Understanding metrics for effective environmental measures under the Agricultural Reform Programme for Scotland

Robin J Pakeman[1], Nikki Baggaley[1], Wakene Negassa[1], Miriam Glendell[1], Cathy Hawes[1], Alison Karley[1], Kenneth Loades[1], Roy Neilson[1], Eric Paterson[1], Marc Stutter[1], Joanne Conington[2], Gemma Miller[2], Philip Skuce[3], Jackie Potts[4]

[1]James Hutton Institute, [2]Scotland's Rural College, [3]Moredun Research Institute, [4]Biomathematics and Statistics Scotland
April 2024

1 Executive summary

To deliver climate change mitigation and adaptation, nature restoration and high quality food production, the Scottish Government produced their vision for agriculture, along with the next steps, to encourage sustainable and regenerative farming in Scotland. A programme of work is underway to reform agricultural payments with a greater emphasis placed on delivering environmental outcomes with a proposed structure of four payment tiers tied to a suite of potential measures that will deliver tangible outcomes.

This study identified the most suitable metrics that could be used to monitor the success of the proposed measures in the agricultural reform programme against environmental outcomes. This includes consideration of cost-effectiveness, practicalities and the skills and capabilities of those tasked with monitoring.

1.1 Findings

We found potential metrics for assessing the success of the measures for all outcomes. Most metrics can already be applied as the methods are available, whilst a small number are under development and could be applied in the near to medium term. These metrics fell into several categories:

- Emissions cannot be measured directly, so we suggest using current farm-level tools to assess GHG emissions, known as carbon audits. A field level, real time GHG emission model is in development as well as a tool for doing this for ammonia.
- Many metrics depend on direct sampling of soil or biodiversity and can't be realistically replaced by proxies or existing data. However, well designed sampling programmes can maximise the efficiency of sampling, e.g. sampling for soil carbon, nutrients, pH and eDNA can be done at the same time.
- The outcomes associated with animal health, nutrition and breeding must be largely monitored through proxy metrics. These are relatively easy to measure and provide useful information directly to the land manager.
- A few metrics, such as pesticide usage data or area under permanent habitat, collected as part of the agricultural census, can be derived from existing data.
- Some of the metrics in development could take advantage of samples/data collected at the start of any monitoring programme (e.g. soil eDNA, acoustic monitoring) and others would come online later (e.g. LIDAR-derived hedge data).
- The measure 'retain traditional cattle' could not be related to the outcomes.
- Deciding on a suitable suite of metrics to assess the benefits of the Agriculture Reform Programme is only one step as there are issues related to design, sample size and data to be considered.

1.2 Recommendations

A full list of suitable metrics for each measure from the Agricultural Reform list of measures is supplied in an accompanying spreadsheet "MeasuresXMetrics.xlsx". The spreadsheet can be filtered to look at what metrics are suitable for each measure, which outcome they relate to, whether the metric is suitable for direct assessment, if it provides additional useful information or if the metric is still in development, whether the metric is suitable against multiple outcomes and who can carry out the monitoring.

Table 1 summarises the spreadsheet by showing which metrics relate to each outcome. The final choice of which metrics to collect will depend on two main factors:

- The availability of resources to carry out any monitoring programme
- The sampling philosophy adopted; whether widespread collection of a few metrics, where data collection could be partly done by land managers, versus a programme designed to give accurate data at the national level by sampling intensely from a representative sample of locations with mainly expert-led sampling.

Combining information on who can do the monitoring and potential likely costs of expert-led monitoring, we suggest the following monitoring philosophy is appropriate:

- All enterprises to assess soil erosion and buffer strip effectiveness.
- All livestock enterprises to record growth rate, milk yields, mortality, conception rates, replacement rates, age at slaughter for sheep and cattle.
- ScotEID to require information on sires.

- All enterprises to use farm tool calculators (carbon audits) to model GHG emissions. Livestock enterprises to model ammonia emissions when a suitable tool is available. The requirement to model might be limited to enterprises above a certain size to reduce costs.
- The remaining outcomes would be best assessed using expert-led monitoring in a sample-based programme similar in philosophy to the Welsh approach. The resources available for monitoring and statistical power analysis would be key inputs into developing a sampling approach with decisions about the trade-off between number of metrics recorded versus sample size needing to be made.

Table 1. Metrics identified as worthy of adoption in future monitoring, listed by outcome - a full list of which metrics are suitable to assess each measure are shown in the spreadsheets MeasureXMetric.xlsx. Metrics are divided into three categories: Suggested metric - a suitable metric for monitoring the relevant outcome(s) that can be applied now; Additional metric – a useful set of additional information or approaches; and Metric in development – analytical methods are still in development, but samples/data can be collected and archived for future analysis. Metrics suitable for use for multiple outcomes are shown in **bold**.

Outcome	Suitable metric	Additional metric	Metric in development
Reducing Soil GHG emissions	**Modelled farm emissions of CH$_4$, CO$_2$, N$_2$O**		Field level, real time emission models.
Increasing soil carbon/organic matter content	**Soil carbon stock (C content and bulk density)** Area under permanent vegetation or other carbon positive management	Soil clay content	**Indicators based on soil FTIR spectroscopy**
Increasing resilience to weather events	**Soil carbon stock** **Water stable aggregates** **Soil bulk density and porosity** **Erosion monitoring**	**Visual Evaluation of Soil Structure (VESS)**	**Indicators based on soil FTIR spectroscopy**
Improving soil nutrient content	**Mineralisable nitrogen and available phosphorus**		**Indicators based on soil FTIR spectroscopy**

Outcome	Suitable metric	Additional metric	Metric in development
Reducing diffuse pollution	**Mineralisable nitrogen, available phosphorus and pH** **Soil bulk density and porosity** **Erosion monitoring and effectiveness of buffer strips**	**Visual Evaluation of Soil Structure** **Detailed monitoring in SEPA catchments to include water quality (nitrate, phosphate etc.)**	**Runoff evaluation using LIDAR derived fine resolution topographic data.**
Improving water and air quality	**Mineralisable nitrogen, available phosphorus and pH** **SEPA regulatory monitoring** **Erosion monitoring and effectiveness of buffer strips**	**Detailed monitoring in SEPA catchments to include water quality (nitrate, phosphate etc.)** Intensive farm-scale monitoring of ammonia emissions in livestock intensive areas	Modelled farm emissions of ammonia
Improving soil water retention and flow	Sub-soil bulk density and porosity **Water stable aggregates** **Erosion monitoring**	**Visual Evaluation of Soil Structure**	
Improving soil biodiversity	Soil surface invertebrates Earthworm functional group abundance	**Pesticide Usage Survey data**	Archive sample for eDNA

Outcome	Suitable metric	Additional metric	Metric in development
Removing drivers for biodiversity loss	Bird, pollinator and plant composition and diversity	Farmland habitat diversity **Pesticide Usage Survey data**	Archive acoustic monitoring files LIDAR derived hedge data
Livestock health	**Growth rate, Milk yields, Mortality, Conception rates, Replacement rates, Age at slaughter**		
Livestock nutrition	**Growth rate, Milk yields, Mortality, Conception rates, Replacement rates, Age at slaughter**	Feed analysis for digestibility/protein	
Livestock genetics	Applications to ScotEID for calf/lamb passports, with requirement for sire details to be included	**Growth rate, Milk yield, Conception rates, Age at slaughter.**	
Livestock methane emissions	**Modelled farm emissions of CH_4, CO_2, N_2O**		
Nutrient management	**Mineralisable nitrogen and available phosphorus** **Effectiveness of buffer strips** **Modelled farm emissions of CH_4, CO_2, N_2O**		Modelled farm emissions of ammonia

Contents

2   Glossary / Abbreviations table

Citizen scientist	Usually used to denote a non-professional scientist. Can range from the public (including land managers) to highly proficient amateur scientists.
FTIR	Fourier-transformed infrared spectroscopy – an analytical technique using infra-red light to identify the chemical composition of materials.
GHG	Greenhouse gases such as CH_4 methane, CO_2 carbon dioxide and N_2O nitrous oxide.
LIDAR	Light Detection and Ranging, is a remote sensing method that uses light in the form of a pulsed laser to measure ranges and hence vegetation structure.
Measure	An action or set of actions employed to reach the outcomes of the Vision for Agriculture.
Method	The processes followed to obtain the data required to produce metrics.
Metric	A quantifiable set of data that can be used to track, compare and assess performance or processes.
RPID	Scottish Government's Rural Payments and Inspections Division

3   Introduction

This report examines the potential metrics for assessing the environmental outcomes of measures identified in the Scottish Agricultural Reform Programme.

3.1   Policy environment

Agriculture is a major contributor to Scottish greenhouse gas (GHG) emissions; currently, it is responsible for c. 20 % of countrywide emissions (Brodie 2023). Agricultural management has also been a major driver of the declines in above- and belowground biodiversity (Walton et al. 2023) and puts significant pressure on Scottish water bodies, preventing them from reaching Good Ecological Status (Environmental Standards Scotland 2022).

Following on from the Scottish Government's Vision for Agriculture, a new Agriculture and Rural Communities (Scotland) Bill has been passed, which will allow for a new framework for future support payments for farmers ("farmer" is used in this report to cover both farmers and crofters), including for environmental goods. This will encourage sustainable and

regenerative farming practices that will help Scotland transition towards net zero, reverse the decline in biodiversity, and improve soil health and water quality.

It is anticipated that there will be a new framework for agricultural payments focused on key outcomes of high-quality food production; climate mitigation and adaptation; nature restoration; and wider rural development alongside a just transition. Greater conditionality will be key, with a transition towards shifting 50% of direct payments to climate action and funding for on-farm nature restoration and enhancement by 2025.

At present a draft list of measures (Appendix A) is being appraised by Scottish Government that covers both land-based and animal-based actions that should lead to improvements in biodiversity, climate, flooding, soil health, water quality and animal health and welfare. However, a system of monitoring and verification is needed to ensure compliance and that the measures are delivering the desired outcomes.

3.2 Aims

The aims of this project were:

- To identify potential metrics that could be used to monitor the success of the proposed measures in delivering the desired environmental outcomes (Appendix B). Those metrics that could be used in practice will have to be cost-effective, practical and within the skills and capabilities of those tasked with the monitoring.
- To take an overview across all the metrics and outcomes to refine the list of metrics to avoid duplication and maximise the usefulness of information collected.

4 Considerations for selecting appropriate metrics

4.1 Introduction

To determine whether any changes over time are the result of direct action through applied measures, it is important to be able to compare areas where measures have been applied with other similar areas that are not in the scheme (control sites). Without this, it is not possible to determine whether any change detected is due to the measures or to other drivers.

It is also possible that even if an improvement is not detected on sites where measures have been applied, the measures might mean that a negative change, that would otherwise have occurred, has been avoided.

A Before-After-Control-Impact (BACI) design is commonly used for monitoring the effect of environmental interventions. However, a difficulty is that areas which are originally selected as controls may join the scheme later. Also, as pointed out by Emmett et al. (2014), it can be difficult to select appropriate controls given the numerous other factors, including field contents, size, and boundary characteristics that would need to be held constant across matched pairs. Even if the areas selected as controls are not part of the current scheme,

they may not be true controls as they may have benefitted from similar environmental measures under legacy schemes.

As a result of these issues, it can be difficult and costly to assess outcomes at the level of individual farms, though overall performance of measures can be assessed through an appropriate monitoring scheme.

4.2 Requirements

Effective monitoring requires an appropriate baseline for measuring outcomes against (Pakeman et al. 2020). A proper baseline gives power to any analysis, as it is detecting change against known values for indicators. For example, agricultural soil monitoring as part of scheme monitoring will need to align with the national soil monitoring programme that is in development.

Similarly, identifying an appropriate sampling design is critical. It needs to cover enterprises in different situations and localities and have the appropriate statistical power to give good evidence on the performance of each measure in at least the medium-term (i.e., to inform revisions to agricultural support schemes). Some outcomes may be detectable quickly, but others, like soil carbon, may take longer to be detectable within realistic sampling regimes (Saby et al. 2008). For other measures it may be difficult to separate the effects of the scheme from market-driven effects, such as the breeding of livestock for reduced methane production, which could be driven by the price of carbon rather than the support from any scheme (Cottle & Conington 2012, 2013).

4.3 Selecting metrics

The selection of metrics depends on several factors, including the design of any monitoring scheme, what is being monitored, for whom and for what purpose, and needs to take account of the trade-offs associated with the approach taken. These can be seen as different aspects of taking either a "broad and shallow" or a "narrow and deep" approach to data gathering for the same amount of effort. Data gathered from a "narrow and shallow" approach will be less detailed and likely less robust, whereas a "broad and deep" approach may be too costly to deploy widely.

4.3.1. Sample or population

Taking a sample of the population and focussing monitoring has the benefit of concentrating resources if it is understood that any sampling design has some measure of uncertainty built in. This type of approach has been adopted in monitoring programmes such as Countryside Survey (e.g., Carey et al. 2008) and the monitoring of the Welsh agri-environment scheme Environment and Rural Affairs Monitoring and Modelling Programme (ERAMMP), which focusses monitoring on 300 1 km x 1 km grid squares and assesses the impact of the scheme using information on how much land in each square is under Glastir funded management (see Section 8.1.1). The approach allows for efficient linkage between changes in different outcomes, but with the proviso that there is uncertainty and that it can only give a national-level picture.

4.3.2. Citizen scientist or specialist

For agriculture, options will include asking the farmer or land manager to gather information, drawing data from wider datasets, or drawing in specialists to sample and process data. There are advantages and disadvantages to asking land managers, as opposed to specialists, to carry out the monitoring. Land managers differ from citizen scientists in other monitoring, e.g., the British Trust for Ornithology's Breeding Bird Survey, which is undertaken by volunteers with a high degree of skill at bird recognition. Expectations would have to be tempered in terms of what can be provided.

Consequently, the advantage of monitoring by the land manager is that it is effectively free, it can be repeated frequently and provides information direct to the land manager. This must be viewed against the benefits of sampling with more accuracy and precision by specialists.

It may be possible to develop hybrid monitoring strategies using the advantages of the different groups, either using land managers to take samples (e.g., soils), which are then sent away for analysis, or deploying monitoring equipment, with the specialists undertaking data analysis. Specialist data analysis is preferable from the point of view of scientific robustness, although monitoring equipment does need expert maintenance, calibration and quality control and is more costly. Alternatively, a tiered approach to monitoring could be followed, with land managers collecting some data whilst more specialised data collection is undertaken on a sample of farms.

4.3.3. Meaningful scales of monitoring

The appropriate scale of monitoring is inherent in what is being monitored. For plants, relatively small areas (a few square metres) tend to be monitored, whilst for butterflies and bees, the area might be a transect 100 m long and 5 m wide, and for birds, the British Trust for Ornithology uses 1 km x 1 km grid squares as the basis of their Breeding Bird Survey.

In consequence, the scale of monitoring for different aspects of the environment and biodiversity will not be the same for all outcomes. There is, therefore, some constraint on the overall approach as it is dependent on finding the most appropriate scale for each outcome.

4.3.4. Who is the monitoring for?

The vision for agriculture includes provision for payments that deliver to defined outcomes. If the aim is to inform management at the farm-scale or smaller, in effect using the results of monitoring in adaptive management, then there may be a benefit to a broad and shallow approach. There is also value in aligning monitoring with appropriate advice and resources for decision making. However, if the monitoring is just aimed at showing which measures are value for money, then a national level focus is more appropriate.

4.3.5. Understanding what is driving change

If measured changes can be linked directly to the impact of targeted funding, or with conditions for an agri-environment scheme, then this is a direct demonstration of the efficacy of the scheme.

However, a narrower set of more detailed monitoring may be better placed to understand more precisely what is driving change as a greater range of measured parameters can be used to examine the processes that lead to change. This improved knowledge might be more useful in developing future schemes and inform adaptive management. A tiered approach to monitoring may deliver the best information.

4.3.6. Can you monitor outcomes, or just activity?

It is possible that suitable methods to measure outcomes at the desired scale are not available or practical. Consequently, it may be that measuring actions or activity remain the only option to assess whether management is driving change in the desired direction. However, there would need to be some form of outcome monitoring at a wider scale to assess overall performance of the scheme.

4.3.7. Does land manager-led monitoring need supervision?

This is a contentious issue, but in other spheres such as sampling for water industry and fish farm compliance there are quality assurance assessments of 'operator collected data'. Some are targeted based on evidence of some kind, but there is a random element to create pressure to conform.

There is a need to consider whether an inspection system is required to ensure there is pressure to maintain high standards of monitoring. Northern Ireland has decided that the best way to obtain robust data for monitoring is to employ people to do the measurement and use techniques such as GPS monitoring to check sample collection protocols are being followed (https://www.afbini.gov.uk/articles/soil-nutrient-health-scheme).

5   Methodology

We used an expert led rapid evidence assessment to look for different ways of assessing the success of each measure against environmental outcomes. This involved a multistep approach to developing appropriate metric recommendations to monitor the environmental outcomes of the new agricultural support system.

5.1 Step 1

For the land-based proposed measures only, we assessed each proposed measure (Appendix B) to identify which of the outcomes it was relevant to. For example, there are nine outcomes listed for In Field – Cultivated Soils, but not all outcomes are relevant for each measure. For example, the outcomes Reducing Soil Greenhouse Gas (GHG) emissions and Increasing soil carbon/organic matter content are unlikely to be affected by

Efficient/Reduced use of synthetic pesticides so it would not be useful to monitor those if this was the sole measure in place.

This step was undertaken by individuals with expertise in each outcome.

5.2 Step 2

For each combination of relevant outcomes and measures, we used expert knowledge and a search of relevant literature to identify potential metrics that could be employed to assess compliance and/or the success of the measure in reaching the desired outcome (Appendix C). These were categorised in the following ways:

- Compliance or outcome-based
- Already collected under the current payment scheme, by agencies or third parties, or if novel data metrics will be required
- Practical for field-level monitoring, holding-based monitoring or for national-scale monitoring only, or unsuitable for routine monitoring.

This step was undertaken with the expertise of the research team backed up by literature searches. However, for the land-based measures, one individual was tasked with identifying appropriate metrics across all measures relevant to a particular outcome to ensure a consistency of approach. In contrast, the livestock-based measures are more holistic and required an expert to consider the actions around these in the round to identify appropriate metrics.

5.3 Step 3

The assessment in Step 2 generated a large list of metrics with associated methods that could be employed to assess the success of the scheme. A series of three workshops was used to consolidate these to ensure that where possible the same method can be employed across as many measures as possible for simplicity and to help in scaling up from individual measures to the success of the whole scheme. This stage delivered a shortlist of metrics that could be used to assess the success of the measures in delivering the desired outcomes, i.e., cost-effective, practical and within the skills and capabilities of those tasked with implementing the metric(s).

5.4 Step 4

This step focussed on identifying data collection approaches for consideration, as well as considering requirements for establishing an initial baseline and for future data collection to assess both compliance/activity and outcomes. Data collected could be integrated into existing data sets, such as the National Soil Inventory of Scotland, to give a longer perspective of change.

6   Potential metrics for each outcome

The outputs from Steps 1 and 2 are presented in Appendices B and C but are summarised below. Step 3 identified a set of metrics that could be employed in monitoring outcomes. This section identifies those metrics that would provide practical and cost-effective information. Potential metrics are categorised into three levels:

- **Suitable metric** – a suitable and available metric for monitoring the relevant outcome(s).
- **Additional metric** – a useful set of additional information or approach.
- **Metric in development** – analytical methods are still in development, but samples/data can be collected for future analysis.

6.1 Reducing soil greenhouse gas (GHG) emissions

6.1.1. The outcome

Greenhouse gas emissions from agriculture are a significant part of the national total. Reducing these emissions is a key goal of the Agricultural Reform Programme and the Climate Change Plan.

6.1.2. Considerations with a metric

Current methods required for direct measurement of GHG fluxes are not suitable for wide-scale use as they are dependent on relatively expensive equipment and a high degree of specialist knowledge to run the equipment.

We suggest that instead of this a modelling approach, based on existing or in development farm/field GHG calculators, is used that would estimate CO_2, N_2O and CH_4 emissions. These are also known as Carbon Audits and are currently funded as part of the Preparing for Sustainable Farming initiative. However, several issues would need considering:

- There are several modelling tools on the market (see section "Reducing Soil Greenhouse Gas (GHG) emissions" in the Appendix), so an updated review (see Leinonen et al. 2019) of their capabilities would be needed to ensure that only suitable products were used, and to ensure consistency of outputs.
- Assistance may be needed, and hence need paying for, in setting up the calculators in the first instance, as in the Carbon Audits in the Preparing for Sustainable Farming initiative.
- Outputs from the calculator depend on the quality of the primary data gathered, which means data quality checks may be a requirement.
- Feed and forage quality might be useful information to feed into the calculators – see section below on Animal health and nutrition.

Land managers will benefit from these whole farm or field-level calculators with the potential to identify cost reductions or increases in productivity through improved forage

and manure management. This could be supported by the soil organic matter and nutrient data collected.

6.1.3. Suggested metrics

Suitable metric: Modelled farm emissions of CH_4, CO_2, N_2O

Metric in development: Modelled gas fluxes in real time at the field scale.

6.2 Increasing soil carbon/organic matter content

6.2.1. The outcome

Increasing the levels of soil carbon through regenerative agriculture can make agricultural land a sink for carbon and facilitate the journey to net zero.

6.2.2. Issues with a metric

Soil organic carbon can be routinely measured. There are different laboratory methods available, all of which work well, but a standardised approach would need to be selected for any scheme. Dry combustion (Dumas method) is widespread in its application and thought of as the best chemical method for soil carbon determination (Chatterjee et al. 2009). In addition, some consideration needs to be given to dealing with soil samples from calcareous soils where inorganic carbon levels are high (mainly carbonates), which though rare do include soils like machair soils. Additionally, by linking soil carbon to clay content (measured when characterising soil texture) a measure of the land parcel's status regarding storing carbon is produced. Thresholds of 13:1, 10:1, and 8:1 clay to soil organic carbon could potentially be applied to arable, arable ley, and woodland systems (Prout et al., 2022).

Laboratory measurement is straightforward, but to calculate stocks, there also needs to be a measurement of soil bulk density (total dry mass per unit volume). Consideration of sampling depth(s) is important as some changes, such as a switch to deeper rooting crops may increase subsoil carbon, while changes in soil tillage might affect the vertical distribution of soil carbon. A standardised sampling protocol needs to take this into account. The approach being taken in Northern Ireland is informative. Every farm and every field are being sampled for carbon and nutrients and soil testing is a precondition of eligibility for environmental payments. Soil carbon stocks are large and are heterogeneously distributed, meaning that quantifying changes over short time periods is seldom possible. For instance, the proposal for a directive on Soil Monitoring and Resilience (Soil Monitoring Law) will require samples to be taken every five years. However, to ensure agronomic management changes will deliver and to identify which ones deliver, actions such as the employment of minimum tillage, use of winter cover crops, inputs of organic wastes and increases in permanent vegetation cover (woodland, hedges, grassland) need to be recorded at the field level alongside actions that will reduce soil carbon such as the removal of permanent vegetation cover and ploughing of grasslands.

Further considerations in developing this sampling include:

• Sampling to be carried out by land manager or by experts. There is a trade-off between cost and reliability but given the range of other soil metrics that need to be sampled to assess other outcomes, we suggest that soil sampling is expert led.

• Should samples from the same field be bulked to reduce costs or should they be analysed separately (expensive) to provide measures of error/heterogeneity and the possibility to statistically assess change at the field level rather than at the farm or national level? For instance, the Soil Nutrient Health Scheme in Northern Ireland analyses a bulked sample of 25 cores but this can miss coldspots and hotspots of nutrients (Hayes et al. 2023). The Welsh Soil Project splits each field into three before the W-shaped sampling is done. There is a direct trade-off between the number of fields that can be sampled and the number of samples per field. We suggest that the most useful information comes from sampling as many fields as possible, so a bulked sample per field would be an appropriate sample to measure. Some within field stratification could be done if there was a clear internal boundary, e.g., between dry slope and wetter flat ground.

• Collecting additional information such as the current and past management and cropping at field level would enhance interpretation.

• Several companies already operate soil testing services. In a competitive market, there is a question regarding how consistency is guaranteed and whether a consistency check should be carried out by a third party. United Kingdom Accreditation Service (UKAS) accreditation would be a minimum standard for participating laboratories.

• Sampling of enclosed land with a single habitat per field is straightforward. However, consideration needs to be given on how to sample from unenclosed land which may contain multiple habitats and a wide range of soil types.

6.2.3. Suggested metrics

- **Suitable metric:** Soil carbon stock, Area under permanent vegetation or other carbon positive management
- **Additional metrics:** Soil clay content
- **Metric in development:** Indicators based on soil FTIR spectroscopy.

6.3 Increasing resilience to weather events

6.3.1. The outcome

Soils are vulnerable to runoff and erosion after heavy rain and to drought. Improving the resilience of soils will safeguard their continuing productivity, reduce their susceptibility to the runoff of water and nutrients, and subsequent downstream impacts on flooding and water quality.

6.3.2. Issues with a metric

Resilience is a synthetic metric and can be best seen as a multi-dimensional concept. In addition, the thresholds for resilience will depend on soil type. Regarding improving soil resilience, mineral soils that have greater soil carbon concentrations tend to retain water and have better soil structure, allowing water flow through them rather than across them. Soils that show water percolating (high permeability) rather than flow across the surface are at lesser risk of runoff and erosion, whereas compacted soils with lesser porosity and greater bulk densities are much more vulnerable to weather events. Compacted soils also restrict water availability and nutrient dynamics impacting crop growth. The presence of water stable aggregates also helps prevent water and wind breaking down the soil and hence lower the risk of erosion. These indicators are covered elsewhere in this report (see sections Increasing soil carbon/organic matter content, Improving soil nutrient content and Reducing diffuse pollution) and hence not covered here in detail.

6.3.3. Suggested metrics

- **Suitable metric:** Soil carbon stock, Water stable aggregates, Soil bulk density and porosity, Erosion monitoring
- **Additional metric:** Visual Evaluation of Soil Structure (VESS)
- **Metric in development:** Indicators based on soil FTIR spectroscopy

6.4 Improving soil nutrient content

6.4.1. The outcome

Maintaining soil nutrient supply to ensure high levels of productivity is important for efficient farming. However, an oversupply of nutrients can lead to losses as emissions of ammonia and nitrous oxide, or as increased nutrient loadings of freshwaters. While Scotland has no widespread and high impact nutrient issues such as Lough Neagh in Northern Ireland, there are localised issues that have been identified through designations such as Nitrate Vulnerable Zones that might be more cost effective/appropriate to measure.

6.4.2. Issues with a metric

The total concentrations of the various soil nutrients are relatively straightforward to sample and analyse and could be combined with sampling for soil carbon. Analysis methods depend on whether a restricted set of macro-nutrients is the focus, or whether micronutrients and heavy metals are also of interest.

Total nutrient levels work well for some nutrients, but there may be an interest in looking at available nutrients where there is an extraction/exchange step to assess what is available to plants and leaching processes. There are standard laboratory methods for this, particularly for nutrients such as potassium and calcium, but phosphate extraction methods have been developed to be specific for different soil acidity levels (pH).

Unfortunately, neither total nutrient levels nor extractable/exchangeable levels work well for nitrogen, as nitrate is very quickly absorbed by roots, leached, or transformed (e.g., to

nitrous oxide). Here, an incubation step is needed, meaning that getting a good understanding of available nitrogen requires sampling, dividing the sample, extracting immediately from one half of the sample, incubating the other half for a set time under standard conditions, and then calculating the release of nitrogen by the soil.

There is an immediate trade-off with adding fertiliser to raise nutrient levels, as excess nutrients can be leached and end up in the aquatic environment, or excess nitrogen can be lost as N_2O. Hence, a balance must be reached where inputs meet plant requirements, while also fostering accumulation of soil organic matter to maximise intrinsic soil nutrient cycling. Current agronomic practice is to apply inorganic fertiliser at rates based on an understanding of plant uptake, but application rates often exceed those which are required as soil-specific variability in supply of nutrients from soil organic matter is usually not accounted for. Tools such as PLANET (Planning Land Applications of Nutrients for Efficiency and the environmenT), a nutrient management decision support tool for farmers and advisers to carry out field level nutrient planning and for demonstrating compliance with the Nitrate Vulnerable Zone (NVZ) rules, could be useful in this regard.

Maintaining optimal pH for crop growth also appears to reduce soil greenhouse gas emissions (Wang et al., 2021; Zhang et al, 2022), but there is a degree of context specificity, and this may not be appropriate for soils of high organic matter content.

6.4.3. Suggested metrics

- **Suitable metric:** Mineralisable N and available P, Soil pH
- **Metric in development:** Indicators based on soil FTIR spectroscopy

6.5 Reducing diffuse pollution

6.5.1. The outcome

Diffuse pollution has severe impacts on freshwater biodiversity and water quality with risks that climate change (low and high flow extreme increases, warmer temperatures) exacerbates effects such that moderate nutrient loading improvements may not lead to improved water quality.

6.5.2. Issues with a metric

Monitoring of diffuse pollution operates across scales, from the field scale, to highlight local improvements, to the catchment scale to understand cumulative effects and impacts (Bieroza et al. 2021). Field-scale predictions and observations of runoff prevalence and pathways, monitoring of soil compaction (measured by soil porosity) and soil chemistry (particularly nitrogen and phosphorus levels) provides an idea of risk, as does monitoring of in-field erosion (Hayes et al. 2023). Management at the edge of fields, e.g., buffer strips are designed to reduce diffuse pollution, but for best effectiveness, their location and design need to be targeted to ensure that they effectively treat converging runoff pathways and critical delivery points to the channel network (Stutter et al. 2021). Similarly, nutrient losses from field drains also need to be monitored as these can only be mitigated by specially designed and strategically located buffer strips.

Water sampling provides integrative evidence of the effectiveness of measures as it reflects management upstream in the catchment. Whilst monitoring of chemistry, biodiversity (invertebrates) and sediment will provide an understanding of upstream issues, it may be difficult to attribute impacts to diffuse or point source pollution (Glendell et al. 2019).

Water quality is closely linked with soil nutrient status, particularly nitrogen and phosphorus status of the soils, so relevant information can be acquired by soil sampling. However, there is also the need to monitor runoff generation and pathways, soil erosion, sediment flows and drainage waters. Monitoring is especially useful during extreme events, including high and low flows. An understanding of pollutant concentration changes over differing flow stages (e.g., inter-storm sampling) brings a wealth of information beneficial to management about source and transport behaviours at field to catchment scales.

We suggest that land managers are given responsibility for assessing erosion and water flow pathways and the subsequent monitoring of erosion and sediment flows, and potentially taking water samples of drainage waters for analysis by specialist laboratories. This would mean farmers assessing whether individual buffer strips were effective at preventing water flows, or whether their design allowed for flow around their edges by visiting them during periods of heavy rain. Future erosion pathways could be identified using fine-scale elevation data from LIDAR to model the flow of water across the surface of land (e.g. Reaney et al. 2019. Aquatic biodiversity requires specialist surveyors and could be done at the same time as the above-ground biodiversity assessment (Section 6.9).

SEPA currently collect a wide range of data from multiple sites. We suggest that it would be of benefit to use the current SEPA monitoring of agricultural catchments as the basis for studies linking agricultural management and water quality, by ensuring studies are joined up. This may mean enhancing the range and/or frequency of measures taken. A nested design could be followed, whereby field- and farm-scale sampling are nested within these catchments representing different land use typologies in Scotland, with water quality being monitored at the catchment outlet. The detailed knowledge from these catchments could be linked to farm-level data to make national estimates of benefits.

Farm-level models for looking at nutrient inputs and losses have been developed for England and Wales, e.g., FARMSCOPER. However, the extent to which it can be applied to the soils, climate and farming systems in Scotland has not been tested and this would need carrying out before it could be recommended as a metric for use in assessing the efficacy of measures.

6.5.3. Suggested metrics

- **Suitable metric:** Mineralisable nitrogen, available phosphorus and pH, Soil bulk density and porosity, Erosion monitoring and effectiveness of buffer strips (including other enhancements e.g., wetlands, wet woodland, sediment traps)
- **Additional metric:** Visual Evaluation of Soil Structure (VESS), Detailed monitoring in representative SEPA and other research catchments for process-based understanding on management impacts

- **Metric in development:** Runoff evaluation using LIDAR derived fine resolution topographic data

6.6 Improving water and air quality

6.6.1. The outcome

Water quality is tightly linked to freshwater biodiversity. However, it also has implications for the cost of water treatment downstream. Air pollution, particularly of ammonia, can also severely impact local biodiversity.

6.6.2. Issues with a metric

There can be a disconnect between actions at the field scale to reduce nutrient loss and water quality as actions can be poorly sited, poorly implemented and miss important routes of pollutant movement. However, there is clear evidence that reduction in soil nutrient status is the most likely route to deliver improvements in water quality, so monitoring for water quality is intrinsically linked to monitoring of soil nutrient status (Hayes et al. 2023).

High-resolution water quality monitoring that would represent the temporal and spatial variability is expensive and the movement of water in catchments may make linking it to the actions of individual farms problematic. Consequently, we suggest a combination of field/farm-level monitoring of soil nutrient status (i.e., soil organic matter, plant available (mineralisable) N, biologically available P and pH) and detailed monitoring of several representative catchment outlets to improve the understanding of processes. These could be based around SEPA's existing catchment observation platforms, with additional investment to maximise the robustness of collected evidence.

Further action to reduce point source pollution, such as slurry pit overflow, farmyards and septic tanks, should not be overlooked (Harrison et al. 2019). Monitoring of this would be in the form of capital spend. Best practice should be followed for digestate and slurry application to land.

Currently available sensors for monitoring ammonia emissions tend to be expensive, require technical expertise and are sensitive to meteorological conditions and other atmospheric gases. Lower cost passive samplers, which could be deployed by non-specialists are less accurate, have lower temporal resolution, and require laboratory analysis (Insausti et al., 2020). A similar approach to that proposed for water quality could be implemented, with intensive monitoring of key areas with intensive livestock production systems, coupled with national scale monitoring utilising the National Ammonia Monitoring Network which monitors atmospheric ammonia concentrations monthly. A farm-level calculator for ammonia emissions is in development as part of the Scottish Government's Strategic Research Programme. This would be the most cost-effective way forward for wide deployment of monitoring.

6.6.3. Suggested metrics

- **Suitable metric:** Mineralisable nitrogen, available P and pH, Soil bulk density and porosity, Erosion monitoring and effectiveness of buffer strips
- **Additional metric:** Intensive farm-scale monitoring of ammonia emissions in livestock intensive areas, Visual Evaluation of Soil Structure (VESS), Detailed monitoring in SEPA catchments for process-based understanding on management impacts
- **Metric in development:** Modelled farm emissions of ammonia

6.7 Improving soil water retention and flow

6.7.1. The outcome

Soil water retention is important in reducing soil erosion and diffuse pollution. If water flows through the soil it is slowed, reducing flood peaks, and there is greater interaction between the soil and water reducing the risk of nutrient loss. In contrast, water flowing across the surface of soils leads to erosion and nutrient runoff.

6.7.2. Issues with a metric

There are several detailed methods available to understand water retention and flow through soils, but they are not appropriate for wide-scale monitoring, apart from their potential use in the detailed monitoring of test catchments. These include detailed measures of soil texture, as well as laboratory measures of hydraulic conductivity. Direct measures of soil compaction with penetrometers suffer from variability due to soil water content, stoniness of the soil and differences between manufacturers. They are not suitable for wide-scale monitoring.

However, a set of straightforward measures are available to assess how soil water behaves. As part of the sampling of soil for soil carbon measurements, bulk density is measured to calculate carbon stocks from carbon concentrations. However, topsoil bulk density can vary seasonally and with respect to management. Subsoil bulk density is an indicator in the draft EU soil monitoring and resilience law and provides a more consistent measure of how the soil is behaving. This is a key parameter for understanding the effect of management on this outcome. However, the additional effort of also recording specific gravity of the soil will allow the calculation of soil porosity, another key parameter that is important for assessing soil water retention.

The Visual Evaluation of Soil Structure (VESS) is a qualitative metric that could also be used to supplement other measures and provide land managers with direct information at the field level on the degree of soil compaction, especially if this included both topsoil and subsoil. For quantitative measures of soil structure, the measurement of Water stable aggregates (WSA) should be considered and removes the potential for subjectivity.

6.7.3. Suggested metric

- **Suitable metric:** Sub-soil bulk density and porosity, Water stable aggregates, Erosion monitoring
- **Additional metric:** Visual Evaluation of Soil Structure (VESS)

6.8 Improving soil biodiversity

6.8.1. The outcome

Maintaining a healthy soil ecosystem is critical to the regulation of key processes, as soil organisms are critical to the cycling of nutrients and to plant growth. For instance, soil animals like earthworms are highly important to water movement in soils.

6.8.2. Issues with a metric

Soil biodiversity, whilst a key soil health indicator (Neilson et al. 2021), is unlikely to be practically assessed by the land manager. Identification of surface-dwelling invertebrates, such as beetles and earthworms, requires specialist taxonomic skills; even for earthworms a total count does not work as all functional groups need to be present for good soil health. Existing data is not available for surface dwelling invertebrates, but data collection methods with pitfall traps are standardised, for example by the [Environmental Change Network](#). However, these methods require at least two visits, so may not be cost-effective. Previous earthworm surveys have been carried out (Boag et al. 1997, Carpenter et al. 2012), we suggest that methodologies should be kept consistent.

Molecular methods have been employed for bacteria, fungi and nematodes. However, methods to characterise complete soil biodiversity using eDNA (environmental DNA) are now emerging. As is typical with emerging technologies, there are issues surrounding data interpretation, thresholds and developing and/or defining baseline comparators. It is, perhaps, too early to suggest using this as a monitoring method, as the science relating molecular data to improvements in soil health is in its infancy. However, as soil sampling is likely to be used to monitor other outcomes, samples could be taken and archived for future use as a baseline to assess change.

Pesticide usage could be a proxy for the pressure on biodiversity, and hence pesticide usage data would be a useful addition to direct monitoring. It is already collected in Scotland, but refining the data to consider impacts on soil organisms and the different application rates would be necessary.

Further consideration needs to be given to:

- Collecting contextual information such as the current and previous crops.
- Whether the optimum times for sampling in spring and autumn coincide with the optimum times for sampling soil carbon and nutrients.

6.8.3. Suggested metrics

- **Suitable metric:** Surface dwelling invertebrates and earthworm functional group abundance
- **Additional metric:** Pesticide usage data
- **Metric in development:** eDNA samples archived as interpretation needs to improve

6.9 Removing drivers for biodiversity loss

6.9.1. The outcome

As much of Scotland is affected by agriculture, sensitive agricultural management is important to delivering the goals of the Scottish Biodiversity Strategy.

6.9.2. Issues with a metric

Biodiversity is intrinsically multi-dimensional, but typical agri-environmental monitoring targets habitat diversity, birds, pollinators and plants, as they give information at different scales.

In most schemes, biodiversity monitoring is done by specialists, as it is the status of priority species that has been the driver for the development of the scheme. However, that is not practical in terms of cost at the farm level, so a choice must be made between:

a. Land manager-led monitoring aided by tools such as report cards and identification guides. Bird surveys could allow different levels of precision from individual species to groups (e.g., finches). Similarly, pollinator surveys could record at the level of group (bumblebee, honeybee, butterfly, hoverfly) or plant surveys, by numbers of different types of flower (e.g., daisy, pea types) in a set area. Alternatively, there is the possibility of sub-contracting to specialists if grant payments included money for monitoring. Land managers setting out acoustic recording devices also fits into this space. The resulting files could be uploaded to a central organisation responsible for analysis. The methodologies for data analysis are still in development, but sound files could be archived for later analysis when the methodologies have matured to deal with high levels of false positive identifications. The biodiversity audit as part of the whole farm plan also falls into this category.

b. Specialist surveys on samples of farms with the sampling design considering the implementation of measures (Pakeman et al. 2020) or being large enough to assess change for most measures, however, they are distributed across the landscape (e.g., the Welsh approach to monitoring Glastir).

There is a clear trade-off here between broad and shallow versus narrow and deep approaches. To enable adaptive management at the farm level, then land manager-led monitoring is important, but there is the risk that the measures deliver higher numbers of generalist species, do not benefit species that are a conservation priority, but the data is incapable of showing this. It may be that a hybrid approach is necessary, so that field/farm-level data is complemented by detailed measures on a sample of land holdings. However,

sample sizes need to be sufficient to confidently assess change. Previous monitoring studies, e.g., Perry et al. (2003), could be used to identify appropriate levels of sampling needed.

Currently collected biodiversity data is not appropriate for agri-environment monitoring for a range of reasons, mainly due to mismatches in scale between land holdings and the specific sampling method used. In the case of breeding bird data, it has been used as a measure of general farmland diversity against which the performance of in-scheme farms has been judged.

Proxies for habitat diversity currently collected by RPID would be useful data, but it only characterises area and has no measure of quality associated. Alternatives include using remote sensing data (e.g., habitat maps or LIDAR derived information on hedgerow extent and conditions) that provides information on land cover and structure, but these are only proxies for biodiversity.

Finally, pesticide usage is a clear driver of biodiversity loss. Usage statistics are already collected using a sampling approach to assess a Scotland-level picture. However, the diversity of chemicals applied, and their different application rates would require methodological developments to combine their usage into meaningful statistics.

6.9.3. Suggested metrics

- **Suitable metric:** Bird, pollinator and plant composition and diversity
- **Additional metric:** Farmland habitat diversity, pesticide usage survey data
- **Metric in development:** Acoustic diversity, LIDAR derived hedge data

6.10 Improving animal nutrition

6.10.1. The outcome

Improving animal nutrition will reduce the time taken to deliver animals to market. This reduces lifetime emissions especially of methane.

6.10.2. Issues with a metric

Improving livestock nutrition leads to increased animal performance and reduced methane, nitrous oxide and ammonia emissions. Monitoring of nutrition can be undertaken through laboratory analysis of feedstuffs. The key analyses are forage digestibility - which can easily be undertaken by many feed companies - and dietary crude protein. There is also an important trade-off already mentioned between optimising nutrition and the increased fertiliser use, leading to greenhouse gas emissions and/or pollution of water courses. However, these are very much business-related metrics, and their collection may not be informative as a means of national monitoring, particularly as silage quality varies between fields, time of year and across years. The need for its collection as part of a national monitoring scheme is, therefore, debateable.

Instead, we suggest that simple measures of animal performance are collected and form part of routine monitoring of flock/herd status. These reflect actual performance rather than inputs into the system and are easier to record.

6.10.3. Suggested metrics

- **Suitable metrics:** Growth rates, Milk yields (Dairy cattle only), Mortality, Conception rates, Replacement rates, Age at slaughter
- **Additional metric:** Feed analysis for digestibility/protein

6.11 Improving animal breeding

6.11.1. The outcome

Focussing on animal breeding can improve the productivity of farming systems and, also, increase the quality of products like meat and milk. In terms of reducing methane production, breeding can directly reduce emissions, but also quicker growing animals will release less over their lifetimes.

6.11.2. Issues with a metric

Selective breeding for improved productivity, improved efficiency or reduced methane emissions could drive permanent and cumulative improvement in performance and/or reductions in methane emissions. Monitoring of selective cattle breeding for specific traits could be undertaken through applications for calf passports to ScotEID, but this would rely upon sire details being recorded on passports (which is currently not mandatory) and on the sire's genetic potential for selected traits being known.

Proxy measures such as growth rates, milk production, conception rates and days to slaughter could also be used to monitor improvements over longer time periods but could be confounded with improvements in nutrition and health.

6.11.3. Suggested metrics

- **Suitable metric:** Sire details included on applications to ScotEID for calf/lamb passports
- **Additional metric:** Growth rates, Milk production, Conception rates, Age at slaughter

6.12 Improving animal health

6.12.1. The outcome

Improved animal health has a direct benefit to animal welfare. However, it also reduces losses during the production process, improving productivity and reducing methane emissions on a lifetime basis.

6.12.2. Issues with a metric

Several endemic (and exotic) diseases and syndromes can impact on the production efficiency and associated GHG emissions of farmed livestock. Some diseases have a direct

impact on individual animals and metrics such as growth rates, reproductive success, and replacement rates. Others have a more indirect impact at herd/flock and national level, through how diseased animals are managed following diagnosis. Data and metrics on the prevalence of key priority diseases and health conditions at a national level are currently not collected, but would be invaluable, if logistically challenging. Eradication may be feasible for some diseases, e.g., Bovine Viral Diarrhoea (BVD), but requires the relevant tools, e.g., vaccines, and diagnostics to be available, in addition to coordination and buy-in across the industry.

The most straightforward was to assess animal health would be to collect a common set of proxy measures, e.g., growth rates, age at slaughter, conception rates, replacement rates and mortality rates will be the most feasible approach to measuring progress on animal health. This approach could also be applied to animal breeding and nutrition.

Recording all these metrics would be useful to both national-level monitoring of performance and for the land manager's care for their livestock. This could also include records of veterinary medicines used to gauge movement towards sustainable prescribing, though this complex topic (Humphry et al. 2021) is outwith the scope of this report.

6.12.3. Suggested metrics

- **Suitable metrics:** Growth rate, Milk yields (Dairy cattle only), Mortality, Conception rates, Replacement rates, age at slaughter

6.13 Methane suppression

6.13.1. The outcome

Methane is a greenhouse gas with much higher global warming potential than carbon dioxide (methane from non-fossil fuel sources has a global warming potential of 27 times that of CO_2 with a 100-year time horizon, IPCC 2021). Enteric methane is released by ruminants such as cattle and sheep as part of the natural digestion of plant material by their associated microbiota. Methane is a significant part of agricultural emissions and so reducing it is key to reaching net zero emissions.

6.13.2. Issues with a metric

Selective breeding for reduced enteric methane emissions/increased animal efficiency (section 6.11) is a long-term strategy. In the short term, feed supplements designed to suppress enteric methane production could be used to drive down emissions. Sexed semen could be used to optimise herd dynamics by reducing numbers of male dairy calves and increasing male beef and dairy-beef calves. Direct measurement of methane emissions depends upon specialised equipment and is therefore not practical at the herd or flock level.

Two potential options are available. Firstly, to monitor the usage of methane-reducing feed supplements and calculate emission reductions based on their reduction factors. However, appropriate reduction factors for all feed products may not be available for all systems. The other option is to use current carbon footprinting tools (e.g., Agrecalc, Cool Farm Tool), but

these need a subscription, may need the help of a consultant to set up and would benefit from information on forage quality and the impacts of feed supplements (so in effect replacing the need for providing information separately on feed supplements). The use of a standard tool across herds/flocks would allow for comparison.

6.13.3. Suggested metrics

- **Suitable metric:** Modelled farm emissions of CH_4, CO_2, N_2O

6.14 Nutrient management

6.14.1. The outcome

Poor nutrient management can lead to the emissions of nitrous oxide, methane and ammonia. It also runs the risk of point source and diffuse pollution into watercourses.

6.14.2. Issues with a metric

Organic manures help recycle nutrients and build soil organic matter. However, there is the potential for them to be a source of ammonia, methane and nitrous oxide, as well as nutrient runoff in water courses. Much can be done to alleviate this, with well-designed and covered manure stores as well as appropriate application techniques. Gaseous emissions are difficult to monitor directly, so these would have to be modelled using a farm calculator. Impacts on soil nutrients and water quality are dealt with in previous sections so a separate metric for nutrient management is not necessary.

6.14.3. Suggested metrics

- **Suitable metric:** Modelled farm emissions of CH_4, CO_2, N_2O, Mineralisable nitrogen and available phosphorus, Effectiveness of buffer strips
- **Metric in development:** Modelled farm emissions of ammonia

7 Coordinated metric collection

This section examines opportunities to synthesise across the required outcomes to minimise the number of metrics to be collected.

7.1 Why is this important?

Any monitoring must be as cost-effective as possible. Consequently, during the design phase decisions should be focussed on making the recording of metrics as straightforward as possible and to build efficiency into any monitoring programme, for example, by sampling multiple metrics on the same visit.

7.1.1. Greenhouse gas emissions

Gas emissions cannot be realistically measured directly. Using current farm-level tools to assess GHG emissions will deliver against multiple outcomes [Reducing Soil Greenhouse Gas (GHG) emissions, Livestock emissions, Nutrient management]. In addition, a tool for

estimating ammonia emissions [Improving water and air quality] is in development, as is a field-level, real time emission model. These will further enhance capability in this area.

7.1.2. Coordinated sampling strategies

Many metrics depend on the direct sampling of soil or biodiversity and can't be realistically replaced by proxies or existing data. However, well designed sampling programmes can maximise the efficiency of sampling, e.g., sampling for soil carbon, nutrients, pH and eDNA can be done at the same time. Even if this were not possible, sampling of soil nutrients, particularly mineralisable nitrogen and available phosphorus, would deliver against multiple outcomes [Improving soil nutrient content, Reducing diffuse pollution, Improving water and air quality, Nutrient management]. Similarly, monitoring soil bulk density is important for multiple outcomes [Increasing soil carbon/organic matter content, Improving soil water retention and flow], as is Water stable aggregates.

7.1.3. Soil monitoring

A range of soil monitoring is already being carried out for different purposes. There is a need to consider how future monitoring could supplement or replace existing work in this area, including:

- Scottish Government's National Test Programme's includes soil carbon measurement in Preparing for Sustainable Farming (PSF).
- The Agriculture and Horticulture Development Board Scorecard for soil health which includes soil structure (using VESS Visual Evaluation of Soil Structure), pH, extractable nutrients (phosphorus, potassium and magnesium), earthworms, soil organic matter.
- Linking Environment And Farming (LEAF).
- Red Tractor.
- Soil Association.

7.1.4. Livestock

The outcomes associated with animal health, nutrition and breeding must be largely monitored through proxy metrics, but these are relatively easy to measure and provide useful information direct to the land manager. However, it would be difficult to disentangle the differing contributions of nutrition, health and breeding on the overall performance of the flock/herd. At present, the separate contributions of improving animal nutrition, improving animal health and improving the genetics of the flock/herd are not easily separated but offer three routes for livestock managers to improve performance and consequently reduce emissions, one or more of which can be followed.

7.1.5. National level data

A few metrics can be based on existing data such as data collected as part of the agricultural census or can be derived from existing data such as satellite habitat maps. These are useful additional data, but do not provide the best metrics to assess the success of outcomes. They include: Area under permanent vegetation or other carbon positive management, Detailed

monitoring in SEPA catchments to include water quality (nitrate, phosphate etc.), SEPA regulatory monitoring, Pesticide Usage Survey data, Farmland habitat diversity. They can be identified by filtering column I in the terrestrial sheet of MeasuresXMetrics.xlsx file.

7.1.6. Metrics currently in development

There are a range of metrics that are in development, some of which could take advantage of samples/data collected at the start of any monitoring programme (e.g., soil eDNA, acoustic monitoring) but others would come online later (e.g., LIDAR derived hedge data).

8 Cost-effective data acquisition strategies

Where this report recommends farmer-led metric recording, then this would provide a whole population value that can be followed through time. However, where only a proportion of the population of farms/fields can be sampled there has to be a statistically sound design adopted. This would include a comparison between areas where measures have been applied with other similar areas that are not in the scheme (control sites). A Before-After-Control-Impact (BACI) design is commonly used for monitoring the effect of environmental interventions. One issue to be addressed is that areas which are originally selected as controls may join the scheme later, so starting with a larger control population may guard against this.

8.1.1. An example - Wales

In Wales, the Glastir Monitoring and Evaluation Programme (GMEP) sample consisted of a stratified random sample 150 "Wider Wales" 1 km squares and 150 targeted at priority areas for the agri-environment scheme. It should be noted, however, that the "Wider Wales" squares do include land which is in the scheme, and that even the targeted 1 km squares contain differing amounts of land where specific management options have been applied. As it was found that the "Wider Wales" squares had considerable coverage of the scheme, in the more recent ERAMMP National Field Survey the aim has been to capture as much in-scheme and counterfactual land as possible within the full set of 300 squares.

To allow sampling effort to be spread across years and provide both temporal and spatial coverage, a rolling monitoring programme was followed by GMEP, in which sites are revisited, for example, every four or five years but different sites are sampled in years two and three. This allows better spatial coverage than if each site was revisited every year, while at the same time providing a more powerful estimate of change over say a five-year period, than not revisiting sites at all. The GMEP scheme uses a four-year rolling monitoring programme. Countryside Survey is also now following a rolling programme. Power analysis for the GMEP scheme (Emmett et al., 2014) suggested that across a variety of metrics, around 45 squares per year was the minimum number that need to be monitored before losing significant power to detect change over a period of 8 years (two cycles of the rolling programme).

8.1.2. Other considerations

Soil nutrients and soil carbon would be most appropriately measured at the scale of fields within farms, as this is the level at which relevant measures are applied. On the other hand, surveys of birds and pollinators, which are mobile over a larger area, might be more appropriately recorded for parcels of land, such as 1 km squares, although it is unlikely that the same measures will have been applied consistently across a whole 1 km square. To provide a common spatial unit across different metrics, the GMEP survey used 1 km squares, but, as it is not possible to sample vegetation and soils over the entire square, five randomly placed plots in each square were used for vegetation monitoring and soil samples were taken from the same plots. Vegetation was also recorded in other plot types, for example, along boundaries and field margins.

If fields rather than squares are to be used for soil monitoring, a representative sample for a particular field or part of a larger field can be obtained by bulking individual cores. For example, in the Soil Nutrient Health Scheme in Northern Ireland samplers follow a 'W' shaped track and take 25 cores. This should give a good estimate of the mean for an individual field but unless replicate cores are analysed individually it does not provide an estimate of the variability within the field. As a result, it is not possible to determine whether a change in a particular field between two sampling occasions is statistically significant. Under the Scottish Government's National Test Programme 20% of arable and improved grassland can claim funding for soil testing each year. If this scheme is continued, it could mean complete coverage of all arable and improved grassland fields after five years. The recommendation of Scotland's Farm Advisory Service (FAS) is that larger fields should be divided into 4 ha units, potentially with the help of the 1:25000 soil map. This approach might provide sufficient replicate samples across a farm as a whole to allow a change to be detected on a specific farm, although it should be noted that unless a suitable control is available it may not be possible to attribute any change to particular measures and that soil carbon changes in response to measures might take longer than five years to be detectable.

9   A note on data

Monitoring across a range of outcomes and metrics will generate a considerable volume of data. This will require a significant investment in design and development of the databases and in the staff required for data curation.

Alongside the technical aspects of database curation and management, consideration should be given to who owns the data – whether the land manager as it concerns their land holding or the taxpayer as they paid for it, who has access to it? – a narrow access regime provides increased security, especially around GDPR, but wider access allows for a broader range of analyses to be carried out. Furthermore, an overall data controller/owner would likely need to be appointed to comply with GDPR. It should be possible to develop data frameworks, where analysis without direct access to locations is possible (similar to medical data where analysis is separated from any data identifying subjects) and comply with Freedom of Information requests. Arguably, data should follow FAIR data principles and be

open access as it has been funded from the public purse, as in the European Soil Observatory.

10 Conclusions

Some metrics will clearly be valuable in identifying the benefits of future agri-environmental management. For example, the collection of data on soil carbon and methane emissions clearly supports the Scottish Government's climate ambitions. Others will support policies regarding sustainability (soil erosion) and the health of Scotland's freshwater resources (reducing diffuse pollution). There is a mix of field data collection, farmer-collected data and modelled information with some usage of existing data.

It should be noted that several metrics have been identified that may only be proxies of the outcome they relate to, such as area of non-farmed habitats or pesticide usage, but they have the advantage of being based on already collected data with the cost savings this brings. Other metrics are still in development but should either be available by the start of the scheme or where samples can be collected for future analysis.

There are several outcomes that are closely related and need consideration together. Improving animal nutrition requires maintaining soil nutrients at a level where protein is not limiting growth. This may require the application of organic and/or inorganic fertilisers. However, excess nutrients can end up as N_2O and ammonia emissions from slurry and the leaching of nitrates into freshwater. Careful management to optimise nutrient use is, therefore, required to reach all the desired outcomes: improving soil nutrient content, reducing diffuse pollution, improving water and air quality, livestock nutrition and nutrient management.

A second set of outcomes are also closely related, those dealing with livestock genetics, health and nutrition, alongside reducing methane emissions. Improved efficiency across the livestock sector should increase margins but at the same time reduce the methane footprint of meat.

Some metrics are not useful in isolation and need to be collected as a set to be useful. This is particularly true for animal health, nutrition and genetics where a range of data on growth rates, milk yields, mortality, conception rates and replacement rates are needed to get a full picture.

The final choice of which metrics to collect will depend on the availability of resources to carry out the monitoring and the type of sampling philosophy adopted. Assembly, curation and analysis of the data will all add costs to metric collection but it is important to get the most out of the data. Data ownership is also a key consideration.

Given the division between farmer-led and expert-led monitoring highlighted in the spreadsheet MeasuresXMetrics.xlsx and in Section 6, we suggest the following:

- All enterprises to assess soil erosion and buffer strip effectiveness as this is highly site specific.

- All livestock enterprises to record growth rate, milk yields, mortality, conception rates, replacement rates, age at slaughter for sheep and cattle.
- ScotEID to require information on sires.
- All enterprises to use farm tool calculators to model GHG emissions. Livestock enterprises to model ammonia emissions when a suitable tool is available. The requirement to model might be limited to enterprises above a certain size to reduce costs.
- The remaining outcomes are best assessed using expert-led monitoring in a sample-based programme similar in philosophy to the Welsh approach. Resources available for monitoring and statistical power analysis would be a key part of how to structure this monitoring. They would also determine whether to focus on a small number of metrics and outcomes and cover a larger sample size, or to cover all outcomes on a smaller sample size. The outcomes monitored in this way include those focused on soils, waters and biodiversity.

11 References

Bieroza, M.Z., Bol, R. & Glendell, M. (2021) What is the deal with the Green Deal: Will the new strategy help to improve European freshwater quality beyond the Water Framework Directive? Science of The Total Environment, 791, 148080. https://doi.org/10.1016/j.scitotenv.2021.148080

Boag, B., Palmer, L.F., Neilson, R., Legg, R. & Chambers, S.J. (1997). Distribution, prevalence and intensity of earthworm populations in arable land and grassland in Scotland. Annals of Applied Biology, 130, 153-165. https://doi.org/10.1111/j.1744-7348.1997.tb05791.x

Brodie, E. (2023) Farm for Scotland's Future: the case for change. A report for Scottish Environment LINK. https://farmforscotlandsfuture.scot/wp-content/uploads/2023/04/Farm-for-Scotlands-Future-full-report.pdf

Carey, P.D., Wallis, S., Chamberlain, P.M., Cooper, A., Emmett, B.A., Maskell, L.C., McCann, T., Murphy, J., Norton, L.R., Reynolds, B., Scott, W.A., Simpson, I.C., Smart, S.M. & Ullyett, J.M. (2008) Countryside Survey: UK Results from 2007. NERC/Centre for Ecology & Hydrology, 105pp. (CEH Project Number: C03259) https://nora.nerc.ac.uk/id/eprint/5191/

Carpenter, D., Sherlock, E., Jones, D.T., Chiminoides, J., Writer, T., Neilson, R., Boag, B., Keith, A.M. & Eggleton, P. (2012). Mapping of earthworm distribution for the British Isles and Eire highlights the under-recording of an ecologically important group. Biodiversity and Conservation, 21, 475-485. https://doi.org/10.1007/s10531-011-0194-x

Chatterjee, A., Lal, R., Wielopolski, L., Martin, M.Z. & Ebinger, M.H. (2009) Evaluation of different soil carbon determination methods. Critical Reviews in Plant Science, 28, 164-178. https://doi.org/10.1080/07352680902776556

Cottle, D.J. & Conington, J. (2012) Breeding for reduced methane emissions in extensive UK sheep systems. Journal of Agricultural Science, 150, 570-583. doi:10.1017/S0021859611000967

Cottle, D.J. & Conington, J. (2013) Reducing methane emissions by including methane production or feed intake in genetic selection programmes for Suffolk sheep. Journal of Agricultural Science, 151, 6, 872-888. doi:10.1017/S0021859612001037

Emmett, B.E., Abdalla, M., Anthony, S., Astbury, S., August, T., Barrett, G., Biggs, J., Botham, M., Bradley, D., Brown, M., Carter, H., Chadwick, D., Cigna, F., Cooper, D., Cooper, J., Cosby, BJ., Creer, S Cross, P., Edwards, F., Edwards, M., Evans, C., Ewald, N., Fitton, A., Garbutt, A., Grebby, S., Greene, S., Halfpenney, I., Hall, J., Harrison, S., Harrower, C., Henrys, P., Hobson, R., Hughes, S., Isaac, N., Jackson, B., Jarvis, S., Jones, D., Keith, A., Kelly, M., Korenko, J., Lallias, D., Leaver, D., Lebron, I., Malcolm, H., Maskell, L., McDonald, J., Moxley, J., Norton, L., O'Hare, M., Owen, A., Pereira, M., Peyton, J., Powney, G., Pywell, R., Rawlins, B., Robinson, D.A., Rorke, S., Rowland, C., Roy, D., Scarlett, P., Scholefield, P., Scott, A., Scott, L., Scott, R., Siriwardena, G., Smart, S., Smith, P., Swetnam, R., Taylor, R., Tordoff, G., Van Breda, J., Vincent, H., Wagner, M., Waters, E., Watkins, J., White, J., Williams, B., Wood, C. & Wright, S. (2014) Glastir Monitoring & Evaluation Programme. First Year Annual Report to

Welsh Government (Contract reference: C147/2010/11). NERC/Centre for Ecology & Hydrology (CEH Project: NEC04780), pp.442. https://gmep.wales/sites/default/files/GMEP-First-Year-Annual-Report-2014-05-17.pdf

Environmental Standards Scotland (2022) Baseline Evidence Review – Water (Strategy and Analysis). https://www.environmentalstandards.scot/wp-content/uploads/2022/09/Baseline-Evidence-Review-Water-2022-09.pdf

Estrada-Carmona, N., Sánchez, A.C., Remans, R. & Jones, S.K. (2022) Complex agricultural landscapes host more biodiversity than simple ones: A global meta-analysis. Proceedings of the National Academy of Sciences, 119, e2203385119. https://doi.org/10.1073/pnas.2203385119

Giles, M.E., Caul, S., King, D., Mitchell, S., Sim, A., Neilson, R. & Paterson, E. (2023). Grass variety selection of microbial community composition is associated with differences in soil CO_2 emissions. Applied Soil Ecology, 190, 104968.

Glendell, M., Palarea-Albaladejo, J., Pohle, I., Marrero, S., McCreadie, B., Cameron, G. & Stutter, M. (2019) Modeling the ecological impact of phosphorus in catchments with multiple environmental stressors. Journal of Environmental Quality, 48, 1336-1346. https://doi.org/10.2134/jeq2019.05.0195

Harrison, S., McAree, C., Mulville, W. & Sullivan, T. (2019) The problem of agricultural 'diffuse' pollution: Getting to the point. Science of the Total Environment, 677, 700-717. https://doi.org/10.1016/j.scitotenv.2019.04.169

Hayes, E., Higgins, S., Geris, J., Nicholl, G. & Mullan, D. (2023) Weighted risk assessment of critical source areas for soil phosphorus losses through surface runoff mechanisms. Catena, 225, 107027. https://doi.org/10.1016/j.catena.2023.107027

Humphry, R., Skuce, P., Hayward, A., Henry, M., Hutchinson, I., Gunn, G. & Costa, M. (2021) Measurement of antimicrobial usage: what can we learn across Livestock Sectors? https://sefari.scot/research/measurement-of-antimicrobial-usage-what-can-we-learn-across-livestock-sectors

Insausti, M., Timmis, R., Kinnersley, R. & Rufino, M.C. (2020) Advances in sensing ammonia from agricultural sources. Science of The Total Environment, 706, 135124. https://doi.org/10.1016/j.scitotenv.2019.135124

IPCC (2021) Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Eds Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R., Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp. doi:10.1017/9781009157896.

Leinonen, I., Eory, V., MacLeod, M., Sykes, AS., Glenk, K., & Rees, RM. (2019). Comparative analysis of farm-based carbon audits. ClimateXChange. https://www.climatexchange.org.uk/media/3584/farm-based-carbon-audits-final.pdf

Neilson, R., Aitkenhead, M., Lilly, A. & Loades, K. (2021) Monitoring soil health in Scotland by land use category – a scoping study. James Hutton Institute. https://doi.org/10.7488/era/1293

Pakeman, R.J., Beaton, J., Fielding, D., Hewison, R., McKeen, M., Potts, J., Quinzo-Ortega, L., Stephen, T. & Stockan, J. (2020) Evaluation of the Biodiversity outcomes of the 2014-20 SRDP Agri-Environment Climate Scheme through a selection of case studies. Report to NatureScot. https://www.nature.scot/doc/naturescot-research-report-1254-evaluation-biodiversity-outcomes-2014-20-srdp-agri-environment.

Perry, J.N., Rothery, P., Clark, S.J., Heard, M.S. & Hawes, C. (2003) Design, analysis and statistical power of the Farm-Scale Evaluations of genetically modified herbicide-tolerant crops. Journal of Applied Ecology, 40, 17-31. https://doi.org/10.1046/j.1365-2664.2003.00786.x

Prout, J.M., Shepherd, K.D., McGrath, S.P., Kirk, G.J.D., Hassall, K.L. & Haefele, S.M. (2022) Changes in organic carbon to clay ratios in different soils and land uses in England and Wales over time. Scientific Reports, 12, 5162. https://doi.org/10.1038/s41598-022-09101-3.

Reaney, S.M., Mackay, E.B., Haygarth, P.M., Fisher, M., Molineux, A., Potts, M. & Benskin, C.M.H. (2019) Identifying critical source areas using multiple methods for effective diffuse pollution mitigation. Journal of Environmental Management, 250, 109366. https://doi.org/10.1016/j.jenvman.2019.109366

Saby, N.P., Bellamy, P.H., Morvan, X., Arrouays, D., Jones, R.J., Verheijen, F.G., Kibblewhite, M.G., Verdoodt, A.N.N., Üveges, J.B., Freudenschuss, A. & Simota, C. (2008) Will European soil-monitoring networks be able to detect changes in topsoil organic carbon content? Global Change Biology, 14, 2432-2442. https://doi.org/10.1111/j.1365-2486.2008.01658.x

Skuce, P., Holland, J., Kenyon, F., Wells, B., McCracken, D., Neilson, R. & Bartley, D. (2022) Anthelmintics and the Environment – opening a whole can of worms? https://sefari.scot/research/anthelmintics-and-the-environment-%E2%80%93-opening-a-whole-can-of-worms

Staggenborg, J. & Anthes, N. (2022) Long-term fallows rate best among agri-environment scheme effects on farmland birds—A meta-analysis. Conservation Letters, 15, e12904. https://doi.org/10.1111/conl.12904

Stutter, M., Costa, F.B. & Ó hUallacháin, D. (2021) The interactions of site-specific factors on riparian buffer effectiveness across multiple pollutants: A review. Science of the Total Environment, 798, 149238. https://doi.org/10.1016/j.scitotenv.2021.149238

Walton, P., O'Brien, D., Smart, J., Burns, F., Basset, D., Bradfer-Lawrence, T., Foster, S., James, B.D., Mancini, F., Mordue, S., Pakeman, R.J., Pescott, O.L., Simkin, J., Stanbury, A.J. & Towers, M. (2023) State of Nature Scotland 2023. The State of Nature Partnership. https://stateofnature.org.uk/wp-content/uploads/2023/09/TP26056-SoN-Scotland-summary-report-v5-1.pdf.pagespeed.ce.EIp-TYaoGQ.pdf

Wang, Y., Yao, Z., Zhan, Y., Zheng, X., Zhou, M., Yan, G., Wang, L., Werner, C. & Butterbach-Bahl, K. (2021) Potential benefits of liming to acid soils on climate change mitigation and food security. Global Change Biology, 27, 2807-2821. https://doi.org/10.1111/gcb.15607

Zhang, H.M., Liang, Z., Li, Y., Chen, Z.X., Zhang, J.B., Cai, Z.C., Elsgaard, L., Cheng, Y., van Groenigen, K.J. & Abalos, D. (2022) Liming modifies greenhouse gas fluxes from soils: A meta-analysis of biological drivers. Agriculture, Ecosystems & Environment, 340, 108182. https://doi.org/10.1016/j.agee.2022.108182

12 List of Appendices

12.1 Appendix A. List of measures under consideration

12.2 Appendix B. Correspondence between land-based measures and relevant outcomes

12.3 Appendix C. Potential monitoring metrics and methods

Click to see: Appendices

If you require the report in an alternative format such as a Word document, please contact info@climatexchange.org.uk **or 0131 651 4783.**

climate✕change

Scotland's centre of expertise connecting climate change research and policy

info@climatexchange.org.uk
+44(0)131 651 4783
@climatexchange_
www.climatexchange.org.uk

ClimateXChange, Edinburgh Climate Change Institute, High School Yards, Edinburgh EH1 1LZ