

オープンソースソフトウェアを用いた 衛星画像の教師なし分類による林況把握

明石 信廣

はじめに

広大な森林に関わる仕事をしていれば、針葉樹人工林にどの程度広葉樹が侵入しているか、病虫害がどの程度広がっているかなど、森林の様子を上空から確かめたい、と思うことがしばしばあるのではないかでしょうか。そこで、古くから航空機からの空中写真や人工衛星からのリモートセンシングの技術が活用され、近年では UAV が活躍しています。それぞれの手法には、活用できる広さや分解能、必要な費用などに違いがあります。

アメリカ地質調査所 (USGS) とアメリカ航空宇宙局 (NASA) が運用している Landsat や、欧洲宇宙機関 (ESA) が運用している Sentinel-2 などの衛星画像は、無償で利用することができます。Landsat は最初の衛星が 1972 年に打ち上げられ、後継機に引き継がれて観測が続けられています。Sentinel-2 は 2015 年に最初の衛星が打ち上げられました。どちらもさまざまな波長を観測するセンサーが搭載され、波長帯によって分解能も異なりますが、植生の分類でよく使われる赤、緑、青、近赤外の波長帯の分解能は Landsat が 30m、Sentinel-2 は 10m です。雲がかかっていると画像が得られないのが弱点ではありますが、撮影後、短期間で画像が公開されるので、分解能が 10m 以上でも判別できる林相や被害であれば、まずは無償の衛星画像で確認してみる、ということが可能になりました。

画像から森林の状態などを分類することで、地図化したり、状態を定量的に示すことも可能になります。近年は、このような衛星画像の活用に役立つオープンソースソフトウェアも充実してきています。オープンソースソフトウェアとは、ソースコードが公開され、無償で自由に利用したり改変したりすることができるソフトウェアのことです。

本稿では、Sentinel-2 の画像と QGIS と GRASS GIS という 2 つのオープンソースソフトウェアを用いて、教師なし分類による林況把握を行った手順の概略を説明しています。詳細な手順については「衛星画像による林況把握方法 オープンソース GIS ソフトによる『教師なし分類』の作業手順（改訂版）」(https://www.hro.or.jp/forest/research/fri/develop/keiei/rinkyohaaku.html) をご覧ください。

教師なし分類とは

衛星画像から土地利用や植生タイプなどの土地被覆分類を行う方法として、「教師付き分類」と「教師なし分類」があります。衛星画像は分解能に応じた画素（ピクセル）で構成され、人工衛星に搭載されたセンサーが観測する波長帯（赤、緑、青、近赤外など）ごとに反射率が記録されています。教師付き分類とは、分類したいカテゴリー（例えば、常緑針葉樹林、落葉広葉樹林、裸地など）に含まれることが分かっている多数の地点を教師として与え、その地点のピクセルの情報から、どのような反射率がどのカテゴリーと対応するかを分析し、その結果にしたがって土地被覆を分類する手法です。一方、教師なし分類は、画像から抽出されたピクセルの反射率のデータをクラスター分類により類似したものをクラス分け（クラスタリング）する手法で、分類されたそれぞれのクラスがどのようなカテゴリーに該当するかを後から対応づけます。教師付き分類では、教師となる情報を適切に与えれば、似たようなところを細かく分類することができますが、教師となる情報を作るための作業が必要です。教師なし分類では、常緑針葉樹と落葉広葉樹などを比較的簡便な手順で分類することができます。

衛星画像を見てみよう

Sentinel-2 の画像は ESA の Web サイト Copernicus Browser (<https://browser.dataspace.copernicus.eu/>) で閲覧できます。データのダウンロードにはユーザー登録が必要ですが、ウェブブラウザ上での閲覧だけなら Anonymously (匿名) でも利用できます。衛星画像がどのように撮影されているのか、まずは見てみましょう。Sentinel-2 は 2 機の衛星で構成され、それぞれが 10 日ごとに同じ場所を撮影しています。2024 年には 3 機目の衛星の運用が開始されました。しかし、地表はしばしば雲に覆われており、画像を入手したい範囲が広くなると、全域に雲のない画像が撮影される機会が少なくなり、年に数回だけという場合もあります。範囲を限定するほど雲のない画像が得られる可能性は高まります。1 枚の画像は 10,980 × 10,980 ピクセル、1 ピクセルは 10m × 10m なので、約 110km 四方が含まれます。

森林に関する地図情報

衛星画像の解析にあたって、森林調査簿に記録されている植栽樹種などの情報を参考にします。近年、森林の林班、小班の区画や植栽樹種などの情報は、オープンデータとして公開されるようになりました。国有林のデータは国土交通省の国土数値情報ダウンロードサイト (<https://nlftp.mlit.go.jp/>) で「国有林野」として公開されています。また、北海道有林は「道有林森林資源情報資料ダウンロードページ」 (<https://www.pref.hokkaido.lg.jp/sr/dyr/DOP.html>)、その他の民有林（一般民有林）は「森林計画関係資料オープンデータ」 (<https://www.pref.hokkaido.lg.jp/sr/srk/A004/B003/>) からダウンロードすることができます（図-1）。

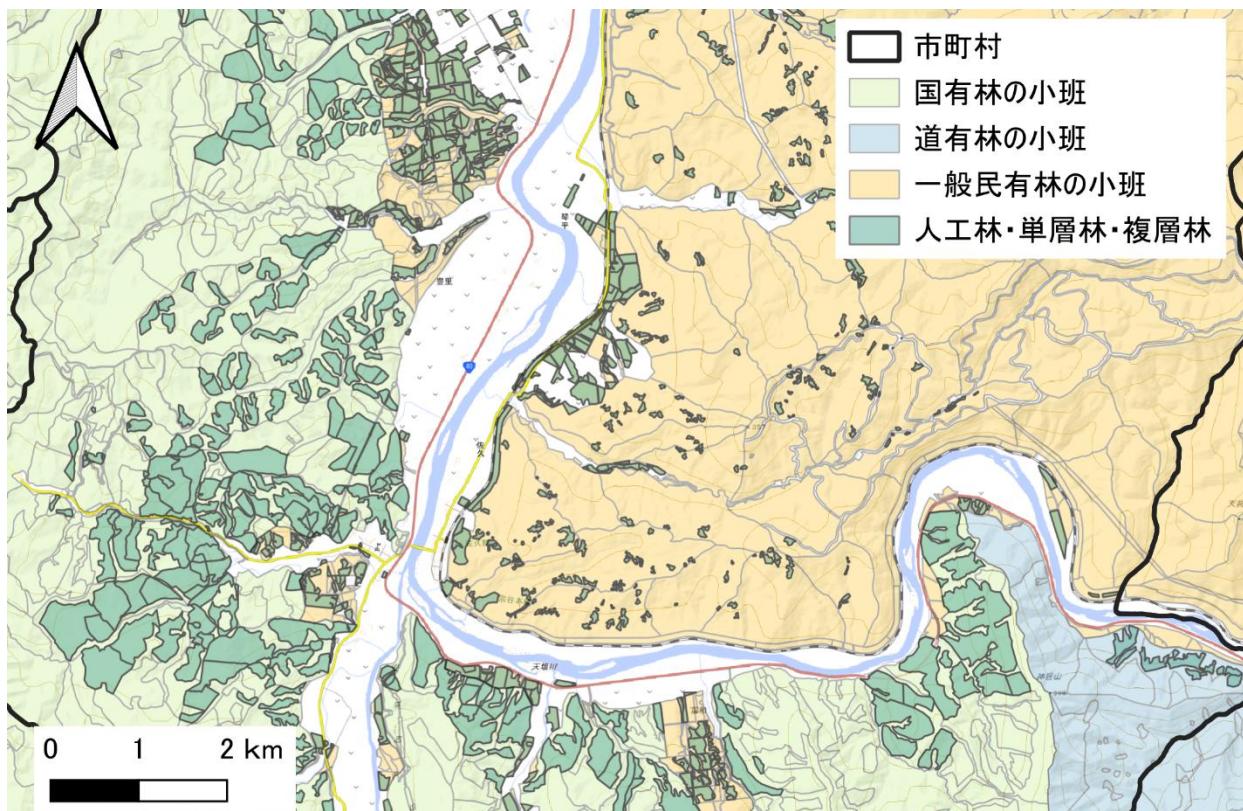


図-1 国有林、道有林、一般民有林の小班区画データ (QGIS で作図)

国有林は国有林 GIS データ (<https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A45.html>)、道有林は「道有林森林資源情報資料ダウンロードページ」 (<https://www.pref.hokkaido.lg.jp/sr/dyr/DOP.html>)、一般民有林は「森林計画関係資料オープンデータ」 (<https://www.pref.hokkaido.lg.jp/sr/srk/A004/B003/>)、背景の地図は地理院タイル標準地図 (<https://maps.gsi.go.jp/development/ichiran.html>) を利用しました。

教師なし分類の手順

教師なし分類は図-2に示した手順で進めます。

(1) 衛星画像の準備

Sentinel-2の画像は、QGISのSentinel 2 Image Downloaderプラグインを使ってダウンロードすることができます。植物は赤色や青色の光を吸収して光合成に利用し、近赤外光はよく反射するため、植物の量などを表す指標として、赤と近赤外の反射率から求められる正規化植生指標（NDVI）がしばしば用いられます。赤、緑、青、近赤外のほか、NDVIもあらかじめ用意されています。

これら5つのバンドのデータをダウンロードしたら、1つのTIFファイルにまとめ、分類したい範囲を切り抜いて保存します。

衛星画像の赤、緑、青のバンドをそれぞれQGISで赤のバンド、緑のバンド、青のバンドに割り当てると、人の目で見た配色に近いトゥルーカラー画像が作成できます（図-3）。

(2) クラスタリング

この画像をGRASS GISに読み込み、i.clusterモジュールを使って教師なし分類のためのクラスタリングをします。クラスタリングとは、値の近いピクセルをまとめてクラスに分ける作業です。画像から一部のデータを抽出し、クラスの初期数を指定して実行します。この結果をもとに、i.maxlikモジュールで分類結果を画像全体に反映させ、ピクセルごとにクラスの番号が付けられた画像を作ります（図-4）。各クラスには1から順に番号が付けられ、クラス数が50個であれば1から50までの50階調になります。

(3) クラスタリング結果の意味づけ

次に、分類結果の画像をQGISで表示させ、それぞれのクラスがどんな状態に対応するのか、意味づけをしていきます。

例えば、図-3では、小班の属性や現地の状況から、濃い緑はトドマツ、やや薄い緑は落葉広葉樹、明るい緑は草地を示していると思われます。このようなピクセルの情報を見て（図-4）、そのクラスのラベルを常緑針葉樹林、落葉広葉樹林、草地などに変更し、画面に表示する色を分類したいカテゴリーに合わせて1つずつ変更します。河川や道路、雲などは、比較的容易に識別できる場合が多いでしょう。

図-4を見ると、常緑針葉樹林や落葉広葉樹林などもそれぞれ複数のクラスに分けられていることが分かります。ある程度妥当と思われたカテゴリーに統合した結果、49のクラスは9個のカテゴリーにまとめられ



図-2 衛星画像の教師なし分類による林況把握の手順

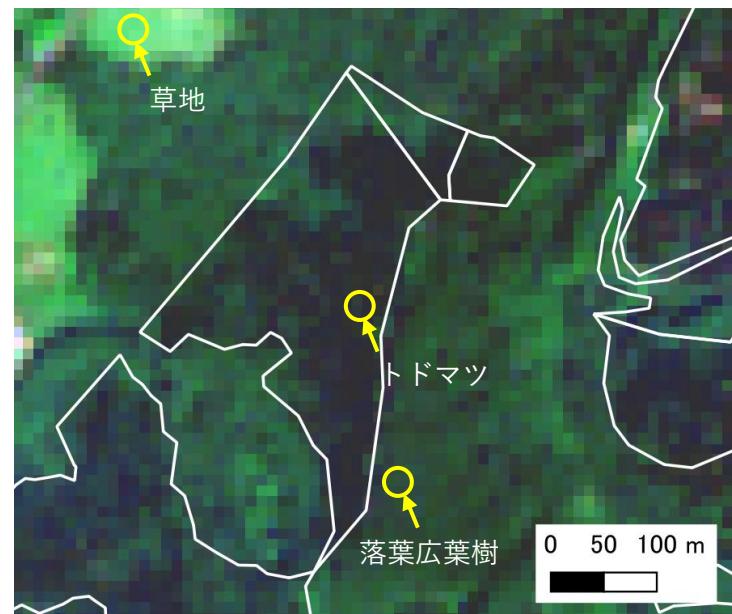


図-3 トゥルーカラー画像 (QGISで作図)

Sentinel-2の2024年7月15日のトゥルーカラー画像。白線は小班。草地、トドマツ、落葉広葉樹の代表的な地点を○印で示す。

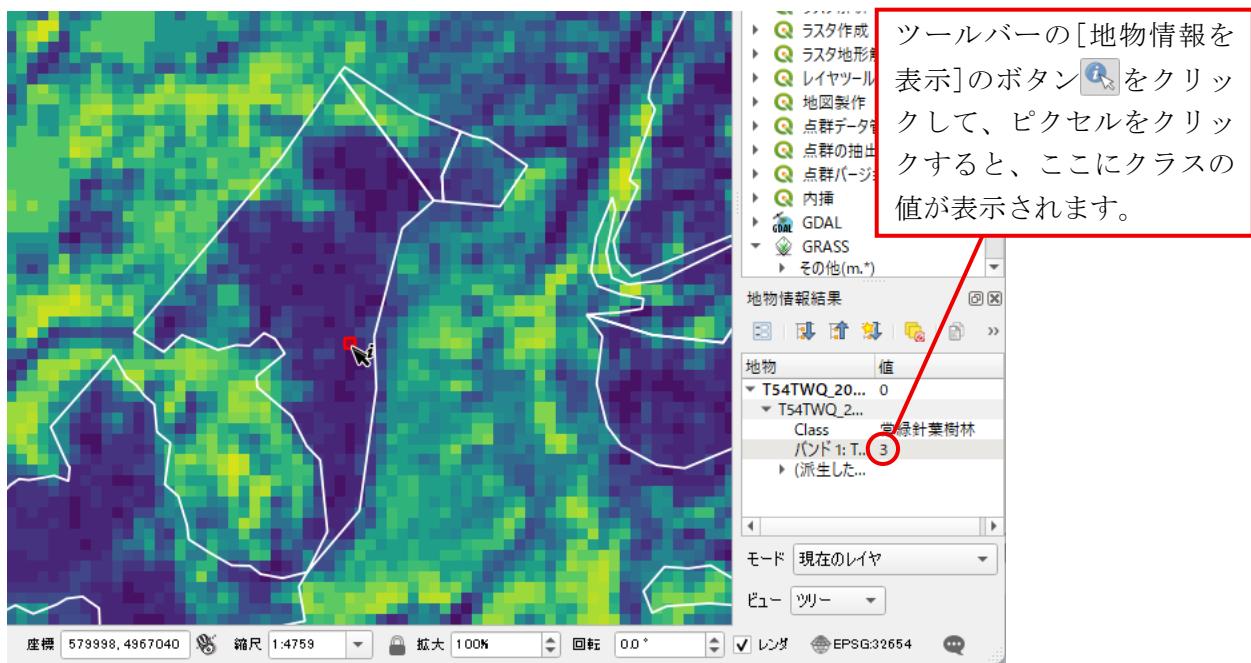


図-4 クラスタリング結果 (QGIS の画面)

Sentinel-2 の 2024 年 7 月 15 日の画像を、クラスの初期数を 50 として分類した画像の一部を、クラスごとに色分けして示した。実際には 49 クラスに分類されている。

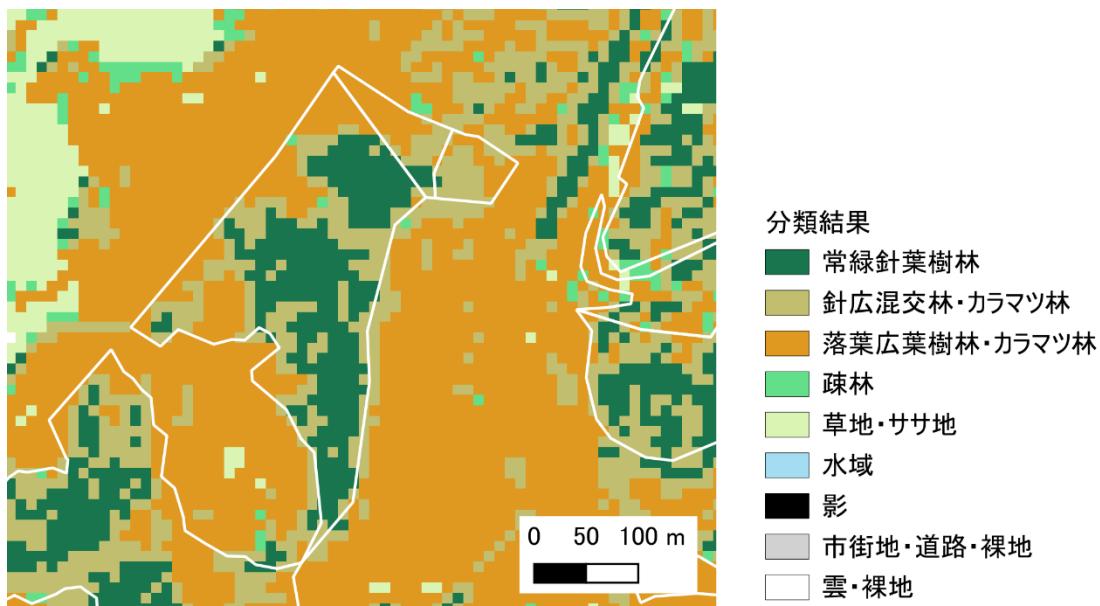


図-5 分類結果 (QGIS で作図)

ました。今回の画像では、落葉広葉樹林とカラマツ林には同じクラスに分類されたピクセルが多く、これらを区別することはできませんでした。また、衛星は毎回ほぼ同じ時刻に撮影するように運用されており、夏至に近い時期に撮影された画像をもとに作成された図-3 には影の部分はありませんが、樹木の北側に影の部分がある場合には、黒っぽい常緑針葉樹との区別がつかないこともあります。

図-6 には、各クラスの赤と近赤外の値と分類されたカテゴリーの関係を示しています。赤は植物によって吸収されるため、雲や裸地、市街地など植物のないところで値が大きく、森林では小さくなっています。赤も近赤外も小さな値のところは河川などの水域や影でした。草地やササ地は赤、近赤外と

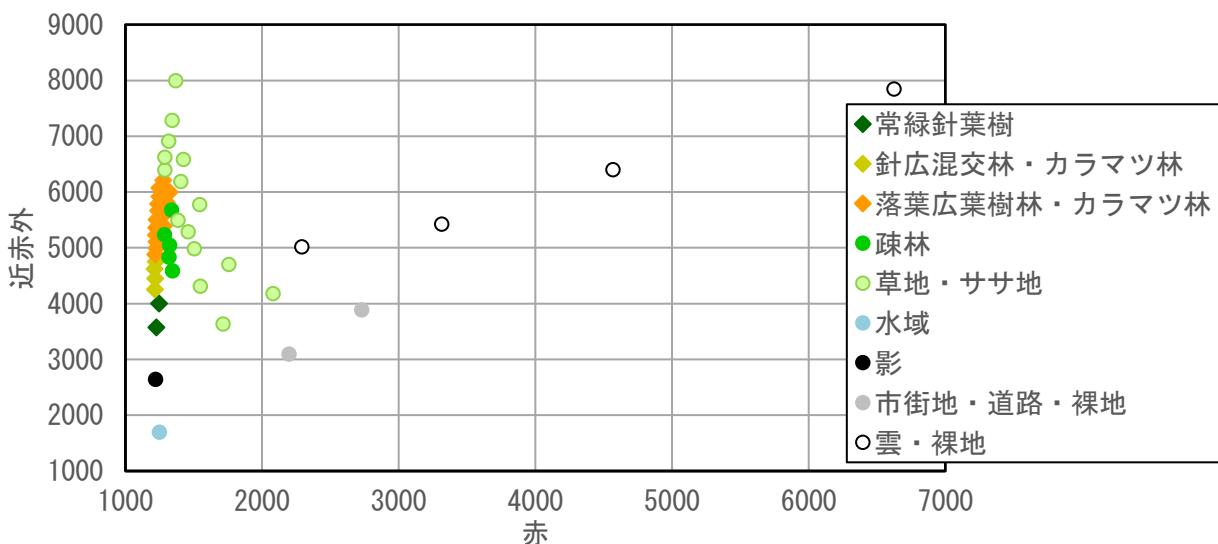


図-6 各クラスの赤と近赤外の値と分類結果

値が大きいほど反射率が高いことを示す。ここでは赤と近赤外を示したが、他の波長でも同様に分類結果との対応を示すことができる。

も森林よりも反射率が高くなっていました。一つのピクセルは10m四方の広がりがあるので、常緑針葉樹と落葉広葉樹が混じったところでは近赤外、草地と森林の境界では赤の値が、両者の中間的な値になっていました。そこで今回は、針広混交林、疎林のカテゴリーを作りました。

このように、各クラスのバンドの値も、意味づけの参考にすることができます。ただし、衛星画像に記録される反射率は大気の状態などによって常に変化します。そのため、同じ範囲が撮影されても、図-6のような関係は画像1枚ごとに異なり、画像ごとに分類の作業を行う必要があります。

教師なし分類の活用

これまでに示してきたように、オープンソースソフトウェアを用いて無償で利用できる衛星画像を教師なし分類することにより、落葉広葉樹林とカラマツ林を区別することはできませんでしたが、これらと常緑針葉樹林を区別することができました。この結果から、例えば広域に分布する常緑針葉樹人工林のなかで植栽樹種が多く生育して今後間伐などが必要と思われる範囲を特定するなど、今後の施業に役立つ情報を得ることができます。衛星画像の撮影時期によっては、落葉広葉樹林とカラマツ林の区別ができることもあるでしょう。また、10m以上の分解能では単木の現象を把握することはできませんが、例えば山火事跡や病虫害などの面的な広がりの把握に活用できると思われます。

QGISの機能を活用すれば、小班ごとに、それぞれのカテゴリーに分類されるピクセルの割合などを集計することも可能ですが(図-7)。今回は森林を常緑針葉樹林、針広混交林、落葉広葉樹林の3段階に区分しましたが、図-6に示したように、衛星画像に記録されている値は常緑針葉樹林から落葉広葉樹林へと連続的に変化するので、林況をさらに細かく区分することもできるでしょう。

現地で標準地調査を行うことにより、その範囲の正確な情報を得ることはできますが、対象林分全体の情報を得るには大きな労力が必要になります。詳しく林分の状態を知るには、衛星画像の教師付き分類や分解能の高い有償の衛星画像の活用、対象範囲が限られる場合にはUAVを活用した資源把握などの技術の活用も考えられます。森林資源の現状把握において、求められる情報の内容や対象範囲の広がりに応じて効果的な手法を選択することが望ましく、教師なし分類もそのための手法の一つとしての活用が期待されます。

なお、ここで説明する手順は、一般的なオフィス用パソコンでも十分に実行することができますが、一度に広い範囲を扱おうとすると、大きなメモリが必要になったり、処理に時間がかかったりします。

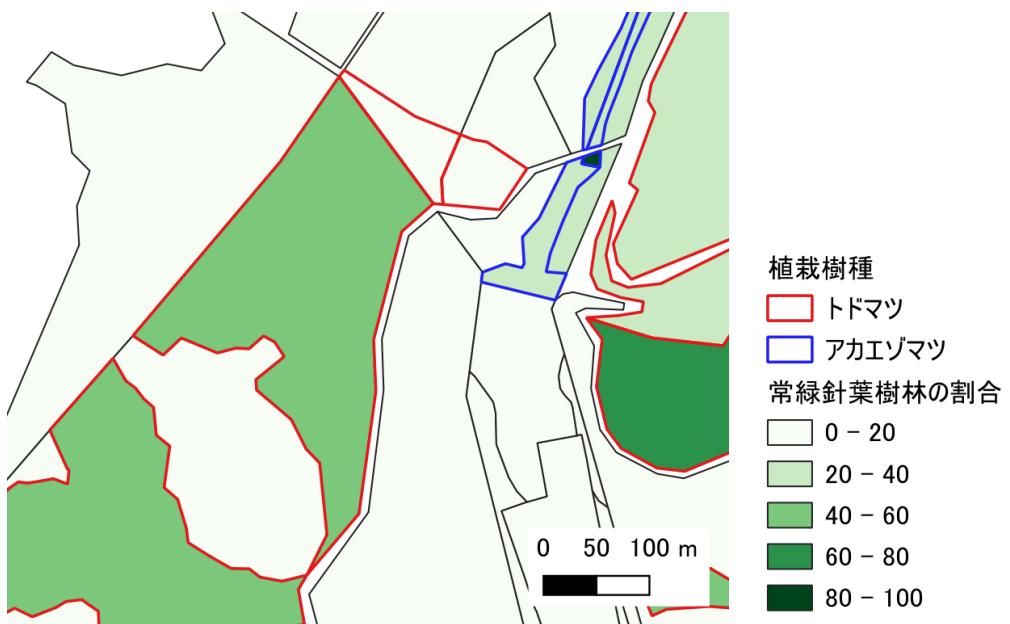


図-7 常緑針葉樹林の割合 (QGIS で作図)

図-5をもとに、小班（図中の区画線）ごとに、常緑針葉樹林に分類されたピクセルの割合を示す。

まずは一辺が数 km の範囲から初めてみてはいかがでしょうか。

（保護種苗部）