Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP)

ERAMMP Report-56: Suitability of Satellite Data and LiDAR for Mapping Hedges

Rowland, C.S., Scholefield, P., O'Neil, A.W. & Marston, C.

UK Centre for Ecology & Hydrology

Client Ref: Welsh Government / Contract C210/2016/2017
Version 1.0
Date: 12-November-2021

Funded by:

Llywodraeth Cymru
Welsh Government

Canolfan Ecoleg a Hydroleg y DU
UK Centre for Ecology & Hydrology

Version History

Version	Updated By	Date	Changes
1.0	Author Team	02/02/2022	Publication

Mae’r adroddiad hwn ar gael yn electronig yma / This report is available electronically at: www.erammp.wales/56

Neu trwy sganio’r cod QR a ddangosir / Or by scanning the QR code shown.

Mae’r ddogfen yma hefyd ar gael yn Gymraeg / This document is also available in Welsh

Series	Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP)
Title	ERAMMP Report-56: Suitability of Satellite Data and LiDAR for Mapping Hedges
Client	Welsh Government
Client reference	C210/2016/2017
Confidentiality, copyright, and reproduction	© Crown Copyright 2022 This report is licensed under the Open Government Licence 3.0.
UKCEH contact details	Bronwen Williams UK Centre for Ecology & Hydrology (UKCEH) Environment Centre Wales, Deiniol Road, Bangor, Gwynedd, LL57 2UW 01248 374500 erammp@ceh.ac.uk
Corresponding author	Clare Rowland, UKCEH clro@ceh.ac.uk
Authors	Rowland, C.S., Scholefield, P., O'Neil, A.W. & Marston, C. UK Centre for Ecology & Hydrology
Contributing authors & reviewers	Bridget Emmett & Bronwen Williams UK Centre for Ecology & Hydrology
How to cite (long)	Rowland, C.S., Scholefield, P., O'Neil, A.W. & Marston, C. (2022). *Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP)*. ERAMMP Report-56: Suitability of Satellite Data and LiDAR for Mapping Hedges. Report to Welsh Government (Contract C210/2016/2017) (UK Centre for Ecology & Hydrology Projects 06297 & 06810)
How to cite (short)	Rowland, C.S. et al. (2022). ERAMMP Report-56: Suitability of Satellite Data and LiDAR for Mapping Hedges. Report to Welsh Government (Contract C210/2016/2017) (UKCEH 06297/06810)
Approved by	

Abbreviations Used in this Report

CHM	Canopy Height Model
CS	Countryside Survey
DCM	Digital Canopy Model
DEM	Digital Elevation Model
DTM	Digital Terrain Model
DSM	Digital Surface Model
EO	Earth Observation
ERAMMP	Environment and Rural Affairs Monitoring & Modelling Programme
GAEC	Good Agricultural and Environmental Conditions
GCPs	Ground Control Points
GIS	Geographic Information System
GMEP	Glastir Monitoring and Evaluation Programme
GSD	Ground Sampling Distance
LiDAR	Light Detection and Ranging
NDVI	Normalized Difference Vegetation Index
NRW	Natural Resources Wales
OS	Ordnance Survey
UKCEH	UK Centre for Ecology & Hydrology
WG	Welsh Government
WLF	Woody Linear Features

CONTENTS

1 SUMMARY

The Welsh Government (WG) uses data on hedges and field boundaries for a variety of purposes including scheme delivery, environmental monitoring and regulatory compliance.

Hedge data are currently acquired mainly through a combination of aerial photography and field visits. Considerable cost savings may be possible, if optimising the use of satellite data enables the number of field visits to be reduced.

This project explored the potential for high resolution satellite data to provide accurate spatial data on hedge location and length.

A number of methods of hedge mapping were tested, including manually digitising hedges and more automated methods using Skysat Imagery Products data from Planet Labs Inc.[1] and LiDAR.

The key findings of the project were:

A. Aerial photography was a better source of data for manually digitising hedges than the Planet Skysat data available for this project. This is because the aerial photography has higher spatial resolution (25cm compared to ~50cm) and the Planet Skysat data used in this project was collected in winter and was badly affected by shadows. Note, for the purposes of this project, only Planet Skysat data collected in the winter were available, however, it is highly likely that multi-temporal Planet Skysat data would improve results further.

B. For two of the three sites, the spatial accuracy of the Planet Skysat data was too low to map hedges, without spending additional effort manually geo-correcting images.

C. Automated methods using LiDAR show promise, but there are issues with producing a final ‘clean’ vector data set that require additional work. LiDAR-based methods could be deployed in a number of ways depending on requirements.

D. Automated methods using LiDAR and aerial photography could be deployed in a number of ways depending on whether the aim is to measure some attributes of hedge condition, or hedge location and length.

E. Methods using automated detection of new hedges using LiDAR data, followed up by manual checking against aerial photography, may be able to capture the best aspects of different approaches.

[1] https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf p.36+

2 INTRODUCTION & PURPOSE OF THE WORK

Opportunity: Welsh Government (WG) currently acquire data on hedges and field boundaries for purposes including scheme delivery, environmental monitoring and ensuring regulatory compliance e.g. GAEC7[2] and EFA hedges[3].

Current methods involve field visits, so considerable cost savings may be possible by optimising the use of satellite data, enabling the number of field visits to be minimised.

The aim of this project is to assess the potential for mapping field boundaries, and specifically hedges, from Planet Skysat satellite data, through two main objectives:

1. Can manual/automatic interpretation of the satellite data meet WG's accuracy criteria for hedge mapping?
2. Does LiDAR data (combined with Planet data) help with discrimination?

Planet Skysat images (< 1m pixel size) were acquired for three test areas through Welsh Government's license with Planet Lab Inc.

The test areas were selected on the basis of available data, specifically:

i. Field-survey data were available with hedge location and length.
ii. The sites fell within Welsh Government's Planet Skysat tasking for winter 2020/21, so high resolution Skysat data had been acquired for them.
iii. The test areas were snow-free, and relatively cloud-free.

Three test areas, distributed across Wales, met these criteria and were acquired.

An initial review of the Planet Skysat data identified some issues, specifically:

1. Spatial accuracy – the spatial accuracy of the Skysat data for two of the three sites was poor, so they were manually geocorrected to correct this issue. The poor spatial accuracy of the data limits its potential for automated processing chains, especially as the spatial accuracy of each image is not listed in the image metadata. This prevents images with poor spatial accuracy being filtered out when searching the Skysat archives. (this is discussed further in Sections 3.1 and 3.2.1).
2. Cloud-masking – the cloud mask was found to miss some areas of cloud (discussed further in Section 3.1 and Section 3.2.2).

[2] https://gov.wales/sites/default/files/publications/2018-01/cross-compliance-2017-landscape-features-gaec-7.pdf

[3] https://gov.wales/single-application-form-saf-2020-rules-booklet

3 DATA SETS AND TEST AREAS

3.1 Planet Skysat Scene Data

The first step was to identify suitable high-resolution Planet Skysat imagery (<1 m pixel size) for hedge mapping.

The Planet Skysat data already held by WG was reviewed to determine whether it was suitable for boundary mapping, however, none of it covered agricultural areas. WG tasked Planet with several further areas, with Planet acquiring Skysat data for those areas during December 2020 and early January 2021. Glastir Monitoring and Evaluation Programme (GMEP) squares (Emmett et al. 2017) were then identified that contained hedges and fell within the tasked areas. This resulted in three good squares with almost complete coverage by fields, and several additional squares with partial coverage by fields, but with some of their extent covered by woodland, or unenclosed upland areas. Searching the Planet Skysat data base, using the QGIS Planet plugin v2[4], resulted in good images for two of the sites, with partial cloud coverage for a third. These three squares are the focus of this report.

Table 3.1: *Dates of the Planet Skysat imagery.*

	Imagery dates
Site 1	24/12/2020
Site 2	25/11/2020
Site 3	29/01/2021

Skysat scene images for three agricultural GMEP squares were acquired (Figure 1). The Skysat images were Skysat scene analytic ortho-rectified images. Before distribution to the user, Planet Labs apply a series of automated processing steps that improve the geometric and radiometric properties of the data set. The Skysat analytic product is distributed as a 4-band data set, with Red, Green, Blue, and Near-Infrared bands, with coefficients to calculate either top-of-atmosphere radiance or Top of Atmosphere Reflectance.

In the absence of a thermal band, Planet's Skysat cloud-masking is based deviation from expected radiance for the time of year. The expected radiance is based on an extensive database derived from Landsat. Planet note that: 'This method is fast and simple, but has limitations: 1. If a region may be covered by snow at a given time of year, clouds are much less likely to be identified. 2. Darker clouds are less likely to be identified. This includes both thin clouds and self-shadowed clouds. 3. Brighter areas, such as desert surfaces, sands, and salt flats, are less likely to be identified as containing clouds. 4. Specular reflection at noon local time are more likely to be marked as clouds.'[5]

The ortho-correction means that the images are corrected with a Digital Elevation Model (DEM) to correct for spatial shifts in the data due to topography. However, this is currently conducted with global-scale DEM products that are coarse compared to the Skysat data (30-90 m post spacing), so this might cause issues in areas with highly variable topography. It is unlikely to affect hedge/boundary mapping but may be an issue for other WG applications.

The Skysat scene images are designed to be used for automatic workflows that require accurately geolocated data. To achieve this, they use automated geocorrection methods that

[4] https://doi.org/10.5285/82c63533-529e-47b9-8e78-51b27028cc7f

[5] https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf p.57

automatically distribute ground control points (GCPs) across the newly acquired image and across a reference image.

The number and the quality of the GCPs affect the quality of the resulting geocorrection. Planet Labs state that 'The accuracy of the product will vary from region to region based on available GCPs', but it is not clear from the Skysat image metadata how many GCPs there are for each of the images or what their accuracy was.

The utility of the Planet Skysat data could be improved by use of higher resolution spatial data sets, for both the ground reference data set and the DEM used for the ortho-rectification, for example, using Welsh Government's existing DEM data. In addition, it would be highly beneficial for Planet to report the RMSE of the GCPs in the metadata, so that users can filter out images with high spatial uncertainty.

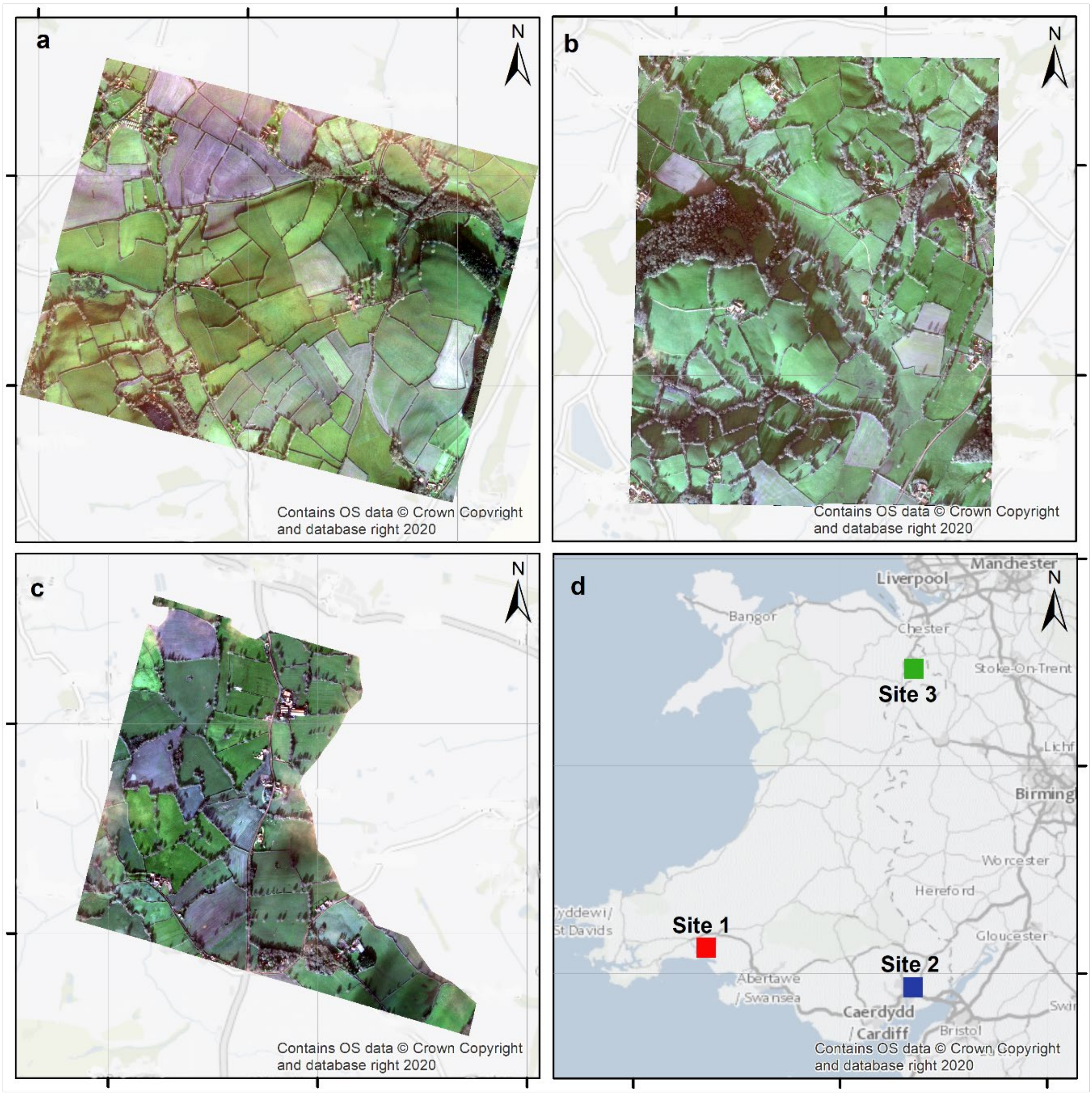


***Figure 1**: Planet imagery for a) site 1; b) site 2; c) site 3; and d) study site locations.*

Table 3.2: Characteristics of Planet Skysat (Planet Labs Inc., 2018)

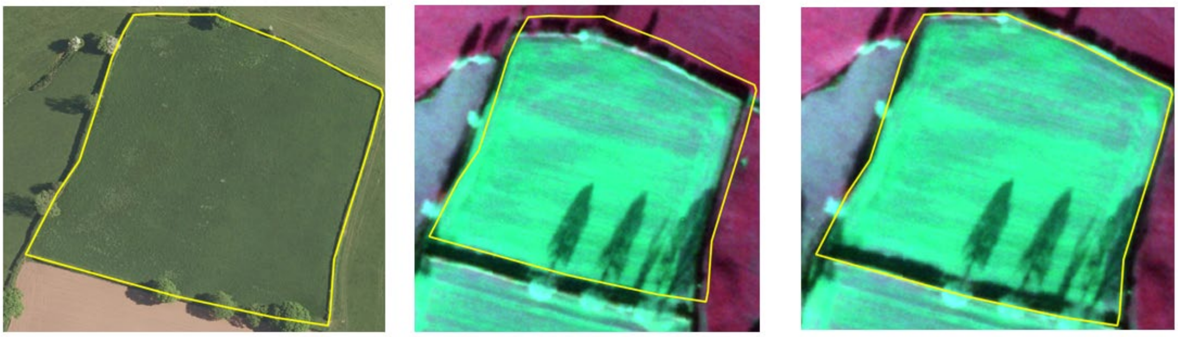
Criteria	Planet Skysat
Bands (wave length)	5 (450–900 nm)
Night-time imagery	No
TIR/SWIR	No
Ground Sample distance	0.65-0.86 m Panchromatic 0.81-1.00 m Multispectral
Pixel Size (orthorectified)	0.5 m
Cadence	Nadir: 28 days per spacecraft; sub-weekly per constellation. Off-Nadir: sub-weekly per spacecraft; intra-daily per constellation.
Spatial accuracy	RMSE < 10 m

Prior to using the Skysat data for boundary mapping, a series of pre-processing steps were applied:

1. For each GMEP square two Skysat collect images were required to provide complete coverage, so the first step was to merge the images and providing complete coverage of the GMEP square.
2. Site 3 was partially covered by cloud, so the cloud-mask supplied by Planet Labs was applied.
3. Images were reprojected to British National Grid coordinates.
4. Spatial accuracy was assessed against aerial photography basemaps and OS open data layers.
5. Manual geo-referencing was applied. Results identified that the spatial accuracy of sites 2 and 3 was poor when compared to other spatial data sets. This meant that boundaries were misplaced. To rectify this manual geo-referencing was applied. Geo-referencing requires the same point to be identified on a reference image and on the image to be corrected. When sufficient common points are identified a geometric correction can be applied to geo-reference the image to be corrected, so that it is in the correct location. The process is time consuming and is a severe constraint on automated processing methods. Sites 2 and 3 were georeferenced in QGIS, using aerial basemaps and OS open layers as reference data sets. For site 2, 21 manual ground control points were and for site 3, 27 manual ground control points were used.

Figure 2 shows the magnitude of the error in the Skysat imagery for site 2.

In Figure 2b) the yellow line around the edge of the field is offset by 12 m, before the geo-referencing, after geo-referencing it is aligned with the top of the field (Figure 2c).



a) Reference field

b) Skysat image for site 2 before geocorrection

c) Skysat image for site 2 after geocorrection

***Figure 2**: Magnitude of spatial error in Skysat data for site 2, with a) showing field outline in Google Maps basemap, b) showing Skysat data before correction and c) Skysat image after geo-correction.*

3.2 Review of Planet Skysat data

3.2.1 Spatial Accuracy

The locational accuracy of the Skysat data was identified as an issue in Section 3.1, especially for sites two and three. A simple assessment was conducted to see how the issue varied across a series of images. Using a Google Maps aerial photography basemap, a field was digitised (red line in Figure 3a) and overlain on three Skysat images (Figure 3b-d), which shows that the field observed moves around in a variety of directions and magnitudes of spatial error.

The yellow line in Figure 3 shows the spatial offset between the edge of the field and its true location, the offsets vary from 3.5 m and 4 m for Figure 3b & c to 12 m for Figure 3c. These spatial errors are within the average spatial accuracies cited by Planet Labs (Table 3.2).

However, critically image specific spatial accuracy data is not part of the Skysat metadata properties, so it does not currently appear possible to automatically filter out images with poor spatial accuracy.

a) Basemap with area of interest

b) 14th December Yellow line 4m

c) 24th December Yellow line 12m

d) 25th December Yellow line 3.5m

***Figure 3**: Variability of spatial accuracy over time.*

3.2.2 Cloud-mask

When cloud is detected in the Skysat images, an automatically created cloud-mask is produced. Site 3 had significant cloud.

Figure 4 shows a portion of the image before and after cloud-masking and highlights that whilst much of the cloud is accurately identified, some cloud does remain. Figure 4 shows the best image, within the tasking period for this site, and highlights the issue of cloud for optical satellite data for Wales (Robinson et al. 2021).

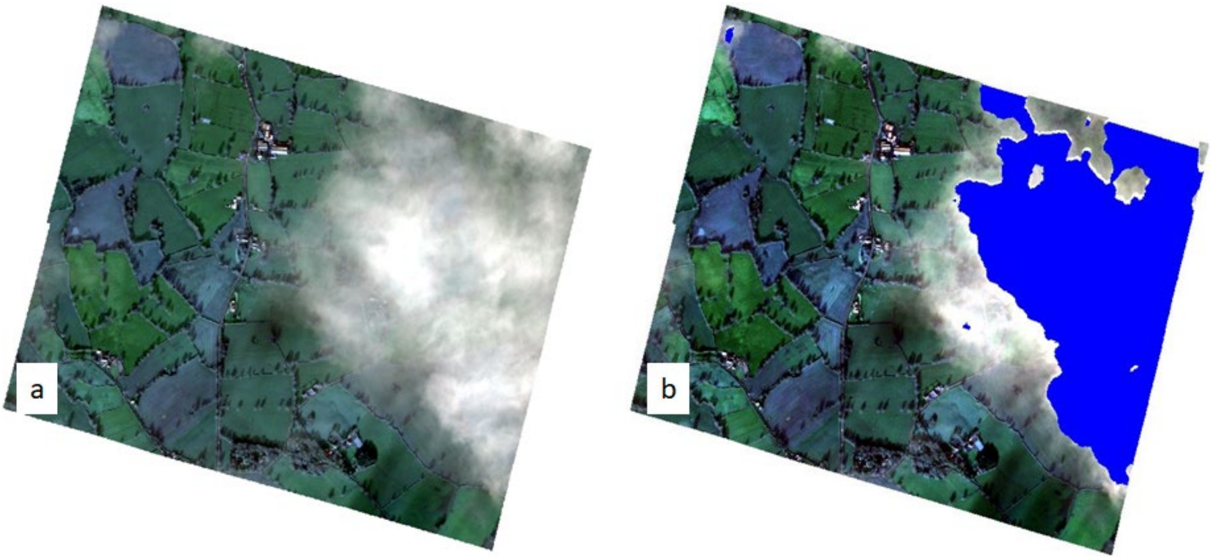


Figure 4: *Cloud presence across site 3: a) Cloud across site 3; b) Skysat cloud mask coverage across site 3.*

3.2.3 Seasonal Variability in the Distinctiveness of Hedges

The Skysat images available to this project were winter images, so most of the hedges will be leaf-off, and as Figures 1b and 2b demonstrate the winter images do show significant shadow.

Planetscope data through the year was reviewed to see when hedges were most distinct (Figure 5). Planetscope at 3m has a lower spatial resolution than the sub-meter Skysat data, but is generally collected more frequently, so provides a better source from which to determine seasonal variability. Visually, the distinctiveness of the hedges does vary through the year, which suggests that if the aim is to identify the hedges purely from the Skysat data then the imagery would need to target appropriate times of the year.

Determining the optimum time of the year would require additional analysis.

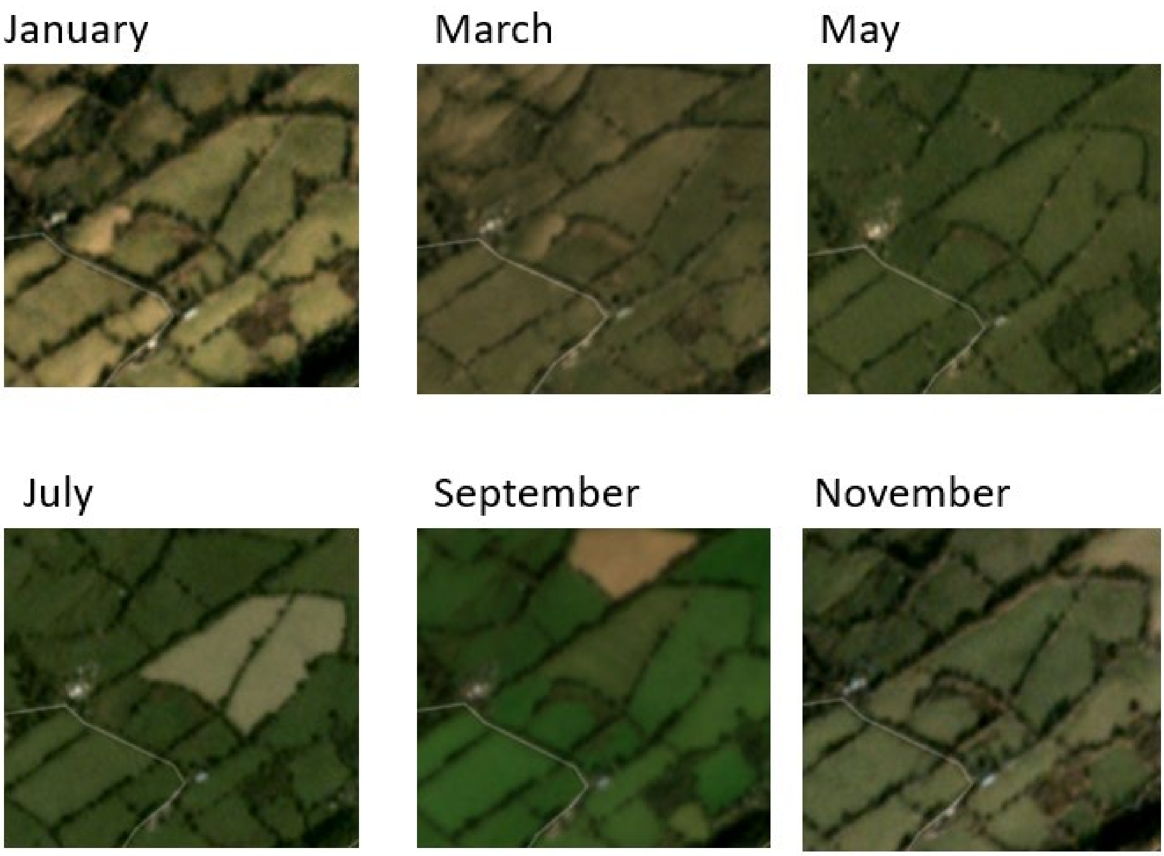


Figure 5: *Variability of hedges through the year.*

3.2.4 Variability of Pixel Size

The spatial resolution of EO data is important for determining the minimum size of object that a satellite will be able detect. The images for one area were reviewed through December 2020 and the reported ground sampling distance (gsd) was found to vary slightly between the different Skysat satellites, with values ranging from 0.65 m for SSC16 to 0.75 m for SSC4. However, if Skysat data were to be used in an operational situation it would be important to get a better understanding of the minimum resolvable object size.

3.3 LiDAR Data

Light distance and ranging (LiDAR) were acquired from the LIe data portal. The portal provides a range of data sets for Wales, including composite LiDAR data sets. The highest resolution LiDAR Digital Surface Model (DSM) and Digital Terrain Model (DTM) data available for each study area was downloaded.

The DSM shows the elevation of the first surface that the LiDAR signal interacts with, whilst the DTM provides the height of the underlying terrain.

For a hedge, the DSM will typically record the top, or near the top, of the hedge, whilst the DTM will record ground elevation. Consequently, the difference between the two will provide an estimate of the height of the hedge (equation 1).

This difference between the DSM and the DTM is known as the Digital Canopy Model (DCM):

$$DCM = DSM - DTM$$

It is important to note that the accuracy of the DCM is dependent upon several factors, including the width and density of the hedge, and the spatial resolution of the LiDAR system. In general, the denser the number of points recorded by LiDAR system the more likely the LiDAR is to record a point near the top of the hedge and produce an accurate estimate of hedge height.

However, for boundary mapping the important factor is that the hedge should register as a raised feature in the landscape and the accuracy of the hedge height is not crucial.

For sites 1 and 3, 1 m LiDAR data was available, although for site 3 the percentage coverage by LiDAR data was limited (Table 3.3), and the area of combined LiDAR and Skysat data was further reduced by the cloud. Site 1 was therefore the site with most potential.

Table 3.3: Summary of LiDAR data sets.

	Imagery date	Resolution	LiDAR coverage
Site 1	April 2009	1m	Complete coverage
Site 2	January 2006	2m	Complete coverage
Site 3	February 2013	1m	~60% of square has LiDAR coverage. Coverage reduces further once Skysat cloud is taken into account.

3.4 Aerial Photography

25 cm Pan Government Agreement (PGA2a) Nextperspectives aerial photography from 2009 was used as a comparison against the satellite data, as it was acquired at a similar time to the LiDAR data.

3.5 GMEP Hedge Data

The GMEP field survey (2013-2016) collected a range of environmental data from across Wales, including Woody Linear Features (WLF) such as hedgerows. Information on hedgerow location and condition was collected for 300 1 km^2 squares across Wales by a team of field surveyors (Maskell et al. 2020).

For linear features a range of attributes are collected, including the location of the hedge and attributes, such as height class, width, DBH category, signs of historic management and evidence of management.

3.6 Data set timing

It is important to note, that there is a large difference in the timing of the data sets, both in terms of year of acquisition and the season of acquisition. The Skysat data and the LiDAR data were both leaf-off, whilst the aerial photography and GMEP field survey were leaf-on. The LiDAR data and the aerial photography for Site 1 were collected in 2009 (the site 2 LiDAR was earlier, but was not used), the GMEP data covered various dates between 2013 and 2016, and the Skysat data were winter 2020/21, Despite this wide variability in the timing of the data sets, the hedges appeared relatively stable over time at the three sites. Ideally, the data would all be collected around the same time, but in this case, for these sites, the impact seems to have been minimal.

4 METHODS

4.1 Manual Digitising

The aim of the manual digitising was to quantify the difference between the accuracy and certainty with which hedges can be digitised from satellite data in comparison to aerial photography. The manual digitising was conducted by opening the respective images in a GIS package and drawing polylines along features that appeared to be hedges. The same process was conducted for both the Skysat images and aerial photography, with the results then compared to the GMEP hedge data.

4.2 Automated Method

Three automated methods were tested.

Method 1: LiDAR + OS field boundary data.

Method 2: LiDAR + satellite data, then classify.

Method 3: LiDAR + satellite data, then mask with OS data and convert to vector.

Method 1 is readily automated, but only detects boundaries that are mapped in the OS data, whereas methods 2 and 3 are potentially able to map features not included in OS MasterMap. However, it is technically complicated to devise a processing chain that correctly translates the raster data into sensible hedge boundaries.

For method 3, an approach has been tested to enable production of a vector-hedge map, but further development would be required if the intention were to implement this operationally and this is discussed further in Section 6.2.

4.2.1 Automated Method 1: LiDAR + OS field boundary data

1. LiDAR (DSM-DTM) is processed in R – to generate delineated tree crowns and tree top locations, above 1 m.
2. The tree top locations were filtered by 2.5 m buffer distance to OS MasterMap boundary features (Feature Code: 10046).
3. The filtered tree tops were then used to identify the delineated tree crowns, and then these tree crowns were used to clip the OS MasterMap boundary features to provide boundary features that have a height above 1 m and are associated with a tree canopy extent.

This method produces a series of vector data sets: Polyline of hedgerows aligned with Mastermap, Delineated tree crowns (Figure 6), Identified tree tops as a point file.

Figure 6: *Comparison of delineated crowns from LIDAR with Aerial Photograph for a survey square using Method 1.*

4.2.2 Automated Method 2: LiDAR + Satellite data, then classify

1. Create a 7-band composite raster from Skysat imagery, NDVI, DCM, and rasterised tree canopy layer (with height values)
2. Create training areas based on aerial interpretation for: Scrub, Urban (developed), Barren (bare soil) Forest, Hedgerow, Planted/cultivated
3. Classify using Random Forest

This method produces a raster-based classification (Figure 7). Further processing would be required to convert it to a vector product.

Figure 7: *Comparison of classified woody cover from satellite with aerial photograph for a survey square using Method 2.*

4.2.3 Automated Method 3: LiDAR + Satellite data, then mask with OS data and convert to vector

1. Threshold LiDAR DCM data, at site 1, a 75 cm threshold was used, so only pixels with a DCM value > 75 cm was considered as a potential hedge. The 2 m LiDAR data for site 2 was found to be too coarse to apply this method further.
2. Extract buildings and woodlands from OS open layer data sets, convert them to raster and create a mask.
3. Add Planet NDVI values to hedge pixels and mask out woodland/ buildings.
4. Threshold NDVI to remove any other non-vegetated areas.
5. Apply function to 'thin' the raster data, reducing a hedge to a single width of pixels. This facilitates conversion of the raster data to a vector line data set.
6. Convert raster data to polyline.

This method produces a line-based vector data set. However, the vectors produced do display artefacts that would require either further developments to the processing method, or manual intervention to resolve.

4.3 Validation

Validation of the results was performed at each of the three study sites using the GMEP field survey data as the reference. For each site, a 'common area' polygon was created, which initially comprised the 1 km^2 GMEP survey area, then applied a sub-setting process excluding firstly areas where field surveying was not possible due to access restrictions.

These areas were identified using the land cover classifications performed during the GMEP field survey, where areas with an unassigned land cover class denoted that surveying was not possible. These areas were buffered so as not to exclude any linear features on the edge of unsurveyed areas (that could have been surveyed and recorded from adjacent areas of the GMEP square).

A second sub-setting stage was then performed to exclude any areas where there was incomplete coverage of the aerial photography or Skysat satellite datasets. The remaining area corresponded to the locations for which data was available from GMEP field survey, and digitisation of both aerial survey and Skysat satellite data, enabling a consistent comparison.

The polyline datasets corresponding to each of the three datasets were then clipped to retain only features falling within the extent of this common area polygon. From these clipped datasets, the total number of features, and the combined length of all the features were extracted for each dataset, for each of the three study sites.

The figures derived from the aerial photography and Skysat satellite digitisation methods could then be directly compared to the reference results obtained in the GMEP dataset.

5 RESULTS

The number of linear features identified along with their total combined length for each site was calculated (Table 5.1). For site 1, this was performed for the GMEP, manual methods and automated methods 1 and 3. Only the manual methods applied to sites 2 and 3, because of the low resolution of the LiDAR data (site 2) and the limited LiDAR coverage (site 3).

Automated Method 2 is not covered by the formal validation, as it did not produce a suitable output.

Table 5.1. *The number and total length of linear features recorded via field survey (GMEP), digitisation of Skysat satellite imagery, and digitisation of aerial photography. These figures were calculated only for areas of GMEP 1 km survey squares surveyed in the field (excluding areas where surveys could not be performed due to access restrictions), and where satellite and aerial survey data was also available.*

Site	Method	Number of features	Total length (m)
1	GMEP	79	12722.6
	Skysat	45	8446.8
	Aerial photography	66	11437.5
	Automated Method 1	215	9114.5
	Automated Method 3	2314	13798.9
2	GMEP	61	9457.9
	Skysat	56	9094.8
	Aerial photography	86	9915.4
3	GMEP	73	8466.6
	Skysat	38	6343.4
	Aerial photography	57	6484.5
1 – 3 combined	GMEP	213	30647.1
	Skysat	139	23885.0
	Aerial photography	209	27837.4

Over all three sites, the GMEP field survey recorded both the highest number of linear features (213) and the highest combined length of these features (30647 m). Aerial photography recorded the next highest feature count (209) and total length (27837 m), with the Skysat satellite data detecting the lowest number (139) and total length (23885 m).

It should be noted that although this pattern was consistent for sites 1 and 3, at site 2 aerial photography did record both a higher number and total length of linear features than the GMEP field survey, although the Skysat satellite data again recorded the lowest number and total length of features.

For site 1, it was additionally possible to compare the results produced by Automated Methods 1 and 3. Both methods produced a far higher number of WLF features – 215 and 2314 respectively – that the manual or GMEP methods, a function of linear features which were previous being recorded as longer single features being split into multiple smaller features. Consequently, the total length of all features for automated method 1 was less than both GMEP and aerial photography methods, albeit higher than produced by the Skysat data.

Automated Method 3, however, did produce the highest total length of all methods. This is likely to be partly a function of the larger number of features mapped, but also that the features produced by Automated Method 3 better represented the localised variability in the path of the linear features, whereas for the digitised methods these were frequently represented by a single straight polyline.

A visual comparison of six WLF products for site 1 collected or derived from GMEP field survey, digitisation of aerial photography, digitisation of Skysat satellite imagery and Automated Methods 1-3 are presented in figure 8.

This demonstrates that some WLF are consistently mapped by all methods. However, there are some locations, such as in the centre of the displayed figures, where there are inconsistencies in the extent of WLF that are detected by the different analysis methods.

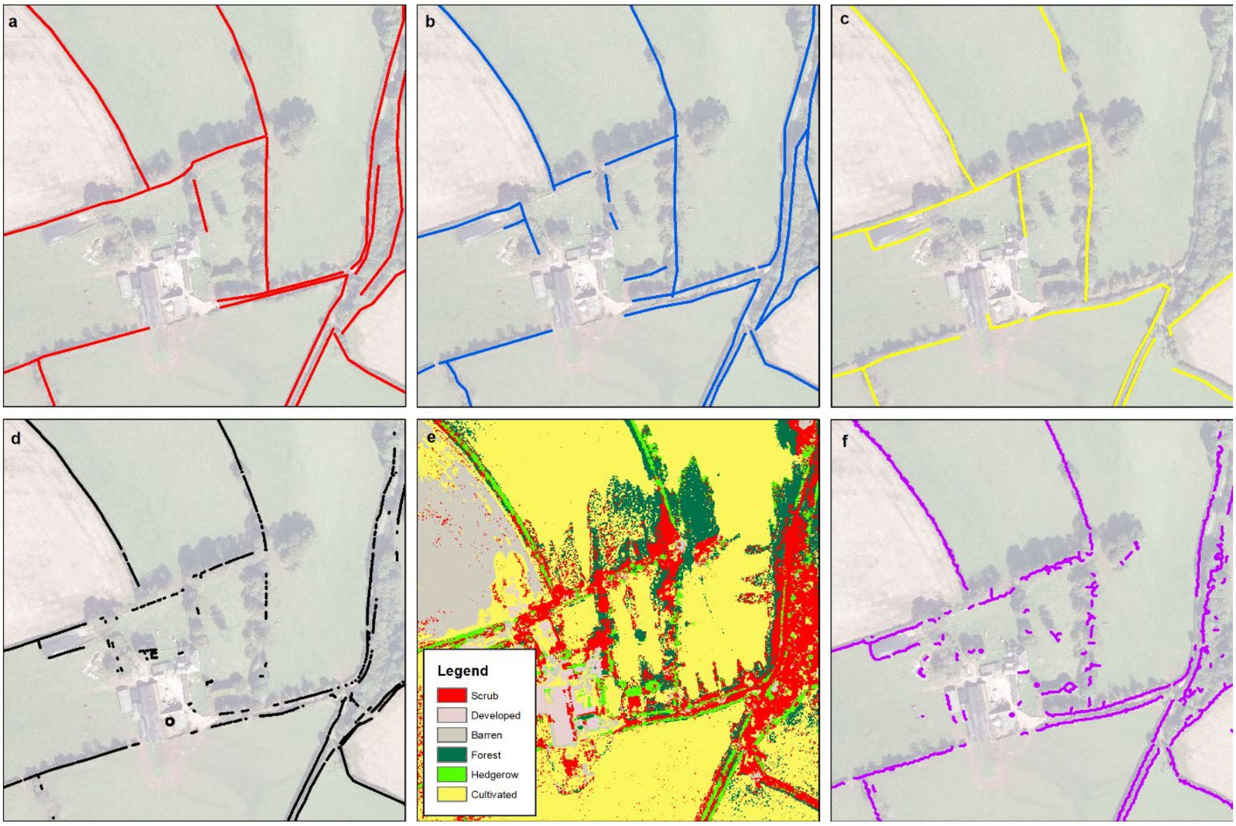


Figure 8: *Woody linear feature outputs from a) GMEP field survey; b) digitisation of aerial photography; c) digitisation of Skysat satellite imagery; d) Automated Method 1; e) Automated Method 2; f) Automated Method 3.*

6 COMPLEMENTARY WORK

Scholefield and Norton (2021) investigated the potential application of LiDAR datasets in combination with existing linear datasets to model the extent of WLF and examined how model outputs compared with field collected data.

Historically, WLF have been mapped during the Countryside Survey field campaigns, with both WLF locations and measures of condition recorded to monitor changes in these features in CS survey squares across the UK. As future CS field survey campaigns will no longer record this information, Scholefield and Norton (2021) investigated whether LiDAR could be used as a data source to continue future monitoring of these features.

The LiDAR data is used to generate a Digital Terrain Model (DTM) describing the elevation of the ground, and a Digital Surface Model (DSM) describing the elevation of features above the ground (trees, buildings etc). By subtracting the DTM elevation values from the DSM elevations, a canopy height model (CHM) is produced quantifying the height of landscape features present.

Filtering procedures were applied to remove non-vegetation features, and further analysis converted the CHM dataset into a series of tree locations (points) and canopy extents (polygons). The tree location and canopy extent datasets were then linked to an existing linear framework, with the Rural Payments Agency (RPA) field boundary data selected due to its higher spatial accuracy with the LiDAR generated data products, compared to either the Land Cover Map spatial framework or OS MasterMap linear frameworks.

Further statistics were generated for each feature, including tree crown area, crown diameter, tree height, and distance between trees, with each tree given a unique ID and the total length of each linear feature calculated. This data could subsequently be spatially joined and compared to CS 2007 attribute data for linear features.

The results showed that of 291 English CS squares, LiDAR data from around only 137 squares were of sufficient quality (or contained enough woody linear features) to estimate total feature lengths based on the LiDAR data alone. Lack of, or incomplete coverage of LiDAR data were the main reason for limits on the number of squares in which this was possible. Wherever LiDAR data was available, a good assessment could be carried out, except for nine squares in which poor alignment of features occurred.

Comparisons were performed between the field collected CS2007 data and the newly created WLF dataset generated using the LiDAR and RPA field boundary data, showing a very good level of agreement between total woody feature extents ($r^2 = 0.9$). Lower levels of agreement were found where the LiDAR acquisition date did not match the data of survey (2007).

However, an important point is that non-RPA field boundary areas are not covered, therefore some supplementation of the framework should be considered for those non-RPA zones.

Additional work by Wood et al. (2015) has also conducted an assessment of the level of habitat data that can be derived from aerial photography, compared to that which can be collected from field surveys.

This assessment determined for six 1 km^2 sites the accuracy and level of detail that could be derived, and the time taken to survey the selected sites to provide an idea of the time costs associated with each method and determine the extent to which point and linear landscape features and Priority Habitats can be assessed using remotely sensed methods.

Field surveys of woody linear features were carried out according to the protocols laid out in Maskell et al. (2008), with all landscape lines, points and areas visited by the habitat surveyor and a range of attributes recorded. These included hedgerow species, linear feature heights, dominant species within areas, field margins, tree species and land use. Aerial photographs were obtained for the corresponding survey squares, although this photography was acquired 3-4 years prior to the field survey.

Feature mapping using the same attributes where possible was again performed based on the aerial photography. The time taken to map the squares using both methods was recorded, with any difficulties of issues noted.

Comparison of the two survey methods showed that field surveys take approximately 12 times longer on average than mapping from the aerial photography, and aerial photography offered 100% coverage of all sites. By comparison, the percentage of sites surveyed during field survey varied between 63% and 100% due to access and logistical restrictions.

However, it was also found that many linear features, and many of the attributes of these features that would be recorded during field survey, were missed when mapping using the aerial photography method. Certain feature types such as hedge types were misidentified, with some features such as banks completely missed.

Additionally, no detailed measurements or condition assessment was possible for any landscape feature, and virtually no species were identified for any of the feature types.

The study concluded that a '*significant underestimate of many features would result if mapping were undertaken from aerial photos on a national level and changes would be difficult if not impossible to assess'*, although is also suggested that aerial photography could potentially be useful in areas where field survey is not possible due to access limitations.

6.1 Advantages and Dis-advantages of Planet Skysat Data

The main advantages are:

1. **The ability to potentially task and acquire imagery rapidly.** However, this is most suitable for environmental events that require rapid, timely data acquisition and intervention, which is unlikely to be true for hedges.
2. **The ability to build up a time-series of data.** This has many benefits, including providing a benchmark against which to measure future changes to the environment in Wales. Plus, the ability to explore methods that use images from multiple dates and also the ability to develop a reference database, as part of a change detection/alert system.
3. **The ability to produce multi-scale solutions**. The routine collection of Planetscope data for Wales provides the opportunity to have an ongoing monitoring system for Wales. Such a system could potentially be designed to trigger a manual/automatic assessment of Skysat data when a certain level of variability was detected in the Planetscope data.

The main disadvantages of the Skysat collect data are:

1. **The poor spatial accuracy of the images.** This required additional processing to correct and adds an additional cost to processing the data. It also limits the potential for using Skysat in automated methods. However, this maybe resolved if Planet can implement bespoke processing for Wales, using higher resolution ground reference data sets.
2. **Limitations of the cloud-mask**. The cloud-mask covers most of the cloud across the Skysat data, however there are areas that remain undetected. This requires additional processing time.

6.2 Potential Methods for Hedge Mapping

Depending on requirements there are a couple of ways that LiDAR data could be incorporated into an operational method for hedge-mapping. Two potential approaches are outlined below.

It is important to note that the level of detail the LiDAR provides about hedgerows will depend on the characteristics and timing of the LiDAR data set (Graham et al. 2019). One of the key issues is being able to separate out raised features that are hedges, from raised features that are walls, or earth banks. The Planet Skysat data, or other very high resolution satellite data, could be very useful for this, if the spatial accuracy can be improved.

6.2.1 Approach 1: Lidar data and boundary data set

Data requirements:

- LiDAR DCM data
- Vector boundary data (e.g. OS MasterMap or existing RPA data)

Key steps in method:

1. Use LiDAR data to create a map of 'raised' features – this will include hedges, individual trees, woodlands, patches of gorse and buildings.
2. Select only 'raised features' associated with field boundaries.

Benefits:

- Can be readily automated.
- Can provide some information on hedge condition (see Section 6)

Limitations:

- Will only map hedges where there is a boundary in the data set, so will not detect new hedges.
- Limited ability to discriminate between hedges/earth banks/stone walls.

6.2.2 Approach 2: LiDAR data, boundary data set and manual intervention

Data requirements:

- LiDAR DCM data
- Vector boundary data (e.g. OS MasterMap or existing Rural Payment Agency boundary data)
- Manual review of potential hedges

Key steps in method:

1. Use LiDAR data to create a map of 'raised' features – this will include hedges, individual trees, woodlands, patches of gorse and buildings.
2. Select only 'raised features' associated with field boundaries.
3. Manually review 'potential hedges' identified in stage 1, against PGA2a aerial photography

Benefits:

- Some stages can be readily automated.
- Can provide some information on hedge condition (see Section 6)
- Can provide benefits of LiDAR data and automation, plus benefits of having a targeted manual review against aerial photography for other areas, as well as potentially targeted field assessment, informed by the manual review.
- More likely to be able to discriminate between hedges/earth banks/stone walls, than Approach 1.

Limitations:

- More expensive than other methods, might require additional training of staff, but could create a high-quality data set.

Note regarding methods 1 and 2 – high resolution satellite data, or aerial photography data could be extracted for the hedges. From this it may be possible to get additional information about hedge condition, although has not been explored. To achieve this would need an extensive database on hedge condition.

7 CONCLUSIONS

This project had two main objectives:

1: Can manual/automatic interpretation of Skysat data meet WG's accuracy criteria for hedge mapping?

Manual interpretation of the Planet Skysat data, compared to aerial photography, and field surveying, showed that the field surveying detected most hedges, followed by the aerial photography, with the Skysat data least suitable for hedge identification.

In addition, the aerial photography was analysis ready and did not require any pre-processing prior to use, whereas the Skysat data did require pre-processing, including time-consuming geocorrections for sites 2 and 3.

The need for manual geocorrection, prior to use, is a significant limitation in the use of Planet Skysat data for Welsh Government activities.

2: Does LiDAR data (combined with Skysat data) help with discrimination?

Combining the LiDAR data with the Planet Skysat data did allow for the potential to filter out non-vegetated areas, but the winter images were not ideal for testing the capabilities of this. Future work could explore the improvements to be gained by using spring/summer images, or multi-date images and more current LiDAR data sets.

The key findings of the project were:

1. Aerial photography was a better source of information for manually digitising hedges than the Planet Skysat data.
2. For two of the three sites, the spatial accuracy of the Skysat data was too low to mapping hedges, without spending additional effort manually geo-correcting images. However, Planet maybe able to use existing WG spatial data sets to improve the spatial accuracy, which could change this finding.
3. The Skysat cloud-mask does not capture all the cloud.
4. The Skysat data would be more useful for automated processing if the Root Mean Square Error for the Ground Control Points (GCP), and the number of GCP's, was reported in the metadata.
5. Automated methods using LiDAR show promise and could be deployed in a number of ways depending on whether the aim to measure some attributes of hedge condition, or hedge location and length.
6. Methods using automated detection of new hedges using LiDAR data, followed up by manual checking against aerial photography or very high resolution satellite data, may be able to capture the best aspects of different approaches.

Future work

If Planet Skysat data for Wales becomes available with higher spatial accuracy, then it would be useful to re-assess the ability to map hedges, by considering the following questions:

1. Which time of year enables the most accurate hedge mapping?
2. Does multi-temporal data improve hedge mapping? If so, which are the key times of year?
3. Can automated change detection be used to automatically identify areas where hedges are being removed (or added)?
4. Can hedges be mapped accurately enough to meet WG's requirements?

8 REFERENCES

Emmett B.E. and the GMEP team (2017) Glastir Monitoring & Evaluation Programme. Final Report to Welsh Government - Executive Summary (Contract reference: C147/2010/11). NERC/Centre for Ecology & Hydrology (CEH Projects: NEC04780/NEC05371/NEC05782). https://erammp.wales/sites/default/files/GMEP-Final-Report-2017.pdf

Graham, L., Broughton, R.K., Gerard, F., Gaulton, R. (2019) Remote Sensing Applications for Hedgerows. In: J. Dover (Ed): The Ecology of Hedgerows and Field Margins. Earthscan/Routledge, Abingdon, UK. pp. 72-89

Maskell, L. C.; Astbury, S.; Burden, A.; Emmett, B. A.; Garbutt, A.; Goodwin, A.; Henrys, P.; Jarvis, S.; Norton, L. R.; Owen, A.; Sharps, K.; Smart, S. M.; Williams, B.; Wood, C. M.; Wright, S. M. (2020) Landscape and habitat area data from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016. NERC Environmental Information Data Centre. https://doi.org/10.5285/82c63533-529e-47b9-8e78-51b27028cc7f.

Maskell, L. C.; Norton, L. R.; Smart, S. M.; Carey, P. D.; Murphy, J.; Chamberlain, P. M.; Wood, C. M.; Bunce, R. G. H.; Barr, C. J. (2008) Countryside Survey. Field Mapping Handbook. NERC/Centre for Ecology & Hydrology, 130pp. (CS Technical Report No.1/07, CEH Project Number: C03259).

Planet Labs Inc. (2018) PLANET IMAGERY PRODUCT SPECIFICATIONS, online at: https://assets.planet.com/docs/Combined-Imagery-Product-Spec-Dec-2018.pdf

Scholefield, P. A.; Norton, L. R. (2021) Can we extend the CS woody linear dataset using LiDAR data? Lancaster, UK, NERC Centre for Ecology & Hydrology, 41pp. (Report for Defra MOA NEC07111).

Wood, C. M.; Norton, L. R.; Rowland, C. S. (2015) What are the costs and benefits of using aerial photography to survey habitats in 1km squares? Lancaster, UK, NERC Centre for Ecology & Hydrology, 23pp. (CEH Project no. C04867) (Unpublished) http://nora.nerc.ac.uk/id/eprint/511330/.

This page is intentionally blank.

ERAMMP Programme Office
UKCEH Bangor
Environment Centre Wales
Deiniol Road
Bangor, Gwynedd
LL57 2UW
+ 44 (0)1248 374500
erammp@ceh.ac.uk

www.erammp.cymru
www.erammp.wales