By Pierre Stratonovich, Agrimetrics

Agrimetrics Agricultural, Woodland, and Landcover algorithms

Feature mapping from ESA Sentinel 1 & 2 timeseries

Agrimetrics developed algorithms to map Agricultural, Woodland, and Landcover features from Sentinel 1 & 2 Earth Observation (EO) timeseries. These algorithms process up to a full year of EO imagery to produce high accuracy maps at 10 m resolution. They share a novel neural network architecture developed by Agrimetrics within the TAiM project.

High levels of cloud coverage are common in satellite imagery. Traditionally, cloudy images are excluded from the analysis which results in irregular time gaps between images. Rather than relying on interpolation to regularise the timeseries or using Recurrent Neural Network architectures which are difficult to train, our algorithms exploit all available observations even when they are at irregular time intervals.

With this architecture, instead of processing each image in isolation, our algorithms process EO timeseries from October to September and refine their predictions as more images are analysed. In Figure 1 we present the Wheat identification algorithm output alongside optical EO at 5 chosen timepoints between October and September. The algorithm is processing a total of 49 images from Sentinel 1 & 2, between the 1st of October 2023 and the 31st of September 2024. Each row in this figure shows the date of a chosen Sentinel-2 acquisition (left column), the corresponding Sentinel-2 image (centre column), and the current output of the algorithm (right column). The output of the algorithm is provided on a grey scale. Darker pixels indicate a low probability of Wheat crops, lighter pixels a higher probability. As more imagery is analysed, the output of the algorithm is refined. The algorithm can ignore temporary changes in appearance caused by snow (17/01/2024) and clouds (13/04/2014 and 23/06/2024).

The overall accuracy of Agricultural algorithms increases from 78 % in October, to 92 % in May and 95 % in September. The accuracy of Barley, Oilseed Rape, and Wheat algorithms per month is provided in Figure 2.

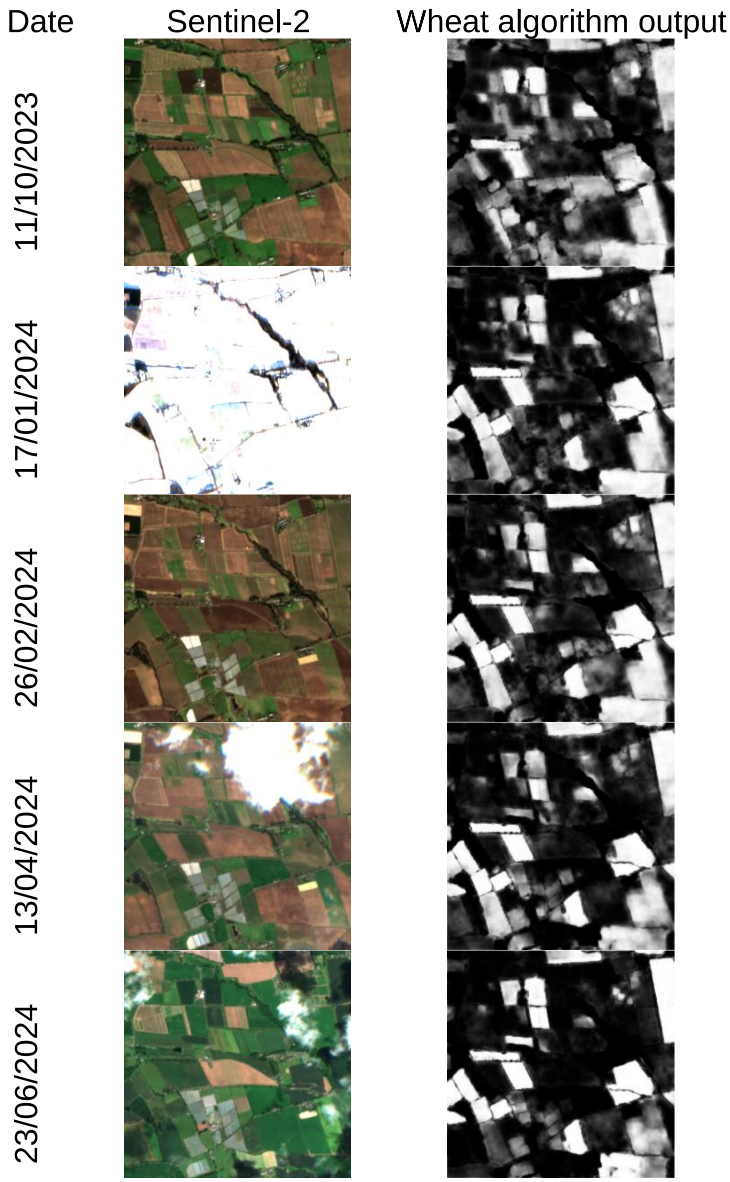

Figure 1: Wheat algorithm output at chosen timepoints during the growing season at Balruddery farm, Scotland, 2024.

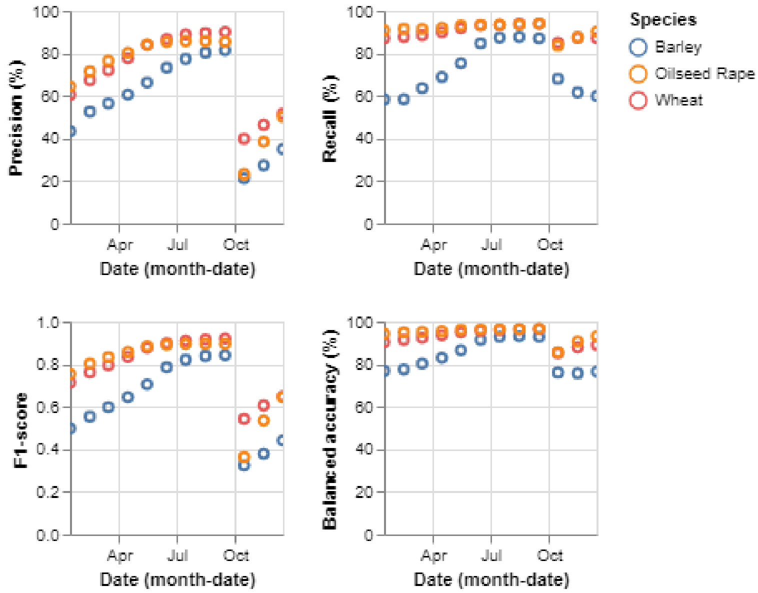

Figure 2: Agricultural algorithm accuracy on test data for Barley, Oilseed Rape, and Wheat, by month of prediction (October to September).

The accuracy metrics presented in Figure 2 were calculated on unseen test data for 4,270 Barley, 2,943 Oilseed Rape, and 8,036 Wheat crops. This unseen test data, obtained from Agronomic records before the TAiM project, is limited to the harvest years 2018 and 2019.

Data collected during the TAiM project supplement the evaluation of the algorithms. The team recorded growing crop species on Balruddery Farm, Scotland. This is very valuable data for validating the Agricultural algorithms because the observations were recorded in 2024 which allow us to assess how the algorithms are performing in different growing conditions. Furthermore, the TAiM data provides an independent assessment of our algorithms.

Accuracy of the agricultural algorithms on this dataset is presented as a confusion matrix in Figure 3. A cell in this confusion matrix summarises the number of times (as a percentage) the algorithms make the right or wrong decision. On the top 3 rows, we compare the observed and predicted crop species for the 3 trained algorithms. Almost all Barley crops were correctly identified by the algorithm (95%). All Oilseed Rape and Wheat crops where correctly identified by the algorithms. One Barley crop (4.5%) was mistakenly identified as a Wheat crop. On the next 3 rows, we compare the observation and predictions for crop species which are not covered by the algorithms (Beans, Potatoes, Vining Peas). The Oilseed Rape algorithm confused Beans with Oilseed Rape crops on

40% of observations. The last 3 rows compare the predicted crops of 3 trial fields at Balruddery Farm for which the actual crop species is unknown. The cereal trials field is identified as a Barley crop whilst other fields are not predicted to have grown species identifiable by the Agricultural algorithms.

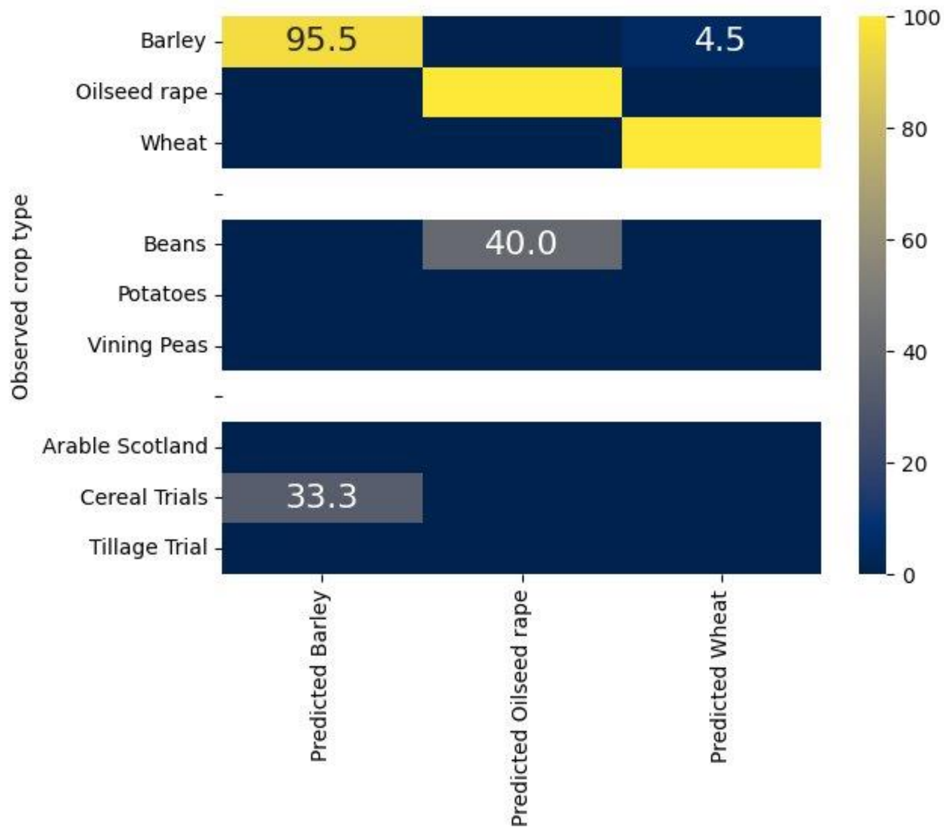

Figure 3: Confusion matrix of agricultural algorithms on Balruddery farm observations.

The Woodland algorithms were developed and evaluated with open-source data. The location of woodlands and their characteristics were extracted from the National Forest Inventory 2020. Using the same architecture as the Agricultural algorithms, Agrimetrics trained and evaluated Woodland algorithms to identify Broadleaved, Conifer, Young trees, and Felled woodlands. The accuracy of these algorithms is presented in Figure 4. They start processing imagery in October and continue until the following September, as the Agricultural algorithms. The length of processing (up to 1 year) and origin (from October) were kept identical to the Agricultural algorithms because the appearance of woodlands changes throughout the year from Autumn to Summer. As for the Agricultural algorithms, we observe a gradual increase in accuracy from October as the number of images accumulates. However, the increase is less important after Spring because the algorithms have processed enough information (Autumn, Winter, Spring) to make an accurate prediction. The overall accuracy of Woodland algorithms increases from 80% in October, to 89% in May and 90% in September.

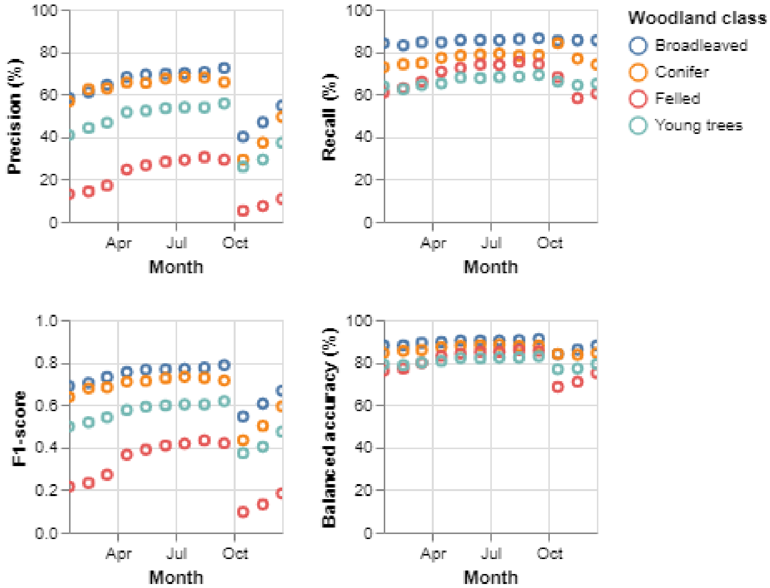

Figure 4: Woodland algorithms accuracy on test data for Broadleaved, Conifer, Felled, and Young trees classes, by month of prediction (October to September).

The Landcover algorithms were developed with OpenStreetMap records. We extracted the location of Bog, Fen, Grassland, Heath, Meadow, Reedbed, Saltmarsh, and Scrub features around the UK. These observations are often inaccurate as they are collected by non-specialist volunteers. However, they are useful in exploring how the developed architecture can distinguish between specific natural habitats. As shown in Figure 5, the resulting algorithms are relatively less accurate than the Agricultural and Woodland algorithms. They follow the same approach as before by processing imagery from October to September. Their overall accuracy increases from 72% in October, to 78% in May and September. The lack of accuracy of OpenStreetMap records affect both the training and validation of these algorithms. The detailed data collected during the TAIM project will provide a fair assessment of these algorithms.

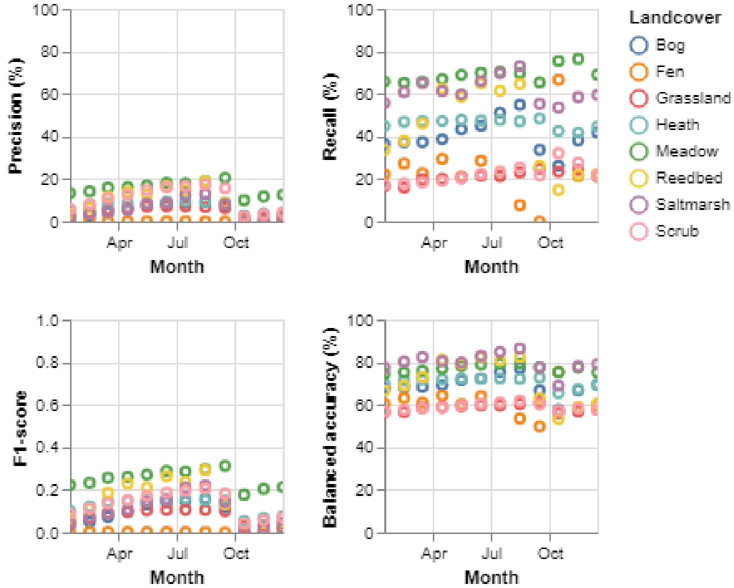

Figure 5: Landcover algorithms accuracy on test data by month of prediction (October to September).

As an example, the Landcover Saltmarsh algorithm produced a plausible classification on the Annan TAIM test site, as shown in

Figure 6.

Figure 6: Optical satellite view of the Annan TAIM test site (Left) and Landcover Saltmarsh algorithm output (Right).