



HAL
open science

Making LiGHT Work of Large Area Survey? Developing Approaches to Rapid Archaeological Mapping and the Creation of Systematic National-scaled Heritage Data

Dave Cowley, Łukasz Banaszek, George Geddes, Angela Gannon, Mike Middleton, Kirsty Millican

► To cite this version:

Dave Cowley, Łukasz Banaszek, George Geddes, Angela Gannon, Mike Middleton, et al.. Making LiGHT Work of Large Area Survey? Developing Approaches to Rapid Archaeological Mapping and the Creation of Systematic National-scaled Heritage Data. *Journal of Computer Applications in Archaeology*, 2020, 3 (1), pp.109-121. 10.5334/jcaa.49 . hal-02567732

HAL Id: hal-02567732

<https://hal.science/hal-02567732>

Submitted on 7 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

POSITION PAPER

Making LiGHT Work of Large Area Survey? Developing Approaches to Rapid Archaeological Mapping and the Creation of Systematic National-scaled Heritage Data

Dave Cowley, Łukasz Banaszek, George Geddes, Angela Gannon, Mike Middleton and Kirsty Millican

The characteristics and form of heritage data are fundamental to its utility in a range of applications, particularly so for heritage agencies who have a remit in management, policy, and the creation and curation of national databases of monuments, sites, and landscapes. Written from the perspective of an archaeological survey function in a national heritage agency, this paper draws on preliminary outcomes from a research and development project that aims to proof protocols for creating systematic data across large areas drawing heavily on remotely sensed data. This recognises that a systemic consideration of the implications of changing technology and data is sometimes desirable, rather than gradual assimilation of developments into existing practice. In particular, the issues being addressed relate to the challenges and opportunities of proliferating remote sensed data and digital workflows. These include the strategic assessment of threat, consideration of fitness for purpose of different datasets relative to landscape characteristics, the documentation of processes and sources of information, the suitability of data structures, and the mechanisms for automating site detection and data creation.

Keywords: Large area survey; Airborne Laser Scanning; LiDAR; orthophotographs; survey scale; heritage management

1 Introduction

This paper presents ongoing work at Historic Environment Scotland (HES) in developing its applications of remote sensing (RS) for archaeological landscape survey (mainly Airborne Laser Scanning or ALS and aerial photographs, but also including scoping of satellite data). HES has many roles as the lead publicly-funded body for the historic environment in Scotland (HES 2019a). However, the focus in this paper is on its functions in identifying, recording, understanding and interpreting the historic environment (Historic Environment Scotland Act 2014), for which archaeological mapping of sites and landscapes is a key publicly-funded role established for over a century (e.g. Dunbar 1992; McKeague & Cowley 2013). This function has benefited from the routine use of traditional aerial photography for over 70 years (Cowley 2016a; Geddes 2014; Maxwell 1983), but the exploration and uptake of other RS data such as ALS/airborne LiDAR data and satellite imagery, for example, have been slow by comparison with some other national heritage agencies (e.g. Crutchley 2013). Consequently, while archaeological use of ALS was common globally by the early 2000s, and there were examples of projects in Scotland applying it (i.e. AOC

2015; Forestry Commission Scotland 2012), it was not until 2016 that HES undertook a landscape-scaled survey of 19 square km of the Kraiknish peninsula on Skye informed by ALS-derived visualisations amongst other sources (HES 2017). This was followed in 2017–18 by a survey of the island of Arran (432 square km) heavily informed by ALS-derived visualisations and designed to test methods for large area rapid archaeological survey (Banaszek, Cowley & Middleton 2018; Cowley & López-López 2017).

The increasing availability of data in Scotland, and the Kraiknish and Arran projects completed in the last few years, have provided the momentum for the structured exploration of approaches to large area survey in Scotland. This has greatly benefited from the many examples of archaeological survey practice informed by RS data in Europe and beyond, and is the context within which HES established the Rapid Archaeological Mapping Programme (RAMP) in 2019 as a two-year research and development (R&D) project that aims to develop protocols for creating systematic archaeological data across large areas. While this approach draws heavily on RS data, the issues being addressed also include the need for explicit documentation of the underlying knowledge-creation processes and the mechanisms for automating site detection. This is explicitly construed as R&D because we recognise the necessity to consider holistically the implications of proliferating data and digital workflows. Central to this

is an assessment of how the effectiveness and cost-benefit of ALS and other RS data varies across Scotland's diverse landscape and how a range of established and emergent data and methods are best deployed in our specific context. Moreover, by examining and testing workflows and considering the character of information derived from different sources and methods, we hope to achieve significant efficiencies and to create information that is systematic and well-documented (i.e., with thorough metadata). The discussion that follows sets out the background to RAMP and some of our preliminary work.

While we are learning from the experiences of others, we hope not to borrow approaches without assessing them critically. In considering approaches to large area survey, the work of English Heritage/Historic England, especially in their National Mapping Programme (Evans 2019; Horne 2011; Winton & Horne 2010), and the ALS-based mapping of Baden-Württemberg in southwest Germany (Bofinger & Hesse 2011; Hesse 2013), have been our main reference points for relevant interesting practice. While these will not be reviewed in detail here, it is worth noting the importance of efficient workflows for mass data handling, exemplified by the work in Baden-Württemberg (Bofinger & Hesse 2011: 165–6), the value of incorporating complementary data sources (i.e. Evans 2019), and the challenges of balancing rapid coverage and detail of information (for example in choosing to create vectors depicting elements of sites, or to create simple site area polygons). These factors all have an immediate impact on resource requirements and thus the scalability of the approach, aspects of which will be considered below through a discussion of levels of survey.

2 Approaching Large Area Rapid Archaeological Survey

HES archaeological survey seeks to address several aims, such as informing detailed analysis of landscapes or site types (e.g., Gannon & Geddes 2015; RCAHMS 2009; Welfare 2011), but is primarily directed to enhancement of the National Record of the Historic Environment (NRHE). The NRHE is readily accessible through 'Canmore' – the online catalogue to Scotland's archaeology, buildings, industrial and maritime heritage (HES 2019b). The national picture is complemented by regional Historic Environment Records, generally maintained by local Councils, or regional administrative bodies. Such databases, in Scotland and beyond (e.g., papers in Larsen 1992; Schut 2009), serve many purposes, informing research, management, strategic planning, community outreach and so on. As products of over a century of recording informed by changing policy and interests, unsurprisingly they show considerable variability in content and coverage (i.e., Banaszek, Cowley & Middleton 2018; Cooper & Green 2015; Cowley 2016b; Verhagen et al. 2016).

The evidence base represented by our heritage information is a foundation for ensuring that the historic environment is understood and valued, cared for and protected, enjoyed and enhanced – what is not known cannot be managed or protected, or inform our understanding of the past. This is a primary imperative to contribute systematic and consistent archaeological information to the

NRHE, which is reinforced by indications that the rate of landscape change may well accelerate soon. Indeed, factors such as responses to climate change and the implications of Britain's withdrawal from the European Union are likely to have impacts that will directly affect agricultural practice and rates of afforestation, for example (e.g., Scottish Government 2019a; Scottish Government 2019b). All such land-use change represents threats to the survival of archaeological remains. Archaeological survey, whether on the ground, aerial or desk-based, has demonstrated significant increases in numbers of sites recorded wherever it has been undertaken, and this supports a reasonable expectation that there are tens of thousands of presently unknown ancient sites scattered across Scotland's landscapes. These 'known unknowns' are the key driver for expedited large area survey, which needs to respond to changing patterns of threats in a timely manner. This expedited approach to survey, which is being explored in a 21st century context by RAMP, follows on from programmes of rapid survey undertaken during the last century, also prompted by the recognition of increased threat from land-use change such as afforestation and improvement of marginal land (i.e., Geddes 2014; Halliday & Ritchie 1992). While these earlier attempts did not endure as priorities, the present need in the face of accelerating land-use change is pressing. At the same time, the proliferation of RS data and the analytical opportunities such data affords present new opportunities to achieve significant upscaling in rates of coverage compared to what has been achieved to date (**Figure 1**).

3 Levels of Survey

Underlying many of the decisions about how to approach a survey is a consideration of objectives and definition of specification, factors that can usefully be approached through defined levels of survey. Such frameworks offer an explicit mechanism to consider the balance between extents of coverage, resource requirements, the range of sources/methods applied, and the character of the data produced – all of which directly inform any desire to achieve significant upscaling in rates of coverage. The HES levels of survey aim to offer transparency for both record creators and users. While different levels of survey (**Table 1**) are likely to be combined in a single project, they are a useful shorthand from which to understand the character and structure of survey outputs.

In this scheme there is a relationship between levels of survey and scale, and this is largely due to a tradition of working within the structure of UK Ordnance Survey (OS) map scales. These step down in a graduated way (e.g., 1:2500, 1:10,000, 1:25,000, 1:50,000, etc.), each with a different level of detail and abstraction of information depending on the nature of the landscape (i.e. urban/rural) and the anticipated usage. There are also linkages between a level of survey and the methodology and source data that may be drawn on for mapping. Thus, for example, aerial photographic and other RS datasets will have their own error budget and mapping tolerances, which differ from survey-grade GNSS, and bear on the metrical accuracy, as one attribute of a survey, that they will support.

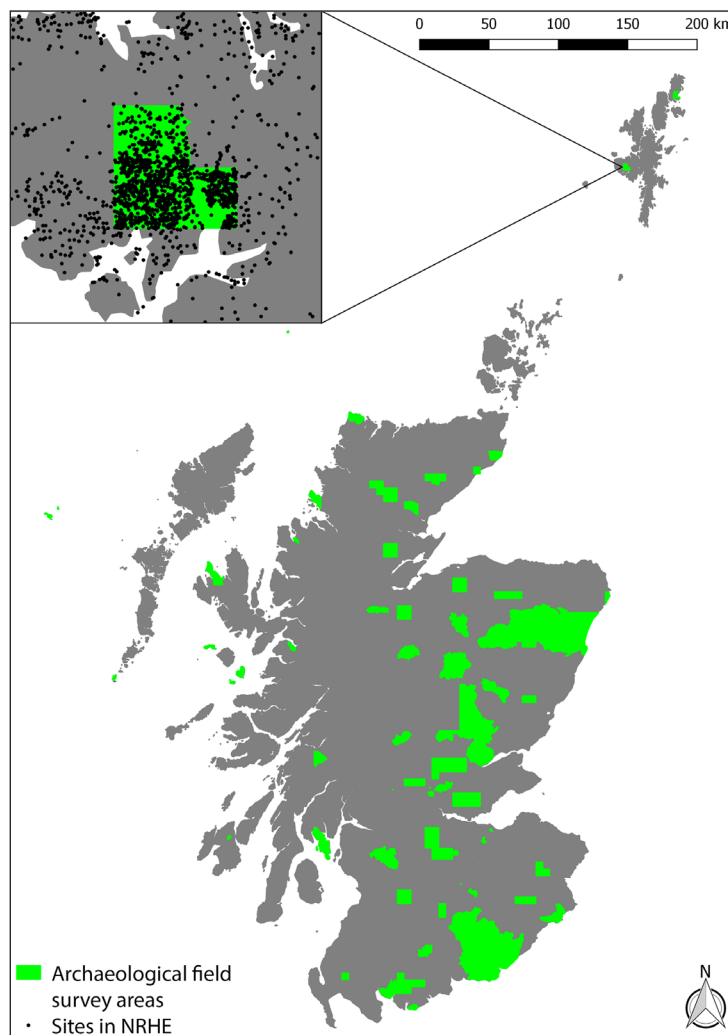


Figure 1: The areas of systematic, extensive field survey to modern standards undertaken over the period from about 1990 to about 2015 have covered some 10% of Scotland’s land mass. The significant addition to the NRHE that systematic area survey can make is illustrated by West Mainland, Shetland (top left), where the discrete area investigated through aerial photographic and pedestrian survey in 2010 is clearly visible in the increased density of site locations. The knowledge that there are many unrecorded sites, and that an increasing rate of land-use change is likely, attaches a high priority to upscaled, considerably more rapid, rates of coverage than those achieved by survey programmes to date. © Historic Environment Scotland.

Table 1: Summary of HES levels of survey with a generalised indication of the suitability for area coverage and resource requirements. In simple terms, the resource requirement by area covered increases from Level 1 to Level 4.

Level of survey	Scale	Map accuracy	Outputs
1: Landscape characterisation (broad-brush assessment)	1:25,000	+/- c. 25m	<ul style="list-style-type: none"> • Location polygons • Classification • Period
2: Core information for NRHE	1:10,000	+/- c. 10m	As Level 1, plus: <ul style="list-style-type: none"> • Project event (description of why and how a project was undertaken, including statement of methodology and accuracy)
3: Basic record	1:10,000 1:2500	+/- c. 10m +/- c. 1m	As Levels 1 and 2, plus: <ul style="list-style-type: none"> • Recording event (description of how a record was created, including source data and personnel/organisation) • Brief written description • Survey at a scale that indicates monument or landscape form
4: Detailed record	1:2500 1:1000 or larger	+/- c. 1m	As Levels 1, 2 and 3, plus: <ul style="list-style-type: none"> • Detailed analysis and interpretation • Survey at a scale suitable to depict character and complexity of the monument • Photographic record as appropriate

Levels of survey strike to the core of the issues that define approaches to rapid large area survey – that there is a challenge in balancing economy and speed of coverage against detail of record. This is an issue in GIS environments, where ‘map scale’ has become somewhat irrelevant, and there is a common tendency to work as a default at the highest level of detail that RS data resolution will support. This temptation is the enemy of rapid large area survey, as it may draw the interpreter into the absorbing detail and away from broader-brush coverage. Addressing this range of issues around explicit survey design, mapped against levels of survey, methodology and source data, to design protocols for systematic and sustainable rapid mapping, is the central concern of the Rapid Archaeological Mapping Programme.

4 Prioritisation, Source Data Assessment, and Methodology

In seeking to explore economic and rapid methods to enrich the NRHE, the need for structured consideration of the following factors is identified: needs (including known and emergent threats and the variable quality of existing data, Section 4.1); systematic assessment of suitability of data to different landscape contexts (Sections 4.2 and 4.3); and development of data analysis and processing routines that are born-digital (Sections 4.4 and 4.5).

4.1 Where and why? Prioritising survey

The considerable variability in coverage and quality of heritage data makes definition of a strategic basis for targeting survey work important. Alongside a structured assessment of the strengths and weaknesses of the currently available data, prioritisation of where to direct survey work will take account of those parts of the landscape that are most vulnerable to change – where the threat to unknown heritage assets is most severe. Land-use change is always happening, but there are also imperatives arising from increased awareness and concern with the impacts of climate change, which the Scottish Government has made

a priority to address (e.g., Scottish Government 2019c). In this agenda there are explicit commitments to encourage more tree planting, including woodland integration and agro-forestry on farms, to promote the multiple benefits of good grassland, and to encourage farmers to invest in renewable energy, including bio-energy (Scottish Government 2019c: 57). These selected examples of initiatives to address climate change all have potentially profound implications for the historic environment and archaeological remains. Strategically improving the knowledge base is one mechanism for ensuring that archaeological assets are considered as part of the process of managed landscape change. A key output of the RAMP development will thus be a flexible matrix for prioritising survey work that accounts for where the most pressing demands for reliable and systematic data are located, drawing in information from regional and national Development and Structure Plans. This will need to be a flexible evolving process of prioritisation, but importantly one that recognises the importance of sound archaeological data as a basic input to ensure that heritage management responses are well-informed.

4.2 Addressing landscape variability

The matching of appropriate survey methods and data sources to landscape types is crucial (Cowley 2011: 44–5; Evans 2019), since the effectiveness of survey and the utility of information produced by it depends on the design and quality of the methodology. Scotland’s 78,000 square km landmass exhibits considerable variability, and this means that a ‘one size fits all’ approach to survey methodology is unlikely to be appropriate. This puts the broad-brush characterisation of the Scottish landscape centre stage – in simple terms allowing assessment of ‘what works well where, and why?’ (Figure 2).

In Scotland, the Historic Land-use Assessment (HLA) dataset (HES 2019c; Watson & Dixon 2018) is a valuable tool to help assess these questions. It provides national spatial data at a scale of 1:25,000 that characterises



Figure 2: The landscapes of Scotland are marked by major variations in topography, land cover and soils, amongst other factors, all of which have impacts on the type of surviving archaeological remains and the suitability of recording methods. SC622986; DP082172; DP056902 ©Historic Environment Scotland.

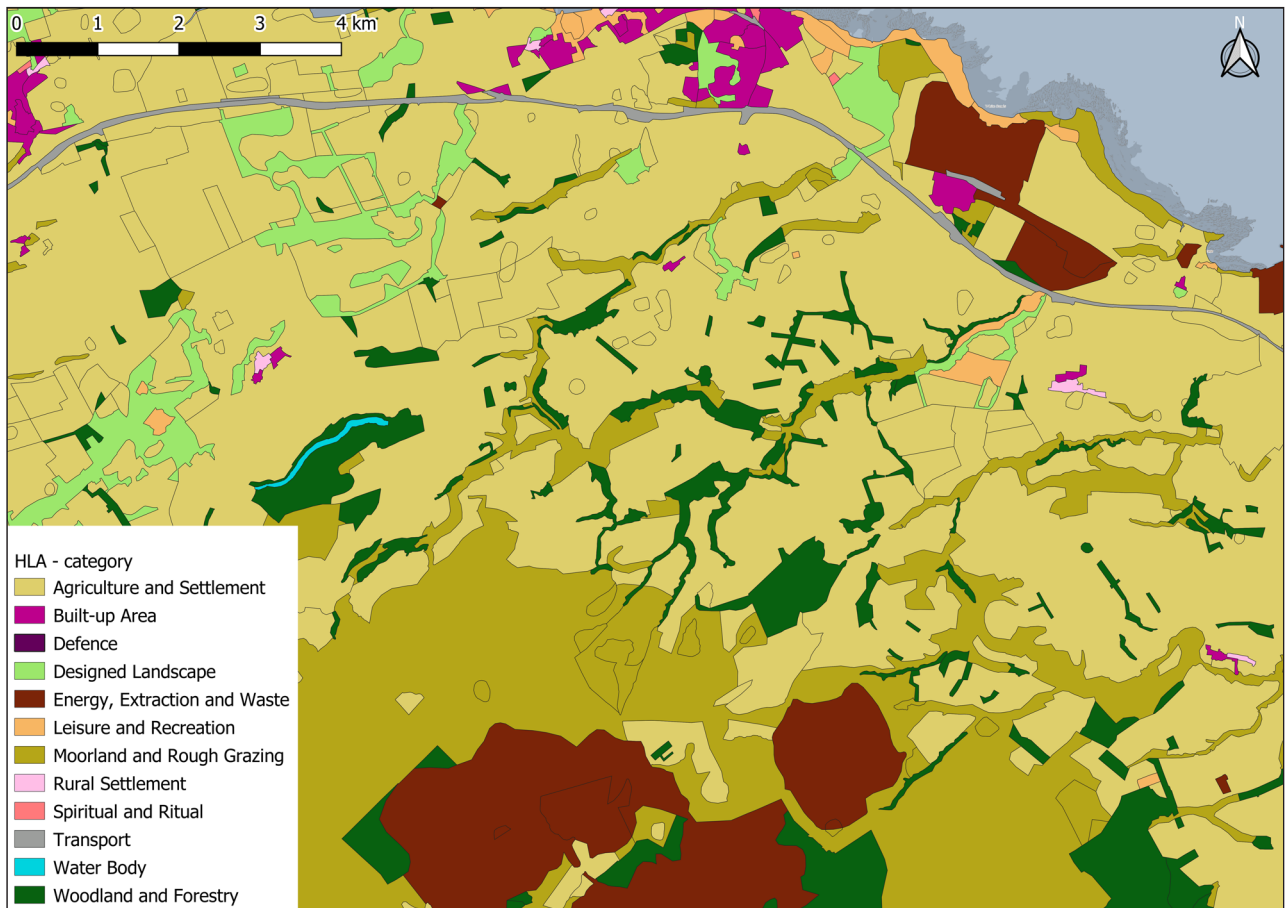


Figure 3: Historic Land-use Assessment data of the eastern part of East Lothian displayed for current landscape character (e.g., built up, recreation, woodland, pasture, etc.). These broad characteristics have implications for archaeological potential and the effectiveness of different survey techniques. All areas also have their relict characteristics recorded as attribute data where possible, indicating, for example, where medieval or earlier landscape features may be visible. © Historic Environment Scotland.

both historic land use and the contemporary landscape for areas greater than one hectare in extent. The HLA dataset has been compiled using current and historical OS maps, the 'All Scotland Survey' of aerial photographs taken in 1988/9, and other aerial images and regional and national archaeological datasets. The data is structured around over 54 current land use types (e.g. coniferous plantations, rectilinear fields and pastures) and 43 purely relict land-use types (e.g., prehistoric settlement, designed landscape) linked to periods of use, and allocated to one of 12 broad categories (e.g. Leisure and Recreation, Transport or Defence), presenting a broad-brush understanding of landscape development over time and the 'past in present' of the contemporary landscape (**Figure 3**). It is also a useful dataset with which to consider the broad relationships between landscape variability, archaeological potential, the suitability of sensors/data sources, and bias in the distribution of archaeological remains.

4.3 Assessing data characteristics and fitness for purpose

The introduction of new methods in archaeology has sometimes been characterised by uncritical applications that foreground technology for the sake of technology.

This is unsurprising when there is a need to showcase advantages and applications (Opitz & Cowley 2013: 7–10) but can run the risk of losing sight of the archaeological imperative by overly-focussing on a technique (see Cowley et al. 2018 for a discussion of this tendency with UAVs). Indeed, some aspects of this propensity can be seen in the use of terms like 'aerial archaeology' or 'satellite archaeology', which are sometimes defined by an uncritical preference for a source of information. Recognising that uncritical attachment to a particular 'flavour of the month' is unwise because it emphasises technology over purpose at hand, the RAMP assessment of data characteristics focusses on fitness for purpose and cost-benefit for the primary aim of generating systematic general-purpose survey information.

4.3.1 3D topographic data: ALS and photogrammetric outputs
ALS data is a key source in the rapid approach to large-scale coverage. RAMP work thus far has largely focussed on where 'general purpose' data is freely available with piecemeal use of bespoke datasets collected locally for HES conservation reasons. This varies greatly in resolution, from a nominal 1 pt/m² to roughly 25 pts/m², and, at the time of writing covers about 20% of the country (**Figure 4**).

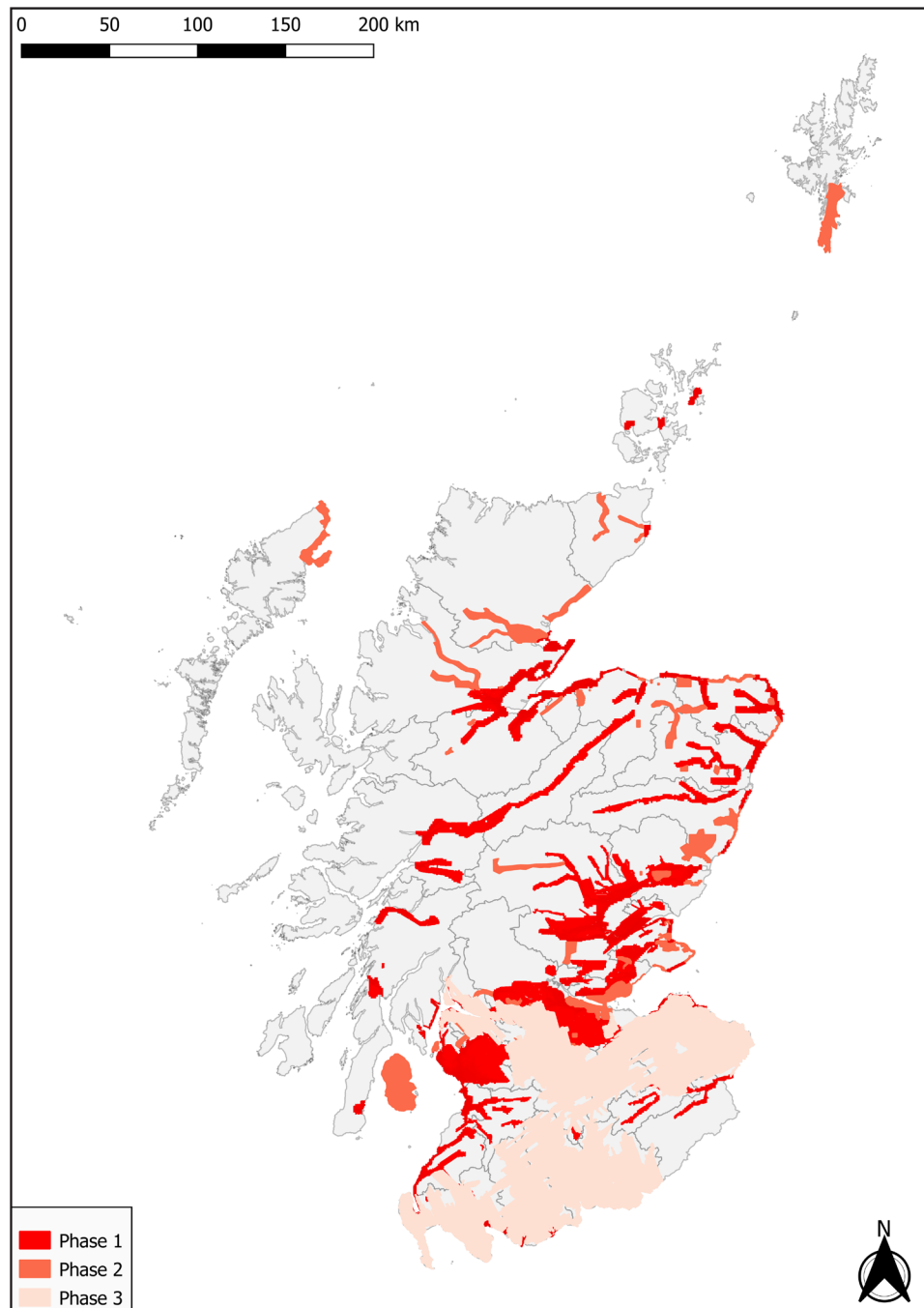


Figure 4: The overall distribution of currently available ALS data in Scotland from three phases of data acquisition. The data varies from 1 pt/m² (Phase 1 and parts of Phase 2) to roughly 25 pts/m² (Phase 3). Depending on the characteristics of the landscapes and archaeological remains, even the lower resolution data is useful. While the very high-resolution data offers considerably more detail, it often does not offer penetration of the dense coniferous canopy of much woodland. © Historic Environment Scotland.

The Level 2 (**Table 1**) survey of Arran (HES 2019d) has generated some metrics with which to assess the utility of ‘medium’ resolution data with an overall point density of 4.67/m² and an average ground point density of 2.95/m². Here desk-based mapping using multiple ALS-derived visualisations, orthophotographs, and 19th and early 20th century maps, was undertaken by nine HES staff, with a focus on the discovery of previously unknown sites. Identifications of sites and targets were delimited with a polygon supported by classifications and a level of confidence in the interpretation (ranked from high (1) to low (3)). Subsequently, in 2018, six weeks of fieldwork by a

team of six were directed to checking the medium and lower confidence identifications, specifically in the hope that field observation would improve the certainty of interpretation. Patterns of walking that explored gaps in coverage were also encouraged, with the routes of most walking documented by GPS recorded tracks (**Figure 5**). During the fieldwork stage, some 500 sites discovered during desk-based mapping were visited, with a further 152 sites discovered solely through field investigation mainly in areas where the ALS survey was suspected to be ineffective. These areas lay primarily within blocks of dense coniferous plantations, with a preponderance of

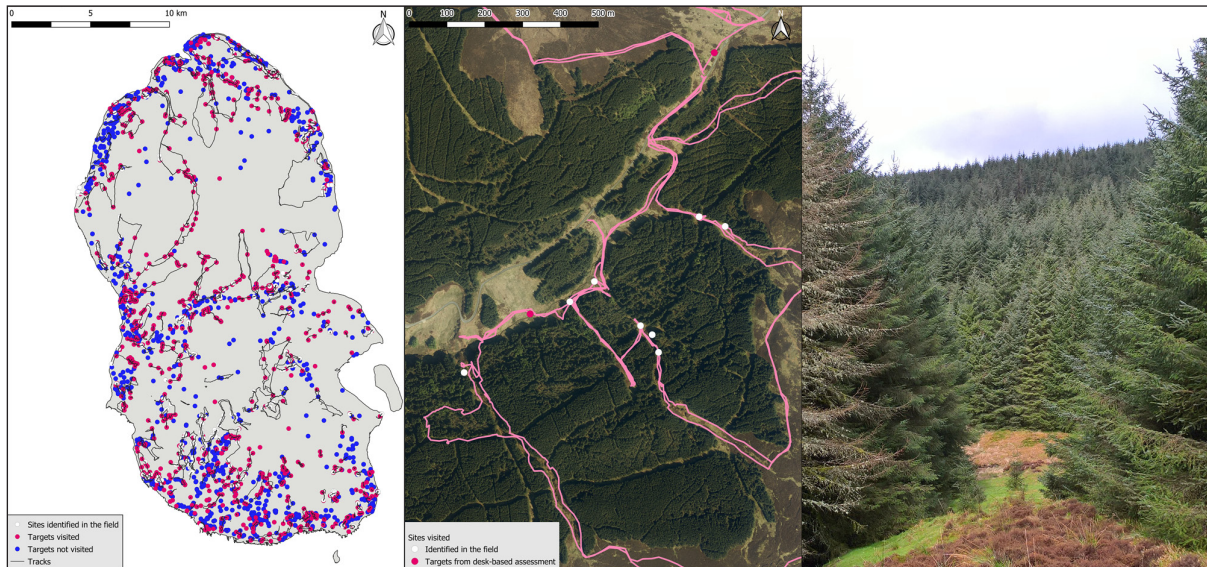


Figure 5: Left – The patterns of desk-based discoveries, sites visited in the field, and sites discovered through pedestrian survey against the GPS recorded tracks of most pedestrian routes. Centre: pedestrian survey was targeted at areas where ALS ground point densities were known to be low because of the very dense coniferous plantations (right). © Historic Environment Scotland; Aerial Imagery © Getmapping plc.

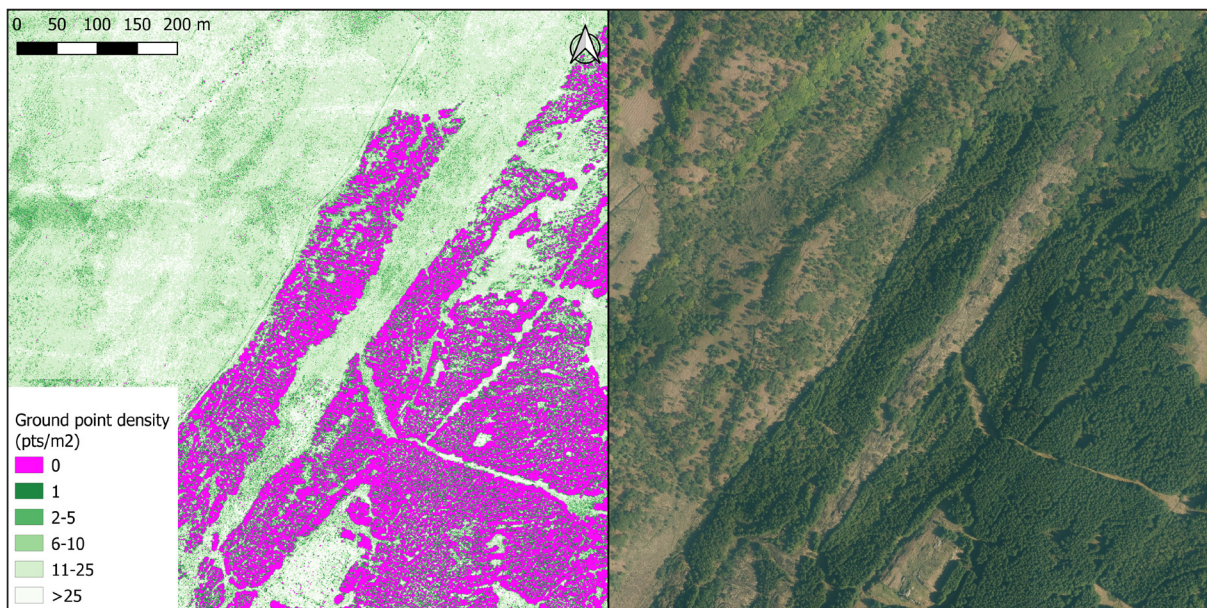


Figure 6: Even the high point densities of the Kilmartin Glen ALS dataset did not penetrate the dense coniferous plantation as illustrated in this plot of areas with no ground points (left). © Historic Environment Scotland; Aerial Imagery © Getmapping plc.

small shieling (transhumant grazing) huts amongst the discoveries from field walking.

These patterns of discovery demonstrate that the medium resolution data is effective for documenting a wide range of remains across a variety of landscape types. However, such documentation needs to be supplemented by targeted field reconnaissance into known areas of lower resolution data. This is because certain monument types may be inadequately captured by the ALS due to their small size.

The benefits of higher resolution data were explored in Kilmartin Glen, during a survey project undertaken in early 2019 that made use of a bespoke dataset with an overall point density of 63.18/m² and an average ground

point density of 24.39/m². The analysis of the survey outcomes is still in progress at the time of writing (January 2020). While the higher resolution data provided additional benefits in detail, assessment of the impact of the higher point densities for areas of coniferous woodland demonstrates that even the high point densities do not penetrate the tightly planted trees that characterize Scottish forestry (**Figure 6**).

This brings into question the cost-benefit of the higher resolution data in this area, where the woodland cover, as in much of Scotland, is dominated by very densely planted conifers. Indeed, this question of the cost-benefit of higher resolution data applies for Level 2 survey across much of the country. This is one of the reasons why an assessment

of the suitability of photogrammetric derived 3D topographic data, which is considerably cheaper than ALS (i.e., 20% of the cost), is being undertaken. This assessment will focus on the differences of the aerial photograph derived topographic data to ALS for the large unwooded areas of the country (see Section 4.2). Here the key question is what added value does relatively expensive ALS represent for large areas of Scotland, where even high-resolution ALS does not penetrate the dense woodland canopy?

4.3.2 Back to basics: utilising aerial photographs

This cost-benefit question of ALS data also bears directly on the assessment of the high-quality orthophotographs (ground sampling distance of 25 cm) that are freely and readily available to HES through a service agreement with the Scottish Government. The utility of such images for archaeological mapping varies enormously depending on lighting and vegetation conditions, especially as none were taken with archaeological imperatives in mind. However, their metrical accuracy is generally well within the tolerances of a Level 2 survey (**Table 1**). If they are serendipitously taken at the right time, they are a highly effective survey data source. The use of simple metadata for the orthophotographs, such as time and date of capture, qualitative sample assessment of archaeological feature visibility (depending on subjective factors such as lighting), and matching to landscape types (Section 4.2, **Figure 2**) will provide a structured appraisal of where suitable imagery is available. Additional imagery collected in an ongoing programme will help formulate a simple matrix to assess the potential utility of imagery as it becomes available. These same general principles apply to the assessment of satellite imagery utility, though currently cost is a major issue with the freely-available data of too coarse a resolution to be useful for anything other than a Level 1 (Characterisation) survey.

The basic assessment of cost-benefit, or the relationships between the resource required to extract information and the extent of that information, will also apply to historic aerial photographs. Aerial photographic coverage extending back to the 1940s can represent unique views of many areas that have since been afforested or subsumed by urban expansion (e.g., Cowley & Stichelbaut 2012). However, as with modern orthophotographs the utility of the imagery for archaeological survey will vary. For example, a rapid assessment of the historic aerial photographic coverage for Arran established that the value of the imagery was compromised by its scale, by the overall dull lighting, and by the poor quality of the available prints. For these reasons, a judgement was made that the time required to inspect 1500 prints, let alone scan and georeference even a subset of them was not justified in the context of a Level 2 survey where area coverage and economy of resource are key considerations. This assessment did not use stereo-viewing of the aerial imagery, and that is recognised as a shortcoming, and indeed a limitation, in the use of orthophotographs and satellite imagery. While stereo-viewing of aerial photographs is undoubtedly best practice, it is time-consuming and bears directly on any cost-benefit assessment. That said, the metadata for the historic aerial photographs will allow quick identification of imagery that may be most

useful (for example, those taken during the winter of upland areas where earthworks may be expected).

4.4 Documentation, data processing, and workflows

The assessment of approaches to survey also includes data processing routines and the mechanisms for adding data to the NRHE. Many of these routines are presently inherently manual, often requiring double-handling and rekeying of data, which takes time and adds to the potential for error. Thus, for the Arran project undertaken in 2018/19, the manual entry of the survey data to the NRHE took several weeks, based as it was on the standard existing practice. Since then, an entirely digital workflow has been designed and is being tested. This combines data downloads of existing records from the NRHE, desk and field-based data collection processes in QGIS, and an ingest mechanism to present a csv file for upload to the NRHE. This latter mechanism will automate the largely manual record creation process and reduce the time taken for this task from several weeks or so to a few days. Terminology for monument classifications and period attribution is enforced throughout from the schemes used in the NRHE.

Underpinning these developments is also a desire to better document survey processes to provide end users with a better idea of how a record was created, who was responsible, and why it was done. Aspects of this are documented through 'project events' (e.g., HES 2017; HES 2019d) and 'recording events'. The project event details why a project was undertaken, its methodology and its general results while, at a record by record level, the recording event documents the source of information (e.g., ALS, aerial photographs, field investigation), the person who created the record (the 'authority'), their role (i.e., archaeologist) and the organisation (e.g., HES). To this structured metadata, we are adding GPS tracks of field investigation that record the ground covered by staff (**Figure 5**). This adds valuable detail to a general statement that 'a survey was undertaken' for anyone wishing to undertake post-hoc analysis of the resultant data (see Cowley 2003 for an example of such analysis for a programme of aerial survey). In considering the underlying patterns that may collectively produce the survey results for a project, we are also documenting interpersonal variability in output, especially from desk-based mapping, recognising that exploration of how and why different staff may see the same data (i.e. Banaszek, Cowley & Middleton 2018: 10–11) is important to the development of a common approach and identifying training needs.

4.5 Automated object detection

The improvement of workflows and the work on the rapid mapping of Arran have demonstrated the extent to which survey processes can be expedited. The extensive use of ALS for the survey of Arran saw the completion of the island in a roughly equivalent amount of time that might have been required to cover a piece of ground half the size or less without the heavy reliance on the visualisations for mapping. The development of an efficient workflow that removes rekeying is projected to reduce many weeks of data input to the NRHE to no more than a week or so. Together with automated data processing (e.g. produc-

tion of ALS visualisations), such measures may improve rates of coverage by up to four-fold compared to previous approaches. These are worthwhile efficiencies, but even so, such improvements in rates of coverage do not scale up to massive areas, such as the 78,000 square km of Scotland's land mass. This is compounded if there is an aspiration to explore the proliferating RS datasets in any way. Most of the aerial photographs and satellite data that have been collected over Scotland during the last 70 years have not been examined for their archaeological potential. For these reasons, a key area of research for RAMP is in machine learning and artificial intelligence (AI) to expedite detection of archaeological objects in landscape datasets (e.g. Bennett, Cowley, & De Laet, 2014; Sevara et al. 2016; Trier, Cowley & Waldeland 2018; Zingman et al. 2016). This is where the real capacity to upscale survey lies.

Preliminary work on automating object detection on Arran (Trier, Cowley & Waldeland 2018) and ongoing work by others (Lambers, Verschoof-van der Vaart & Bourgeois 2019; Landauer & Hesse 2019; Verschoof-van der Vaart & Lambers 2019) is proving promising, but has also thrown up a series of issues including a general concern about how an AI works. There is a need to understand the performance of neural networks, including factors such as how they are trained, how they learn, how transferable systems are from one area to another, and how and why outputs from one network may differ from another. For example, the influence of pre-training of neural networks on down-the-line outputs is currently unclear (i.e., Gallwey et al. 2019; Trier, Cowley & Waldeland 2018: 9–10), or how the character of inputs (i.e., which visualisation, what processing parameters, etc.) may influence the results. These are expressions of wider concern with better understanding of how processes work, or how data is structured, and in the context of AI and machine learning, how such approaches are integrated into work patterns.

For HES, as a national body seeking more extensive, reliable, and systematic archaeological data to inform management and research, a key issue is bringing AI to bear in exploring the growing mass of remote-sensing data (Cowley 2020). If we are to do more than scratch the surface of large volumes of data, we need ways of detecting objects of interest that can work tirelessly and systematically and need to understand the character of the outputs. For example, how reliable are they? Or, how “competent” was the system that produced them? Of course, there are also questions of how human resourcing (with the experience and knowledge that comes with it) and approaches to automated detection are balanced, a balance that will vary according to the purpose at hand. Nevertheless, for scaling to achieve large area coverage, there is no question that a broad umbrella of automated detection approaches is a key component of a developing survey toolkit.

5 Discussion: Towards Expedited, Systematic and Explicit Heritage Data

This paper has presented questions and preliminary results from a research and development project in Scotland that is exploring methods to create extensive systematic heritage data in an economical way. The focus here is on the creation of general-purpose heritage data

that can support a range of down-the-line applications including heritage management and research. As such, the emphasis may differ from projects that have a primary focus in research. When exploring approaches to rapid archaeological mapping, a key step is to comprehensively and systematically review survey methods. This is especially important considering dramatic developments in the availability and character of digital datasets and their implications for workflows. Often novel survey methods gradually assimilate into routine practice, but such an approach may not necessarily produce the greatest dividend. For this reason, the implications of proliferating RS data are subject to wholesale review, with the goal of creating an approach and workflows that are fit for use in a Scottish context. Such a review will move forward the best of established practice in a framework of workflows that are designed to make the best of digital data.

5.1 Fitness for purpose

In the framework of this JCAA Special Collection, it is worth stressing the necessity to fit data sources and survey methods (i.e. Levels of Survey) to the purpose at hand – in this case, the economical and rapid creation of large area ‘baseline’ survey. In this context, ALS is certainly a game-changer in a country littered with earthworks, but it is not a panacea. For example, high-resolution ALS cannot penetrate the often-dense Scottish forest canopy, which means that the technique adds little value to exploration of many afforested areas. For large area survey, including both woodland and other areas, this makes ALS an expensive option compared to height data derived from photogrammetry, which may cost 80% less, or to orthophotographs taken in appropriate conditions. Here there is a very basic value for money assessment: for large areas of the country, the cost-effective solution will not always be ALS, and certainly not high-resolution ALS. Indeed, for the type of survey considered in this paper, what is characterised as medium resolution ALS data is entirely suitable and brings with it the benefits of reduced computational and storage demands. This general consideration also applies to ‘detail’, which may often be uncritically regarded as necessarily ‘good’, but which depends on the purpose at hand.

5.2 Detail and scale

In any survey work there is a tension between the level of detail that can be achieved, and the resource required. By framing our development of approaches to a large area or national survey within explicit levels of survey, we hope to address the ambiguity that sometimes attaches to survey scale. This is especially true in a GIS world where scale can vary endlessly with the roll of a mouse wheel. Thus, for example, during desk-based mapping, a 1000x1000m grid is deployed to help interpreters manage their progress, with the grid also used to define a ‘viewing scale’ that iteratively appears to work well in balancing a view of context against too much detail. Interpreters are also recommended not to zoom in beyond roughly 1:1500 to avoid focusing too much on detail, recognising that this may contribute to missing small features. Zooming in too much is also discouraged so that pixilation does not become dominant.

Some framework for working scale is important if large area coverage is to be maintained, and discipline is important given the tendency of some interpreters to favour detail over generalisation. Such an approach needs to bring with it an expectation that some features may be missed, and that more detailed work will almost certainly bring other features to light. Yet, this relationship between how comprehensive a survey may be and its operational scale is not easily quantified. It does, however, highlight the importance of explicit documentation of why and how a survey was conducted.

5.3 Survey outputs: attributable and explicit

Heritage data in Scotland, as elsewhere, suffers from bias and a lack of metadata through which users can assess the utility of datasets for their purposes. For these reasons, the explicit documentation of processes, and attribution of the sources of information is central to data creation. All too often, the processes that have contributed to the creation of survey data are opaque and must be taken on trust, which is a poor basis for informed use of such data. Moreover, good structured documentation of survey data provides the means for *post hoc* analysis of outputs, which can inform the identification of training needs and the development of quality assurance routines.

5.4 Sites and landscapes

The textured landscape view that RS data can provide is challenging the suitability of our record structure. The NRHE originates in a record created by the Ordnance Survey Archaeology Division as a card index to the antiquities shown on printed maps. Unsurprisingly, this had a clear focus on traditional archaeological sites – cairns, forts, and settlements. Over time, as the definition of what constitutes ‘archaeology’ has developed, for example, to include cold war military archaeology (e.g. Cocroft, Thomas & Barnwell 2005) and contemporary archaeology (Graves-Brown, Harrison & Piccini 2013), and as a broader interest in landscape archaeology has become more influential, the traditional definition of a site (and by extension how this maps across to ‘records’ within the NRHE) has become less well-defined. Thus, to a degree, the structure of and processes behind the record have not kept up with recent and current developments.

As an example, the increasing availability of RS data is highlighting the extent of often fragmentary remains of relict land use, such as short lengths of medieval and post-medieval field boundaries protruding beyond the limits of later field systems, scattered blocks of cultivation remains, and large areas of peat cutting and trackways (**Figure 7**). Using the traditional ‘sites and monuments’ framework

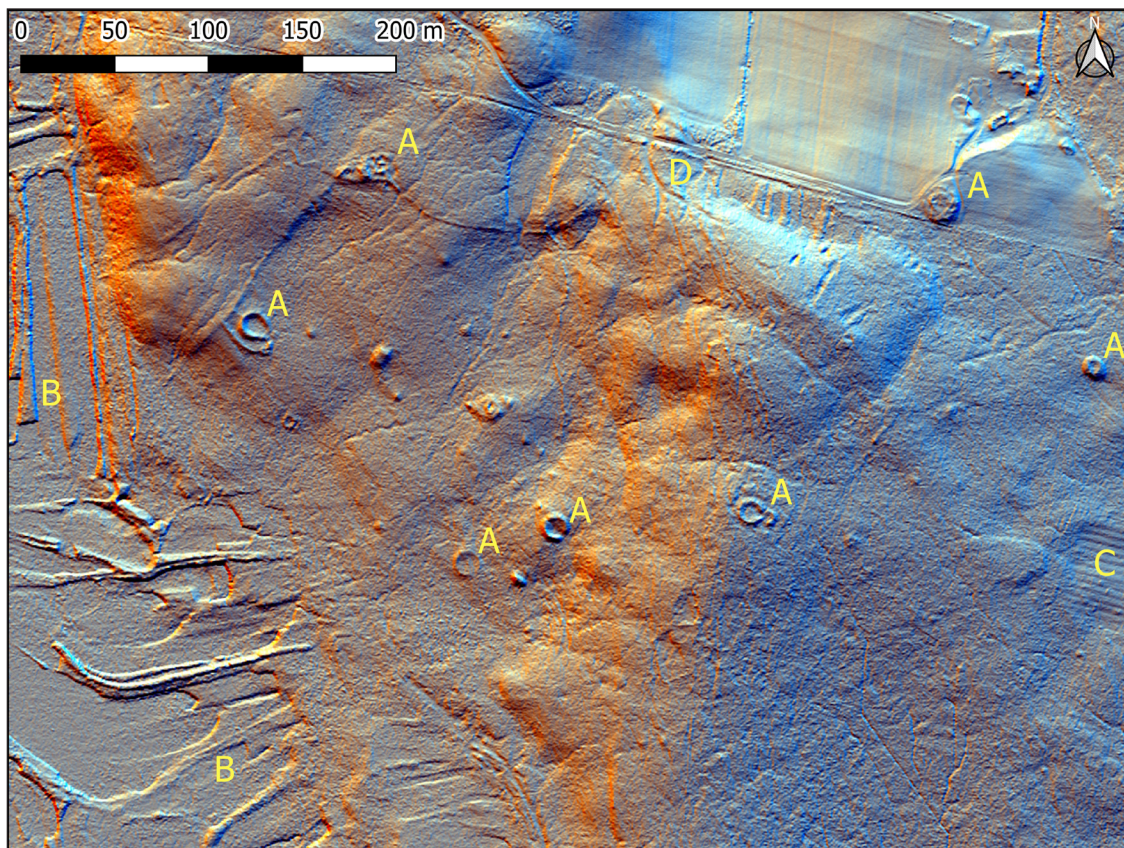


Figure 7: Remote sensed datasets such as ALS are highly effective in documenting landscapes such as at Machrie Moor on Arran. This landscape contains a range of traditional ‘sites’ (**A**) including prehistoric roundhouses and burial cairns of Neolithic and Bronze Age date that sit comfortably within the ‘sites and monuments’ structure of much of the NRHE. However, such an image also records the fragmentary remains of past land use, including peat and turf cutting (**B**), remains of ploughing (**C**) and trackways (**D**). Some, by virtue of the area that they occupy, are recorded in the HLA, but others are not. A challenge to traditional record structure is how to usefully record such remains because the current structure tends to disaggregate them in to discrete areas. This is an issue that may require additional data structure to deal with effectively. © Historic Environment Scotland.

of the NRHE and a Level 2 survey specification for such remains disaggregates these manifestations of extensive past land use, and this seems unhelpful. Rather, in many cases, these types of remains seem best dealt with at the broad-brush scale of a Level 1 survey, such as the Historic Land-use Assessment (HLA, see Section 4.2; HES 2019d). Indeed, the HLA has already mapped some aspects of relict land-use and settlement as seen on aerial photographs and historic maps, but this has a minimum area threshold of one hectare. This creates a tension in how to deal with small areas of cultivation remains and other fragmented remains of past land-use in the NRHE. The default is to create site records, but that disaggregates the record of past land-use in the landscape. Equally, the minimum area criteria of one hectare in the HLA does not help to deal with smaller patches of fragmented remains. This foregrounds the question of how the land-use/landscape textures of the historic environment are best recorded for users – especially as the proliferation of RS data is documenting features such as cultivation remains on a massive scale. One approach to resolving this tension is to create a land-use layer that documents such remains, irrespective of area, bridging the divide between the broad-brush of the HLA and the ‘sites and monuments’ framework of the NRHE, and delivering information to end users in a coherent well-structured way. Resolving this issue is a work in progress. It does, however, highlight the implications of RS datasets for rethinking what we consider a ‘record’ and how we structure historic environment data.

6 Conclusions: Making LiGHT Work of Large Area Survey

This paper has reviewed the implications of RS datasets from the perspective of an archaeological survey function in a national heritage agency. Drawing on preliminary outcomes from the Rapid Archaeological Mapping Programme at Historic Environment Scotland, the value of a systematic consideration of the implications of changing technology and data is highlighted. By bringing us back to the challenges of developing approaches to systematic large area survey, while there is no question that ALS is a game-changer in making lighter work of such survey, it is not a panacea. Rather, alongside other approaches and data sources, its utility requires considered assessment of the purpose at hand, its fitness for that purpose, and the cost-benefits of particular approaches. The wider implications of the proliferation of remotely sensed data and digital workflows are also clear. While such developments have often been gradually assimilated into existing practice, with RAMP, we are finding the root and branch reassessment of our workflows valuable for considering how to create systematic large area archaeological information that does not suffer from a multiplicity of formats.

The volume of available RS data has already outstripped our capacity to engage with it, highlighting the need to develop approaches that exploit the richness of RS data and address the challenges of efficiently extracting information from complex digital sources. This is needed to establish survey methods that are rooted in the reality of the major developments in digital data over the last decade, and that exploit the opportunities they afford. Here

there are broader implications for the need to develop economical approaches to large area surveys, as many other parts of the world do not have the baseline data that Scotland has, including areas of the globe that face ever more rapid and unprecedented landscape change.

Acknowledgements

Our thanks to colleagues for contributing to the various strands of work reported on here, and to Rog Palmer, Robin Turner and the anonymous reviewers for their valuable comments on the text. A special thanks to Anna Cohen, Damian Evans, and Sarah Klassen for their initiative in organising the ‘Paris Dialogue on Archaeological Lidar’ at which aspects of this paper were presented. Publication funds were provided by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 639828).

Competing Interests

The authors have no competing interests to declare.

References

- AOC.** 2015. *A Window on the Hidden Bronze Age Landscape of Caithness*. <http://bronzeagecaithness.aocarchaeology.com/> [Last accessed 13 October 2019].
- Banaszek, Ł, Cowley, D and Middleton, M.** 2018. Towards national archaeological mapping. Assessing source data and methodology – a case study from Scotland. *Geosciences*, 8(8): 272. DOI: <https://doi.org/10.3390/geosciences8080272>
- Bennett, R, Cowley, D and De Laet, V.** 2014. The data explosion: tackling the taboo of automatic feature recognition in the use of airborne survey data for historic environment applications. *Antiquity*, 88(341): 896–905. DOI: <https://doi.org/10.1017/S0003598X00050766>
- Bofinger, J and Hesse, R.** 2011. As far as the laser can reach... Laminar analysis of LiDAR detected structures as a powerful instrument for archaeological heritage management in Baden-Württemberg, Germany. In: Cowley, D (ed.), *Remote Sensing for Archaeological Heritage Management, Proceedings of the 11th EAC Heritage Management Symposium*, Reykjavík, Iceland, 25–27 March 2010. Brussels: Archaeolingua. pp. 161–171.
- Cocroft, W, Thomas, R and Barnwell, P.** 2005. *Cold War: Building for nuclear confrontation 1946–1989*. London: English Heritage.
- Cooper, A and Green, C.** 2015. Embracing the Complexities of ‘Big Data’ in Archaeology: the Case of the English Landscape and Identities Project. *Journal of Archaeological Method and Theory*, 23: 271–304. DOI: <https://doi.org/10.1007/s10816-015-9240-4>
- Cowley, D.** 2003. A case study in the analysis of patterns of aerial reconnaissance in a lowland area of southwest Scotland. *Archaeological Prospection*, 9(4): 255–265. DOI: <https://doi.org/10.1002/arp.199>
- Cowley, D.** 2011. Remote sensing for archaeology and heritage management – site discovery, interpretation and registration. In: Cowley, D (ed.), *Remote Sensing for Archaeological Heritage Management*,

- Proceedings of the 11th EAC Heritage Management Symposium*, Reykjavík, Iceland, 25–27 March 2010. Brussels: Archaeolingua. pp. 45–57. DOI: <https://doi.org/10.1080/00293652.2012.721794>
- Cowley, D.** 2016a. Creating the cropmark archaeological record in East Lothian, southeast Scotland. In: Crellin, R, Fowler, C and Tipping, R (eds.), *Prehistory without Borders: Prehistoric Archaeology of the Tyne-Forth Region*. Oxford: Oxbow. pp. 59–70.
- Cowley, D.** 2016b. What do the patterns mean? Archaeological distributions and bias in survey data. In: Campana, S and Forte, M (eds.), *Digital Methods and Remote Sensing in Archaeology – Archaeology in the Age of Sensing*. New York: Springer. pp. 147–170. DOI: https://doi.org/10.1007/978-3-319-40658-9_7
- Cowley, D.** 2020. Scotland in miniature: automating archaeological survey on Arran. *British Archaeology*, January–February 2020, 44–8.
- Cowley, D** and **López-López, A.** 2017. Developing an approach to national mapping – preliminary work on Scotland in miniature. *AARGNews*, 55: 19–25.
- Cowley, D, Moriarty, C, Geddes, G, Brown, GL, Wade, T** and **Nichol, CJ.** 2018. UAVs in context: archaeological airborne recording in a national body of survey and record. *Drones*, 2(1): 2. DOI: <https://doi.org/10.3390/drones2010002>
- Cowley, D** and **Stichelbaut, B.** 2012. Historic Aerial Photographic Archives for European Archaeology. *European Journal of Archaeology*, 15(2): 217–236. DOI: <https://doi.org/10.1179/1461957112Y.0000000010>
- Crutchley, S.** 2013. Using lidar data – drawing on 10 year's experience at English Heritage. In: Opitz, R and Cowley, D (eds.), *Interpreting archaeological topography: airborne laser scanning, 3D data and interpretation*. Oxford: Oxbow. pp. 136–146. DOI: <https://doi.org/10.2307/j.ctvh1dqdz.16>
- Dunbar, JG.** 1992. The Royal Commission on the Ancient and Historical Monuments of Scotland: the first eighty years. *Transactions of the Ancient Monuments Society*, 36: 13–77.
- Evans, S.** 2019. *Historic England Aerial Investigation & Mapping (formerly National Mapping Programme). Standards Technical Review*. Swindon: Historic England.
- Forestry Commission Scotland.** 2012. *Archaeological Measured Survey on Scotland's national forest estate*. Available at <https://forestry.gov.scot/publications/70-archaeological-measured-survey-on-scotland-s-national-forest-estate/download> [Last accessed 13 October 2019].
- Gallwey, J, Eyre, M, Tonkins, M** and **Coggan, J.** 2019. Bringing Lunar LiDAR Back Down to Earth: Mapping Our Industrial Heritage through Deep Transfer Learning. *Remote Sensing*, 11(17): 1994. DOI: <https://doi.org/10.3390/rs11171994>
- Gannon, A** and **Geddes, G.** 2015. *St Kilda: The Last and Outmost Isle*. Edinburgh: HES.
- Geddes, G.** 2014. Archaeology in the margins – RCAHMS emergency survey in the 1950s. *Proceedings of the Society of Antiquaries of Scotland*, 143: 363–392.
- Graves-Brown, P, Harrison, R** and **Piccini, A.** 2013. Introduction. In: Graves-Brown, P, Harrison, R and Piccini, A (eds.), *The Oxford Handbook of the Archaeology of the Contemporary World*. Oxford: Oxford University Press. pp. 1–23. DOI: <https://doi.org/10.1093/oxfordhb/9780199602001.013.024>
- Halliday, S** and **Ritchie, G.** 1992. The Afforestation Land Survey: Royal Commission on the Ancient and Historical Monuments of Scotland. In: Macinnes, L and Wickham-Jones, CR (eds.), *All Natural Things: Archaeology and the Green Debate*. Oxford: Oxbow Books. pp. 169–175.
- Hesse, R.** 2013. The changing picture of archaeological landscapes: lidar prospection over very large areas as part of a cultural heritage strategy. In: Opitz, R and Cowley, D (eds.), *Interpreting archaeological topography: airborne laser scanning, 3D data and interpretation*. Oxford: Oxbow. pp. 171–183. DOI: <https://doi.org/10.2307/j.ctvh1dqdz.19>
- Historic Environment Scotland Act.** 2014. <http://www.legislation.gov.uk/asp/2014/19/section/2/enacted> [Last accessed 6 February 2020].
- HES.** 2017. *Kraiknish Archaeological Survey*. Available at <https://canmore.org.uk/project/1013793> [Last accessed 13 October 2019].
- HES.** 2019a. *Historic Environment Scotland*. Available at <https://www.historicenvironment.scot/> [Last accessed 13 October 2019].
- HES.** 2019b. *Canmore*. Available at <https://canmore.org.uk/> [Last accessed 13 October 2019].
- HES.** 2019c. *HLAmap*. Available at <https://hlapmap.org.uk/> [Last accessed 27 October 2019].
- HES.** 2019d. *Arran Archaeological Survey*. <https://canmore.org.uk/project/1036806> [Last accessed 28 October 2019].
- Horne, P.** 2011. The English Heritage National Mapping Programme. In: Cowley, D (ed.), *Remote Sensing for Archaeological Heritage Management, Proceedings of the 11th EAC Heritage Management Symposium*, Reykjavík, Iceland, 25–27 March 2010. Brussels: Archaeolingua. pp. 143–151.
- Lambers, K, Verschoof-van der Vaart, W** and **Bourgeois, Q.** 2019. Integrating Remote Sensing, Machine Learning, and Citizen Science in Dutch Archaeological Prospection. *Remote Sensing*, 11(7): 794. DOI: <https://doi.org/10.3390/rs11070794>
- Landauer, J** and **Hesse, R.** 2019. Machine learning for large area archaeological feature detection. Applying transfer learning to airborne lidar data. In *Proceedings of the 24th Conference on Cultural Heritage and New Technologies*. Vienna: CHNT.
- Larsen, C.** 1992. *Sites and Monuments: National Archaeological Records*. Copenhagen: The National Museum of Denmark.
- Maxwell, GS.** 1983. Recent aerial survey in Scotland. In: Maxwell, GS (ed.), *The impact of aerial reconnaissance on archaeology*. York: Council for British Archaeology. pp. 27–40.
- McKeague, P** and **Cowley, D.** 2013. From paper to digital, and point to polygon – the application of GIS in a

- national body of survey and record. *International Journal of Heritage in the Digital Era*, 8: 678–94. DOI: <https://doi.org/10.1260/2047-4970.2.4.677>
- Opitz, R and Cowley, D.** 2013. Interpreting archaeological topography: lasers, 3D data, observation, visualisation and applications. In: Opitz, R and Cowley, D (eds.), *Interpreting archaeological topography: airborne laser scanning, 3D data and interpretation*. Oxford: Oxbow. pp. 1–12. DOI: <https://doi.org/10.2307/j.ctvh1dqdz.6>
- RCAHMS.** 2009. *In the Shadow of Bennachie: A Field Archaeology of Donside, Aberdeenshire*. Edinburgh: RCAHMS.
- Schut, P.** 2009. *Listing Archaeological Sites, Protecting the Historical Landscape*. EAC Occasional Paper 3. Brussels: Archaeolingua.
- Scottish Government.** 2019a. *Scottish Agricultural Census: June 2019*. <https://www.gov.scot/publications/final-results-june-2019-agricultural-census/> [Last accessed 6 February 2020].
- Scottish Government.** 2019b. *Scotland's Forestry Strategy 2019–2029*. <https://www.gov.scot/publications/scotlands-forestry-strategy-20192029/> [Last accessed 6 February 2020].
- Scottish Government.** 2019c. *Protecting Scotland's Future – The Government's Programme for Scotland 2019–20*. Available at <https://www.gov.scot/publications/protecting-scotlands-future-governments-programme-scotland-2019-20/pages/2/> [Last accessed 27 October 2019].
- Sevara, C, Pregesbauer, M, Doneus, M, Verhoeven, G and Trinks, I.** 2016. Pixel versus object – a comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data. *Journal of Archaeological Science: Reports*, 5: 485–498. DOI: <https://doi.org/10.1016/j.jasrep.2015.12.023>
- Trier, ØD, Cowley, DC and Waldeland, AU.** 2018. Using deep neural networks on airborne laser scanning data: Results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland. *Archaeological Prospection*, 26: 165–175. DOI: <https://doi.org/10.1002/arp.1731>
- Verhagen, P, Vossen, I, Groenhuijzen, MR and Joyce, J.** 2016. Now you see them, now you don't: Defining and using a flexible chronology of sites for spatial analysis of Roman settlement in the Dutch river area. *Journal of Archaeological Science: Reports*, 10: 309–321. DOI: <https://doi.org/10.1016/j.jasrep.2016.10.006>
- Verschoof-van der Vaart, WB and Lambers, K.** 2019. Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands. *Journal of Computer Applications in Archaeology*, 2(1): 31–40. DOI: <https://doi.org/10.5334/jcaa.32>
- Watson, F and Dixon, P.** 2018. *A History of Scotland's Landscapes*. Edinburgh: Historic Environment Scotland.
- Welfare, A.** 2011. *Great Crowns of Stone: The Recumbent Stone Circles of Scotland*. Edinburgh: RCAHMS.
- Winton, H and Horne, P.** 2010. National archives for national survey programmes. In: Cowley, D, Standring, R and Abicht, M (eds.), *Landscapes through the Lens: Aerial Photographs and Historic Environment*. Oxford: Oxbow. pp. 7–18.
- Zingman, I, Saupe, D, Penatti, O and Lambers, K.** 2016. Detection of fragmented rectangular enclosures in very high resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8): 4580–4593. DOI: <https://doi.org/10.1109/TGRS.2016.2545919>

How to cite this article: Cowley, D, Banaszek, Ł, Geddes, G, Gannon, A, Middleton, M and Millican, K. 2020. Making LiGHT Work of Large Area Survey? Developing Approaches to Rapid Archaeological Mapping and the Creation of Systematic National-scaled Heritage Data. *Journal of Computer Applications in Archaeology*, 3(1), pp. 109–121. DOI: <https://doi.org/10.5334/jcaa.49>

Submitted: 12 November 2019

Accepted: 24 March 2020

Published: 21 April 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Journal of Computer Applications in Archaeology is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS