, supporting nature-based solutions (NbS) projects for clients such as Apple, Shell, Laconic and Equinor, and leading forest carbon developers.

PRODUCT

HabitatMapper^{TM}

Our HabitatMapper^{TM} technology delivers highly accurate maps of land cover over any ecosystem on Earth. We use satellite data from a variety of sensors to produce a time series of maps that indicate changes over time, allowing for the assessment of deforestation (or other land cover changes) and forest regrowth across project sites or entire countries.

- 10 to 25m resolution
- >90% Overall Accuracy
- 6+ land cover classes



Space Intelligence 2023 landcover map over Indonesia

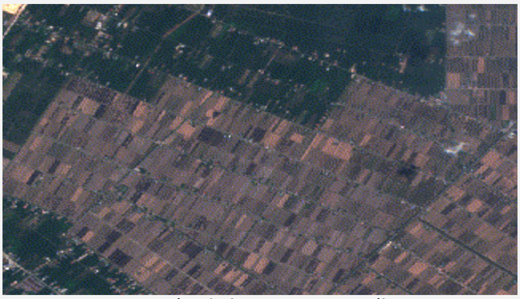
Data inputs to our land cover maps

Our land cover maps are produced using a dense stack of satellite data from multiple sources, using machine learning algorithms that are trained and tested with expert ecological input. Each map involves data from a single calendar year, with the central date of the satellite mosaics detailed in the product specifications.

Optical data

Optical satellite sensors use multiple bands throughout the electromagnetic spectrum to help differentiate various types of land cover, and are useful for spotting visual changes in vegetated and non-vegetated surfaces[1]. At Space Intelligence, we use:

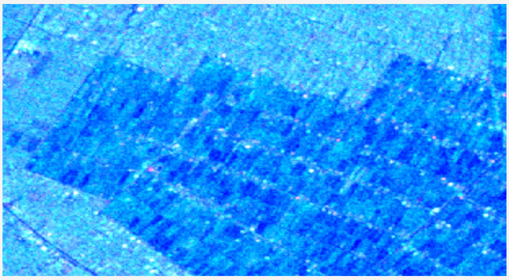
- Sentinel-2: 10m resolution, operated by ESA/EU. Two or more satellites in orbit, with data every ~5 days everywhere in the tropics.
- Landsat 8 & 9: 30m resolution, operated by USGS/NASA. Two satellites in orbit, with data every ~8 days everywhere in the tropics.


Optical data over Brazil

Synthetic Aperture Radar (SAR) data

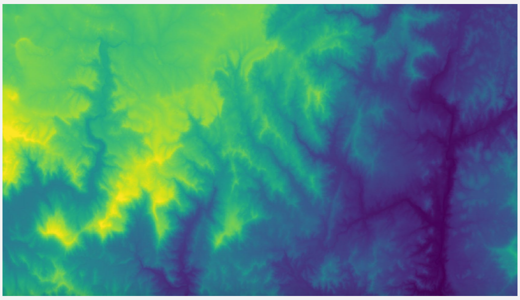
SAR satellite sensors emit microwave signals that can see through clouds, providing information on the orientation, density and water content of structures on the Earth's surface (e.g., tree branches and trunks). At Space Intelligence, we use:

- Sentinel-1: 10m resolution C-band (6cm wavelength), operated by ESA/EU. Two or more satellites in orbit, with data every ~12 days everywhere in the tropics.
- ALOS-2 PALSAR-2: 10 to 100m resolution L-band (23cm wavelength), operated by JAXA. Data every ~42 days.


SAR data over Brazil

Topography data

Topography datasets from Copernicus DEM (digital surface model that includes terrain info on slope, elevation, and aspect) are vital as they impact the type of landcover that flourishes in a region. These datasets help our machine learning algorithms correct for topography-related effects that might otherwise affect the map accuracy.


Elevation data over Brazil

[1]Nomura, K and ETA Mitchard. 2018. More than meetings the eye: Using Sentinel-2 to map small plantations in complex forest landscapes. *Remote Sensing* 10, 1693. doi: 10.3390/rs10111693.

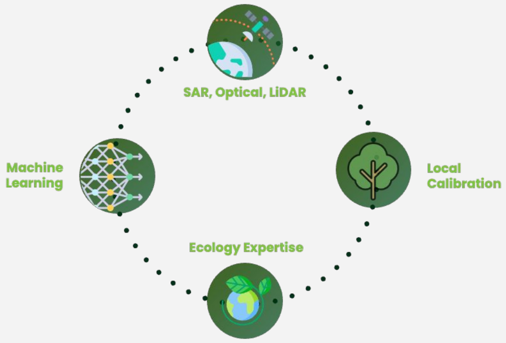
OVERVIEW

Space Intelligence HabitatMapper™

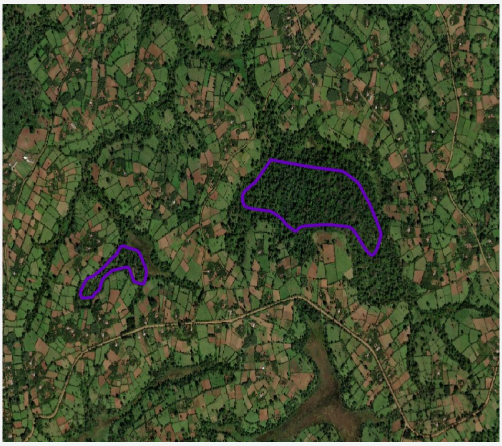
Making the maps

Data preparation

The satellite datasets are combined into data cubes after the application of smart techniques to bring out seasonality in the data, which is key in the delineation of tricky land cover classes. We also add layers derived from the satellite data to the data cubes, mostly vegetation indices, to further help our machine learning models with nuanced differences between the land cover classes.



Cloud and cloud shadow are removed from the optical satellite data, and all dataset are terrain and radiometrically corrected and aligned. The resulting data cube has many tens of bands and is at 10 m x 10 m pixel size, meaning hundreds of billions to trillions of satellite data points over a country.



Truth points and polygons

We combined multiple data sources: field data, open geographically located photos, existing mapping datasets, and our experts looking at very high resolution (<1 m) satellite data, to create a set of polygons that reliably represent specific land cover classes.

This includes creating high quality truth polygons making the best use of the expert knowledge of our world class ecology team and cutting edge machine learning techniques. These truth polygons used to train our machine learning models are critical to separating these classes, which often superficially look similar from space.

Machine learning

Our data scientists then train and run a machine learning algorithm combining the truth polygons and the satellite data cube, to predict maps for the requested time period. These are tested against the independent test dataset. Based on the results and visual inspection of the map by our experts, we improve the map by re-running it, making changes to the truth polygons, the satellite datasets used, or the machine learning algorithm parameters, until we have a final map for a country that meets our accuracy requirements.

Accuracy assessment process

We take the quality of our maps very seriously at Space Intelligence and assess the accuracy and uncertainty of our maps in a statistically rigorous way, following good practice described in the scientific literature[1] and international standards[2,3]. We set targets for performance based on the requirements of a specific contract.

We report the accuracy of our maps using three different but important accuracy metrics: (1) commission or user's accuracy; (2) omission or producer's accuracy; and (3) overall accuracy. We consistently aim to an accuracy level of >95% in both commission and omission accuracy.

Summary

Accuracy is nuanced, and we are experts in that nuance.

We believe a single number is not an accurate assessment of accuracy, and a more detailed analysis is the only way to truly assess this metric.

We ensure our maps meet the needs and expectations of our partners, and that we provide full information on performance and the expected rate of inevitable uncertainties and errors across our map.

Overview of approaches to accuracy assessment

Accuracy can be tested for in a number of ways but the **only approach that gives a meaningful assessment of the accuracy of a random pixel in the output map is approach 4.** This is the approach used by Verra in the new Consolidated REDD+ Methodology (VM0048) and recommended by IPCC good practice guidance[1-3], but is not always followed by companies keen to parade high accuracy statistics.

1 A comparison of input training data (normally polygons drawn by eye with classes such as 'forest' or 'non-forest') compared to the output map.

2 A similar comparison of the output map with polygons, but using independently produced polygons not used to train the map. Normally the unit used is the proportion of pixels correctly classified ('overall accuracy'), with individual pixels being the unit of assessment, not the whole polygons.

3 Using data collected in-situ (field data).

4 The independent assessment of a set of isolated points, placed over the entire output map using a statistically valid sampling approach (usually a grid or a stratified random sample).

---

[1] Olofsson, P., et al. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2014.02.015

[2] The International Panel on Climate Change (IPCC). 2003. Good Practice Guidance for Land Use, Land-use Change and Forestry. https://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_contents.html

[3] GFOI. 2020. Methods and Guidance from the Global Forest Observations Initiative (MGD). Edition 3.0. https://www.reddcompass.org/mgd/resources/GFOI-MGD-3.1-en.pdf

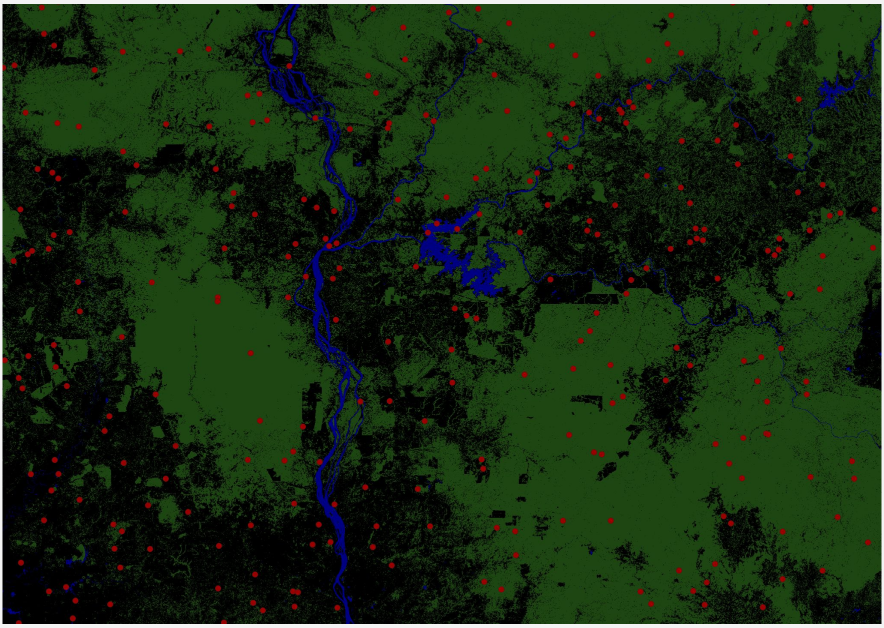
Case study: Cambodia National Maps

Cambodia experiences high deforestation rates, primarily from small-scale clearing for various crops, including palm oil, rubber, and betel nuts. Its forests comprise a complex mix of dry, moist, and wet areas, including swamps and mangroves.

As a case study, here are the findings from a 2020 land cover map by Space Intelligence, part of a 13-year analysis of the country's land cover dynamics. The maps demonstrate high accuracy, effectively distinguishing forest from non-forest and successfully excluding timber plantations, tree crops, and rubber plantations.

	Commission accuracy (User's accuracy)	**Omission accuracy** (Producer's accuracy)	**Overall Accuracy**
Forest	96 ± 2 %	94 ± 2 %	-
Non-forest	96.2 ± 1.6 %	97.5 ± 1.3 %	-
-	-	-	96.1 ± 1.3 %

Thematic accuracies for Space Intelligence's Cambodia forest data for 2023, assessed independently using a probabilistic sampling design, showing high overall accuracy but also low amounts of confusion between forest and non-forest. All values include the 95th confidence interval.



A subset of Space Intelligence's Cambodia forest data, showing forest in green, non-forest in black and water in blue. Red dots are some of the thousands of random accuracy assessment points used to assess the accuracy of the map.

Appendix: Assessment of Accuracy Approaches

Accuracy can be tested for in a number of ways but the only approach that gives a meaningful assessment of the accuracy of a random pixel in the output map is approach 4. This is the approach used by Verra in the new Consolidated REDD+ Methodology (VM0048) and recommended by good practice guidance[1-3], but is not always followed by companies keen to parade high accuracy statistics.

Approach 1: A comparison of input training data (normally polygons drawn by eye with classes such as 'forest' or 'non-forest') compared to the output map. Normally the unit used is the proportion of pixels correctly classified ('overall accuracy'), with individual pixels being the unit of assessment, not the whole polygons.

Assessment: *While common, this method is not scientifically appropriate as there is no independence between test and training datasets.*

A further issue is the use of pixels not polygons as the basis of comparison. Neighbouring pixels (for example all pixels making up a field cleared in a forest) are treated as independent samples, whereas in fact they are neighbours and share many more characteristics than two random pixels in the output *map.*

For both these reasons, it will inevitably overstate accuracy.

Approach 2: A similar comparison of the output map with polygons, but using independently produced polygons not used to train the map. Normally the unit used is the proportion of pixels correctly classified ('overall accuracy'), with individual pixels being the unit of assessment, not the whole polygons.

Assessment: This shares the issue above of the use of treating non-independent neighbouring pixels as independent, overstating accuracy

Approach 3: Using data collected in-situ (field data).

Field data is often proposed as the 'gold standard' of validation data, and assumed of higher accuracy than other sources of data.

Assessment: However, because of accessibility and its elevated cost as compared to, for instance, interpretation of high resolution remote sensing data, the small sample size and narrow proportion of the area available for sampling means its results cannot be extrapolated reliably to large regions.

It also tends to be spatially biassed, which can be a problem when assessing land cover products, since accessible areas are typically more disturbed.

Approach 4: The independent assessment of a set of isolated points, placed over the entire output map using a statistically valid sampling approach (usually a grid or a stratified random sample).

Assessment: this approach provides a reliable assessment of the accuracy of different classes of a map, and the confidence intervals of that assessment. More points can be added in a statistically valid way until confidence intervals on the accuracy assessment are sufficiently narrow to meet requirements.

---

[1] Olofsson, P., et al. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2014.02.015

[2] The International Panel on Climate Change (IPCC). 2003. Good Practice Guidance for Land Use, Land-use Change and Forestry. https://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_contents.html

[3] GFOI. 2020. Methods and Guidance from the Global Forest Observations Initiative (MGD). Edition 3.0. https://www.reddcompass.org/mgd/resources/GFOI-MGD-3.1-en.pdf