Scoping for developing an integrated digital data approach for land-use statistics in Scotland

Prepared for RESAS in the context of the SEFARI Fellowship on the *'Provision of spatially-referenced data relevant to land-based policy evolution in Scotland'*

Zisis Gagkas

The James Hutton Institute

Aberdeen

October 2021

Table of Contents

Summary

This report provides a review of land and environmental monitoring and modelling programmes in England and Wales and draws lessons to build the case for developing a Scottish integrated approach for environmental and land-use statistics, to support evidence-based decision and policy making for the land and agricultural sectors. Review findings and relevant literature are synthesized to provide recommendations for the development of the different components of a Scottish integrated approach to environmental and land-use data, regarding its governance options, utilisation of existing ICT infrastructure, methods, and datasets, alignment and coordination with strategic research programmes and the role of emerging digital and spatial technologies for new data collection and monitoring. This report also provides an indicative timescale for implementation and delivery. Finally, data-related issues are highlighted that pose potential risks to the delivery of the integrated approach for land-use statistics and suggestions are made for minimising their impacts.

1 Introduction

1.1 Background & rationale

The Scottish Government (SG) is advocating a post-Covid 19 green recovery based on green stimulus measures aligned with climate objectives, aiming to tackle the climate crisis and to strengthen the transition to a new social and economic model that will be climate neutral, resilient, sustainable, and inclusive. A green recovery depends greatly at considering Scotland's natural capital and looking at how to build a sustainable economy making the most of these natural assets. Climate change is altering the way that pressures on our environment interact with natural assets and the benefits or services they provide. Scotland has legislated to set a target of net zero greenhouse gas (GHG) emissions by 2045 and similarly the UK by 2050. Land use change has been recognised as a significant element of a holistic strategy to deliver these challenging targets because the way we manage land and plan land-use change plays an important role in meeting targets for climate emission reductions and for the mitigation of and adaptation to climate change impacts.

In this complex and fast evolving policy landscape, there is an increasing need for reliable information regarding the extent and condition of land systems and natural capital assets to support policy needs and objectively evaluate policy options and their impacts (both in time and space). At the same time, new digital capabilities hold more potential than ever before to inform policy making and benefit businesses, communities and individuals and transform the foundational sectors of the economy. Advances in digital technology have led to a rapid increase in the volume of data being captured, curated, and processed on a daily basis. Combining digital technologies and environmental data offers new capabilities for global sustainability and resilience. For example, spatial and other emerging digital technologies, datasets, and infrastructure can play an integral role in supporting the monitoring of land resources and for improving land-based evaluations.

In this context, the development of an integrated approach to collecting, analysing and re-using data for Scotland is proposed, to provide evidence-based support in developing, testing, implementing, and evaluating policy options in the land and agricultural sectors. An integrated digital data approach can support a wide array of SG policies related to land-use and digital technologies, including the Land Use Strategy; Environmental Strategy; Climate Change Plan; Regional Land Use Partnerships (RLUPs); Digital Strategy, and wider UK policies and international initiatives (e.g., UK Geospatial Strategy).

An integral part of the integrated approach for land-use statistics should be a comprehensive, Scotland-wide environmental census of the extent and condition of terrestrial ecosystems. This would be used to establish a baseline, against which progress towards environmental targets and goals, defined by relevant policies, can be measured. This is critical in determining whether the environment is improving, static or deteriorating further. Currently, progress cannot be measured until a baseline is established because current available information from a variety of datasets and sources are incomplete, fragmented, or of partial spatial coverage and do not have a common baseline collection date. The baseline census would aim at identifying and measuring the extent and condition of terrestrial ecosystems and land resources. Conducting the baseline census could form the first step towards implementing an integrated framework for assessing pressures and the associated goods and services provided by terrestrial ecosystems. New datasets of selected indicators generated by the baseline census could fill important data and knowledge gaps and would enable us to pilot approaches to understand how best we can identify the condition of environments and ecosystems, identify detrimental changes and improvements, and to test and improve policy.

Examples of policy drivers that can greatly benefit from the development of an integrated data approach include:

- Evaluation of the climate change mitigation and adaptation potential of nature-based solutions in Scotland, notably peatland restoration and new tree planting, and evaluation of the multiple impacts of policy decisions regarding a range of services delivered by restored peatlands and new woodlands, which include carbon sequestration, flood risk mitigation, biodiversity, access, health and wellbeing, and provision of timber and wood products.
- Evaluation of the effectiveness of post-CAP agri-environmental schemes for farm subsidies, related to the introduction of the 'public money for public goods' principle at the heart of spending decisions.
- Assessment of natural capital assets and evaluation of the effectiveness of policies designed to halt biodiversity loss and decline.

Overall, evidence provided by a solution developed in Scotland to provide an integrated data approach could be used to improve the understanding and management of cross-cutting environmental issues related to climate change adaptation, healthy soils, farming and food security and sustainable development.

1.2 Aims & objectives

The overall aim of this report is to present findings from a review of environmental monitoring and assessment programmes currently in place in England and Wales with the objective to a) identify components of these programmes that are relevant to an integrated data approach in Scotland for land-use statistics, b) draw lessons from the development and operation of these programmes in relation to developing a similar approach in Scotland, and c) synthesise findings to provide recommendations and highlight considerations regarding the development of a Scottish integrated data approach for land-use statistics. The focus of this review is the recently launched Natural Capital and Ecosystem Assessment (NCEA) programme by Defra in England[1] and the Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP) in Wales[2]. Section 2 of this report presents findings from the review regarding a) information related to baseline environmental assessments and environmental monitoring networks, and b) modelling platforms developed to run scenarios for the evaluation of policies in the land sector. Section 3 uses review findings from Section 2, and from relevant literature and consultation with SG analysts, researchers, and stakeholders to a) provide recommendations for developing different components of a Scottish digital data approach for land-use statistics and propose an outline for its implementation, and b) identify potential risks for the delivery of the proposed integrated approach related to issues of data quality, data standards and spatial infrastructure.

[1] https://www.edie.net/news/11/Defra-unveils-natural-capital--assessment--funding-as-post-Brexit-green-recovery-shapes-up

[2] https://erammp.wales/en

2 Review of UK systems

2.1 Overview

The NCEA is a new Defra Group evidence programme which, in the long term, aims to deliver high quality national and local evidence to assess the state and condition of biodiversity, ecosystems, and natural capital assets across terrestrial, freshwater, and marine environments in England. NCEA pilots approaches to understand how best to assess environmental conditions, calculate the value of interventions and to develop, test, implement and evaluate policy. An example of a current pilot project is the England Peat Map (EPM), which is a Defra 'Nature for Climate' funded project within the NCEA Programme, to deliver new improved mapping of England's peat resource. The mapped outputs, derived from a combination of field survey, earth observation, and modelling, will describe the extent, depth, and condition of peat across the country. This information will be used to support a range of purposes including estimating carbon stocks and GHG emissions from peat, targeting peat restoration activities and wider nature recovery.

Specific information for planned pilot NCEA projects is not always readily available. Hence, for the purpose of this review, information published from the Natural Capital Committee (NCC) were used, whose recommendations have been used to establish the case for the NCEA programme and underpin the development of its core components. There is a particular focus on NCC recommendations for an environmental census of the stock of natural capital assets for England[3], which are highly relevant to making the case for taking an integrated approach for land-use data in Scotland.

Similarly, the overall aim of ERAMMP is to deliver a programme of environmental monitoring and modelling which collects data across the Welsh landscape and links any changes to their impacts on a wide range of benefits including their economic consequences. The rational for setting up ERAMMP came from the realisation by the Welsh Government (WG) that the environment supports significant economic sectors, including agriculture, fisheries, tourism, and forestry, that are of importance to other policy areas including health and well-being, energy, and infrastructure. Therefore, WG decided that it was necessary to build a robust monitoring and modelling programme in order to develop policies that build social, economic and environmental resilience and to evaluate programme implementation.

In this context, ERAMMP's main objective is to build a coherent and strategic view of interactions at the landscape or sectoral scale, through large-scale and linked-up modelling.

[3] Natural Capital Committee (NCC), 2019. The Natural Capital Committee's advice on an environmental baseline census of natural capital stocks: essential foundation for the government's 25 Year Environment Plan: https://www.gov.uk/government/publications/natural-capital-committee-advice-on-developing-an-environmental-baseline-census

ERAMMP has been operational since 2018 and the programme is designed not only to be a key source of data for future editions of the State of Natural Resources Report (SoNaRR), but to also undertake modelling, such as for post-Brexit impacts on the farming sector, and for the design and evaluation of programmes delivering to WG's Natural Resources Policy.

2.2 Environmental Monitoring & Modelling

2.2.1 Review & integration of existing monitoring datasets

NCC recommends that previous large-scale environmental monitoring programmes should be reviewed for lessons learned ahead of designing the baseline environmental census for England. For this reason, NCC undertook a high-level review of existing datasets to consider their suitability for the baseline measurement of the stock of natural capital assets, which revealed a distinct lack of robust baseline against which to assess changes in the environment. NCC recommends filling knowledge gaps using information from the Countryside Survey (CS)[4], which was a GB-wide 'audit' of the natural resources of the UK's countryside conducted in 1978, 1998 and 2007 by the Centre of Ecology and Hydrology (CEH). For example, NCC proposes using measurements from the CS's soil component (e.g., carbon content, bulk density, pH, nitrogen, Olsen P) to complement the lack of comprehensive datasets that can be used to assess the condition of soils in England. This approach requires significant effort for standardising and harmonising datasets from different sources but provides the advantage of building a baseline that can be used to detect temporal trends and changes in the extent and condition of natural capital assets.

A similar approach to the one recommended by the NCC has been adopted by ERAMMP, which builds on the CS and the Glastir Monitoring and Evaluation Programme (GMEP). The GMEP was commissioned to establish a baseline of environmental condition at the start of the Glastir Scheme used to deliver payments for environmental goods. Land survey locations in GMEP were selected within 300 1km^2 squares selected randomly to represent Welsh landscape, which provided good coverage of all major land classes defined by characteristics such as climate, geology, and topography. This approach ensured that the survey captured a representative set of the major farm, woodland and land managed for habitat in Wales, as these land uses are closely linked to these fundamental landscape characteristics. The survey approach also recognised that natural resources are inter-dependent and impact on each other by co-locating many measurements within the same 1 km^2 survey squares.

Field survey methods used were those previously developed for CS; this approach was taken to allow for the detection of long-term trends across Wales due to a wide range of pressures such as fertiliser price, legacy of past land management schemes and climate change. Additional methods for the surveys done in the GMEP project were developed to capture co-

[4] https://countrysidesurvey.org.uk

located baseline data on the diversity of birds and pollinators, landscape quality, status and threats of cultural features and public rights of way. Approximately half of all data collected by GMEP was representative of land inside and half outside of Glastir to ensure there was sufficient data to provide a backdrop against which Glastir impact could be evaluated.

2.2.2 Scope of baseline monitoring

Regarding new environmental data collection, a key part of ERAMMP is the National Field Survey (NFS) in Wales, which is scheduled to run in 2021/2022. NFS is intended to be an annual rolling snapshot aiming to map the condition of the natural environment across Wales and to provide information for the evaluation of the Glastir land management scheme. This work will involve re-surveying 130 of the 300 1km^2 squares that were sampled during the GMEP (2012-2016) when baseline data was collected. Surveyors will record many aspects of the Welsh countryside including its habitats, woodland, plant and animal species, freshwater (headwater streams and ponds), soils and soil erosion, cultural features, and footpath condition. An interesting aspect of ERAMMP's NFS is that it uses flexible and adaptable sampling designs for different natural capital assets within the same 1 km^2 squares, e.g., for vegetation sampling the survey uses a combination of random and targeted (such as Priority Habitats) sampling to ensure that it captures the most sensitive or rare components of the landscape or environment.

The NCC proposes a comprehensive baseline environmental census for England that is not only limited to land or terrestrial ecosystems but covers all types of natural capital assets. However, because of its broad scope, the NCC proposes that this environmental census is based on the integration of existing datasets, while the focus of new data collection should be to fill data and knowledge gaps. Their proposal is to allocate new data collection across eight themes with a focus on collection by asset type, with suggested themes covering: atmosphere; freshwater; soils; ecological communities; land and coasts; species; urban natural capital assets, and oceans. NCC recommends that new data collection should be targeted on data gaps, especially for soils, marine ecosystems, biodiversity, ecological communities, and urban areas. These existing data gaps can be supplemented with new data, including data collected by the public, and should also allow for new data capture resulting from the use of emerging technologies (e.g., Sentinel satellites and Landsat imagery, and smart phone apps). For this reason, NCC recommends that new data collection should employ methodologies which are relatively simple to use and yet provide robust data.

2.2.3 Spatial coverage

As mentioned previously, ERAMMP has opted for a regular grid approach comprising of 1km^2 grid squares building on the same network used by the CS and GMEP programmes, which has the advantage of providing baseline information for the same locations that can be used to identify spatial and temporal changes or trends. However, as the NCC points out, choosing a reasonable spatial scale to record natural capital assets or habitat conditions is one of the biggest challenges for every environmental census because measurements taken need to be detailed enough to a) identify local pressures and provide local assessments of important environmental features, and b) be scaled-up to provide a national picture. The different sampling designs employed by ERAMMP surveys within the same 1 km^2 grid squares seem to have been designed to provide the capacity for both local and national assessments. Furthermore, the NCC stresses that working at the national scale alone may not always be appropriate for an environmental baseline census of assets, for example, working at smaller land units as at catchment level may be more appropriate for new data collection. Therefore, the NCC's recommendation is that measurements should follow an agreed full spatial coverage for the different asset types across all of England and not just focus on priority areas, and that the scale used for individual measurements should enable a systems-based and integrated approach to be taken in the data analysis.

2.2.4 Governance & Coordination

The NCC has recognised that clear leadership is needed to gather good quality, accurate data. It considered several approaches for coordinating data collection and integration of datasets and advised the UK Government to consider a single governing body or group being charged with coordinating the census, while Government agencies could be responsible for delivering aspects of the NCEA programme. For example, currently Natural England delivers elements of the terrestrial and marine parts of the NCEA Pilots building on skills and capacity in earth observation, data management, analysis and modelling and specialist expertise (habitats, marine, landscape, soils, evaluation, natural capital etc).

NCC's recommendations are similar to the approach adopted by ERAMMP, which involves a large consortium of partners that makes best use of existing and ongoing activities across the monitoring and modelling community. The ERAMMP consortium consists of the WG and Natural Resources Wales (NRW) and a wide array of external project partners (research institutes, Universities, public bodies, consultancies and industry) led by the UK CEH[5]. The programme also benefits from aligned funding from CEH.

The governance and leadership aspect of ERAMMP was explored in the 'Future Options' project (2016)[6], which explored options for an integrated Natural Resources Monitoring Framework for Wales and was the precursor to ERAMMP. Findings from this project

---

[5] Project partners include: ABPMer, ADAS, Bangor University, Brecon Beacons National Park Authority, British Geological Survey, British Trust for Ornithology, Cardiff University, Cranfield University, eftec, Forest Research, Institute for European Environmental Policy, National Botanic Garden of Wales, National Trust, Office of National Statistics, Pembrokeshire Coast National Park Authority, Public Health Wales, Ricardo, Snowdonia National Park Authority, Staffordshire University and Swansea University.

[6] Emmett, B.A., Bell, C., Chadwick, D., Cheffings, C., Henrys, P., Prosser, H., Siriwardena, G., Smart, S., Williams, B., (2016) Options for a New Integrated Natural Resource Monitoring Framework for Wales; Phase 1, Report Summary to Welsh Government (Contract reference: C147/2010/11; Agreed Additional Work Requirement Dated 8th March 2016). NERC/Centre for Ecology & Hydrology (NERC CEH Project: NEC05945): https://erammp.wales/en/r-futureoptions

recognised the need for increased benefits with reduced overall costs to be realised through greater partnership working which exploits the joint resources and expertise across WG, NRW, industry, third sector and research sectors in Wales. Recommendations from the 'Future Options' project were eventually accepted by WG, which led to the creation of ERAMMP and of the respective consortium of project partners.

2.2.5 Cost-effectiveness

The NCC has recognised from an early stage that building a credible baseline of natural capital assets for the whole of England will require the appropriate level of investment from the government. NCC has advocated that there will most likely be modest costs associated with building a baseline of natural capital assets that are related to accessing existing datasets (including the need to purchase intellectual property rights), collecting data where there are gaps, supporting the necessary governance group and creating or adapting a database to store the collected information. They have also highlighted that the baseline census has the potential to address multiple environmental data needs across a range of policy areas, which should increase the cost-effectiveness of the programme[7]. ERAMMP has been built around a similar approach, by trying to follow the 'collect once - reuse often' principle and by being a key source of data and evidence underpinning a variety of policies within the WG.

NCC offers certain recommendations for designing the baseline census (and the NCEA programme overall) in a way that minimises costs and considers environmental data needs across the (UK) Government. These include integrating and 'synchronising' existing datasets and establishing clear leadership to ensure a joined-up approach in environmental asset data collection, thereby preventing the current (what the NCC considers) large-scale duplication and waste of public funds in England. They also suggest collecting new data only for filling the gaps, incurring only incremental costs in doing so by making the most use of new technologies such as Earth Observation, drones, and Artificial Intelligence.

2.2.6 Citizen Science

The NCC's view is that a citizen science element to developing the environmental baseline is an incredibly powerful way for key stakeholders and the public to engage with the environment and enable large numbers of citizens to be involved in some aspect of the collection and analysis of environmental data. Citizen science is also an excellent way to ground truth new technologies used to collect environmental data. Ground-truthing offers a means to increase confidence in, and possibly add granularity to, data collected over a wider scale or using imprecise assessment tools. For example, citizens could be asked to verify satellite identified habitats using mobile apps with enabled geotagging or could use censors to monitor local air quality to supplement national database datasets.

---

[7] Natural Capital Committee advice on government's 25 Year Environment Plan and progress reports: https://www.gov.uk/government/publications/natural-capital-committee-advice-on-governments-25-year-environment-plan

2.2.7 Modelling

This section focuses on the modelling capabilities of ERAMMP, an operational programme compared to modelling in the context of NCEA Pilot programmes, which is largely under development.

GMEP models were previously used to explore the likely outcome of a range of Glastir options at a national scale for climate change mitigation, water quality and biodiversity. Within ERAMMP, a more ambitious programme of modelling is planned that uses advances in monitoring, modelling, and decision support tools to assist a wide range of policy portfolios related to natural resource management. This modelling framework consists of two components:

- 'Quick Start' which explores the environmental consequences of potential responses by farmers and other land managers to different trade deals from the WG EU exit policy team. These responses are converted to land-use change across Wales using a rule base. Models for climate mitigation, recreation, biodiversity, water quality, carbon storage and air quality for public health are then run. These model outputs are then presented both as changes to natural resources and changes to 'public goods'. These in turn can often be converted to changes in economic values.
- An Integrated Modelling Platform (IMP) is also being developed which is directly driven by external socio-economic drivers, considers feedbacks between process models and includes climate sensitivity. Models and outputs are as for the Quick Start modelling but also include crop-yield and profitability. A web-based user interface allows for different options to be explored, including the impacts of different public payment scheme options and climate change scenarios, and their associated environmental and economic consequences.

The IMP comprises a chain of specialised, state-of-the-art models developed by the research consortium and customised (as best as possible) with Welsh data. IMP has been developed as an integrated system of 11 inter-connected models that have been linked together by establishing data-flows between models across a model chain (Figure 1). These models cover agriculture, forestry, land use allocation, biodiversity, and a range of ecosystem services (including water quality, air quality, greenhouse gas emissions/carbon sequestration) and their valuation. This integrated approach recognises that policy effects in one sector have indirect effects in other sectors. In this way, the IMP explicitly accounts for biophysical and socio-economic interactions between sectors. The IMP operates at various spatial resolutions depending on what scale is most appropriate for the indicator being simulated (e.g., sub-farm, farm, catchment). The finest spatial resolution that is used for simulating farm type and land use transitions is the Decision Making Unit (DMU), which is sub-farm scale (often field-scale) defined as a managerially homogenous cluster of soil type, rainfall, and land cover. The modelling outputs are generally presented graphically on maps, while summary level data is available in addition to analysis at sub-national level.

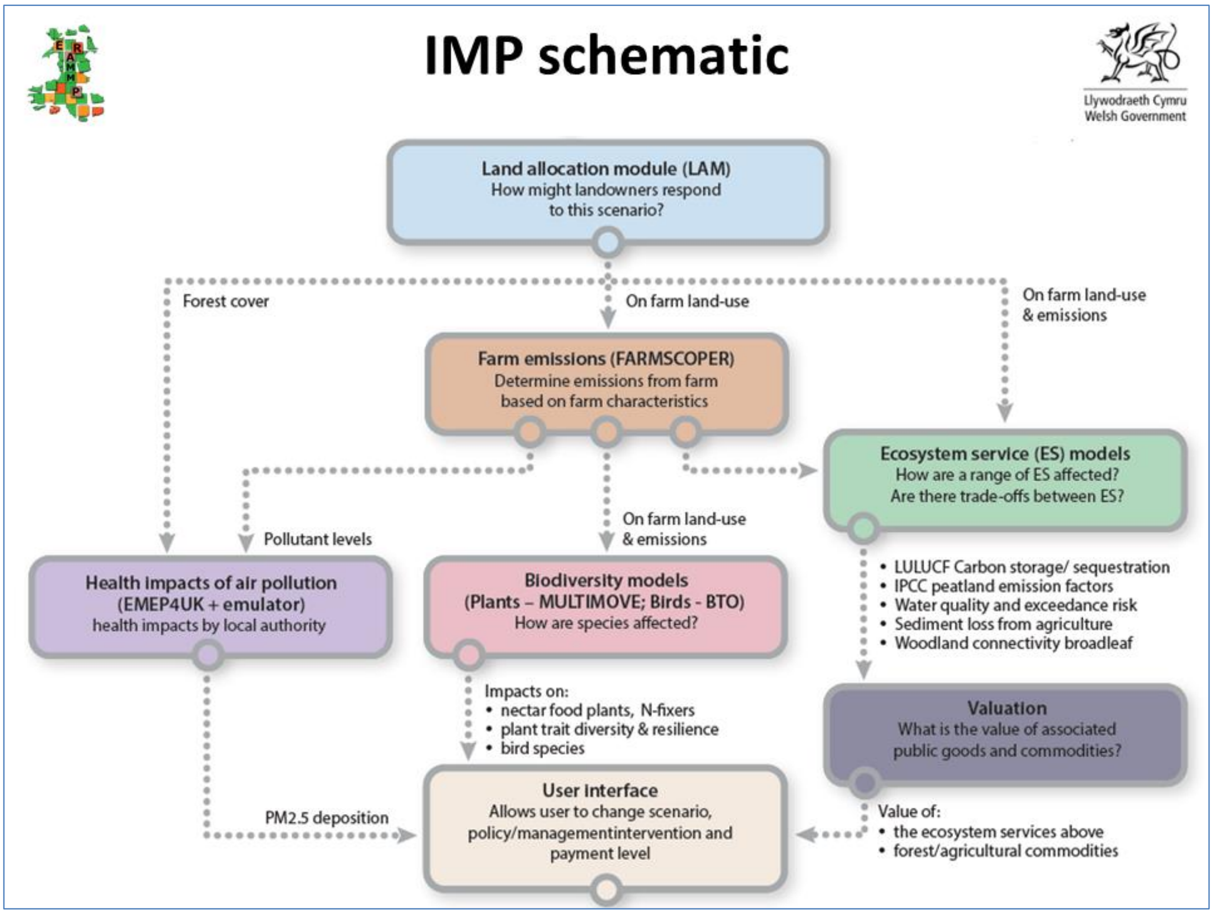


Figure 1. Schematic of the ERAMMP IMP showing linkages between inter-connected models[8].

Below, a quick description is given of three (3) examples of ERAMMP modelling conducted to answer policy questions and evaluate policy scenarios that are also of Scottish relevance, related to EU exit impacts and evaluation of programmes delivering to the natural resources policy:

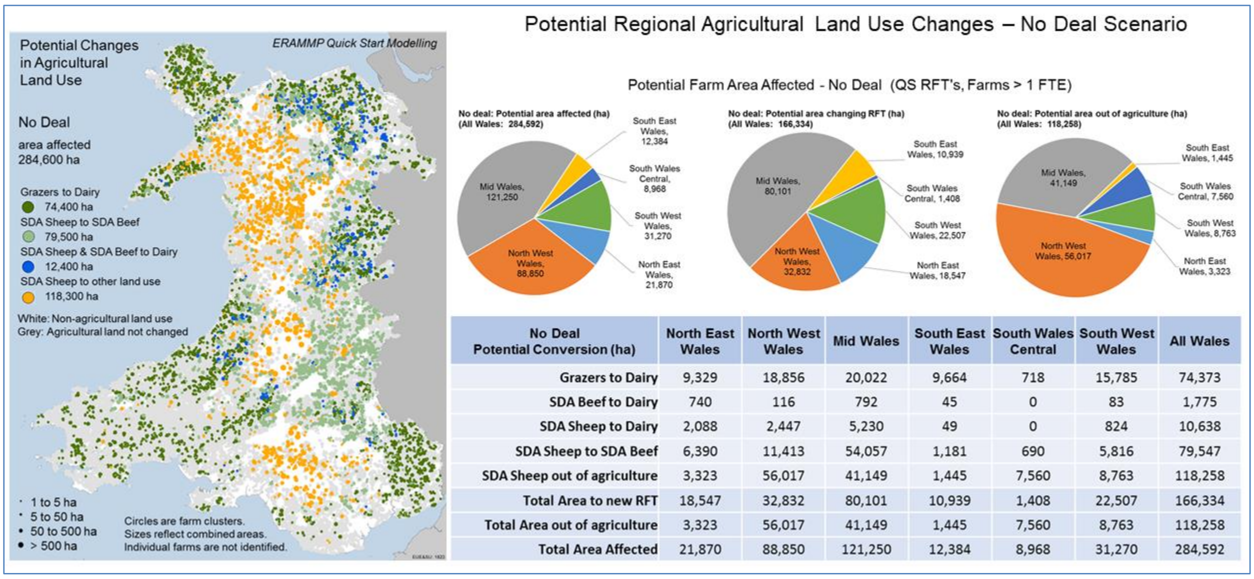
- In 2019 ERAMMP was asked to undertake rapid modelling to explore the potential impacts of different Brexit trade deals on the agricultural sector and other environment outcomes[9]. Three Brexit trade scenarios were developed by WG: i) EU Deal (EU-based free trade agreement); ii) No Deal (WTO rules apply); and iii) Multilateral Free Trade Agreements (MFTA). The Evidence and Scenarios Roundtable Sub-Working Group translated the trade scenarios into potential shifts within and between the principal 'Grazing Livestock' sectors in Wales (dairy, beef and sheep) and the principal 'Small Sectors' livestock producers (commercial pork and poultry) in response to changing market demand for dairy and meat products. The ERAMMP 'Quick Start' programme undertaken by the research consortium then converted the

[8] ERAMMP Report 42: IMP overview: https://erammp.wales/en/imp-models

[9] Cosby, B.J., Thomas, A., Emmett, B.A., et al. (2019) ERAMMP Report 12: QuickStart-1. Report to Welsh Government (Contract C210/2016/2017) (CEH NEC06297): https://erammp.wales/en/reactive-modelling

potential livestock shifts into the potential changes in agricultural land-use needed to manage and support the livestock shifts (see Figure 2).

- The Sustainable Farm Scheme (SFS) Evidence Pack Review[10] is a set of reports and technical annexes based on ERAMMP analysis and published in 2019 and 2020. They summarise current academic thinking on a range of agricultural and environmental specialist topics and were created to help inform the WG as it seeks views on proposals to support farmers after Brexit.
- The National Forest Evidence Review[11] was commissioned by WG from ERAMMP to provide key evidence of potential benefits and disbenefits of woodland creation, woodland expansion and managing undermanaged woodland, to provide an evidence base to inform the development of a National Forest for Wales.



No Deal Potential Conversion (ha)	North East Wales	North West Wales	Mid Wales	South East Wales	South Wales Central	South West Wales	All Wales
Grazers to Dairy	9,329	18,856	20,022	9,664	718	15,785	74,373
SDA Beef to Dairy	740	116	792	45	0	83	1,775
SDA Sheep to Dairy	2,088	2,447	5,230	49	0	824	10,638
SDA Sheep to SDA Beef	6,390	11,413	54,057	1,181	690	5,816	79,547
SDA Sheep out of agriculture	3,323	56,017	41,149	1,445	7,560	8,763	118,258
Total Area to new RFT	18,547	32,832	80,101	10,939	1,408	22,507	166,334
Total Area out of agriculture	3,323	56,017	41,149	1,445	7,560	8,763	118,258
Total Area Affected	21,870	88,850	121,250	12,384	8,968	31,270	284,592

Figure 2 Potential regional agricultural land use changes for the ‘No Deal’ EU exit scenario based on ERAMMP’s ‘Quick Start’ modelling.

[10] Emmett, B.A. et al. (2019). Report-10A: Integrated Analysis. ERAMMP Report to Welsh Government (Contract C210/2016/2017) (CEH NEC06297): https://erammp.wales/en/r-sfs-evidence-pack

[11] Beauchamp, K., et al. (2020). ERAMMP Report-32: National Forest in Wales - Evidence Review. Report to Welsh Government (Contract C210/2016/2017) (UKCEH 06297): https://erammp.wales/en/r-forest-evidence

3 Recommendations

3.1 Recommendations for taking an integrated approach to data development

3.1.1 Governance

It is crucial that a governance group or coordination board is established to lead and coordinate the overall effort for developing an integrated digital data approach, that is representative of both evidence providers and users. Different governance options and opportunities for Scotland could be explored by an equivalent to WG's 'Future Options' project that could incorporate findings from workshops and extensive consultation between SG and relevant evidence providers, data users and stakeholders.

The overall aim of the governance body should be to deliver an adaptive approach to monitoring, increase efficiencies, improve partnership working, and help guide future management decisions. For this reason, the governance body should be responsible for advising on the optimisation and targeting of the collective survey, monitoring, analytical and interpretation resources in Scotland. Based on NCC's recommendations for England, the governance body could be tasked with a) Defining the assets or land indicators to be included in the integrated data approach; b) Defining the body/organisation responsible for each asset or set of indicators; c) Defining the scale at which information about each asset will be collected and/or analysed; d) Organising collection and analysis activities (e.g. who will be responsible for collecting information for which assets/indicators); and e) Ensuring that an integrated systems-based analysis is achieved.

We also need to highlight the great potential and existing capacity that exists in Scotland for forming a partnership, like ERRAMP's in Wales, that can be used for designing and implementing an integrated approach for land-use statistics in Scotland:

- SG Divisions and Units: analysts, scientific advisors and policy makers and managers from RESAS and the Rural Payments & Inspections Division (RPID).
- National agencies: NatureScot, SEPA, Forestry and Land Scotland, Scottish Forestry, and Centres of Expertise (CoE) (Climate Change, Waters and Biodiversity (*forthcoming*)).
- Research in the land and agricultural sectors: SEFARI (Scottish Environment, Food and Agriculture Research Institutes), Scottish academia/Universities.
- Private sector: environmental consultancies and earth-observation (EO) and digital data innovation industry.

Realising greater partnership and collaborative working, which aims to exploit the joint resources and expertise across SG and national agencies, SEFARI, academia and industry sectors in Scotland could deliver increased benefits with reduced overall costs and make more effective use of people and funding.

3.1.2 Strategic alignment

As proposed earlier, the purpose of an integrated digital data approach should be to provide robust evidence to support decision and policy making in the land and agricultural sectors in Scotland. To achieve this, the integrated approach would need to align with the SG's strategic vision and policy priorities about land-use in Scotland. For example, developing the integrated approach needs to consider that land-use and land-use change decisions need to be fully integrated and joined-up across sectors and scales, and that these decisions need to take full account of the wider social, economic and environmental benefits which come from land use. In addition, it is important to highlight that land-use decisions need to consider multiple objectives, e.g., changes in land-use would need to be optimised to reduce GHG emissions alongside meeting various environmental goals. Therefore, the new framework needs to be adaptive and responsive to policy priorities and emerging risks, whilst maintaining a systematic approach towards monitoring the extent and condition of land resources and developing the baseline.

In addition, the integrated approach would greatly benefit by aligning with the next Environment, Natural Resources and Agriculture Strategic Research Programme (SRP 2022-2027) funded by RESAS/SG, which has the objective to "*support research that is relevant, respected and responsive to Scotland's environment, communities, its people and to the rural economy*"[12]. The SRP is a large-scale, multidisciplinary programme with a budget of around £48 million a year that covers a wide spectrum of topics including: Plant and animal health; Sustainable food system and supply; Human impacts on the environment; Natural resources; and Rural Futures. Research topics of particular relevance to an integrated digital data approach for land-use statistics are Land Use (incl. mapping), Climate Change, Agricultural GHGs and Large-Scale Modelling (under Theme C: Human Impacts on the Environment) and Biodiversity, Natural Capital, and Soils (under Theme D: Natural Resources). Alignment and coordination between the integrated approach and SRP could facilitate effective synergies and co-benefits (including improving cost-effectiveness) and presents a unique opportunity to use research findings from the SRP to support the development and implementation of the integrated approach.

An example of the opportunities and benefits that could arise from a potential alignment between the integrated digital data approach and SRP is the proposed development of a new Soils Monitoring Network under the Large-Scale Modelling topic. Monitoring and assessment of soils is inherently linked with the monitoring and assessment of land resources because soil properties and functions provide supporting services that underpin the functioning,

[12] RESAS Invitation Tender for Grant Funding – Strategic Research Programme 2020-2027 (Restricted access)

condition, and provision of services from habitats and ecosystems. For example, the success of peatland restoration projects, which play a crucial part in climate change mitigation and for meeting climatic targets, depends greatly on restoring soil wetness characteristics, while recent studies have shown that soil properties need to be carefully considered when planning the creation of new woodland for optimising the potential for carbon sequestration from new tree planting[13]. In addition, international mechanisms such as the FAO Global Soil Partnership (GSP)[14] promote the sustainable management of soil resources to support the provision of essential ecosystem services towards food security and improved nutrition, and climate change adaptation and mitigation. The systematic soil monitoring framework for Scotland is proposed to focus initially on vulnerable soils (including peatland soils) and will be based on best use of available and novel data and by using robust methods to enhance interoperability between existing datasets collected previously for different purposes. As with recommendations for the integrated approach, the soils monitoring framework aims to be flexible enough and be readily adaptable to act as a resource for helping to address emerging policy and management questions, and should make use of feasible scenarios for e.g., land use and climate change, and also link to work on soil indicators and metrics. Therefore, it is evident that co-development of the integrated data approach for land-use statistics, especially its baseline component, and soil monitoring programmes could create great co-benefits for decision and policy making regarding the provision of robust and integrated evidence.

3.1.3 Design & development scale

A fundamental decision regarding the implementation and delivery of a Scottish integrated approach for land-use statistics is whether this would be designed and implemented from the start as a framework providing national coverage (as in the case of ERAMMP) or whether it would be more feasible to trial approaches in the context of pilot (i.e., Pathfinder) projects.

The ERAMMP approach is obviously more comprehensive and very attractive but requires significant funding and investment, especially due to differences in geographical extents between Scotland and Wales. Alternatively, Pathfinder projects could be used to explore resourcing models, data capture and data analysis and interpretation activities. For example, they could be used to explore a census pilot where small scale local projects undertake full baselining activities of land resources in a defined area. The initial design of these smaller/local scale projects should ensure that these are not disjoint geographically and that their scale is sufficient to support both the scaling-up and down of the data for multiple uses.

---

[13] Friggens, N.L., Hester, A.J., Mitchell, R.J., Parker, T.C., Subke, J-A., Wookey, P.A. (2020). Tree planting in organic soils does not result in net carbon sequestration on decadal timescales., Global Change Biology. 26, 5178-5188: https://doi.org/10.1111/gcb.15229

[14] http://www.fao.org/global-soil-partnership/en

Overall, setting-up pilot (Pathfinder) projects should support the development of a clear and comprehensive strategy to steer the collection, management, use and dissemination of its data, information, and evidence. To achieve this, there needs to be a clear and transparent approach for selecting these pilot projects to optimise their impact and efficiency and avoid duplication with other similar projects. For example, selection could prioritise pilot projects that can provide evidence needed to fill knowledge gaps that hinder the implementation of high priority policies, and hence increase the likelihood of meeting respective policy targets (i.e., climatic targets). Pathfinder projects could also be aligned with initiatives relevant to the land use policy, such as the Regional Land Use Partnerships (RLUPs) that are currently being piloted to help develop Scotland's approach to land-use in support of the green recovery and transition to net-zero. There is also the potential for developing hybrid approaches, such as using measurements from local projects to supplement national data collection activities for certain land indicators or developing specific modelling tools for certain policy priorities in specific geographical regions and extrapolating the approach and findings nationally at a later stage.

3.1.4 Emerging digital technologies

Emerging digital technologies have led to a rapid increase in the volume of data being captured and processed and have enabled a step-change in global capacity for integrated monitoring, analysis, modelling, and visualisation of the natural environment at potentially transformative spatial and temporal scales. Currently, there is an overabundance of digital technologies (Figure 3) that provide an excellent opportunity to create a digitally-enabled environment through more integrated networks of sensors (*in-situ* and remote sensing based), together with methodologies and tools for assessing, analysing, monitoring, and forecasting the state of the natural environment at higher spatial resolutions and finer temporal scales than previously possible. Building a digital environment has the potential to deliver the capacity to improve the understanding and modelling of long-term environmental change and provide evidence to support both decision-making and operational activities within government departments.

In this context, a recent SEFARI Fellowship on 'Enabling digital social innovations in environmental monitoring in support of Scotland's green recovery'[15] provided a comprehensive review of the components of different digital technologies used in environmental modelling (e.g., geospatial dashboards, satellite and drone remote sensing data, in-situ sensor data) and recent innovations. It also provided recommendations for the digital automation of data pipelines and their central role in the creation, use, and testing of digital environments for environmental monitoring and how it can support a green and technology-led recovery in Scotland.

[15] Macleod, K., Eardley, B., Hallard, M., (2020) Enabling digital social innovations in environmental monitoring in support of Scotland's green recovery and the SDG 2030 agenda: a suggested strategy 2021-2025, SEFARI Fellowship report.

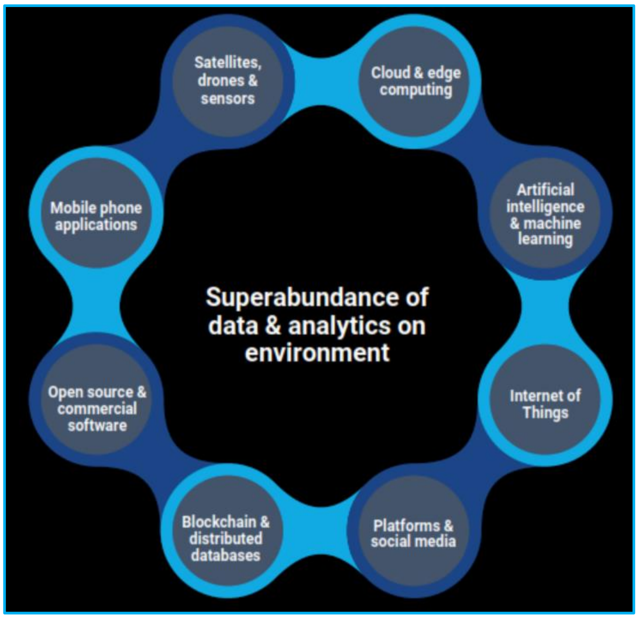


Figure 3 Schematic of digital technologies ecosystem for environmental application[16].

Scotland is exceptionally well-placed to harness the potential generated by emerging digital technologies related to big data geographical analysis, cloud-based computing and remote sensing and earth observation (EO). This is evidenced by the significant expertise in geospatial and EO analysis that exists within the public and private sectors and Scottish academic and research institutions and the plethora of high-quality and innovative spatial datasets and web mapping applications that have been produced in Scotland.

For example, the Edinburgh International Data Facility (EIDF)[17], built and managed by the University of Edinburgh's supercomputing centre, is a high-powered analytics and storage service that supports research and data-driven innovation in the Edinburgh and South-East Scotland region. The EIDF underpins the Data-Driven Innovation (DDI) initiative and plays a key role in establishing Edinburgh as the Space Data Capital of Europe. In particular, the Data SlipStream, hosted on the EIDF is a first step towards creating a system to enable the processing and analysis of large volumes of satellite imagery and other geospatial data. Its aim is to accelerate the speed of space research and product development whilst giving companies and organisations access to the University of Edinburgh's network of scientists and data experts, facilitating the development of new algorithms that produce useful information from satellite data.

In addition, recent developments of new innovative approaches combining spatial and EO technologies, machine learning (ML) models and cloud-based artificial intelligence (AI) platforms have produced new datasets for land resources in Scotland. For example, Space

[16] Jensen, D., (2020) Building a 10-year action agenda for harnessing data and digital transformation to accelerate environmental sustainability UN Environment Programme Coordinator, Digital Transformation Task Force 16 November 2020.
[17] https://ddi.ac.uk/about-us/eidf/

Intelligence Ltd partnered with NatureScot and the Scottish Wildlife Trust in 2019 to leverage latest developments in AI applied to big data from space in order to support the updating of Scotland's Natural Capital Asset Index (NCAI); this partnership led to the production of a high spatial resolution habitat mapping system (freely available online) that can help landowners and policy makers with their land-use decision making, including working out where to restore habitats[18]. In a similar application, analysts from RESAS, working in collaboration with EDINA at the University of Edinburgh and the Joint Nature Conservation Committee (JNCC) used agricultural survey data from 2019, images from the European Space Agency (ESA) Copernicus Satellite Programme and ML algorithms to develop a Scottish Crop Map[19] that recognised different crop types growing in nearly 400,000 fields in Scotland. Moreover, combining imagery from different satellite sensors has been shown to be successful at mapping and monitoring the condition of peatlands in four pilot areas in the UK uplands[20]. Overall, there is currently a variety of projects and initiatives that aim to harness space data for generating new land resources or environmental datasets, such as the 'Mapping the Nation: Towards a National Land-Use Map for Scotland' project from Strathclyde University[21].

The main advantages of EO applications for land resources mapping and monitoring is that they can provide complete, nationwide geographical coverage in a short period of time; they are spatially and temporally consistent, available at a range of spatial and temporal scales and delivered through a variety of means (e.g., satellite, aircraft, drone); a wide range of EO data is freely available and relatively easy to access; and EO data can accurately and easily map and detect changes between core land cover classes over both space and time. However, it is important to stress that the wider use and applicability of EO-derived land products depends greatly on their ground-truthing, as their quality can suffer from misclassifications due to mixed spectral signals or issues with cloud coverage or low predictive modelling accuracies for specific land cover types. Therefore, it is recommended that the design of new (ground-level) surveys and data collection activities organised by an integrated approach for land-use statistics in Scotland should also consider the ground-truthing and validation of EO-derived land resources mapping. There is also potential for the validation of maps produced using EO data from citizen science, for example recording habitat types using smartphone apps with geotagging from hillwalkers in remote areas.

[18] https://www.space-intelligence.com/2021/04/01/weve-just-published-the-first-ever-scotland-wide-high-resolution-habitat-maps-for-free

[19] https://www.gov.scot/news/mapping-scotlands-crops

[20] Williamson, J., Morton, D., Artz, R., Burden, A., Rowland, C., Tornero, L., O'Neill, A., Poggio, L., Khomik, M., Donnelly, D., Evans, C.D. (2018). The role of earth observation in an integrated framework for assessing peatland habitat condition and its impact on greenhouse gas accounting. Final report to Defra, Project number MI07. Centre for Ecology & Hydrology, Bangor.

[21] McGrath, C., Keenan, V. (2021). Mapping the Nation: Towards a National Land-Use Map for Scotland. University of Strathclyde. Online workshop, Nov 12-13 2020: https://pureportal.strath.ac.uk/en/publications/mapping-the-nation-towards-a-national-land-use-map-for-scotland

3.1.5 Mapping evidence needs - Review methods & datasets

Land resources underpin several sectors including agriculture, forestry and tourism, while they also make a significant contribution across government policies including the health and well-being agenda. An integrated approach for land-use statistics should be able to service a wide range of evidence needs across many international, UK and domestic policies and drivers. Therefore, one of the first tasks towards developing this integrated approach should be to map evidence needs across relevant SG departments and national agencies based on main policy drivers and priorities to identify activity overlaps and evidence gaps.

Mapping evidence needs would then need to be supplemented by a comprehensive review of available methods and existing datasets related to land resources and land and agricultural applications in Scotland, including a robust gap analysis, ahead of any decisions being made regarding the design and development of an integrated digital data approach. The review should also include existing monitoring programmes to ensure they are fit for purpose (in terms of sampling, what is being measured, the spatial scale used and affordability) and existing environmental assessments and relevant indicators[22] to inform decisions on the selection and suitability of target indicators. The review should also ensure that a future integrated framework should build-on than rather duplicate already existing work, e.g., new digital platforms of the Land Capability for Agriculture (LCA) and for Forestry (LCF) developed to inform policy for land use adaptation to climate change[23], and interactive models developed to support land use planning such as for woodland expansion and associated multiple benefits[24]. This data review is crucial in the effort to build a baseline assessment of the extent and condition of land resources in Scotland, that can then be supplemented by new data collection and be used to detect and assess long-term changes and trends. As suggested by the NCC, following this approach should be the most cost-effective option for delivering the integrated digital data approach.

Scotland has the advantage of having a wealth of existing information on methods, datasets and models needed to build the integrated data approach for land-use statistics, which (in most cases) is also readily available from various sources. These include previous SRP and CoE projects; surveys and monitoring programmes from Scottish agencies (e.g., peatland depth and condition surveys from Nature.Scot), GB/UK programmes (e.g., Countryside Survey) and EU programmes (e.g., the Land Use/Cover Area frame Survey (LUCAS)); and other nationally and internationally-funded projects implemented by research and academic institutions and industry in Scotland that have produced various outputs/products relevant to an integrated approach for land-use statistics. The comprehensive review would need to identify most of the available information on datasets and methods and filter and select those most relevant

[22] State of the Environment Report https://www.environment.gov.scot/our-environment/state-of-the-environment
[23] https://www.climatexchange.org.uk/research/areas-of-expertise/climate-change-adaptation
[24] RLUP: Woodland Expansion | The James Hutton Institute

for filling knowledge gaps and building the integrated approach. A similar review exercise of multiple sources at a smaller scale has been conducted to identify datasets that contain baseline and/or resurvey data for Scottish soils[25], which identified 41 existing datasets representing a range of scales (local to national), dataset size, and land-use.

Moreover, SG databases and web mapping platforms should be reviewed to identify relevant and suitable datasets that can be included within the integrated approach for land-use statistics, e.g., for supporting the environmental baseline census or as inputs to modelling. Recent work from RESAS analysts has identified and filtered spatial datasets from various sources related to land resources that are stored in SpatialData.gov.scot Metadata Portal[26], Scotland's catalogue of spatial data. This process resulted in the production of two R Shiny apps that were designed to support work for RLUPs: a) a Data Catalogue app developed by David Cook that can be used to easily filter more than 100 available spatial datasets related to land resources in Scotland based on various recorded information organised in a consistent set of fields (e.g., filtering by theme, dataset type, land cover type and provision of related ecosystem services[27]), and b) a Spatial Visualisation app developed by Mehdi Walji that provides a map interface for visualising one or more spatial datasets such as habitat, land-use and soils maps, and can produce land cover statistics for single datasets or for dataset combinations. Both apps provide excellent examples of how reviews of existing datasets can assist in supporting policies related to land-based evaluations.

3.1.6 ICT infrastructure & modelling

Digital data infrastructure and informatics toolsets are needed to support the development of an integrated digital data approach and the delivery of its associated services and outputs (results of statistics analyses, graphs, maps). A recent workshop on environmental infrastructure by UKRI's 'Constructing a Digital Environment Strategic Priorities Fund'[28] programme highlighted requirements for observation, simulation, and data infrastructure. This included the need for distributed networks of environmental sensors, additional forms of autonomous data collection, a cyber-secure infrastructure, data integration across different existing data centres, and citizen science. Another recent workshop of the 'Mapping the Nation: Towards a National Land-Use Map for Scotland' project[29] highlighted the importance of investing on cloud-based solutions for developing land-use analysis and

[25] Neilson, R., Lilly, A., Aitkenhead, M., Artz, R.R.E., Baggaley, N., Giles, M.E., Holland, J., Loades, K., Ovando Pol, P., Rivington, M., Roberts, M., Yeluripati, J. (2020). Measuring the vulnerability of Scottish soils to a changing climate., ClimateXChange Report, 42pp. https://era.ed.ac.uk/handle/1842/37248

[26] https://spatialdata.gov.scot/geonetwork/srv/eng/catalog.search#/home

[27] Translation of datasets to ecological indicators was done by Peter Philips (Head of Natural Capital Land Management SG) and Zisis Gagkas (The James Hutton Institute) using expert judgement.

[28] https://digitalenvironment.org

[29] McGrath, C., Keenan, V. (2021). Mapping the Nation: Towards a National Land-Use Map for Scotland. University of Strathclyde. Online workshop, Nov 12-13 2020: https://pureportal.strath.ac.uk/en/publications/mapping-the-nation-towards-a-national-land-use-map-for-scotland

mapping systems. Key workshop findings were that a cloud solution would allow data analysis to be performed online, avoiding restrictions on individual organisation's hardware capabilities, which would become more important in the future as data becomes more numerous and more detailed. They also suggested having a central system for processing to improve transparency, by allowing version control and processes to be stored, replicated, and understood.

In this context, it is recommended that a future integrated approach for land-use statistics should build on existing infrastructure, such as Scotland's Environment Web (SEWeb)[30], which is a web-portal with a mapping interface that brings together environmental information and data in one place, making it easier to search, discover, analyse, and interpret. All existing and newly collected datasets should be captured and integrated in a centralised, open-source database and web-portal that could be maintained and updated by the governance body. Where it is appropriate to do so, data should be published under the Open Government Licence (OGL) and make it accessible to as wide an audience as possible without restriction. To ensure a successful data integration, datasets related to the integrated approach's context from various data hosting centres within the SG/public sector and other organisations would need to be identified, standardised, and harmonised. In addition, the programme should aim to build capacity within Scotland in terms of skills, expertise, and technology e.g., through training programmes and partnership work with industry, in software and hardware design, data analytics and sensor and sensing technologies for environmental monitoring.

Like ERAMMP, the integrated approach for land-use statistics should include a modelling and scenario testing component to underpin data interpretation, develop a predictive capacity and enable rapid feedback to policy and management. This will support the ongoing development of more robust policies for optimising the social and economic benefits derived from land resources and ecosystems in the long term. Models are critical tools for integrating and upscaling data and for allowing exploration of scenarios for current and future land-use policy, e.g., for understanding the limits and possibilities of various interventions to reduce GHG emissions or for enhancing carbon storage in the agricultural sector. The modelling platform could utilise a combination of several models, provided by consortium partners, and should make effective use of existing and new monitoring data. Also, the modelling platform should be integrated in the centralised web portal and, like ERAMMP, should provide a web-based user interface for running different scenarios.

Moreover, the integrated modelling framework should adopt a systems-based approach to inform land-use and environmental policy making, which utilises holistic approaches to conceptualise problems and characterise interdependencies, trade-offs, synergies, unexpected outcomes, risks, and opportunities from combined interventions. Systems

[30] https://map.environment.gov.scot/sewebmap

models enable evidence syntheses to be integrated with each other, alongside other sources of evidence including, for example, computational models and qualitative data. For example, a systems-based approach has been used to assess the impact of different land interventions on the overall utility of the English uplands and the provision of specific ecosystem services[31]. These approaches can guide evidence-based environmental policy and management and enable policy analysts to understand complex interactions and relationships more easily.

3.1.7 Mapping funding opportunities

Integration of existing datasets, use of existing infrastructure, establishment of clear leadership and alignment and co-development (to a certain extent) with other SG-funded programmes, such as the next SRP 2022-2027 and CoE programmes, should help minimise (at least) the initial costs required for the development of the integrated approach. It is recommended that a list of potential funding opportunities is generated from consultation with SG and stakeholders/project partners during the process of building the case for developing the integrated approach.

For example, the integrated approach could benefit from aligned national, UK and international funding from project partners (public bodies, academia, and industry), which is directly related to the integrated approach's components and applications. Examples of funding mechanisms include the 'Constructing a Digital Environment'[32] programme, under which NERC provides funding for a range of research projects that seek to bring together the 'Digital Environment' community. These projects aim to deliver increased benefit from existing and new sensor networks technology and their associated infrastructure and are exploring methodologies and tools for assessing, analysing, monitoring, and forecasting the state of the natural environment at high spatial and temporal resolutions. Examples of funded projects in Scotland include developing a soil carbon sequestration monitoring system to meet farmer and policy maker information needs[33] and developing a digital environment for water resources to improve the next generation of environmental regulations for waters[34]. In addition, the Bayes Centre, which is the University of Edinburgh's innovation hub for Data Science and AI, aims to act as a catalyst for growth through its funding of the growing Scottish space sector and its support through mentorship and training.

3.1.8 Timescale

Table 1 presents an indicative timescale for the implementation and delivery of the integrated digital data approach and its components consisting of six (6) overlapping Phases within a 5-

[31] Stewart, G.B., Glendell, M., McMorran, R., Troldborg, M., Gagkas, Z., Ovando, P., Roberts, M., Maynard, C., Williams, A., Clay, G., Reed, M.S. (2021) Uplandia: making better policy in complex upland systems., Defra and Natural England Report, 68pp.: https://eprints.ncl.ac.uk/274791

[32] https://digitalenvironment.org

[33] https://digitalenvironment.org/home/digital-environment-projects/#NE/V003259/1

[34] https://digitalenvironment.org/home/digital-environment-projects/#NE/T005564/1

year period. The timescale assumes that setting-up pilot (Pathfinder) projects will be the preferred option for delivering the integrated approach for land-use statistics. The timescale prioritises identifying options and developing agreed recommendations for development (similar to the 'Future Options' project/ precursor to ERAMMP) including setting-up the governance body (Phase 1). A comprehensive review of methods, existing datasets and ICT and spatial infrastructure should be completed in the first year (Phase 2) that will run parallel to and feed into the selection and design of the pilot (Pathfinder) projects (Phase 3). The delivery of the selected Pathfinder projects is scheduled for Years 2 and 3 (Phase 4), while development of the unified online data portal is scheduled for Years 2 and 3 (Phase 5). The framework for the integrated modelling platform should be finalised by Year 5 (Phase 6).

Table 1. Indicative timescale for implementation

Delivery Phases	Year 1	Year 2	Year 3	Year 4	Year 5
Phase 1: Development options & governance	■				
Phase 2: Review (methods, data, infrastructure)	■				
Phase 3: Selection of Pathfinder projects	■				
Phase 4: Delivery of Pathfinder projects		■	■		
Phase 5: Development of online data portal		■	■		
Phase 6: Development of integrated modelling platform				■	■

Phases 1, 2 and 3 are low risk and low cost, and their successful delivery depends on clear leadership, cooperation and coordination within SG and programme partners. The objective of Phase 4 should be to trial and test new data collection designs, test and evaluate models, identify weaknesses in methods and explore extrapolation of project findings at regional and/or national scales. Phase 5 should focus on developing and/or upgrading the web-portal and its functionality and on inputting existing datasets and new data generated from the Pathfinder projects in Phase 4. Phase 6 should focus on the delivery of the integrated modelling platform by developing protocols for the integration and harmonisation of existing models and new models developed in Phase 4 models and their outputs and exploring upscaling options.

3.2 Data concerns & considerations

3.2.1 Overview

Issues related to datasets relevant to land-based evaluations pose risks to the successful delivery of the integrated approach for land-use statistics. From an early stage, the programme would need to define the strategy and governance arrangements for new data capture and storage, management, quality control, sharing and dissemination of new and existing datasets. In this section, potential issues with new and existing datasets are identified and recommendations are suggested for minimising their potential impact on programme delivery. The focus is on spatially-referenced datasets that form the basis for generating land-use statistics, such as new georeferenced data collected from on-ground surveys and geospatial layers (maps) derived from difference sources (e.g., from digitising paper maps or for satellite imagery analysis). Issues covered include duplication and fragmentation, accessibility and licence restrictions and the importance of common data standards and of spatial infrastructure and human capacity.

3.2.2 Data duplication & fragmentation

Modelling and analysis of land-use and land-use change requires accurate spatial data of land resources characteristics. Datasets used to assess the extent and condition of different land cover types (habitats, agriculture, and forestry) in Scotland are developed and maintained by a wide range of scientific and public bodies; examples include Nature.Scot's data on Protected areas, the Habitat Map of Scotland (HabMoS), the National Vegetation Classification (NVC) and Phase 1 survey data, and the recent Scotland Habitat and Land cover maps (SLAM-MAP) produced in partnership with Space Intelligence; UKCEH's Land Cover Maps (LCM); Land Cover Scotland 1988 (LCS88, Macaulay Institute/The James Hutton Institute); various datasets from the Forestry Commission or Forest Research (e.g., the National Forest Inventory, the Caledonian Pinewood Inventory and tree suitability for planting maps); Integrated Administration and Control System (IACS) data managed by SG RPID; and datasets and maps from the Countryside Survey and EU LUCAS and CORINE programmes.

Most of these datasets have been identified and are included in the Data Catalogue and Spatial Visualisation apps produced by RESAS analysts to support work on RLUPs. However, the process of developing those apps revealed that there is significant replication and duplication of datasets along with both spatial and temporal fragmentation. In particular, identified issues included:

- Use of multiple classification systems for mapping the same land cover and/or habitat types: there seems to be a confusion or overlap between land cover and land-use mapping, while habitats are mapped using different systems (e.g., NVC or EUNIS). This means that datasets are not always compatible or directly comparable, and hence they are difficult to harmonise or jointly interpret.
- Use of different methodologies (e.g., ground surveying vs remote sensing) for producing maps at different spatial scales or spatial resolutions and in different dataset format (e.g., vector/polygons vs raster/grid cell): this produces spatial inconsistencies when different spatial layers need to be combined, i.e., boundaries of same land cover or use (e.g., forestry) delineated by different datasets for the same areas do not match-up.
- Partial coverage vs national coverage: some datasets provide detailed mapping within designated areas or at regional or catchment scale whereas others provide coarser mapping at national scale.

- Datasets are often spatially and temporally fragmented and focus on individual land types rather than a more systematic sampling or mapping at a regular interval; this means that they cannot be used to detect temporal change or trends.

The above issues highlight the importance for the integrated approach for land-use statistics to be based on a single, dynamic Land-Use map of national coverage, based on a commonly-agreed classification system that is produced at a high spatial resolution and accuracy and is updated at regular temporal intervals (e.g., annually). This is approach towards a National Land-Use Map for Scotland has been recently explored in a scoping survey and workshop[35], which suggested that land-use mapping (although harder to accurately categorise) should be prioritised over land cover mapping because is much more useful for land-based evaluations and environmental risk assessments.

Probably the most appropriate (and cost-efficient) approach for producing a national land-use map fit for the integrated approach's purposes would be to combine and integrate existing land datasets with EO-derived data and new data collected from surveys, aiming to extract and combine the most detailed available information for different land-use types from each source. For example, detailed crop type information and field boundaries from the Land Parcel Identification Scheme (LPIS) within cultivated areas could be integrated with detailed boundaries of forestry plantations and seminatural habitat mapping at high spatial resolution in more remote and less accessible areas.

For this purpose, significant effort would be needed for the systematisation, harmonisation, and integration of large-scale spatial land-use datasets from a range of sources. This approach has been trialled for case study locations in Scotland's two National Parks and the River Dee catchment in Aberdeenshire, where integrated spatial datasets were generated via systematic unification of different land cover and land-use (e.g., LCS88, LCM and IACS)[36]. The SLAM-MAP provides an excellent example of an EO-based mapping approach that uses a consistent modelling methodology of high classification accuracy that can be applied at different temporal scales enabling the detection of temporal change; for example, SLAM-MAP has produced land cover maps for 2019 and 2020 products that were then used to assess change in land cover at a national scale. In addition, new data collected from surveying could be used for filling knowledge gaps or supplementing or ground-truthing/ validating any EO-based mapping.

---

[35] McGrath, C., Keenan, V. (2021). Mapping the Nation: Towards a National Land-Use Map for Scotland. University of Strathclyde. Online workshop, Nov 12-13 2020: https://pureportal.strath.ac.uk/en/publications/mapping-the-nation-towards-a-national-land-use-map-for-scotland

[36] Hewitt, R., Macleod, C.J.A., Baggio Compagnucci, A., Castellazzi, M., Miller, D.G., Gimona, A., (2018) Maps of land use data and ecosystem services for Scotland: examples applied to the National Parks and Aberdeenshire River Dee. The James Hutton Institute. Report to RESAS, 1.4.3 Objective D Deliverable D3.

3.2.3 Accessibility and licence restrictions

Although all or most land-related datasets and maps produced with SG funding are freely-available under OGL, there are challenges for data sharing between different sectors (public, academia and private) or of certain datasets that are linked to personal data (e.g., RPID datasets linked to crop subsidies). Findings from a recent study that interviewed people from different sectors working with geospatial data and their analysis in Scotland[37] highlighted that there are no straight-forward connections and no basic data sharing agreements implemented between different sectors. Participants in this study also mentioned that the process of arranging data sharing through contracts and licenses is often extremely time-consuming, involving lots of time and resources, and often delaying the start of projects. Similarly, accessing data through request submissions can be a very long and bureaucratic process, while participants from public institutions found data sharing challenging because their IT security systems are very high and are not allowing for the use of conventional data sharing platforms. The study also identified that data generators and collectors generally tend to be protective of their data, whether it is due to commercial reasons, concerns over IP or just fear of being judged because of their data quality.

Hence, it is evident that issues related to licencing and data sharing that restrict data accessibility and use could be problematic for the smooth operation of the integrated approach for land-use statistics, especially if a large number of partners from different sectors was involved in its governance body. To reduce duplication and increase efficiencies, a clear approach to efficient and effective data sharing should be adopted to enable the conversion of data into robust evidence products. Therefore, it is advised that some form of standardised protocols for licence and data sharing purposes are developed at the early stages of the programme, such as standardised license templates for certain types of geospatial datasets, which would be helpful in improving data accessibility and flows.

3.2.4 Importance of data standards & metadata

Adhering to agreed data standards is vital for maximising the reuse of data and the potential for data integration. Data standards are created so that data attributes and associated metadata are exposed, and an understanding of the underlying data structure is made as simple as possible and common across different sources. An example of such data standards is the EU INSPIRE Directive, which aims to enhance the sharing of environmental spatial information and better facilitate public access to spatial information. Data standards are also used to ensure consistency across data, which can be crucial for integrated analysis or presentation of evidence across multiple data layers. In addition, setting agreed data standards for data collection could enable a large amount of planned new data collection to be added to the central open access database.

[37] Scottish Science Advisory Council (SSAC), (2021) Future Landscapes: Report on Geospatial Knowledge. SSAC: [Future Landscapes Report on Geospatial Knowledge.pdf (scottishscience.org.uk)]

Therefore, the integrated approach for land-use statistics should aim at setting agreed data standards for data collection that abide by sets of guidelines for good practice on data management, such as the FAIR[38] principles. The FAIR acronym stands for Findable, Accessible, Interoperable and Reusable and promote data management practises that are based on transparency, reproducibility and reusability and apply to data, metadata and supporting infrastructure (algorithms, tools, and workflows). A recent SEFARI Fellowship on digital social innovations in environmental monitoring[39] highlighted the importance for environmental data to meet the FAIR principles. They proposed a plan for data-information knowledge, based on environmental monitoring data flows, that is findable and accessible to whoever needs it, interoperable with existing and new applications and workflows/data pipelines, and that it is widely reused from individual to national level decision making to improve the climate change, biodiversity loss, and unsustainable use of natural resources challenge situations as part of a green technology-led recovery. In addition, a new SEFARI fellowship with Environmental Standards Scotland (ESS) is expected to provide new insights on standards for environmental data; the Fellowship will inform the establishment of post-Brexit environmental governance arrangements in Scotland and help ESS with developing their approach to monitoring compliance with, and effectiveness of, environmental law. Findings from both SEFARI Fellowships could provide the integrated approach for land-use statistics with feedback and guidance related to setting agreed data standards.

An important aspect of good practice on data management is recording metadata. In a recent study[40], all interviewees working with geospatial data agreed that data standards improve their experience of data use, workflow, productivity and provide confidence on the reliability of the dataset. Metadata remains absolutely essential for discoverability and access and provide benefits for both the data collector and data users in the long term. However, this study also highlighted that writing metadata is one of the main barriers to making data publicly available, particularly if aspiring to follow the INSPIRE Directive, the FAIR principles, or other standards with long documentation, which takes a substantial amount of time to write and might be challenging to understand and to translate into particular actions. In the context of the integrated approach for land-use statistics, providing training to staff from project partners could help with improving the writing metadata or SG and partners could consider whether it is possible to simplify some metadata fields needed by considering the programme's needs and the needs of end users.

[38] GO FAIR initiative: Make your data & services FAIR (go-fair.org)

[39] Kit Macleod, K., Eardley, B., Hallard, M., (2020) Enabling digital social innovations in environmental monitoring in support of Scotland's green recovery and the SDG 2030 agenda: a suggested strategy 2021-2025, SEFARI Fellowship report.

[40] Scottish Science Advisory Council (SSAC), (2021) Future Landscapes: Report on Geospatial Knowledge. SSAC: Future Landscapes Report on Geospatial Knowledge.pdf (scottishscience.org.uk)

3.2.5 Infrastructure & Capacity

In a previous Section we recommended building the integrated approach for land-use statistics on existing web portals and web mapping infrastructure, such as SEPA's SEweb, and making use of existing data hosting platforms and databases within SG. However, if this integrated approach were to replicate the capabilities and functionalities of ERAMMP, then further investment might be needed to upgrade existing web platforms, such as in spatial data infrastructure (SDI) for storing, sharing, viewing, and analysing geospatial and spatially-referenced data over the web, coupled with cloud-based services for the analysis of large volumes of data (e.g., from satellite imagery). Other potential challenges include building protocols and routines that enable efficient data sharing between programme partners and facilitate linkages and inter-connections between different models and the integration and harmonisation of model outputs, as well as deploying advanced cybersecurity measures to protect online systems from information disclosure, theft of or damage to software and digital data. Finally, the SG and project partners should map skills and expertise in their respective institutions to assess whether enough capacity exists for supporting the development of the different components of the integrated approach.

4 Conclusions

Report findings demonstrate the importance of developing an integrated digital approach to collecting, analysing, and re-using data to provide evidence-based support in developing, testing, implementing, and evaluating policy options in the land and agricultural sectors in Scotland. They highlight the need for a comprehensive, Scotland-wide environmental census of the extent and condition of terrestrial ecosystems for establishing a baseline against which progress towards environmental targets and goals can be measured, and for determining whether the environment is improving, static or deteriorating further. The report also highlights the opportunities arising from emerging spatial and other digital technologies to support the development of an integrated approach for assessing, analysing, monitoring, and modelling the state of the natural environment at high spatial resolutions and fine temporal scales.

A wealth of land and other environmental datasets and scientific evidence exist that can be used to support the development of an integrated approach for land-use statistics, by identifying and filling-up knowledge gaps to help build a comprehensive assessment of the baseline condition. Coupled with the high capacity and expertise that exists in Scotland in the environmental and digital technology sectors (Government agencies, Research/Academia, Private sector) and existing infrastructure means that Scotland is well-placed and at a strong starting position for building a new integrated digital data approach for land-use statistics. However, successful delivery of the integrated digital data approach would depend greatly on establishing clear leadership and effective coordination between partners, identifying knowledge gaps, mapping evidence and prioritising policy needs, and identifying shortcomings in analytical resource and existing systems and related infrastructure.

In addition, developing the integrated digital approach should consider that land-use planning and land-use change decisions need to be fully integrated and joined-up across sectors and scales, meaning that the impact of these land decisions on the wider environment also needs to be fully considered. For example, how land is managed affects soil health, surface and groundwater quality, the capacity of habitats to moderate hydrological extremes (floods, droughts) and GHG emissions from land. Therefore, the integrated digital approach needs to identify linkages between sectors and natural capital assets and be flexible, adaptive and responsive to policy priorities and emerging risks related to effects of land-use change on the extent and condition of soils, water resources and on pollutant emissions and air quality.

5 Acknowledgements

This research was funded through the SEFARI Fellowship on the 'Provision of spatially-referenced data relevant to land-based policy evolution in Scotland'. I would like to thank colleagues for their inputs and collaboration in various aspects of the Fellowship and their valuable feedback on this report: Mehdi Walji and David Cook, Fast Stream statisticians in RESAS, Alastair McAlpine and Kerstin Hinds from RESAS, Peter Phillips and Ross Johnston from Natural Capital Land Management SG, Alice Hunter (Head of Regional Land Use Partnerships, SG) and other people from RESAS, RPID and wider SG. I would also like to thank a wide range of colleagues from various stakeholders (SEPA, Nature.Scot, Forestry and Land Scotland) and from the James Hutton Institute who provided their time to discuss about the integrated approach for land-use statistics and give valuable feedback and suggestions.

6 Appendix

The proposed recommendations and considerations presented in Section 3 are summarised in the Table below.

Report Section	Recommendation	Why	How
1. Recommendations for developing an integrated approach to land-use data			
3.1.1	Establish governance body	Clear leadership and effective coordination are needed to develop an integrated approach to land-use statistics	Build partnership from SG/agencies, SEFARI, academia, industry to facilitate collaborative working
3.1.2	Strategic alignment with policy and relevant research	Alignment with the SG's strategic vision and policy priorities about land-use and land-use change in Scotland is needed to develop an approach that is fully integrated and joined-up across sectors and scales	Provide robust evidence to support decision and policy making in the land and agricultural sectors by aligning with strategic (SRP) and other policy-related research (e.g., CoE projects)
3.1.3	Pilot (Pathfinder) projects	Support the development of a clear and comprehensive strategy to steer the collection, management, use and dissemination of data, information, and evidence	Develop transparent project selection criteria based on filling knowledge gaps and helping to meet policy targets
3.1.4	Use of emerging digital technologies	Most cost-effective option for land resources mapping and monitoring at national geographical coverage and at fine spatial and temporal scales	Use existing expertise in spatial and EO/ML/AI technologies to develop innovative applications for land monitoring and mapping
3.1.5	Mapping evidence needs and data and methods review	Integrated approach needs to be able to service a wide range of evidence needs across many international, UK and domestic policies	Conduct comprehensive review of evidence needs and existing methods and datasets to identify knowledge gaps and build baseline assessment
3.1.6	Utilisation of existing ICT infrastructure	Most cost-effective option for building the unified open-access database, web portal and integrated modelling platform	Review specifications and functionalities of existing ICT infrastructure
3.1.7	Mapping funding opportunities	Reduce costs for developing the integrated approach for land-use statistics	Use aligned funding from partners to contribute to components of the integrated approach
2. Recommendations for issues related to data			
3.2.2	National digital Land-Use map	Need for a single, dynamic, regularly updated Land-Use map of national coverage at high spatial resolution to generate accurate land-use statistics	Harmonise and integrate existing land datasets and combine with EO-derived data and new data collected from surveys

3.2.3	Improve data accessibility and sharing	Quick and easy data sharing between partners from different sectors is essential for the operation of the integrated approach for land-use statistics	Develop standardised data licence and data sharing procedures based on programme and user needs
3.2.4	Adoption of data standards	Common data standards are needed to maximise the reuse of data and the potential for data integration	Adopt FAIR principles for environmental data and simplify metadata recording
3.2.5	Upgrading existing ICT infrastructure	The integrated approach for land-use statistics requires a complex framework linking components such as the unified open-access database, web-mapping tools and integrated modelling platform	Invest in upgrading/ improving existing SDI infrastructure, cloud-based services, and cybersecurity measures