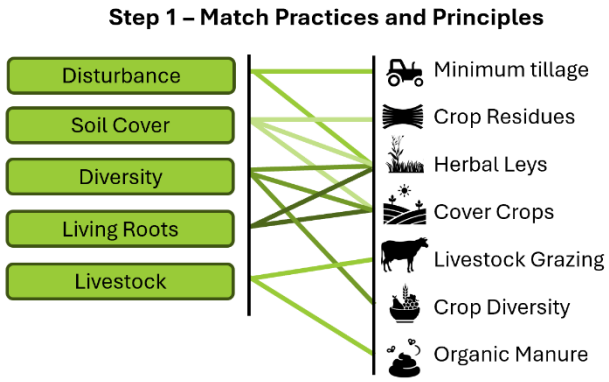
Supplementary Material: Measuring the Socio-Economic and Environmental Outcomes of Regenerative Agriculture across Spatio-Temporal Scales

By Katherine Berthon, Ruth N Wade, Pippa J. Chapman, Coline C. Jaworski, Jonathan R. Leake, Tim Daniell, Niamh M. McHugh, Lisa Collins, Yu Zhao, Penelope J. Watt, Bob Doherty, Peter Jackson, Lynn V. Dicks

Explanation of the H3 Scoring System

Step 1 – Match Principles to Practices

The first step in applying the H3 scoring system is to decide on a list of practices that are applicable for your farming context. In the H3 project, an initial list of nine practices was given by farmers (Berthon et al. 2024), but the scoring described in this paper includes seven of these practices. This is for two reasons: 1) spring cropping was not treated as a stand-alone practice as it is naturally applied at the same time as cover cropping and would be counted under crop diversification; and 2) detailed information on the percentage of area trafficked as a result of reduced compaction techniques was not available when compiling this score.



Practices in Full:

1. Use no-till / minimum non-inversion tillage,;
2. Retain crop residues in field;
3. Introduce herbal leys in the rotation
4. Use multi-species cover and catch cropping in the rotation;
5. Introduce livestock to arable rotations;
6. Introduce crop diversity in the rotation (especially legumes);
7. Use organic manures/green compost/ digestate/ compost;

Once practices are defined for a farming context, they are matched to the relevant regenerative principles. In our example, minimising tillage achieves the principle minimising disturbance. Maintaining soil cover is achieved by retaining crop residues and use of herbal leys, or cover crops over winter. Diversity is achieved maximising crop diversity in the rotation, including through use of mixed species herbal leys or cover crops. An additional score can be given at the field level and calculated based on rotational diversity to ensure the same fields are not being cropped consistently with the same crops but we have simplified the scoring here. Maintaining living roots year round is achieved through the use of cover crops and herbal leys. Livestock integration is achieved through grazing livestock on cover crops, leys or wheat over winter, and/or the occasional application of manure.

Step 2 – Calculate the frequency of practice implementation based on the cropping history

Once the alignment of principles and practices is decided, the frequency of each practice can be calculated at either the field or farm level as appropriate to its realistic use in the rotation. In this way, each practice gets a score for how effectively it is being applied on a farm, scaled between 0 and 1, where 0 is complete absence of the practice and 1 represents consistent and frequent application. Some practices have multiple 'levels' and the relative score can be weighted accordingly. For example, the goal for the H3 project was zero/no tillage, but minimum tillage has been shown to have benefits compared to plough systems (Li et al. 2020). To recognise this nuance, instead of using a binary, with 0 indicating the field was ploughed, and 1 indicating direct drilled, we allowed half scores for fields that had minimum tillage applied. The frequency score then represents more of an average intensity of tillage than the consistent application or not of tillage on a given field. While overall, the relative frequency of each practice is determined from the area of fields in which it is employed, it must be scaled relative to the maximum fields available or the expected frequency of application. For example, herbal leys are only present in one or a few fields per year and are given a binary score. Similarly, crop diversity is scaled relative to the number of fields present so as to not

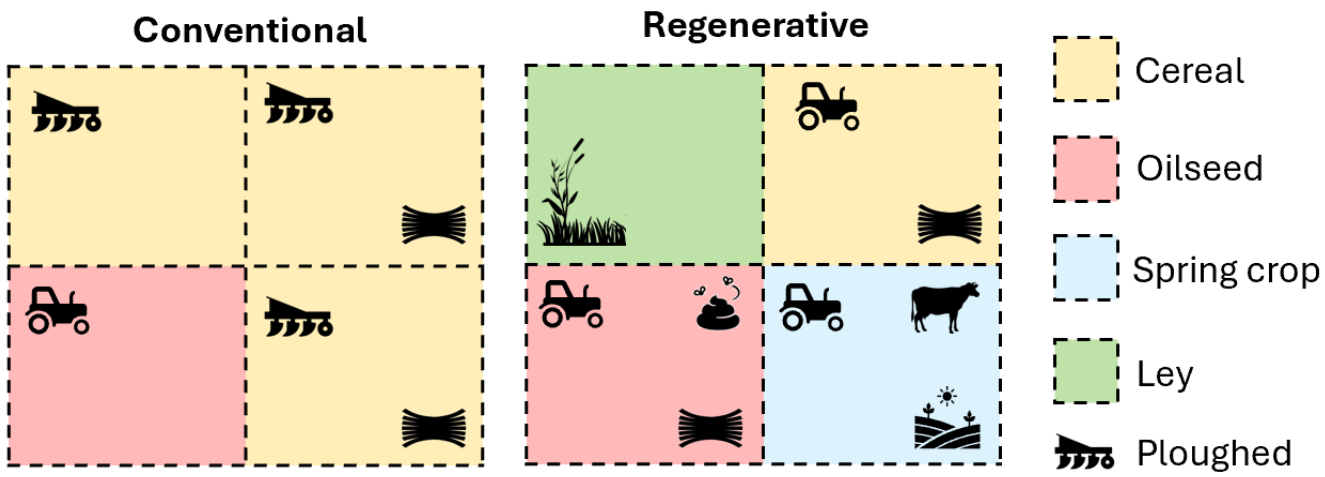
penalise smaller farms. Table 1 provides a full list of the practices and how their scores are calculated for a given year.

Table 1. Metrics for calculating individual practice scores for a subset of the practices listed in the paper.

Practice	**Calculation**	**Explanation**
Min Till (MT)	$MT = \frac{sum(Tillage\ binary\ per\ field * field\ area)}{Total\ field\ area}$	A score of 1 is for when all fields are direct drilled for a given year – score of half for when min tilled and 0 when all ploughed
Herbal Leys (HL)	$HL = if(any\ field), yes = 1, no = 0$	A maximum Herbal ley score is given when there is at least one herbal ley on the farm for that year – 0 if not. The score can be adjusted for the quality of the cover crop or ley, relating to its growth during a given season.
Cover Crops (CC)	$CC = \frac{Area\ with\ cover\ crop}{Total\ field\ area}$	A maximum score if every cash crop field contains a cover or catch crop. The score can be adjusted for the quality of the cover crop or ley, relating to its growth during a given season.
Crop Residue (CR)	$CR = \frac{Area\ with\ residue\ retained}{Total\ field\ area}$	A maximum score is given where all cash crop fields have straw retained, or have living mulch present instead
Crop Diversity (CD)	$CD = \frac{Number\ of\ crops}{Total\ fields}$	CD – a maximum score, if all fields are different crops in a given year, scaled by the number of possible crop types available. If the number of fields is less that the minimum number of crops, the number of fields is taken to scale the overall diversity score. If there are more fields than are possible to have in different crops (i.e. large farms with >10 fields), the total possible crops can be used to average across fields instead. In the H3 dataset, over the last five years farmers have grown 10 different types of crops, wheat, barley, oats, oilseed, linseed, Beans, Peas, Sunflowers, and Poppies and Herbal Leys, and the blocks contain up to 9 fields, so the number of fields was used in our example.
Livestock Integration	$LI = \frac{area\ grazed}{Total\ field\ area}$	A maximum score is given if every field either has livestock grazing applied
Organic Manure	$OM = \frac{area\ manure\ added}{Total\ field\ area}$	A maximum score is given if every field has had addition of organic manure applied in the last three years

Consider the following farms, one conventional and one regenerative, and their differential implementation of the seven practices above. The conventional farm contains three fields of wheat which are ploughed and one OSR field that is min tilled during establishment. They retain straw on two of the wheat fields but bale the third. The regenerative farm grows a different crop in each field, contains a herbal ley in the rotation, grows cover crops behind their spring crop, and uses minimum tillage during planting. The herbal ley is also untilled, and in application, the

scoring would be related would be included as a min till field for real scoring (as in Figure 4). All fields are the same size here for ease of calculation, but the score would be relatively adjusted for the area in which the practice was applied, as a proportion of the farm area as in Table 1.



Frequency of each practice (no weighting added as in Table 1):

Conventional	0.25	0	0.5	0	0	0.5	0
Regen	0.75	0.25	0.5	0.5	0.25	1	0.5

Step 3 - calculate five principle scores on the basis of practice implementation

For each cropping season, you can then generate a score for each principle from the relative frequency (i.e. the average scores) of its constituent practices, where 0 is when no practices are used i.e. the principle is not being implemented, and 1 means all practices are being applied consistently, and the principle is being effectively implemented at the farm scale. The relative weighting and collation of practices is slightly different for each principle, as there is large functional overlap between practices.

Principle 1 – minimise disturbance

Practices: Min till or no till (MT) and herbal leys (HL)

$$P1 = \frac{MT + HL}{2}$$

Notes- the minimum tillage score is a cumulative score of the weighted tillage score for each field

Principle 2 – Soil covered

Practices: Cover Crops (CC), Crop Residues Retained (CR), Herbal Ley (HL)

$$P2 = \frac{CR + (CC\ OR\ HL)}{2}$$

Notes - Cover crops and Herbal Leys are functionally the same for keeping the soil covered, so farms that implement either a herbal ley or a cover crop to keep the soil covered are scored the same.

Principle 3 – Diversity

Practices: Cover Crops (CC), Crop Diversity (CD), Herbal Ley (HL)

$$P3 = \frac{CC + CD + HL}{3}$$

Notes – to avoid a maximum score being achieved with fields that have the same crops in the same fields (but a different crop in every field), scoring should include a measure of rotational diversity, or in-field diversity to ensure diversity at the field level as well as the farm level. We have simplified this metric for demonstration.

Principle 4 – Living Roots

Practices: Herbal Leys, Cover Crops

$$P4 = \frac{CC + HL}{2}$$

Notes – living mulches can also promote living roots, and can be added as an additional practice here

Principle 4 – Living Roots

Practices: Livestock Grazing (LI), Organic Manure Additions (OM)

$$P5 = \frac{LI + OM}{2}$$

Notes – organic matter addition may not happen on the same fields that livestock are grazed (it is an either or) but the organic matter score should represent the inclusion of organic matter on that field at an appropriate application rate

Step 4 - **monitor changes across time by building a cumulative score across years and across principles**

The final scoring for each principle depends on the practices chosen and the temporal span of the data collated. This can be achieved by aggregating the principles into a unified score across farming seasons, or by relating individual principle scores to outcomes. While the former may be more useful from a policy or certification standpoint, it is likely that the latter is more informative from a mechanistic perspective as some principles are more relevant to achieving certain outcomes. Regenerative agriculture is a process with no defined end point, and the scoring system allows for relative tracking of the implementation of regenerative practices across time using a rolling 3 to 5 year window. Alternatively, the scoring system to see the cumulative benefits of implementation of regenerative farming in the longer term with more consistent implementation across years resulting in higher scores. In both cases, a separate score can be calculated for each principle for each harvest year then averaged across years.

The consideration of practices under multiple principles naturally gives higher weight to practices that may achieve multiple regenerative goals but does not weight them relative to the extent of their impact on outcomes. For example, herbal leys and cover crops contribute to diversity, soil cover and living roots, but some cover crops or leys may be more diverse mixes of species and provide greater soil benefits. Modification of the scoring could weight cover crops or leys according to their species mix and their purported benefits (e.g. grass vs legume leys) and users should be especially careful if using combined scoring across principles.

Next Steps – Testing the effect of different principle scores on environmental outcomes

The H3 experiment will use the scoring system to assess the impact of transitioning to regenerative practices, by correlating changes in the relative farm scores to changes in environmental outcomes for biodiversity and soil health. This can be achieved by aggregating the principles into a unified score across farming seasons, or by relating individual principle scores to outcomes (**Figure 4,** main text). Methods for collection of earthworm data are given in Berthon et al. 2024a). To test the impact of principle scores on outcomes, we used a five-year weighted score using historical data from 2018-2022 for H3 farms. Scores were weighted using the following formula to achieve a score standardised between 0-1:

$$FiveYrWeight = \frac{Score[2018] * 1 + Score[2019] * 2 + Score[2020] * 3 + Score[2021] * 4 + Score[2022] * 5}{15}$$

We performed a series of linear regressions testing the impact of each principle score, and a combined overall score with equal weighting for each of the principles. Future work could consider weighting the principles in an overall score to better match with the evidence for their impact on different outcomes. Data and code to run the models are provided as separate supplementary material files.

The model outputs behind Figure 4 for regression of principles vs worm density are given in the below table:

Model	**Intercept**	**Est (β)**	**SE**	**statistic**	**p-value**
Soil Cover	7.543	17.322	6.554	2.643	0.0165
Minimise Disturbance	7.833	12.064	5.430	2.222	0.0394
Living Roots	15.307	-3.356	6.188	-0.542	0.594
Diversity	16.267	-3.500	5.872	-0.596	0.559
Livestock Integration	14.599	-1.345	11.695	-0.115	0.91
Combined Score	10.823	1.979	2.131	0.929	0.365