By Iain Cameron, EOLAS Insight

Translating field surveys to satellite pixels

Understanding how spatial resolution affects map validation against field data.

There are several satellite-derived habitat maps, and typically these have spatial resolutions that are based upon the pixel size of the satellite sensor used. While there are some satellites with sub-metre pixel sizes, perhaps the most widely used for habitat mapping are Sentinel-2 and Landsat 8/9 with pixel sizes of 10m – 30m. While high resolution in terms of global mapping, 10 – 30m is coarse in the context of habitats and presents challenges for using field survey data to validate satellite maps.

Figure 1 shows a 25cm aerial photograph that has been degraded to 10m (Sentinel-2) and 30m (Landsat 8/9). In the aerial photo we can identify many of the features that an ecologist would see on the ground: individual trees, patches of bracken, heather, even cattle. Boundaries between these features are clearly identifiable, and the resolution is sufficient to show changes in both vegetation structure and colour. With 10m and 30m pixels boundaries between features are indistinct, and much of the structural information is lost. However, differences in colour can still be seen, allowing the bracken and tree stands to be identified at 10m and 30m scales.

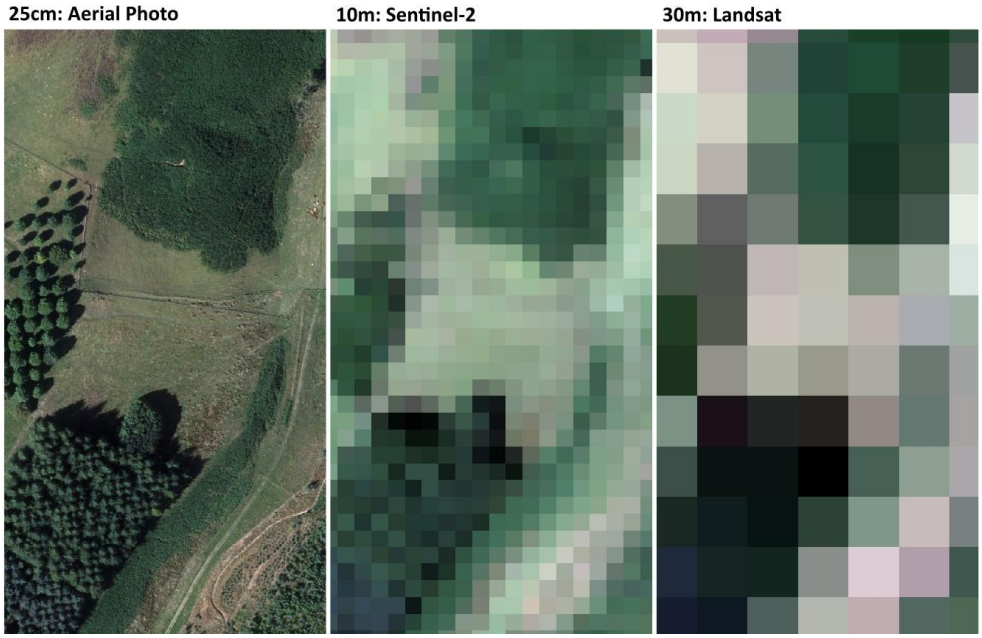

Figure 1: True colour aerial photo and satellite images over the same area containing woodland, grassland and bracken habitats

The effect of resolution on map validation

This reduction in spatial information reduces the accuracy of identifying both the true extent and nature of habitats from satellite imagery before we even consider additional errors arising from the selected classification algorithm. It also presents questions for how best to use high resolution field survey data to validate lower resolution satellite derived mapping.

Figure 2 shows UK-HAB Level 4 field survey polygons collected during this project on the left, and the same data naively resampled to 20m (analogous to the UK-CEH or Space Intelligence Landcover Map resolutions) on the right. Using the field survey polygons to directly validate the 20m data we achieve an overall accuracy of 89%, and this is purely due to resolution differences as there is no classification algorithm error at play.

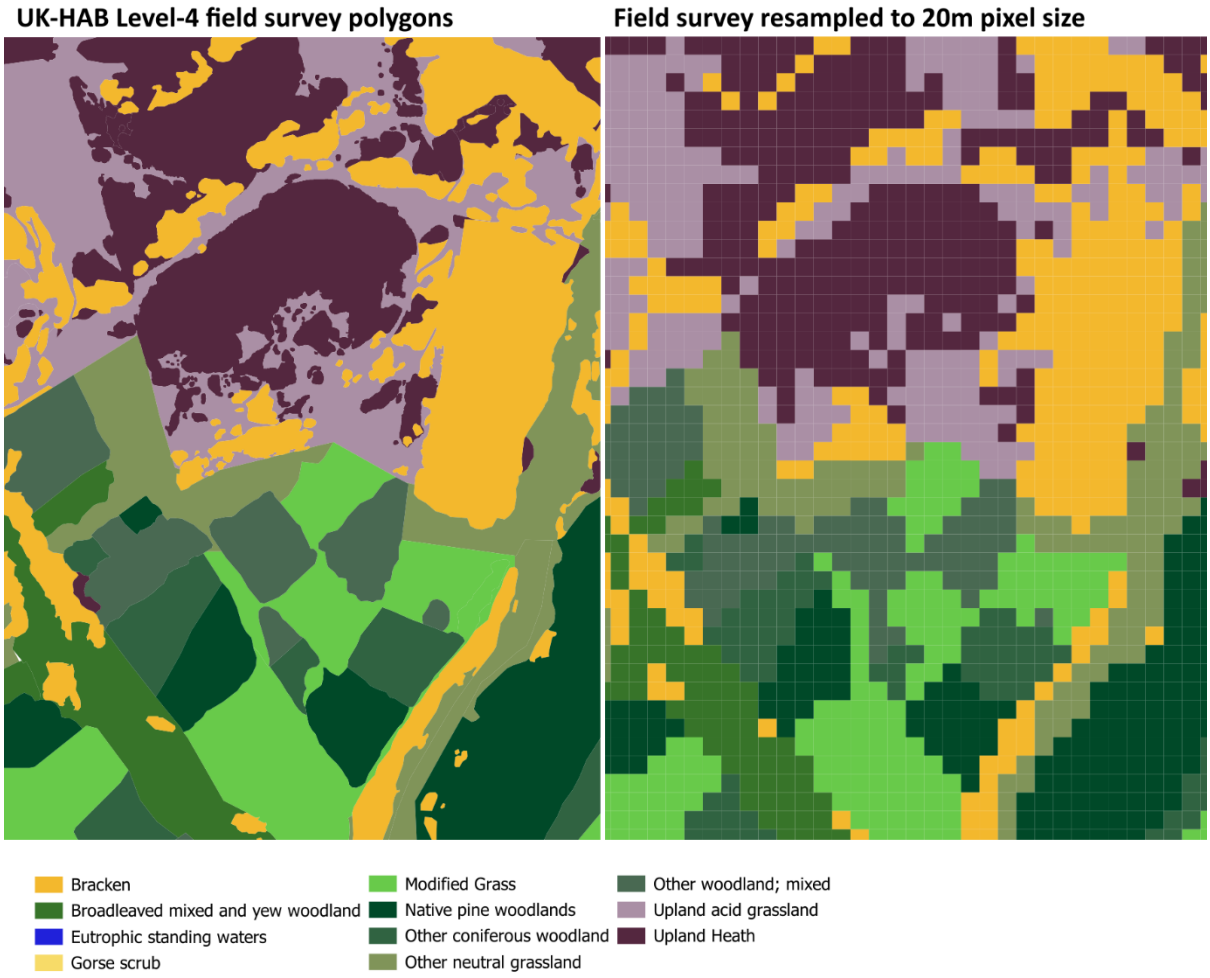

Figure 2: UKHAB Level-4 field survey polygons (left) translated to 20m pixels (right)

Translating field survey data to satellite resolution

In the above example we have translated field survey polygons to 20m by simply calculating the dominant habitat class by area for each pixel. However, additional factors such as how many classes fall within each pixel and the percentage coverage by the dominant class, can also be calculated.

Figure 3 shows two scenarios for using these parameters to filter the pixels used for validation. If we consider only pixels that contain a single class we achieve a 100% accuracy, but this comes at a significant cost of rejecting almost 50% of the pixels. Choosing pixels where the dominant class covers 80% of the pixel area, we achieve an accuracy of 97% and reject 27% of the pixels.

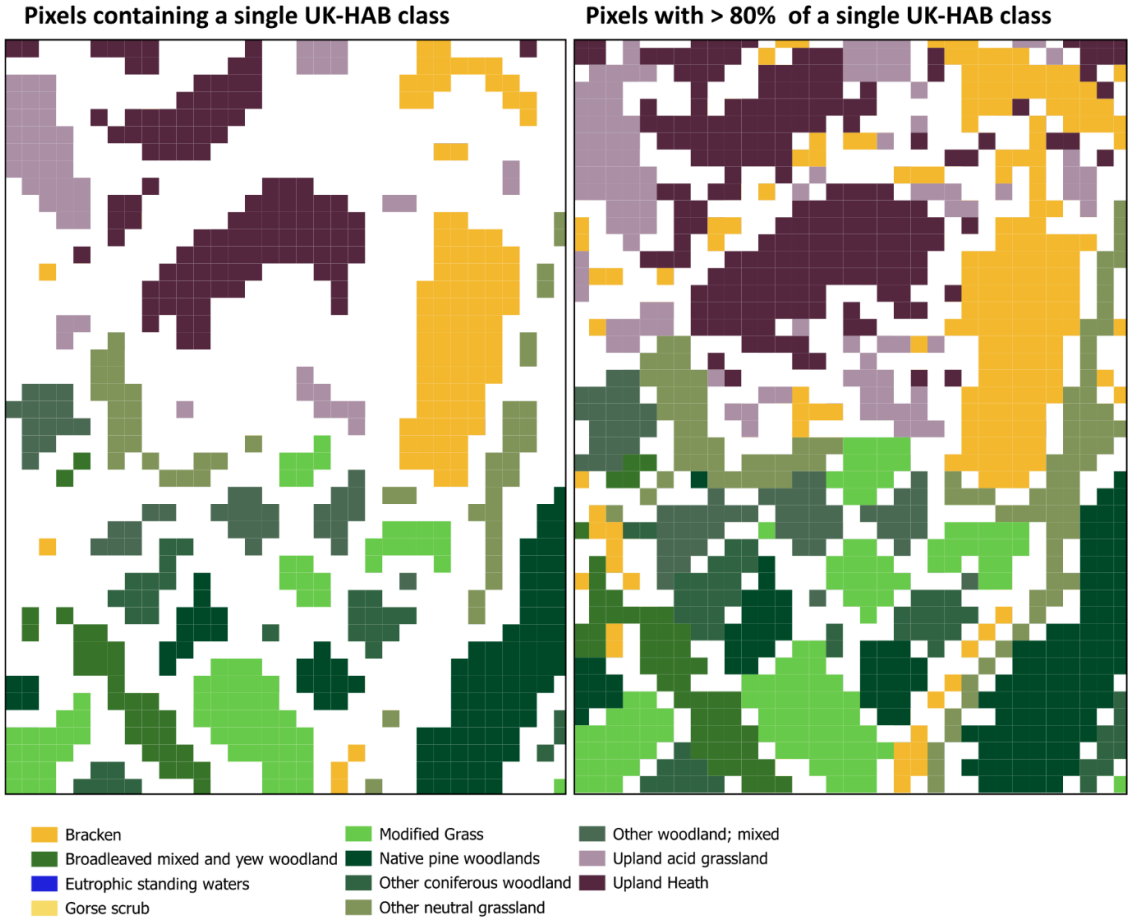

Figure 3: Field survey data resampled to 20m and filtered to show pixels containing only a single habitat class (left) and pixels where the dominant class covers greater than 80% of the pixel (right).

In both cases the rejected pixels are in areas containing complex habitat mosaics that are likely to be ecologically meaningful, but the 80% threshold seems a reasonable compromise between coverage and resolution-induced error.

Summary

These examples illustrate that we need to carefully consider how we use high accuracy field data to validate satellite derived maps, as direct comparison can paint an unduly negative picture regarding the accuracy of satellite mapping. We can also calculate metrics that provide insight into how field data survey and satellite resolution compare and use this directly in our map evaluation.

We also need to consider the capability of the satellite sensor itself and ensure that we are seeking features that are resolvable. For example, a hedgerow classification is impossible to deliver using Landsat. For this reason, satellite derived habitat maps typically operate at a broad habitat level such as EUNIS / UK-HAB level 2 and above.