



# Rotational grazing and multispecies herbal leys increase productivity in temperate pastoral systems – A meta-analysis

Matthew W. Jordon<sup>a,\*</sup>, Kathy J. Willis<sup>a</sup>, Paul-Christian Bürkner<sup>b</sup>, Gillian Petrokofsky<sup>a,c</sup>

<sup>a</sup> Department of Zoology, University of Oxford, Oxford OX1 3SZ, United Kingdom

<sup>b</sup> Cluster of Excellence SimTech, University of Stuttgart, Stuttgart, Germany

<sup>c</sup> Oxford Systematic Reviews LLP, Oxford OX2 7DL, United Kingdom

## ARTICLE INFO

### Keywords:

Sustainable Intensification  
Agroecology  
Grazing management  
Legume  
Diverse swards  
Systematic review

## ABSTRACT

Reducing greenhouse gas emissions associated with ruminant livestock production is important for climate change mitigation. Regenerative Agriculture (RA) practices are increasingly promoted to improve forage production and livestock performance in temperate livestock systems. These practices include i) rotational grazing (RG) of livestock around multiple subunits of pasture to achieve ungrazed periods of ‘rest’, and ii) herbal leys (HL), where perennial forbs such as chicory, lucerne and trefoils are included as components in multi-species swards. While there are plausible mechanisms for adoption of these practices to improve agricultural productivity, quantitative syntheses of their impacts are required. Here, we conduct a systematic review and meta-analysis of the effects of RG and HL practices on herbage dry matter (DM) production, animal daily live-weight gain (DLWG), and sheep wool growth in temperate oceanic regions. We use quantitative predictors in our Bayesian hierarchical models to investigate the role of rest period and stocking density in RG systems, and specific plant traits and sward diversity in HL. We found that herbage DM increased by 0.31 t.ha<sup>-1</sup> over a growing season as the proportion of rest in an RG grazing system increased from 0 to 1. Stocking density significantly moderated the effect of rest period on sheep and cattle DLWG; at higher stocking densities, longer rest periods were required to maintain livestock growth rates. In HL studies, herbage DM yielded 1.63 t.ha<sup>-1</sup> more per metre of increased sward root depth and a sward entirely comprised of legumes yielded 2.20 t.ha<sup>-1</sup> more than when no legumes were present. Sheep DLWG increased by 3.50 g.day<sup>-1</sup> per unit increase in leaf nitrogen concentration (mg.g<sup>-1</sup>), but we could not determine an effect of leaf condensed tannin content on animal performance. Although there remain differences between the RG and HL study treatments meta-analysed here and RA in practice, our results provide empirical support for some of the mechanisms attributed to increased pasture and livestock productivity following adoption of selected RA grazing practices.

## 1. Introduction

Global food production is responsible for 30% of anthropogenic greenhouse gas emissions, causing climate change (Clark et al., 2020). Ruminant livestock significantly contribute to these emissions, particularly through methane produced via enteric fermentation (Gerber et al., 2013; Godfray et al., 2018; Herrero et al., 2011). The short-lived nature of methane in the atmosphere has led to calls for reduced livestock production in order to deliver rapid emissions reduction as a quick win for climate change mitigation (Smith and Balmford, 2020). Alternatively, or in addition, avenues to reduce emissions associated with ruminant-sourced foods include: enhancing livestock health and fertility

to reduce livestock numbers required for a given level of food produced, using feed additives to inhibit enteric methane production, increasing growth rates to reduce lifetime emissions, and improved feed quality and digestibility which lowers methane production (Herrero et al., 2016; Hristov et al., 2013). These latter two options can be pursued at least in part through adopting management practices which improve productivity in sheep and cattle grazing systems.

Simultaneously, Regenerative Agriculture (RA) practices are rapidly gaining attention as a means to improve pasture and livestock productivity in temperate grazing systems through enhancing soil health and promoting ecosystem functioning. These include rotational grazing (RG), which can be defined as “the movement of livestock between two

\* Corresponding author.

E-mail address: [matthew.jordon@zoo.ox.ac.uk](mailto:matthew.jordon@zoo.ox.ac.uk) (M.W. Jordon).

<https://doi.org/10.1016/j.agee.2022.108075>

Received 14 February 2022; Received in revised form 13 June 2022; Accepted 20 June 2022

Available online 28 June 2022

0167-8809/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

or more subunits of pasture such that alternating periods of grazing and no grazing (rest) occur within a single growing season” (Briske et al., 2011b).<sup>1</sup> Further, herbal leys (HL) incorporate perennial forbs and legumes into pasture swards, including chicory (*Cichorium intybus*), birdsfoot trefoils (*Lotus corniculatus* and *L. pedunculatus*), lucerne (*Medicago sativa*), plantain (*Plantago lanceolata*), sainfoin (*Onobrychis viciifolia*), sulla (*Hedysarum coronarium*) and yarrow (*Achillea millefolium*) e.g. (Barry, 1998; Li and Kemp, 2005, 1996; Douglas 1986; John & Lancashire 1981; Stewart 1996) with the aim of benefiting from certain properties of these species. These practices occur on spectra of grazing management and sward diversity, respectively (Fig. 1). Conventional lowland livestock management practices in temperate countries such as the United Kingdom (UK) often comprises continuous grazing or ‘set stocking’ of low diversity pastures comprising a small number of grass species, predominantly perennial ryegrass (*Lolium perenne*), timothy (*Phleum pratense*), cocksfoot (*Dactylis glomerata*) and fescues (*Festuca spp.*), in addition to white and red clovers (*Trifolium repens* and *T. pratense*, respectively) (Wilkins and Jones, 2000). More ‘regenerative’ pasture management increases sward diversity to include a mix of grasses, legumes and forbs (either by sowing these or creating the conditions for them to persist or re-establish naturally) and increases the proportion of the growing season that pastures are rested (usually accompanied by higher stocking density when grazed), i.e. moving further along the conceptual axes in Fig. 1.

Interest in RA is relatively recent (Giller et al., 2021), but RG and incorporation of forbs into pastures (HL) have been promoted for over a century in temperate regions, for example to restore degraded rangeland in North America and as part of ley farming systems in Western Europe, respectively (Voisin, 1959; Briske et al., 2011b; Elliot, 1908; Turner, 1951).

RG is purported to increase livestock productivity (e.g. animal live-weight gain, milk production, or wool growth) and pasture carrying capacity via enhancing forage productivity. Periods of rest following removal of livestock grazing pressure are thought to allow enhanced root development which enables rapid regrowth of herbage following infrequent defoliations, compared to smaller root systems and permanently low leaf photosynthetic area under continuous grazing pressure (Sanderman et al., 2015; Hacker, 1993; Voisin, 1959; Savory and Butterfield, 2016). Further, decreasing the area available to livestock at any one time through subdivision of pasture into paddocks may reduce selectivity of animal grazing, preventing less palatable species from dominating the sward and deteriorating forage quality, plus achieving more uniform distribution of livestock manure and urine across the pasture (Briske et al., 2008; Norton, 1998). These effects are believed to be accentuated under longer rest periods, typically corresponding with shorter grazing periods and higher stocking densities to match forage availability with livestock requirements. However, although these mechanisms are plausible and convincing to many land managers, they are predominantly based on scientific theory, and the currently available scientific evidence has been found to be inconclusive for USA rangelands (Briske et al., 2008; Briske et al., 2011a; Briske et al., 2011b).

HL (and greater sward diversity more broadly) may enhance livestock productivity both ‘indirectly’ via improved pasture productivity but also directly via specific properties of the herbage of individual species, compared to grass-only swards. Pasture productivity is increased predominantly through enhanced niche complementarity due to greater species diversity, with mechanisms including i) greater

<sup>1</sup> We use ‘rotational grazing’ throughout as a catch-all term for all grazing systems that fall within this definition, including mob grazing, cell grazing, paddock grazing, controlled grazing, holistic planned grazing, strip grazing or precision grazing. We use proportion rest period and stocking density as quantitative predictors in our meta-analysis (see Methods, Section 2.2.1), as the simple categories of continuous vs rotational grazing are inadequate to capture the diversity of practices covered by this term.

variation and depth in rooting structures conferring drought tolerance, thus stabilising forage quality and supply throughout the grazing season (Cranston et al., 2015), and improving nutrient uptake through accessing different soil profiles (Li and Kemp, 2005; Stewart, 1996), ii) legumes fixing atmospheric nitrogen (Luscher et al., 2014) and in turn increasing N availability to other sward components (Suter et al., 2015), iii) increased resilience of production across varying growing conditions (Skinner and Dell, 2016; Sanderson et al., 2005), and iv) functional redundancy (Weisser et al., 2017). Herbage from these swards can also have a higher nutritive value, including increased crude protein content (Cranston et al., 2015; Luscher et al., 2014), improved palatability leading to higher uptake (Wilkins and Jones, 2000; Burke et al., 2002), and enhanced mineral content (Barry, 1998), which may contribute to increased livestock productivity.

Some herb species also contain secondary metabolites such as condensed tannins (CT) which can protect protein from degradation in the rumen, thus potentially increasing uptake in the animals’ intestines. Although plausible, this mechanism may be limited in practice because i) CTs vary in their bioactivity and impact on forage palatability and digestibility, mediated by their concentration, molecular structure and the wider dietary composition, ii) recent studies indicate CTs may simply shift excretion of dietary nitrogen excretion from urine to faeces rather than increase net N uptake, iii) CTs may only confer a nutritional benefit when dietary protein exceeds animal requirements in which case other nutrient deficits may limit performance, and iv) CT-rich species tend to compete poorly in swards resulting in limited inclusion in the diet (Grosse Brinkhaus et al., 2016; Mueller-Harvey et al., 2019; Loza et al., 2021). Nevertheless, CTs can reduce intestinal parasitic worm burdens, and methane and nitrous oxide emissions associated with livestock production (Fox et al., 2018; Luscher et al., 2014; Mueller-Harvey et al., 2019).

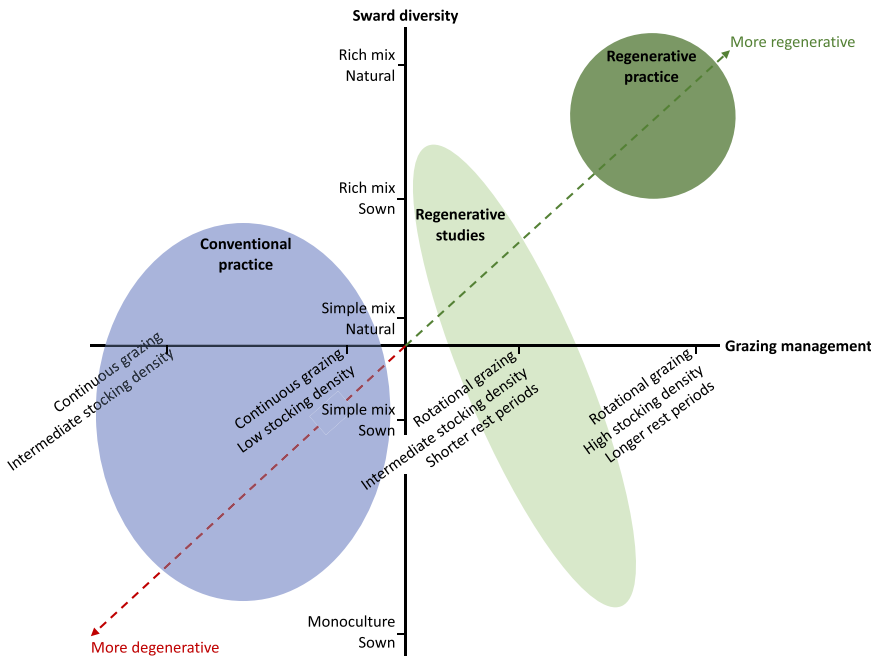
Despite compelling and plausible mechanisms for these regenerative grazing practices to increase productivity, there is currently a paucity of quantitative syntheses that test these relationships (Briske et al., 2011b; Conant et al., 2017). Here, we provide a meta-analysis of data presented in previously published studies to examine the impacts of RG and HL on pasture and livestock productivity in temperate oceanic regions (Köppen-Geiger Cfb), conducting a systematic review to assemble a database of relevant studies. We test the hypothesis that these regenerative practices increase plant and animal productivity, using quantitative predictors in our Bayesian hierarchical analyses to evaluate possible mechanisms for this. We aim to establish whether sufficient evidence exists to promote adoption of selected regenerative grazing practices in temperate regions to deliver the benefits currently attributed to these.

## 2. Methods

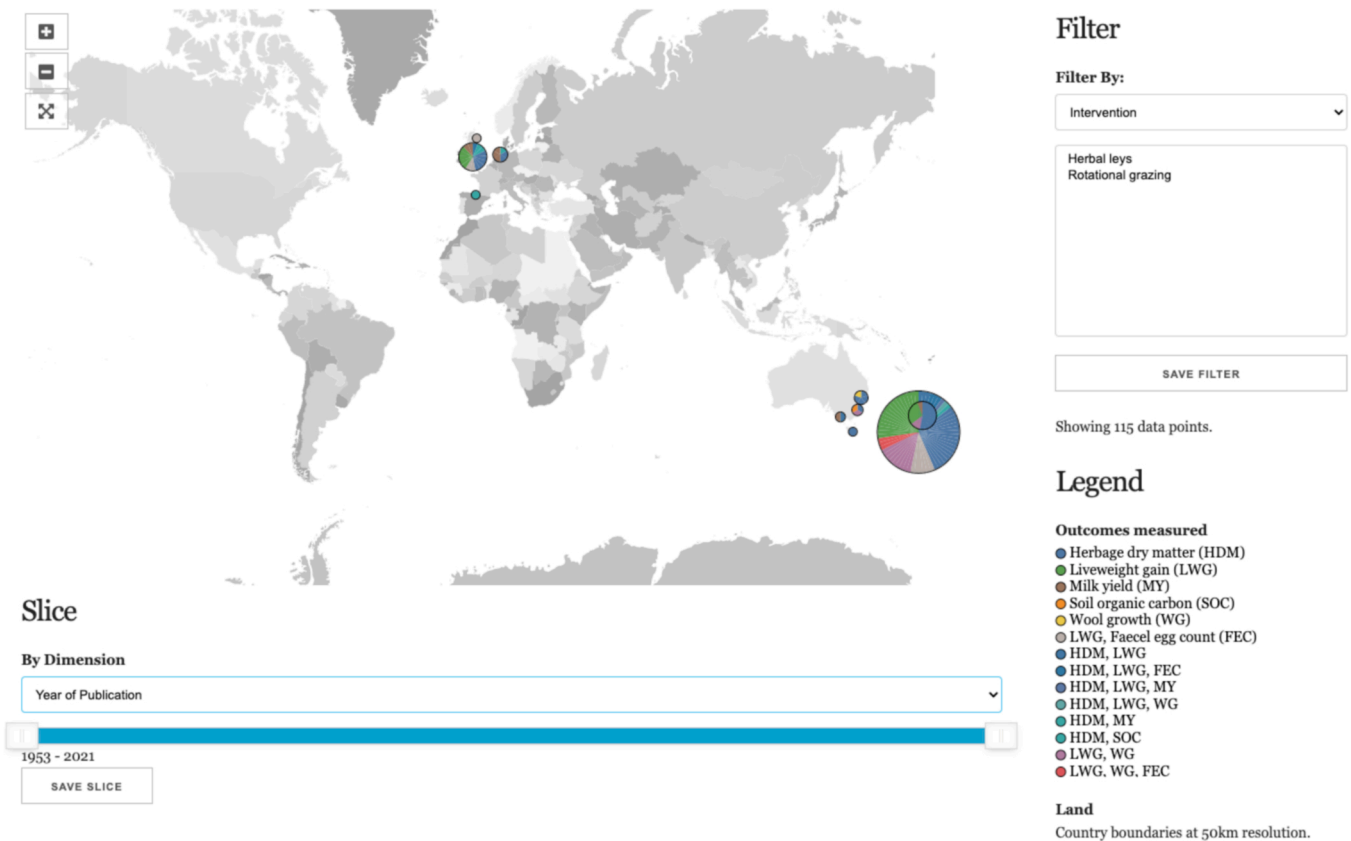
### 2.1. Systematic review

We followed the Collaboration for Environmental Evidence guidelines (CEE, 2018) to address the systematic review question “What are the impacts on soil carbon and farm productivity from adopting rotational grazing practices and incorporating perennial forbs into pastures (herbal leys) in temperate oceanic sheep and cattle farming systems?”, using the Population, Intervention, Comparator, Outcome and Location (PICOL) framework (Table S1). Full details of our systematic review following the Reporting standards for Systematic Evidence Syntheses (ROSES) framework (Haddaway et al., 2017b) are given in the Supplementary Materials. All data extracted from relevant studies of RG and HL are provided in Supplementary Data 1 and 2, respectively, and further supplementary files and R code are available online in the Zenodo repository (Jordon, 2022).

We conducted searches in Web of Science, CAB Abstracts and Scopus (details in Table S2, Supplementary Methods 1.1) and undertook ‘reverse snowballing’ of citations from reference lists of included articles



**Fig. 1.** Conceptual diagram of spectra of grazing management intensity and pasture sward diversity. Rotational grazing and herbal ley practices fall along the x and y axes, respectively. These interact particularly in the lower left and upper right quartiles, where continuous grazing pressure can reduce sward diversity by removing palatable or sensitive species, and mob grazing can increase the natural sward diversity through unselective grazing and periods of rest, respectively. Note that the management implemented by Regenerative Agriculture practitioners is often of a higher intensity on these axes compared to published studies of rotational grazing and herbal leys analysed here. Although not displayed, the positive or negative effects of these practices are likely to saturate or plateau with time.



**Fig. 2.** Evidence map. 115 relevant studies identified by systematic review process for inclusion in meta-analysis, created using the Thalloo framework (Martin, 2018). Position of pie charts reflects study locations (degrees decimal coordinates), size of pie charts is proportional to the number of studies in that region (or the site when zoomed in online), and the colour of the chart segments shows the number of studies of each intervention (see legend). Inset shows southern Hemisphere studies. An interactive version of this evidence map with the accompanying study database is available online at [https://oxlel.github.io/evidencemaps/agricultural\\_productivity/](https://oxlel.github.io/evidencemaps/agricultural_productivity/).

until no additional relevant records were returned (Table S3, Supplementary Methods 1.1.1). We screened records at title, abstract and full text levels using pre-determined inclusion and exclusion criteria (Table S4, Supplementary Methods 1.2). Data from relevant studies were extracted to a spreadsheet (Supplementary Data 1&2, Supplementary Methods 1.3), with any extra calculations to convert data to the required format documented in this datasheet and conducted in individual Excel files for each article, available online (Jordon, 2022). Each study was assigned a critical appraisal score reflecting the quality of experimental design (Table S5). As it is well-established in systematic reviewing that not all evidence is of equal quality and therefore validity (CEE, 2018), we adapted a critical appraisal scoring approach from Haddaway et al. (2017a) as a means of testing the sensitivity of our meta-analysis results to studies with the highest risk of bias (low or unclear validity, Supplementary Materials 1.3 & 2.2). This is not an assertion that results from such studies are invalid, but rather to ascertain the strength of conclusions from our analysis. Where desired data was missing from articles (Tables S6 and S7), we attempted to contact the corresponding author with a request for data (Supplementary Methods 1.4). We received responses from authors of five articles, of which three were able to provide additional information and are thanked in the Acknowledgments.

Our systematic review resulted in a database of 84 articles containing 115 studies across 9 countries with temperate oceanic regions (Fig. 2 and S1). From this, we extracted 101 observations from studies of rotational grazing vs set stocking, and 485 observations from studies of herbal vs conventional leys (Supplementary Data 1&2, respectively) for quantitative meta-analysis or narrative synthesis (Table 1). Where values for required predictor variables were not provided in articles and could not be retrieved from study authors, these observations were excluded from meta-analyses, resulting in fewer observations analysed than present in the dataset (number of observations specified in results table for each analysis).

**Table 1**

Number of studies that measured each outcome for rotational grazing and herbal leys identified by systematic review. Values in brackets are number of studies where standard errors are presented for estimates. The sum of outcomes listed here is greater than the number of studies identified by the systematic review because some studies present multiple outcomes.

Outcome	Livestock type	Unit	Intervention	
			Rotational grazing	Herbal leys
Soil organic carbon	n/a	g 0.100 g <sup>-1</sup> t.ha <sup>-1</sup>	–	1 (1)
Productivity measure	Herbage dry matter	n/a	2 (2)	1 (1)
	Daily liveweight gain	Sheep	13 (3)*	42 (14)
				*
			5 (2)*†	50 (28)
				*
			3 (3)*†	1 (1)
Milk yield	Sheep	kg.day <sup>-1</sup>	–	1 (1)
	Cattle		4 (0)	5 (1)
	Sheep	mg.cm <sup>-2</sup> .day <sup>-1</sup>	–	7 (5) *
Wool growth	Sheep	g.day <sup>-1</sup>	3 (0)	7 (6) *

- No studies report this intervention-outcome combination.

\* Intervention-outcome combinations where sufficient data available to conduct quantitative meta-analyses.

† Daily liveweight gain for cattle and sheep under rotational grazing vs set stocking was analysed together, with livestock type as a fixed effect in the analysis.

## 2.2. Meta-analysis

We fitted Bayesian hierarchical (i.e. random effects) models using the *brms* package in R version 4.0.3 (Bürkner, 2017; Bürkner, 2018; Stan Development Team, 2019; R Core Team, 2020). We analysed herbage dry matter (t.ha<sup>-1</sup>), livestock daily liveweight gain (g.day<sup>-1</sup>) and wool growth (mg.cm<sup>-2</sup>.day<sup>-1</sup> and g.day<sup>-1</sup>) data extracted from studies, as these were the systematic review outcomes (Table S1) with sufficient data available for meta-analysis (Table 1). The data points we analysed (rows in our dataset) were outcome means per treatment per study in our systematic review. We did not compute comparative effect size metrics (i.e. between treatments, or between control and treatment), because there are substantial differences in treatment interventions both within and between studies. Instead, the comparison was directly encoded via the hierarchical structure of our Bayesian models. Further, the interventions analysed here are best expressed using continuous rather than categorical predictors (see Sections 2.2.1 and 2.2.2), enabling a more informative analysis. Within each response variable, data across all treatments and studies are directly comparable on the same scale and our approach has the advantage that outputs from the model are readily understood. In addition, we centred all continuous predictors around their respective means so that the model output intercept was biologically meaningful rather than corresponding to predictor values of zero.

Our R code is available online (Jordon, 2022). Details of model sampling are given in Table S8 and model summary outputs are given in the Supplementary Materials. We used default priors of *brms*, provided online (Jordon, 2022), which are either weakly informative or uninformative to reduce the risk of incorrectly specified priors biasing model outputs. We checked model convergence using the Rhat diagnostic and ensured effective sample size measures were sufficiently large (Vehtari et al., 2020). Model non-convergence was remedied by increasing the number of iterations for sampling and divergent transitions were addressed by decreasing the sampler step size (Stan Development Team, 2020). We assessed the statistical significance of fixed effect model predictors based on whether their 95% credible intervals included zero and used Bayes R<sup>2</sup> to estimate the proportion of variation explained by the overall model and fixed effects only (Gelman et al., 2019). Both between- and within-study heterogeneity was modelled in the form of corresponding standard deviation parameters (across studies, and across effect sizes within studies, respectively). We imputed standard errors of study effect sizes where these were missing and conducted sensitivity analyses to test the robustness of our methodology (Supplementary Materials 2.2.).

We plotted the conditional effects of model predictors on productivity outcomes, showing regression lines for individual predictors and interaction terms where all other model predictors are at the reference category, with 95% Credible Intervals. We also displayed the raw productivity data from the underlying studies for each intervention using the *forestplot* package (Gordon and Lumley, 2020) (Supplementary Materials 2.2). To test for possible publication bias, we conducted Egger's regression test for funnel plot asymmetry using the *regtest* function in *metafor* (Viechtbauer, 2010) for productivity data with errors present (EP) and present the results from this along with funnel plots of herbage dry matter and DLWG effect sizes against their standard errors (Supplementary Methods 2.5). However, we did not interpret the results from these further, as there are multiple potential sources of funnel plot asymmetry in our data (most notably, substantial heterogeneity between studies and treatments), of which publication bias is only one possible explanation, making this test potentially misleading in either direction.

### 2.2.1. Rotational grazing

The key aspects of grazing systems that are likely to affect pasture productivity and livestock performance are the duration of grazing and recovery periods, and the stocking density during the grazing period (Voisin, 1959; Briske et al., 2008; Briske et al., 2011a; Techio Pereira et al., 2018). Simply analysing grazing practice as a predictor with two levels (rotational vs continuous grazing) is inappropriate for the diversity of management practices included within RG (Sanderman et al., 2015). Therefore, we extracted quantitative parameters of treatments from each study to include in our analyses. Where studies provided a range of values for grazing and rest periods in RG treatments due to management flexibility to match grass growth, we used the median value of this range in analysis. Continuous grazing treatments have no rest period, informatively captured in the analysis as zero, but it is difficult to meaningfully represent their grazing period due to i) differences in grazing season length between studies, and ii) grazing and rest

---


$$\text{Herbage DM} \sim \text{Root depth} * \text{Legume} + \text{Species richness} + \text{Species evenness} + \text{Latitude} + (1|\text{Unique study ID})$$

$$\text{Liveweight gain} \sim \text{Root depth} + \text{Leaf N} + \text{Species richness} + \text{Species evenness} + \text{Latitude} + (1|\text{Unique study ID})$$


---

periods in RG treatments typically sum to a total rotation length shorter than the grazing season such that there are multiple rotations per season, making comparison of grazing periods with continuous treatments difficult. Further, grazing and rest periods in our dataset are negatively correlated for RG treatments, making it challenging for our meta-analysis model to identify both predictors. Therefore, we instead expressed grazing practices as the proportion of rest period in the system. This captures information about both the grazing and rest periods, although reduces this to a relative relationship thus sacrificing some information about the absolute magnitude of rest periods.

We fitted the following models as specified in R formula syntax:

$$\text{Herbage DM} \sim \text{Rest period} + \text{Latitude} + (1|\text{Unique study ID})$$

$$\text{Liveweight gain} \sim \text{Rest period} * \text{Stocking density} + \text{Livestock type} + (1|\text{Unique study ID})$$

$$\text{Livestock type} + (1|\text{Unique study ID})$$

where.

- *Herbage DM*, study treatment measurement of herbage dry matter ( $\text{t. ha}^{-1}$ ) accounting for its sample size (see [Supplementary Materials 2.2](#)),
- *Liveweight gain*, study treatment measurement of sheep daily liveweight gain ( $\text{g.day}^{-1}$ ), accounting for its standard error,
- *Rest period*, proportion of time within rotation that paddock rested between grazing in rotational system, automatically zero for set stocking or continuous grazing treatments,
- *Stocking density*, Livestock Units per ha, harmonised using coefficients based on animal feed requirements (Defra, 2010; Sac Consulting, 2020), for the area being grazed at any one time (i.e. individual paddock for RG vs whole field for set stocking),
- *Livestock type*, two-level factor (sheep or cattle),
- *Latitude*, absolute Latitude of the study site (decimal degrees),
- *Unique study ID*, included as a random intercept to account for the hierarchical structure of the data

### 2.2.2. Herbal leys

The purported productivity benefits of including perennial herbs in pasture swards in addition to grasses and clovers are attributed to

specific properties of these species, including root depth, nitrogen fixation and leaf crude protein and CT contents (see Introduction). Different herbs commonly included in HL differ in these properties, and studies differ in the herb composition and seed mixes included in their experimental treatments. Because analysing productivity outcomes using herb presence/absence as a categorical predictor would not indicate which features of herbs, if any, influenced productivity, we extracted traits values for root depth, plant nitrogen fixation and leaf nitrogen content from the TRY database (Kattge et al., 2020), and calculated aggregated trait values in addition to species richness and evenness (inverse Simpson's diversity index) for multi-species swards ([Supplementary Materials 2.1](#)). Leaf CT content was not available from TRY for any forage plant species in our dataset so we analysed this separately ([Section 2.2.2.1](#)).

We fitted the following models:

$$\text{Wool growth} \sim \text{Leaf N} + (1|\text{Unique study ID})$$

where.

- *Herbage DM*, study treatment measurement of herbage dry matter ( $\text{t. ha}^{-1}$ ) accounting for its standard error,
- *Liveweight gain*, study treatment measurement of sheep daily liveweight gain ( $\text{g.day}^{-1}$ ), accounting for its standard error,
- *Wool growth*, study treatment measurement of wool growth ( $\text{mg. cm}^{-2}.\text{day}^{-1}$  or  $\text{g.day}^{-1}$ ), accounting for its standard error,
- *Root depth*, aggregate root depth for sward (m), weighted average by relative abundance of species present,
- *Leaf N*, aggregate leaf nitrogen content per leaf dry mass ( $\text{mg.g}^{-1}$ ) of sward, weighted average by relative abundance of species present,
- *Legume*, aggregate score for nitrogen fixation capacity, corresponding to abundance-adjusted proportion of sward that is leguminous (values 0–1),
- *Species richness*, number of species present in sward or seed mix,
- *Species evenness*, inverse Simpson's diversity index for sward or seed mix,
- *Latitude*, absolute Latitude of the study site (decimal degrees),
- *Unique study ID*, included as a random intercept to account for the hierarchical structure of the data

We included fewer predictors in the meta-analysis of wool growth data due to less observations of this productivity outcome ([Table 1](#)). We only analysed liveweight gain measured for sheep, because the eight cattle observations would be insufficient to identify a *Livestock type* predictor, as in the RG analysis.

To test whether livestock productivity was being influenced by forage availability or other features of the sward, we fitted an additional model for studies that measured both herbage dry matter and livestock daily liveweight gain:

$$\text{Liveweight gain} \sim \text{Herbage DM} + \text{Leaf N} + (1|\text{Unique study ID})$$

where the variables are the same as above, apart from *Herbage DM* which does not account for measurement standard error due to the difficulty of representing this within a predictor and *Liveweight gain* is weighted by

sample size (n) rather than standard errors (Supplementary Materials 2.2).

**2.2.2.1. Condensed tannins.** Although leaf CT content per leaf dry mass ( $\text{mg.g}^{-1}$ ) trait data was not available from the TRY database, 15 studies in our dataset measured the tannin content of forage available in their treatments (i.e. whole-forage analysis rather than only CT-rich components of sward). We used these data firstly to investigate the effect of CTs on sheep liveweight gain by fitting the following models:

$$\text{Liveweight gain} \sim \text{Leaf CT} * \text{PEG} + \text{Latitude} + (1|\text{Unique study ID})$$

$$\text{Liveweight gain} \sim \text{Leaf CT} * \text{Leaf N} + \text{Latitude} + (1|\text{Unique study ID})$$

where

- *Liveweight gain*, study treatment measurement of sheep daily liveweight gain ( $\text{g.day}^{-1}$ ), accounting for its standard error,
- *Leaf CT*, measured leaf condensed tannin content per leaf dry mass ( $\text{mg.g}^{-1}$ ) of sward,
- *Leaf N*, aggregate leaf nitrogen content per leaf dry mass ( $\text{mg.g}^{-1}$ ) of sward, weighted average by relative abundance of species present,
- *PEG*, polyethylene glycol (binds and inhibits CTs), administered to animals (yes/no),
- *Latitude*, absolute Latitude of the study site (decimal degrees),
- *Unique study ID*, included as a random intercept to account for the hierarchical structure of the data

Studies that measured herbage CT content but did not contain PEG treatments were coded as “no” for the PEG predictor. We fit these models separately to test different hypotheses, i.e. i) does CT inhibition moderate the effect of high-tannin forages on DLWG, and ii) does leaf CT concentration moderate the effect of leaf protein content on DLWG. Secondly, we used these data to calculate average CT contents for the species or mixes available (Table 2). To further understand the implications of plant CT content on animal performance, we investigated their influence on internal parasite burden. Sixteen studies of daily liveweight gain in sheep on HL in our dataset included faecal egg counts (FEC,  $\text{eggs.g}^{-1}$ ), a measure of internal parasitic worm burden such as

**Table 2**

Estimates of total tannin content (not extractable, protein-bound or fibre-bound), offered to livestock (not selected). For tannin measurements with standard errors, we fitted an intercept-only model in *brms* that generated ‘meta-estimates’ of tannin content accounting for these standard errors. However, because not all studies presented error terms for tannin measurements, and we anticipate the disadvantage of a smaller sample size outweighs the benefit of accounting for standard error in the average score, we also calculated the arithmetic mean and used this in our analysis model. We present estimates from both approaches for comparison, with 95% Credible Intervals provided for meta-estimates.

Species	Model	Tannin estimate ( $\text{g.kg}^{-1}$ DM)	n	
			Obs	Studies
Birdsfoot trefoil ( <i>Lotus corniculatus</i> )	Meta-estimate	25.52 [25.05, 25.96]	13	5
	Mean	26.35	13	5
Greater birdsfoot trefoil ( <i>Lotus pedunculatus</i> )	Meta-estimate	–	0	0
	Mean	34.44	8	6
Lucerne ( <i>Medicago sativa</i> )	Meta-estimate	0.30 [0.27, 0.34]	8	3
	Mean	0.50	13	5
Sulla ( <i>Hedysarum coronarium</i> )	Meta-estimate	45.10 [44.61, 45.61]	5	2
	Mean	28.53	10	4
Perennial ryegrass, white clover ( <i>Lolium perenne</i> , <i>Trifolium repens</i> )	Meta-estimate	1.25 [0.61, 1.87]	4	2
	Mean	1.38	4	2

*Nematodirus* and *Trichostrongylus*. We extracted 73 FEC observations from these studies and fitted the following model with FEC as a predictor to verify the effect of worm burden on sheep growth rates:

$$\text{Liveweight gain} \sim \text{FEC} + \text{Latitude} + (1|\text{Unique study ID})$$

where

- *Liveweight gain*, study treatment measurement of sheep daily liveweight gain ( $\text{g.day}^{-1}$ ), accounting for its standard error,
- *FEC*, sheep faecal egg count ( $\text{eggs.g}^{-1}$ ),
- *Latitude*, absolute Latitude of the study site (decimal degrees),
- *Unique study ID*, included as a random intercept to account for the hierarchical structure of the data

We then investigated whether leaf CT content reduced the worm burden of sheep grazing that herbage, for treatments with sward species compositions present in Table 2:

$$\text{FEC} \sim \text{Leaf CT} * \text{Wormer} + \text{Latitude} + (1|\text{Unique study ID})$$

where

- *FEC*, study treatment measurement of faecal egg count ( $\text{eggs.g}^{-1}$ ), accounting for its sample size (n),
- *Leaf CT*, average leaf condensed tannin content per leaf dry mass ( $\text{mg.g}^{-1}$ ) of sward, using values from Table 2,
- *Wormer*, anthelmintic treatment administered to animals (yes/no),
- *Latitude*, absolute Latitude of the study site (decimal degrees),
- *Unique study ID*, included as a random intercept to account for the hierarchical structure of the data

We included an interaction term between the leaf CT and Wormer predictors in the model. This is because if leaf CT content did effect FEC, we would expect this impact to be overridden by anthelmintic treatment. Studies that measured FEC and for which leaf CT were available but did not contain information on sheep anthelmintic treatment were coded as level “no” in the Wormer predictor.

### 3. Results and discussion

#### 3.1. Rotational grazing

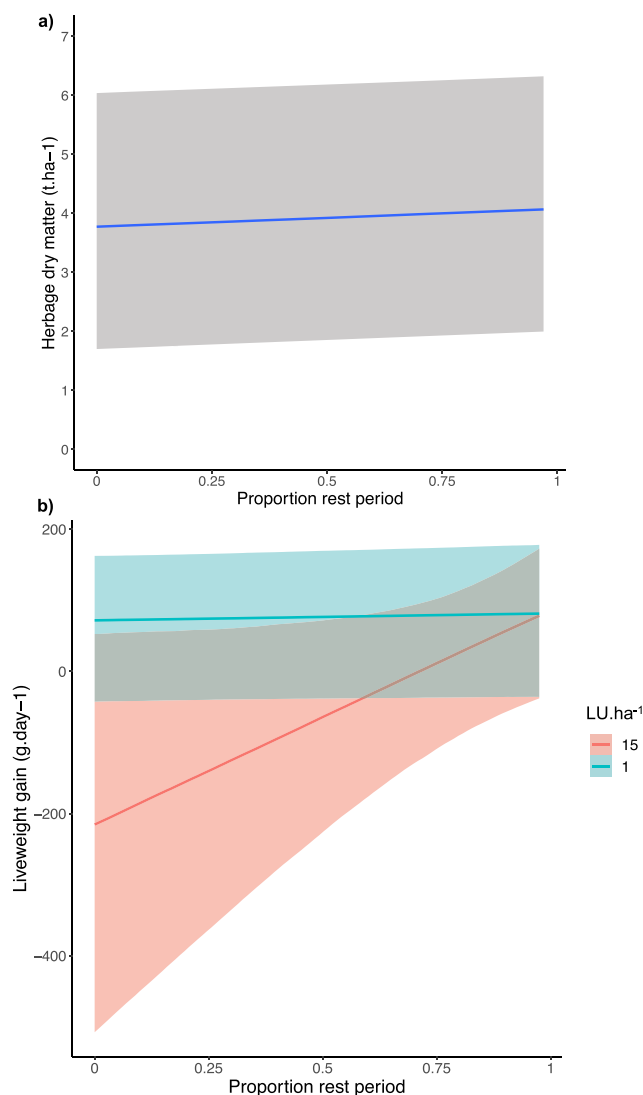
We found that increasing the proportion of the grazing season that pasture is rested as a result of RG improves herbage dry matter (DM) production, with an increase of  $0.31 \text{ t.ha}^{-1}$  (95% Credible Intervals, CI, [0.28, 0.33], Table 3, Fig. 3a) in DM between rest proportions of 0 and 1, corresponding to continuous grazing and continuous rest, respectively. This effect was robust to removal of studies with low or unclear validity (Table S9), and accords with results from previous meta-analyses (Badgery and Michalk, 2017; McDonald et al., 2019). However, our result should be treated with caution as too few studies presented measurement standard errors to enable us to account for these in the analysis (Supplementary Materials 2.2) and the fixed effects accounted for very little variation in the data ( $R^2$  0.0751, Table 3) with most heterogeneity explained by between-study variation. The effect of rest period on livestock daily liveweight gain (DLWG) changes with stocking density; at low stocking densities the proportion of rest period has little influence on DLWG, but at higher stocking densities rest period positively predicts DLWG (Fig. 3b). This interaction is significant at the level of 95% Credible Intervals, [0.78, 41.6] (Table 3), but should also be treated with caution as most studies analysed were of low or unclear validity (too few studies of high validity to conduct a sensitivity analysis) and the beneficial effect of rest was not preserved in the sensitivity analysis of studies that presented standard errors (Table S9).

Nevertheless, our DLWG findings accord with expectations from theoretical mechanisms of pasture productivity under RG and our results

**Table 3**  
Model estimates of herbage dry matter (DM, t.ha<sup>-1</sup>) and sheep and cattle daily liveweight gain (DLWG, g.day<sup>-1</sup>) for studies of rotational grazing. 95% Credible Intervals are given in square brackets, with \* denoting where these do not overlap with 0. The - symbol denotes predictors not included in that model. All continuous predictors were centred before analysis. Full model summary outputs presented in [Supplementary Materials 2.3.1](#).

Outcome	Model†	Intercept	Proportion rest period	Stocking density	Rest period*stocking density	Cattle	Latitude	Standard deviation parameters		R <sup>2</sup>	Fixed effects	Data points	Studies
								Within studies	Between studies				
Herbage DM (t.ha <sup>-1</sup> )	WN	3.99 [1.89, 6.21]	0.31 [0.28, 0.33]	-	-	-	-0.07 [-0.40, 0.31]	0.33 [0.33, 0.34]	3.59 [2.28, 5.82]	0.882	0.0751	19	9
DLWG (g.day <sup>-1</sup> )	EI	-39.4 [-179, 80.1]	221 [2.18, 448]	-11.1 [-22.4, -0.40]	20.8 [0.78, 41.6]	787 [596, 978]	-	14.1 [5.40, 25.4]	108 [61.9, 197]	0.985	0.911	35	8

† WN: meta-analysis weighted by estimate sample size because too few standard errors available, EI: standard errors imputed where missing, see [Supplementary Materials 2.2](#) for details.



**Fig. 3.** Productivity effects of rotational grazing. Conditional effects of a) rest period (proportion of grazing season) on herbage dry matter production (DM, t.ha<sup>-1</sup>), and b) rest period (proportion of grazing season) and stocking density (LU.ha<sup>-1</sup>) on sheep and cattle daily liveweight gain (g.day<sup>-1</sup>). Error bars show 95% Credible Intervals.

from the herbage dry matter analysis. At low stocking densities, a period of rest is expected to have little influence on DLWG due to a sufficient forage availability not limiting animal growth, whereas at higher stocking rates longer rest periods are required to allow herbage recovery to meet livestock requirements in the next grazing period. Although not tested here, it is also possible that livestock performance could begin to reduce in systems with the longest rest periods due to declining forage quality with age, mediated by a less favourable ratio of soluble to structural cell components in older tissues (Briske et al., 2008). A previous synthesis from New Zealand identified a disjoint between improved pasture performance not being reflected in improved animal growth (Brougham, 1971), which our results suggest can potentially be explained by the stocking density under consideration. Our findings contrast with results from a recent global meta-analysis and rangeland research trials which typically find that herbage production and animal performance under RG is worse or no better than continuous grazing (Hacker, 1993, Norton, 1998, Briske et al., 2008; Mcdonald et al., 2019). This may be explained by differences in forage type and climate changing the influence of rest on forage and livestock productivity (Briske et al., 2011a). For example, temperate oceanic regions in our

systematic review typically have relatively consistent precipitation throughout the grazing season, allowing forage to utilise periods of rest to recover, compared to more arid rangeland environments where rest periods can achieve only limited forage recovery at low rainfall times of year (Briske et al., 2008). It has also been advocated that to obtain the greatest productivity benefit from RG, paddock rest periods need to be varied according to the season to account for changes in forage growth rate (Voisin, 1959), although many studies in our dataset applied the same grazing and rest periods across the whole grazing season.

The increase in DM production with lengthened rest period, and the implication from Fig. 3 that livestock growth rates can be maintained as stocking density increases if rest periods are sufficiently long, appears to suggest that RG can support higher stocking densities than continuous grazing, as has been found in Australia (Badgery and Michalk, 2017) although disputed in a rangeland context (Hacker, 1993; Briske et al., 2008; Hawkins, 2017). In our analysis, however, rest period and stocking density are non-independent. Experimental studies in our dataset often allocated equal total pasture area and animal numbers to RG and continuous grazing (CG) treatments. The RG treatment pasture area is then sub-divided into paddocks with the livestock confined to one of these at a time on rotation, compared to having access to the whole CG pasture area. Rest period is therefore positively correlated to stocking density across RG treatments in our database, as all else being equal increasing rest period requires creating more paddocks by reducing individual paddock size, thus increasing stocking density for a given paddock (Voisin, 1959), rather than overall increasing the livestock carrying capacity of the whole grazing platform. It is therefore difficult to conclude whether high stocking densities on small paddocks and frequent moves confer any actual increase in livestock carrying capacity over lower stocking densities on larger paddocks with less frequent moves, from the analysis conducted here.

We were unable to fully explore suggested mechanisms for RG practices to increase productivity with our dataset. Although it was necessary to use a proportion to capture grazing and rest period information in our analysis, this prevented us from testing the effect of specific grazing duration on forage and livestock productivity. RA practitioners frequently highlight the importance of frequent moves (<1–3 days) to prevent forage regrowth being cropped in the same grazing period, which is believed to negatively impact plant recovery (Voisin, 1959; Savory and Butterfield, 2016) although some have suggested that the importance of this has been overstated (Hacker, 1993). In addition, RA practitioners often apply a heuristic principle that retaining at least 50% of herbage prevents root death as a plant stress response following grazing, thus enabling more rapid herbage recovery from this maintained root and leaf architecture (Savory and Butterfield, 2016; Brown, 2018). However, we were not able to extract information on the herbage and root biomass before and after each grazing period in rotational systems to determine the influence of the proportion of biomass left as residual on herbage recovery.

### 3.2. Herbal leys

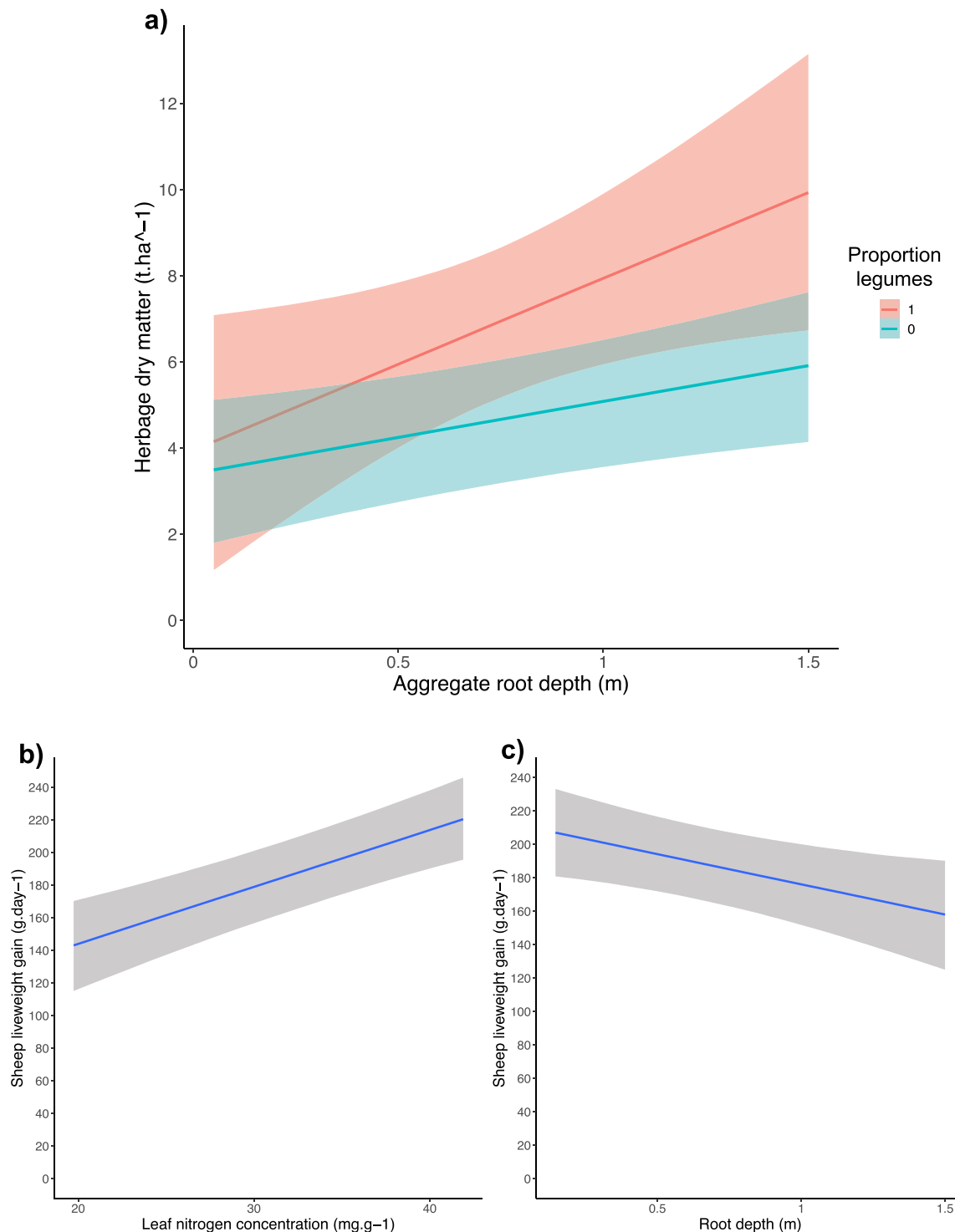
We found herbage dry matter increased by 1.63 t.ha<sup>-1</sup>, 95% CI [0.36, 2.89], per metre increase in sward average rooting depth. There was also an increase of 2.20 t.ha<sup>-1</sup>, 95% CI [1.26, 3.14], when the sward was entirely comprised of legumes compared to legumes not present (Table 4), although note this is purely in terms of interpreting the model result. In practice, there are agronomic limitations to swards entirely comprised of legumes and diversity also positively predicted production (next paragraph). These findings confirm expectations from previous reviews (Cranston et al., 2015; Luscher et al., 2014; Li and Kemp, 2005; Stewart, 1996). Further, the rate of yield increase with greater rooting depth appears to be enhanced when legumes are present (Fig. 4a), which would be expected from the simultaneous removal of two key limiting factors on plant growth (water and nitrogen). However, we were not able to confirm this positive interaction in our analysis, 95% CI [- 0.61,

**Table 4**  
Model estimates of herbage dry matter (DM, t.ha<sup>-1</sup>), sheep daily liveweight gain (DLWG, g.day<sup>-1</sup>) and sheep wool growth (mg.cm<sup>-2</sup>.day<sup>-1</sup> and g.day<sup>-1</sup>) for studies of herbal leys. 95% Credible Intervals are given in square brackets, with \* denoting where these do not overlap with 0. The - symbol denotes predictors not included in that model. All continuous predictors were centred before analysis. Full model summary outputs for herbage DM and DLWG presented in Supplementary Materials 2.3.1.

Outcome	Model†	Intercept	Root depth	Legume	Root‡legume	Leaf N	Herbage DM	Species richness	Species evenness	Latitude	Standard deviation parameters		R <sup>2</sup>	Number of		
											Within studies	Between studies		Whole model	Fixed effects	Data points
Herbage DM (t.ha <sup>-1</sup> )	EI	7.97 [6.05, 9.86]	1.63 [0.36, 2.89] *	2.20 [1.26, 3.14] *	2.33 [-0.61, 5.27]	-	-	0.13 [-0.04, 0.29]	2.76 [0.95, 4.56] *	-0.55 [-0.90, -0.20] *	1.44 [1.21, 1.69]	5.77 [4.59, 7.30]	0.931	0.232	174	41
	EI	188 [166,210]	-36.3 [-64.4, -7.28] *	-	-	3.50 [2.17, 4.82] *	-	4.72 [-12.5, 22.2]	-12.48 [-91.54, 64.9]	5.43 [0.74, 10.17] *	46.0 [40.3, 52.5]	70.8 [54.5, 91.0]	0.640	0.199	184	46
Wool growth	WN	260 [211,306]	-	-	-	2.80 [2.21, 3.39] *	-4.62 [-5.89, -3.35] *	-	-	-	38.2 [36.8, 39.7]	74.4 [45.4, 122]	0.697	0.0958	34	9
	EP (mg.cm <sup>-2</sup> .day <sup>-1</sup> )	0.80 [-0.04, 1.71]	-	-	-	0.01 [0.00, 0.01] †	-	-	-	-	0.04 [0.01, 0.09]	0.56 [0.06, 2.15]	0.828	0.163	11	3
EP (g.day <sup>-1</sup> )	EP (g.day <sup>-1</sup> )	16.0 [11.6, 20.41]	-	-	-	0.23 [0.03, 0.42] *	-	-	-	-	1.76 [1.01, 3.03]	5.71 [3.10, 10.8]	0.835	0.162	17	6

† EI: standard errors imputed where missing, EP: standard errors presented in study, WN: meta-analysis weighted by estimate sample size because too few standard errors available, see Supplementary Materials 2.2 for details.  
‡ Although marginally significant, divergent transitions when model ran so results not reliable





**Fig. 4.** Productivity effects of herbal leys. Conditional effects of a) sward average root depth (m) and legume abundance on herbage dry matter production (DM, t.ha<sup>-1</sup>), and b) leaf nitrogen concentration per leaf dry mass (mg.g<sup>-1</sup>) and (c) sward average root depth (m) on sheep daily liveweight gain (g.day<sup>-1</sup>). Error bars show 95% Credible Intervals.

5.27] (Table 4), and the positive impact of root depth was not preserved in sensitivity analyses (Table S10). Although HL are purported to stabilise forage production throughout the grazing season by providing better growth in dry conditions (Cranston et al., 2015), we were unable to consider any differences in timing of growth between study treatments in our analyses due to few studies providing sufficient temporal resolution of herbage dry matter measurements.

Species evenness (inverse Simpson's diversity index) of the sward

also positively predicts dry matter production, with an increase of 2.76 t.ha<sup>-1</sup>, 95% CI [0.95, 4.56], between indices of 0 and 1. This aligns with evidence of 'overyielding' in mixtures compared to monocultures from plant biodiversity studies in both experimental plots (Weisser et al., 2017; Tilman et al., 2001; Hector et al., 1999) and agricultural grasslands (Nyfeler et al., 2009; Finn et al., 2013), but was not robust to sensitivity analyses here (Table S10). We did not detect a significant increase in herbage production with higher sward species richness, 95%

CI [- 0.04, 0.29] (Table 4). This result could be due to i) the high number of monoculture treatments in our dataset including particularly high-yielding species such as lucerne (Douglas, 1986), and ii) many of the studies in our database are of short duration, so may not have been exposed to the environmental stresses that result in mixes performing better on average (Sanderson et al., 2005). Herbage production is

negatively related to absolute latitude (Table 4), as would be expected within temperate oceanic regions with shorter growing seasons and cooler temperatures further from the equator.

Livestock DLWG increased by  $3.50 \text{ g}\cdot\text{day}^{-1}$ , 95% CI [2.17, 4.82] (Table 4, Fig. 4b), for each milligram increase in leaf nitrogen content per gram of leaf dry mass ( $\text{mg}\cdot\text{g}^{-1}$ ). This is expected given dietary crude

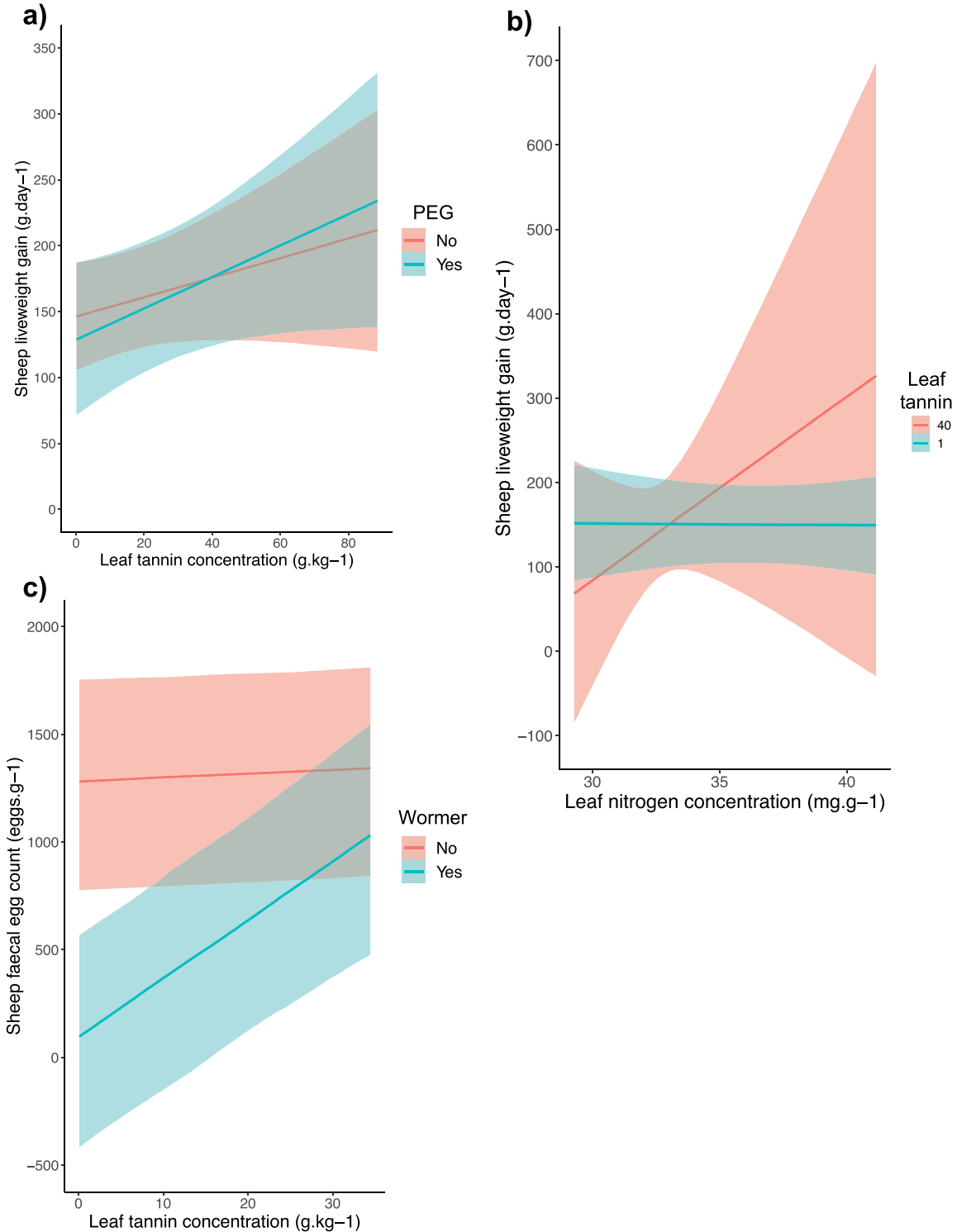


Fig. 5. Sheep growth and internal parasite effects of leaf condensed tannins (CT). Conditional effects of a) interaction between leaf CT concentration per forage dry mass ( $\text{g}\cdot\text{kg}^{-1}$ ) and polyethylene glycol (PEG, inhibits tannins) treatment on sheep daily liveweight gain ( $\text{g}\cdot\text{day}^{-1}$ ), b) interaction between leaf nitrogen concentration per leaf dry mass ( $\text{mg}\cdot\text{g}^{-1}$ ) and leaf CT concentration per forage dry mass ( $\text{g}\cdot\text{kg}^{-1}$ ) on sheep daily liveweight gain ( $\text{g}\cdot\text{day}^{-1}$ ), and c) interaction between leaf CT concentration per forage dry mass ( $\text{g}\cdot\text{kg}^{-1}$ ) and anthelmintic treatment (wormer) on sheep daily liveweight gain ( $\text{g}\cdot\text{day}^{-1}$ ). Error bars show 95% Credible Intervals.

protein content boosts animal growth rates (Cranston et al., 2015; Luscher et al., 2014), although i) leaf nitrogen concentration analysis is not a direct proxy of the complete protein content of forage, ii) animal performance is influenced by protein that reaches the small intestine for absorption (i.e. protected from degradation in the rumen) rather than simply ingested protein, and iii) forage needs to contain the correct protein:carbohydrate ratio to enable optimal animal performance. A positive relationship with leaf N was also observed in the limited data available on wool growth ( $\text{mg.cm}^{-2}.\text{day}^{-1}$ ,  $\text{g.day}^{-1}$ , Table 4, Fig. S2), potentially mediated by dietary essential amino acid availability (Barry, 1998). It is important to note that our result supports use of high leaf-nitrogen forages to promote livestock growth, rather than HL per se; some forage species monocultures that achieve the highest sheep DLWGs are already widespread in conventional sward mixes, e.g. white clover (Fig. S3). DLWG decreased in deeper rooting swards by  $36.3 \text{ g.day}^{-1}$ , 95% CI [- 64.1, - 7.62] (Table 4, Fig. 4c), per metre of sward average rooting depth. This is arguably unexpected given the enhanced mineral content of some deep-rooted herb species (Li and Kemp, 2005; Stewart, 1996), but is potentially explained by the lower forage quality of some deeper rooting grasses such as cocksfoot (*D. glomerata*) compared with high-quality but shallow rooted species such as clovers (Fig. S3). These effects of leaf N and root depth were robust to sensitivity analyses (Table S10).

Further, species richness and evenness did not appear to affect livestock growth rates (Table 4). Although the benefits of plant diversity to forage production may not be anticipated to translate directly to livestock DLWG, there are several reasons to expect species mixtures to deliver better animal performance than monocultures. Most forage species have disadvantages when sown in monoculture (Stewart, 1996; Burke et al., 2002) which can be mitigated when grown in mixtures, for example by providing a more optimal dietary carbohydrate to protein ratio or plants containing CTs such as trefoils reducing instances of clover-induced bloat in livestock (Luscher et al., 2014). However, in well-designed seed mixes the benefits of such species complementarity are likely to saturate at a low total species richness which could explain why our results do not support hyper-diverse mixes strictly for improving livestock performance. Forage digestibility, as influenced by sward management such as grazing regime, likely masks any effect of species richness and evenness, and lower diversity swards are in general easier to manage to maintain digestibility. Indeed, improved milk yield when grazing diverse swards has been linked to higher palatability and therefore greater intake rather than higher forage quality per se (Jonker et al., 2019; Loza et al., 2021). Although highly diverse swards are widely promoted by RA practitioners, this is often to deliver multiple objectives which include increasing yield stability, promoting soil microbiological activity and building soil carbon (see Section 3.3) rather than further enhancing forage productivity or liveweight gains per se. The subset of our data that measured both livestock DLWG and herbage dry matter for the same treatments identified a negative relationship between these two outcomes; DLWG decreased by  $4.62 \text{ g.day}^{-1}$  for every extra  $\text{t.ha}^{-1}$  of dry matter production, 95% CI [- 5.89, - 3.35] (Table 4). Taken together, these results confirm the well-established notion that forage quality (for which we used leaf nitrogen content as a proxy) drives livestock growth when forage quantity (herbage dry matter) is not limiting. Finally, absolute Latitude positively predicts DLWG (Table 4). Although this could potentially be due to reduced animal heat stress in cooler regions, we restricted our meta-analyses to studies from temperate oceanic regions (Cfb, Köppen-Geiger classification) in an effort to control for the effect of climate on productivity outcomes and this finding could be spurious given it was not supported in the EP and CA sensitivity analyses (Table S10).

We did not find a significant effect of leaf CT concentration ( $\text{mg.g}^{-1}$  of leaf dry mass) on sheep DLWG. This relationship did not change when tannin action was inhibited through livestock treatment with polyethylene glycol (PEG) which bind CTs (Fig. 5a, Table 5). However, CT concentration alone is known to be a poor indicator of bioactivity and

**Table 5**  
Model estimates of sheep daily liveweight gain (DLWG,  $\text{g.day}^{-1}$ ) and sheep faecal egg count (FEC,  $\text{eggs.g}^{-1}$ ) for studies of herbal leys which measured either leaf condensed tannin concentration of treatment forages or sheep internal parasite burden. 95% Credible Intervals are given in brackets, with \* denoting where these do not overlap with 0. The - symbol denotes predictors not included in that model. All continuous predictors were centred before analysis. Full model summary outputs presented in Supplementary Materials 2.3.1.

Outcome	Model <sup>1</sup>	Intercept	Tannin	PEG	Tannin*PEG	Leaf N	Tannin*leaf N	FEC	Wormer	Tannin*Wormer	Latitude	Standard deviation parameters			R <sup>2</sup>	Fixed effects	Number of
												Within studies	Between studies	Whole model			
DLWG ( $\text{g.day}^{-1}$ )	EI	158 [120,197]	0.74 <sup>‡</sup> [- 0.38, 1.86]	-10.14 [- 54.5, 32.4]	0.45 [- 1.15, 2.06]	-	-	-	-	-	-28.1 [- 55.3, - 0.90]*	53.7 [39.9, 72.3]	63.1 [36.0, 101]	0.629	0.235	45	15
	EI	170 [110,231]	1.41 <sup>‡</sup> [- 1.94, 4.90]	-	-	7.73 [- 9.64, 25.5]	0.57 [- 0.46, 1.61]	-	-	-	-	52.9 [36.4, 77.3]	69.8 [37.2, 114]	0.645	0.119	32	13
	EI	184 [146,222]	-	-	-	-	-	-0.04 [- 0.05, - 0.02]*	-	-	4.20 [- 2.67, 11.0]	45.2 [36.6, 56.0]	70.2 [43.0, 109]	0.722	0.378	73	16
FEC ( $\text{eggs.g}^{-1}$ )	WN	1300 [802,1770]	1.89 <sup>§</sup> [0.35, 3.45]*	-	-	-	-	-	-818 [- 928, - 706]*	25.4 [17.8, 33.0]*	-4.19 [- 80.6, 70.1]	397 [386,409]	807 [529,1290]	0.826	0.204	34	13

<sup>1</sup> EI: standard errors imputed where missing, WN: meta-analysis weighted by estimate sample size because too few standard errors available, see Supplementary Materials 2.2 for details.

<sup>‡</sup> Tannin values measured in studies.

<sup>§</sup> Tannin values from meta-estimate of data in (predominantly) other studies.

we were unable to account for the molecular structure of CTs in different forage species which mediates their impact (Mueller-Harvey et al., 2019). Furthermore, we analysed CT data from eight species monocultures or mixtures, so other differences in composition, digestion or intake between these forages could mask any impact of tannins on livestock performance (Mueller-Harvey et al., 2019). Although the EP sensitivity analysis found that leaf CT concentration positively predicted DLWG (Table S11), we do not find mechanistic support for this in our dataset. CTs are expected to protect forage protein from rumen degradation thus increasing uptake in the small intestine (Luscher et al., 2014; Wilkins and Jones, 2000; Barry, 1998), but this has yet to be linked to improved animal growth rate (Mueller-Harvey et al., 2019) and leaf CT content did not have a significant moderating effect on the relationship between leaf nitrogen content and DLWG in our analysis (Fig. 5b, Table 5 & S11). Instead, our result accords with recent empirical work which found that consumption of CT-rich forages simply shifted livestock N excretion from urine to faeces but did not increase animal retention (Grosse Brinkhaus et al., 2016; Azuhwi et al., 2013). The expanding Credible Intervals in Fig. 5b also indicates that we have insufficient observations of high leaf CT contents in our dataset to identify this relationship, and therefore our result should be treated as absence of evidence rather than evidence of absence of an effect.

Regarding the impact of average leaf CT content (Table 2) on livestock internal parasite burdens, we identified a significant interaction between herbage CT content and wormer treatment (Table 5). If CTs in forage have anthelmintic properties then we would expect to find a negative effect of increased leaf tannin content on FEC, which would be attenuated when sheep are dosed with de-wormer. Our analysis in fact identified no effect of leaf CT content on FEC when no wormer treatment was applied, but positively predicted FEC when wormer was applied (Fig. 5c). This unintuitive result suggests that tannins may be related to other factors not accounted for in our analysis which increase FEC count at higher leaf CT contents. When analysed separately, sheep FEC was negatively related to DLWG as expected, reducing livestock growth by  $40 \text{ g.day}^{-1}$ , 95% CI [- 50, - 20], per 1000 eggs per gram of faecal matter (Table 5). Although reduced intestinal worm burden on HL is frequently attributed to the anthelmintic properties of CT-rich herb species in the sward, other possible mechanisms include interrupting the parasites' life cycle by a combination of i) leaf shape and growth habitat of herbs, in addition to taller sward heights at grazing, impairing the ability of larvae to climb sufficiently far up forage to be ingested by sheep compared to short grass-majority swards, and ii) RG practices (frequently adopted on HL in order to preserve more sensitive sward components by allowing recovery time) preventing larvae hatched from eggs deposited in one grazing period surviving until the next grazing period to reach a new host.

We used trait data from the TRY database (Kattge et al., 2020) to estimate trait values for forage species and mixtures in our analysis. However, it is well-established that many traits are to some extent plastic and thus differ depending on site-specific conditions (Fromm, 2019; Sultan, 2000). We were unable to verify the generalisability of trait values from this database to the study sites in our systematic review, nor do we account for the differences in trait values between naturally occurring wild-type varieties (potentially measured in TRY) versus agriculturally improved cultivars selected for specific properties (often sown in the agricultural research trials analysed here). However, given studies in our systematic review did not measure values for these traits (except for CTs, not available in TRY), this was a necessary approximation to test widely cited mechanisms for the potential of HL to increase productivity. Although trait values likely differ between study sites and vary across cultivars, we feel that the TRY database gives acceptable estimates of the relative magnitude of traits between species and thus argue that the overall trends identified by our analysis are valid. We also tested the sensitivity of our results to the quality of trait data available from TRY which preserved the findings discussed above (TSV, Table S10).

Aggregating traits across species present in a sward is arguably an oversimplification of the ecological dynamics that could affect productivity. For example, although increased rooting depth confers drought tolerance and greater access to nutrients, complementarity in rooting structures between sward components is also likely to be important. In addition, seed mixes, underpinning aggregated trait values for 95 observations of herbage dry matter and 23 observations of DLWG (Table 4 & S10), provide only weak inference of i) species abundance in the mature sward, particularly due to seed size, sowing method and management non-randomly influencing species establishment and persistence, ii) forage biomass available per species, and iii) forage consumed by livestock, particularly given livestock are known to selectively graze both species and plant components in mixtures. However, analyses that tested the sensitivity of our results to this trait aggregation process and use of seed mix data for weighting trait values supported our findings for DLWG and the influence of legumes on herbage dry matter production, although the positive effects of root depth and species evenness on herbage DM were not preserved (AV and SMV, Table S10). Finally, although trait or seed mix data was missing for some sward components, excluding these observations only affected our finding of species evenness promoting herbage DM production (MD, Table S10).

### 3.3. Soil carbon

Enhancing soil organic carbon (SOC) stocks is widely promoted as a key benefit of adopting RA (Moyer et al., 2020; Burgess et al., 2019; Newton et al., 2020). Regenerative practices in temperate arable systems have recently been demonstrated to increase SOC concentration without reducing crop yields (Jordon et al., 2022). However, our systematic review only identified two studies of SOC under RG and HL each (Table 2), which is insufficient for meta-analysis and is in line with the findings of Conant et al. (2017) and Byrnes et al. (2018) that there remains a paucity of evidence to assess the effect of these practices on SOC in temperate pastoral systems.

From the results of our systematic review, Otálora et al. (2021) identified significantly higher carbon stocks in the top 10 cm of soil under 'regenerative' (1–2 days grazing, 24 days rest) compared to 'conventional' (6–10 days grazing, 15 days rest) RG after six years in Basque Country, Spain, while (Orgill et al., 2018) did not detect any difference in 0–30 cm SOC stocks under native pasture after five years of cell grazed vs set stocked in New South Wales, Australia. There are multiple potential mechanisms for RG to increase SOC stocks compared to continuous grazing. Firstly, rest periods which allow forage to accumulate and leaf photosynthetic area to increase will likely promote root growth (as plant belowground biomass is typically in proportion to aboveground biomass) and increase root exudates into the soil. This promotes soil microbiological activity which may stabilise a portion of these organic inputs in soil aggregates. Secondly, mob-style grazing systems, which aim to allow substantial forage accumulation and then trample a proportion of this during short grazing events, results in a layer of dead and decomposing vegetation at the soil surface which may increase plant residue inputs to the soil (Jones and Donnelly, 2004; Eyles et al., 2015; Rumpel et al., 2015; Piñeiro et al., 2010). Despite being credible, these mechanisms have yet to be supported by empirical findings in temperate regions. A recent global meta-analysis found greater SOC under RG compared to continuous grazing in other climatic regions (Byrnes et al., 2018), whereas temperate studies of grazing systems typically find no effect on SOC (Techio Pereira et al., 2018). This could be due in part to the inherent difficulty of detecting small changes in SOC stocks when substantial heterogeneity exists in baseline soil properties both within and between experimental pastures (Sanderman et al., 2015), and increased carbon inputs to soil not necessarily resulting in SOC accumulation, for example through concurrent changes in carbon stabilisation or decomposition (Eyles et al., 2015; Whitehead et al., 2018). In addition, there are several confounding factors which are possibly responsible for a disjoint between anecdotal practitioner

reports of increased SOC following adoption of RG and existing empirical evidence, including: i) stocking density rather than grazing system (rotational vs continuous) per se impacting mechanisms for building SOC, ii) management intensity, which is typically higher under RG compared to simpler set-stocking, may lead to simultaneous adoption of other beneficial management practices which influence SOC, and iii) baseline, as previously degraded soils are likely to show improvements in SOC when improved management practices are adopted regardless of grazing system (Jones and Donnelly, 2004; Rumpel et al., 2015; Conant et al., 2017; Briske et al., 2008; Abdalla et al., 2018).

Regarding HL, the two studies identified by our systematic review (Cong et al., 2014; Savage et al., 2019) concur with findings from the Jena biodiversity experiment, field trials in the USA and a recent synthesis across biomes in China that plant species richness is positively related with increased SOC (Prommer et al., 2020; Skinner and Dell, 2016; Chen et al., 2018). This is likely due to higher levels of root exudation and reduced evaporation from the topsoil due to denser vegetation in diverse plant communities promoting metabolic activity of soil microorganisms (Lange et al., 2015). In addition to diversity generally, properties of certain plants typically included in HL can specifically benefit soil carbon accumulation. For example, enhanced rooting depths can deliver root exudates to lower soil horizons, which may promote microbial activity and soil aggregation, and therefore SOC stabilisation (Whitehead, 2020; Dodd et al., 2011). Furthermore, the presence of legumes increases both below- and above-ground productivity through nitrogen fixation. This can deliver more plant residue inputs to the soil and increase availability of N in the soil for humus formation (Rumpel et al., 2015; Conant et al., 2001; Luscher et al., 2014). Although HL typically require rejuvenation through reseeding every four or five years as the more sensitive or palatable species are lost from the mixture, this can be achieved by direct drilling approaches rather than soil cultivation, which risks degrading SOC by disrupting soil aggregates and exposing the organic matter within these to microbial degradation. In addition, practices further along the regenerative 'spectrum' (Fig. 1) would seek to create and maintain species-diverse permanent pastures through grazing management rather than reseeding or overseeding approaches.

### 3.4. Limitations and future directions

We provide an initial quantitative synthesis of the mechanisms by which regenerative grazing management claims to increase productivity. There was insufficient evidence to analyse the impact of these practices on milk production (Table 2), resulting in our findings predominantly applying to lamb and beef production systems rather than dairy, although inclusion of herbs in multispecies swards has been found to increase dairy cow milk yield in a recent meta-analysis (Mccarthy et al., 2020). Although our analysis identifies several statistically significant relationships between quantitative features of management interventions and productivity outcomes, we cannot demonstrate causation using our approach. It therefore remains possible that the relationships observed here could be mutually caused by an unidentified third variable (i.e. a confounder), but our results have value by indicating that mechanisms cited by proponents of RA at least predict productivity outcomes. Future work should seek to verify these mechanisms through detailed experimental manipulations, for example through factorial treatments of different rest periods and stocking densities in RG systems, or comparing HL sward mixes with multiple cultivars of the same forage species that differ in key traits of interest such as rooting depth.

However, there are several key differences between the management interventions analysed here and RA in practice, which highlight important caveats to the applicability of our findings. Firstly, studies in

our systematic review typically implement *prescriptive* experimental treatments, which contrasts with the more holistic and *adaptive* management of RA practitioners. This has been summarised in a rangeland context by Briske et al. (2008): "Reduced flexibility in grazing experiments removes many sources of potential variation, but at the risk of becoming unrealistically abstracted from management applications. This is very likely the most serious limitation of the experimental data assessing the efficacy of rotational grazing." Future experimental work should compare prescriptive treatments with more adaptive management to see if this influences outcomes (Briske et al., 2011b), as has been done to some extent in North America (Lacanne and Lundgren, 2018; Rowntree et al., 2020). Secondly, there are likely to be differences in other management practices and philosophy between the studies synthesised here and RA practitioners which are not included in our analyses and may influence productivity outcomes. Studies in our dataset frequently applied RG or HL treatments using a 'conventional' agriculture approach of soil tillage to reseed pastures, growing forage species monocultures, applying synthetic fertilisers, and focusing purely on herbage or stock production as outcomes, which would not be considered truly regenerative by some practitioners. Conversely, key management objectives in RA include promoting soil microbiological activity, restoring ecosystem functioning, and increasing system resilience, hence why actions such as maximising sward diversity may be prioritised even if this potentially offers limited or no direct livestock production benefit as found in our analysis. Thirdly, we were unable to consider the interaction between RG and HL, despite these clearly being non-independent (Fig. 1). This was because no studies implemented both interventions factorially and few studies considered the impact of grazing management on sward diversity. Ironically, the one study which considered RG and HL in combination was excluded from our analysis because the interventions were confounded (Zaralis and Padel, 2019), yet this is most likely what RA would entail in practice. Future studies that compare 'conventional' with 'regenerative' management systems could help address this, either via experimental treatments or observations of practitioners' farms. Finally, it is important to note that RA as implemented on temperate livestock farms may include other practices, in addition to those considered here, including silvopastoral agroforestry or integration of livestock into arable cropping rotations. Our results should therefore not be used without caveats when assessing the overall impact of regenerative grazing. We are not claiming our results are universally generalisable, but that they are indicative of potentially significant benefits.

On a related note, the results of meta-analyses such as this aim to identify general relationships across studies and therefore contexts. While this is useful for informing policymakers and identifying directions for future academic work, our results are not suitable for providing management recommendations to practitioners in their specific context, although we did restrict our systematic review to a temperate oceanic climate to maximise the generalisability of results within these regions. Furthermore, we do not account for the potential disadvantages of adopting these interventions in practice. For RG, these include costs of labour, fencing and water provision infrastructure to divide existing fields into smaller paddocks for rotation. HL seed mixes can be expensive and unreliable to establish, and often contain species which are unsuited to winter grazing and require RG to ensure persistence, reducing farm-scale management flexibility. Future work should also consider the drivers and barriers for farmers to adopt these practices, which may mitigate any potential productivity benefits. In addition, there are other potential benefits of transitioning to RA in pastoral systems, including enhanced biodiversity and improvements in water quality and water flow regulation, which we do not consider here but should be accounted for in future work to assess the overall environmental impact of regenerative grazing practices.

#### 4. Conclusions

We provide empirical support for mechanisms by which rotational grazing and increasing sward diversity to include perennial forbs (i.e. herbal leys) can increase forage production and sheep and cattle growth rates in temperate oceanic regions. We find that increasing the proportion of the grazing season that a pasture is rested promotes herbage DM production and livestock DLWG (at higher stocking densities). Inclusion of deep-rooted species and legumes into the sward, and higher species evenness, enhances DM production. Higher forage nitrogen concentration (a proxy for protein content) promotes livestock growth rates. The currently available evidence does not support assertions that tannins promote livestock growth rate, enable better dietary protein utilisation or reduce internal parasite burdens. Further work is required to determine the influence of forage CT content on livestock performance. Although additional research is necessary to verify the mechanisms investigated here beyond predictive relationships, and there remains a disjoint between experimental treatments in published studies and actions of RA practitioners, our results suggest that further adoption of these RA practices on temperate grazing land has potential to reduce the greenhouse gas emissions associated with ruminant farming by increasing livestock productivity.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

We would like to thank L. Epelde, S. Orgill and T. Kristensen for providing additional information on their studies. We would also like to thank Leo Petrokofsky for generating the online evidence map. We would like to thank the anonymous reviewers for their comments, which greatly improved the manuscript. This work was supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC), UK [grant number BB/M011224/1]. PCB would like to acknowledge funding by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy [EXC 2075 - 390740016]. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The datasets generated and analysed during this study are available in the Zenodo repository, <https://doi.org/10.5281/zenodo.5786039>, (Jordon, 2022), and are also attached in the Supplementary Materials. R code used for analysis and figure plotting is also available in the Zenodo repository, <https://doi.org/10.5281/zenodo.5786039>, (Jordon, 2022).

#### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.agee.2022.108075](https://doi.org/10.1016/j.agee.2022.108075).

#### References

Abdalla, M., Hastings, A., Chadwick, D.R., Jones, D.L., Evans, C.D., Jones, M.B., Rees, R. M., Smith, P., 2018. Critical review of the impacts of grazing intensity on soil organic carbon storage and other soil quality indicators in extensively managed grasslands. *Agric. Ecosyst. Environ.* 253, 62–81.

Azuhwi, B.N., Hertzberg, H., Arrigo, Y., Gutzwiller, A., Hess, H.D., Mueller-Harvey, I., Torgerson, P.R., Kreuzer, M., Dohme-Meier, F., 2013. Investigation of sainfoin (*Onobrychis viciifolia*) cultivar differences on nitrogen balance and fecal egg count in artificially infected lambs. *J. Anim. Sci.* 91, 2343–2354.

Badgery, W.B., Michalk, D.L., 2017. Synthesis of system outcomes for a grazing-management experiment in temperate native pastures. *Anim. Prod. Sci.* 57, 1869–1876.

Barry, T.N., 1998. The feeding value of chicory (*Cichorium intybus*) for ruminant livestock. *J. Agric. Sci.* 131, 251–257.

Briske, D.D., Derner, J.D., Milchunas, D.G. & Tate, K.W. 2011a. An Evidence-Based Assessment of Prescribed Grazing Practices. In: Briske, D.D. (ed.) *Conservation Benefits of Rangeland Practices: Assessment, Recommendations, and Knowledge Gaps*.

Briske, D.D., Derner, J.D., Brown, J.R., Fuhlendorf, S.D., Teague, W.R., Havstad, K.M., Gillen, R.L., Ash, A.J., Willms, W.D., 2008. Rotational grazing on rangelands: reconciliation of perception and experimental evidence. *Rangel. Ecol. Manag.* 61, 3–17.

Briske, D.D., Sayre, N.F., Huntsinger, L., Fernandez-Gimenez, M., Budd, B., Derner, J.D., 2011b. Origin, persistence, and resolution of the rotational grazing debate: integrating human dimensions into rangeland research. *Rangeland Ecol. Manag.* 325–334.

Brougham, R.W., 1971. Frequency and intensity of grazing and their effects on pasture production. *Proc. N. Z. Grassl. Assoc.* 32, 137–144.

Brown, G., 2018. *Dirt to Soil. One family's journey into Regenerative Agriculture*. Chelsea Green Publishing, White River Junction, Vermont.

Burgess, P.J., Harris, J., Graves, A.R. & Deeks, L.K. 2019. *Regenerative Agriculture: Identifying the Impact; Enabling the Potential*. Bedfordshire, UK: Cranfield University: Report for SYSTEMIQ.

Burke, J.L., Waghorn, G.C., Chaves, A.V., 2002. Improving animal performance using forage-based diets. *Proc. N. Z. Soc. Anim. Prod.* 62, 267–272.

Bürkner, P.-C., 2017. brms: an R package for bayesian multilevel models using stan. *J. Stat. Softw.* 80.

Bürkner, P.-C., 2018. Advanced bayesian multilevel modeling with the R package brms. *R. J.* 10, 395–411.

Byrnes, R.C., Eastburn, D.J., Tate, K.W., Roche, L.M., 2018. A global meta-analysis of grazing impacts on soil health indicators. *J. Environ. Qual.* 47, 758–765.

Chen, S., Wang, W., Xu, W., Wang, Y., Wan, H., Chen, D., Tang, Z., Tang, X., Zhou, G., Xie, X., Zhou, D., Shanguan, Z., Huang, J., He, J.S., Wang, Y., Sheng, J., Tang, L., Li, X., Dong, M., Wu, Y., Wang, Q., Wang, Z., Wu, J., Chapin 3RD, F.S., Bai, Y., 2018. Plant diversity enhances productivity and soil carbon storage. *Proc. Natl. Acad. Sci. USA* 115, 4027–4032.

Clark, M.A., Domingo, N.G.G., Colgan, K., Thakrar, S.K., Tilman, D., Lynch, J., Azevedo, I.L., Hill, J.D., 2020. Global food system emissions could preclude achieving the 1.5° and 2° C climate change targets. *Science* 370, 705–708.

Conant, R.T., Paustian, K., Elliott, E.T., 2001. Grassland management and conversion into grassland: effects on soil carbon. *Ecol. Appl.* 11, 343–355.

Conant, R.T., Cerri, C.E.P., Osborne, B.B., Paustian, K., 2017. Grassland management impacts on soil carbon stocks: a new synthesis. *Ecol. Appl.* 27, 662–668.

Cong, W.-F., Van Ruijven, J., Mommer, L., De Deyn, G.B., Berendse, F., Hoffland, E., Lavorel, S., 2014. Plant species richness promotes soil carbon and nitrogen stocks in grasslands without legumes. *J. Ecol.* 102, 1163–1170.

Cranston, L.M., Kenyon, P.R., Morris, S.T., Kemp, P.D., 2015. A review of the use of chicory, plantain, red clover and white clover in a sward mix for increased sheep and beef production. *J. N. Z. Grassl.* 77, 89–94.

CEE, 2018. *Collaboration for Environmental Evidence. Guidelines and Standards for Evidence synthesis in Environmental Management. Version 5.0*. Pullin, A.S., Frampton, G.K., Livoreil, B. & Petrokofsky, G. (eds.). Online, [www.environmentalevidence.org/information-for-authors](http://www.environmentalevidence.org/information-for-authors).

Defra 2010. *Definitions of Terms used in Farm Business Management*. 3rd edition ed. UK.

Díaz De Otálora, X., Epelde, L., Arranz, J., Garbisu, C., Ruiz, R., Mandaluniz, N., 2021. Regenerative rotational grazing management of dairy sheep increases springtime grass production and topsoil carbon storage. *Ecol. Indic.* 125.

Dodd, M.B., Crush, J.R., Mackay, A.D., Barker, D.J., 2011. The "root" to more soil carbon under pastures. *Proc. N. Z. Grassl. Assoc.* 73, 43–50.

Douglas, J.A., 1986. The production and utilisation of lucerne in New Zealand. *Grass Forage Sci.* 41, 81–128.

Elliot, R.H., 1908. *The Clifton Park System of Farming and laying down land to grass - a guide to landlords, tenants and land legislators*. Faber and Faber Limited, London.

Eyles, A., Coghlan, G., Hardie, M., Hovenden, M., Bridle, K., 2015. Soil carbon sequestration in cool-temperate dryland pastures: mechanisms and management options. *Soil Res.* 53.

Finn, J.A., Kirwan, L., Connolly, J., Sebastia, M.T., Helgadottir, A., Baadshaug, O.H., Bélanger, G., Black, A., Brophy, C., Collins, R.P., Cöp, J., Dalmannsdóttir, S., Delgado, I., Elgersma, A., Fothergill, M., Frankow-Lindberg, B.E., Ghesquiere, A., Golinska, B., Golinski, P., Grieu, P., Gustavsson, A.-M., Höglind, M., Huguenin-Elie, O., Jørgensen, M., Kadziulienė, Z., Kurki, P., Llurba, R., Lunnan, T., Porqueddu, C., Suter, M., Thumm, U., Lüscher, A., Wilsey, B., 2013. Ecosystem function enhanced by combining four functional types of plant species in intensively managed grassland mixtures: a 3-year continental-scale field experiment. *J. Appl. Ecol.* 50, 365–375.

Fox, N.J., Smith, L.A., Houdijk, J.G.M., Athanasiadou, S., Hutchings, M.R., 2018. Ubiquitous parasites drive a 33% increase in methane yield from livestock. *Int. J. Parasitol.* 48, 1017–1021.

Fromm, H., 2019. Root Plasticity in the Pursuit of Water. *Plants* 8.

Gelman, A., Goodrich, B., Gabry, J., Vehtari, A., 2019. R-squared for Bayesian Regression Models. *Am. Stat.* 74.

Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, A. & Tempio, G. 2013. *Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities*. Rome, Italy: FAO.

Giller, K.E., Hijbeek, R., Andersson, J.A., Sumburg, J., 2021. *Regenerative Agriculture: An agronomic perspective*. Outlook Agric. 50, 13–25.

Godfray, H.C.J., Aveyard, P., Garnett, T., Hall, J.W., Key, T.J., Lorimer, J., Pierrehumbert, R.T., Scarborough, P., Springmann, M., Jebb, S.A., 2018. *Meat consumption, health, and the environment*. Science 361.

Gordon, M. & Lumley, T. 2020. *forestplot: Advanced Forest Plot Using 'grid' Graphics*. R package version 1.10.1.

- Grosse Brinkhaus, A., Bee, G., Silacci, P., Kreuzer, M., Dohme-Meier, F., 2016. Effect of exchanging *Onobrychis viciifolia* and *Lotus corniculatus* for *Medicago sativa* on ruminal fermentation and nitrogen turnover in dairy cows. *J. Dairy Sci.* 99, 4384–4397.
- Hacker, R. A brief evaluation of time control grazing. Proceedings of the Eighth Annual Conference of the Grassland Society of New South Wales, 1993. 82–89.
- Haddaway, N.R., Macura, B., Whaley, P. & Pullin, A.S. 2017b. ROSES for Systematic Review Reports. 1.0 ed.
- Haddaway, N.R., Hedlund, K., Jackson, L.E., Kätterer, T., Lugato, E., Thomsen, I.K., Jørgensen, H.B., Isberg, P.-E., 2017a. How does tillage intensity affect soil organic carbon? A systematic review. *Environ. Evid.* 6.
- Hawkins, H.-J., 2017. A global assessment of Holistic Planned Grazing™ compared with season-long, continuous grazing: meta-analysis findings. *Afr. J. Range Forage Sci.* 34, 65–75.
- Hector, et al., 1999. Plant diversity and productivity experiments in European grasslands. *Science* 286, 1123–1127.
- Herrero, M., Gerber, P., Vellinga, T., Garnett, T., Leip, A., Opio, C., Westhoek, H.J., Thornton, P.K., Olesen, J., Hutchings, N., Montgomery, H., Soussana, J.F., Steinfeld, H., McAllister, T.A., 2011. Livestock and greenhouse gas emissions: The importance of getting the numbers right. *Anim. Feed Sci. Technol.* 166–167, 779–782.
- Herrero, M., Henderson, B., Havlík, P., Thornton, P.K., Conant, R.T., Smith, P., Wiersma, S., Hristov, A.N., Gerber, P., Gill, M., Butterbach-Bahl, K., Valin, H., Garnett, T., Stehfest, E., 2016. Greenhouse gas mitigation potentials in the livestock sector. *Nat. Clim. Change* 6, 452–461.
- Hristov, A.N., Oh, J., Lee, C., Meinen, R., Montes, F., Ott, T., Firkins, J., Rotz, A., Dell, C., Adesogan, A., Yang, W., Tricarico, J., Kebreab, E., Waghorn, G., Dijkstra, J., Oosting, S., 2013. Mitigation of greenhouse gas emissions in livestock production – A review of technical options for non-CO<sub>2</sub> emissions. In: Gerber, P.J., Henderson, B., Makkar, H.P.S. (Eds.), *FAO Animal Production and Health Paper No. 177*. FAO, Rome, Italy.
- John, A., Lancashire, J.A., 1981. Aspects of the feeding and nutritive value of *Lotus* species. *Proc. N. Z. Grassl. Assoc.* 42, 152–159.
- Jones, M.B., Donnelly, A., 2004. Carbon sequestration in temperate grassland ecosystems and the influence of management, climate and elevated CO<sub>2</sub>. *N. Phytol.* 164, 423–439.
- Jonker, A., Farrell, L., Scobie, D., Dynes, R., Edwards, G., Hague, H., Mcaliffie, R., Taylor, A., Knight, T., Waghorn, G., 2019. Methane and carbon dioxide emissions from lactating dairy cows grazing mature ryegrass/white clover or a diverse pasture comprising ryegrass, legumes and herbs. *Anim. Prod. Sci.* 59.
- Jordon, M.W., 2022. Data for "Rotational grazing and multispecies herbal leys increase productivity in temperate pastoral systems – a meta-analysis (1.0.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.5786039>.
- Kattge, J., Bonisch, G., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Tautenhahn, S., Werner, G.D.A., Aakala, T., Abedi, M., Acosta, A.T.R., Adamidis, G.C., Adamson, K., Aiba, M., Albert, C.H., Alcantara, J.M., Alcazar, C.C., Aleixo, I., Ali, H., Amiaud, B., Ammer, C., Amoroso, M.M., Anand, M., Anderson, C., Anten, N., Antos, J., Apgaua, D.M.G., Ashman, T.L., Asmara, D.H., Asner, G.P., Aspinwall, M., Atkin, O., Aubin, I., Baastrup-Spohr, L., Bahalkeh, K., Bahn, M., Baker, T., Baker, W.J., Bakker, J.P., Baldocchi, D., Baltzer, J., Banerjee, A., Baranger, A., Barlow, J., Barneche, D.R., Baruch, Z., Bastianelli, D., Battjes, J., Bauerle, W., Bauters, M., Bazzato, E., Beckmann, M., Beekman, H., Beierkuhnlein, C., Bekker, R., Belfry, G., Belluau, M., Belouin, M., Benavides, R., Benomar, L., Berdugo-Lattke, M.L., Berenguer, E., Bergamin, R., Bergmann, J., Bergmann Carlucci, M., Berner, L., Bernhardt-Romeremann, M., Bigler, C., Bjorkman, A.D., Blackman, C., Blanco, C., Blonder, B., Blumenthal, D., Bocanegra-Gonzalez, K.T., Boeckx, P., Bohlmann, S., Bohning-Gaese, K., Boisvert-Marsh, L., Bond, W., Bond-Lamberty, B., Boom, A., Boonman, C.C.F., Bordin, K., Boughton, E.H., Boukili, V., Bowman, D., Bravo, S., Brendel, M.R., Broadley, M.R., Brown, K.A., Bruelheide, H., Brummich, F., Bruun, H. H., Bruy, D., Buchanan, S.W., Bucher, S.F., Buchmann, N., Buitenwerf, R., Bunker, D. E., Burger, J., et al., 2020. TRY plant trait database - enhanced coverage and open access. *Glob. Chang Biol.* 26, 119–188.
- Lacanne, C.E., Lundgren, J.G., 2018. Regenerative agriculture: merging farming and natural resource conservation profitably. *PeerJ* 6, e4428.
- Lange, M., Eisenhauer, N., Sierra, C.A., Bessler, H., Engels, C., Griffiths, R.I., Mellado-Vazquez, P.G., Malik, A.A., Roy, J., Scheu, S., Steinbeiss, S., Thomson, B.C., Trumbore, S.E., Gleixner, G., 2015. Plant diversity increases soil microbial activity and soil carbon storage. *Nat. Commun.* 6, 6707.
- Li, G., Kemp, P.D., 2005. Forage chicory (*Cichorium intybus* L.): a review of its agronomy and animal production. *Adv. Agron.* 88, 187–222.
- Loza, C., Reinsch, T., Loges, R., Taube, F., Gere, J.I., Klus, C., Hasler, M., Malisch, C.S., 2021. Methane emission and milk production from jersey cows grazing perennial ryegrass–white clover and multispecies forage mixtures. *Agriculture* 11.
- Lüscher, A., Mueller-Harvey, I., Soussana, J.F., Rees, R.M., Peyraud, J.L., 2014. Potential of legume-based grassland-livestock systems in Europe: a review. *Grass Forage Sci.* 69, 206–228.
- Jordon, M.W., Willis, K.J., Bürkner, P.C., Haddaway, N.R., Smith, P., Petrokofsky, G., 2022. Temperate Regenerative Agriculture practices increase soil carbon but not crop yield – a meta-analysis. *under review, preprint available (Research Square)*. 10.21203/rs.3.rs-1064515/v2.
- Martin, A.C. 2018. *Thaloo Evidence-Mapping: A Jekyll Theme* [Online]. Available: (<https://github.com/AndrewIOM/thaloo>) [Accessed April 2020].
- Mccarthy, K.M., Mcaloon, C.G., Lynch, M.B., Pierce, K.M., Mulligan, F.J., 2020. Herb species inclusion in grazing swards for dairy cows-a systematic review and meta-analysis. *J. Dairy Sci.* 103, 1416–1430.
- Mcdonald, S.E., Lawrence, R., Kendall, L., Rader, R., Brudvig, L., 2019. Ecological, biophysical and production effects of incorporating rest into grazing regimes: a global meta-analysis. *J. Appl. Ecol.* 56, 2723–2731.
- Moyer, J., Smith, A., Rui, Y. & Hayden, J. 2020. *Regenerative Agriculture and the Soil Carbon Solution*. Rodale Institute.
- Mueller-Harvey, I., Bee, G., Dohme-Meier, F., Hoste, H., Karonen, M., Kölliker, R., Lüscher, A., Niderkorn, V., Pellikaan, W.F., Salminen, J.-P., Skot, L., Smith, L.M.J., Thamsborg, S.M., Totterdell, P., Wilkinson, I., Williams, A.R., Azuhwi, B.N., Baert, N., Brinkhaus, A.G., Copani, G., Desrués, O., Drake, C., Engström, M., Frygasas, C., Girard, M., Huyen, N.T., Kempf, K., Malisch, C., Mora-Ortiz, M., Quijada, J., Ramsay, A., Ropiak, H.M., Waghorn, G.C., 2019. Benefits of Condensed Tannins in Forage Legumes Fed to Ruminants: Importance of Structure, Concentration, and Diet Composition. *Crop Sci.* 59, 861–885.
- Newton, P., Civita, N., Frankel-Goldwater, L., Bartel, K., Johns, C., 2020. What is regenerative agriculture? A review of scholar and practitioner definitions based on processes and outcomes. *Front. Sustain. Food Syst.* 4.
- Norton, B.E., 1998. The application of grazing management to increase sustainable livestock production. *Anim. Prod. Aust.* 22, 15–26.
- Nyfelser, D., Huguéin-Elie, O., Suter, M., Frossard, E., Connolly, J., Lüscher, A., 2009. Strong mixture effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive overyielding. *J. Appl. Ecol.* 46, 683–691.
- Orgill, S.E., Condon, J.R., Conyers, M.K., Morris, S.G., Alcock, D.J., Murphy, B.W., Greene, R.S.B., 2018. Removing grazing pressure from a native pasture decreases soil organic carbon in Southern New South Wales, Australia. *Land Degrad. Dev.* 29, 274–283.
- Piñeiro, G., Paruelo, J.M., Oesterheld, M., Jobbágy, E.G., 2010. Pathways of grazing effects on soil organic carbon and nitrogen. *Rangel. Ecol. Manag.* 63, 109–119.
- Prommer, J., Walker, T.W.N., Wanek, W., Braun, J., Zezula, D., Hu, Y., Hofhansl, F., Richter, A., 2020. Increased microbial growth, biomass, and turnover drive soil organic carbon accumulation at higher plant diversity. *Glob. Chang Biol.* 26, 669–681.
- R Core Team 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rowntree, J.E., Stanley, P.L., Maciel, I.C.F., Thorbecke, M., Rosenzweig, S.T., Hancock, D.W., Guzman, A., Raven, M.R., 2020. Ecosystem impacts and productive capacity of a multi-species pastured livestock system. *Front. Sustain. Food Syst.* 4.
- Rumpel, C., Crème, A., Ngo, P.T., Velásquez, G., Mora, M.L., Chabbi, A., 2015. The impact of grassland management on biogeochemical cycles involving carbon, nitrogen and phosphorus. *J. Soil Sci. Plant Nutr.* 15, 353–371.
- Sac Consulting 2020. *The Farm Management Handbook 2020/21*, Scotland.
- Sanderman, J., Reseigh, J., Wurst, M., Young, M.A., Austin, J., 2015. Impacts of rotational grazing on soil carbon in native grass-based pastures in Southern Australia. *PLoS One* 10, e0136157.
- Sanderson, M.A., Soder, K.J., Muller, L.D., Klement, K.D., Skinner, R.H., Goslee, S.C., 2005. Forage mixture productivity and botanical composition in pastures grazed by Dairy Cattle. *Agron. J.* 97, 1465–1471.
- Savage, J., Woodcock, B.A., Bullock, J.M., Peyton, J., Hulmes, S., Hulmes, L., Nowakowski, M., Pywell, R.F., 2019. Impact of grassland management on biomass production and nutritional quality, invertebrate communities, and soil health in Berkshire (UK) 2009-2012. *NERC Environ. Inf. Data Cent.*
- Savory, A., Butterfield, J., 2016. *Holistic Management: A Commonsense Revolution to Restore Our Environment*. Island Press.
- Skinner, R.H., Dell, C.J., 2016. Yield and soil carbon sequestration in grazed pastures sown with two or five forage species. *Crop Sci.* 56, 2035–2044.
- Smith, P., Balmford, A., 2020. Climate change: 'no get out of jail free card'. *Vet. Rec.* 71.
- Stan Development Team 2019. *Stan Modeling Language Users Guide and Reference Manual. Version 2.27*.
- Stan Development Team. 2020. *Brief Guide to Stan's Warnings* [Online]. Available: (<https://mc-stan.org/misc/warnings.html>) [Accessed].
- Stewart, A.V., 1996. Plantain (*Plantago lanceolata*) – a potential pasture species. *Proc. N. Z. Grassl. Assoc.* 58, 77–86.
- Sultan, S.E., 2000. Phenotypic plasticity for plant development, function and life history. *Trends Plant Sci.* 5, 537–542.
- Suter, M., Connolly, J., Finn, J.A., Loges, R., Kirwan, L., Sebastia, M.T., Lüscher, A., 2015. Nitrogen yield advantage from grass-legume mixtures is robust over a wide range of legume proportions and environmental conditions. *Glob. Chang Biol.* 21, 2424–2438.
- Techio Pereira, L.E., Da Silva, S.C., Matthew, C., López, I.F., Fischer Shbrisia, A., 2018. Grazing management for sustainable grazing systems. In: Marshall, A., Collins, R. (Eds.), *Improving grassland and pasture management in temperate agriculture*. Burleigh Dodds Science Publishing, Cambridge, UK.
- Tilman, D., Reich, P.B., Knops, J., Wedin, D.A., Mielke, T., Lehman, C., 2001. Diversity and productivity in a long-term grassland experiment. *Science* 294, 843–845.
- Turner, F.N., 1951. *Fertility Farming*. Faber & Faber Ltd, London, UK.
- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., Bürkner, P.C., 2020. Rank-normalization, folding, and localization: An improved Rhat for assessing convergence of MCMC. *Bayesian Anal.* 1–28.
- Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48.
- Voisin, A., 1959. *Grass Productivity*. Crosby Lockwood & Son Ltd, London.
- Weisser, W.W., Roscher, C., Meyer, S.T., Ebeling, A., Luo, G., Allan, E., Bessle, H., Barnard, R.L., Buchmann, N., Buscot, F., Engels, C., Fischer, C., Fischer, M., Gessler, A., Gleixner, G., Halle, S., Hildebrandt, A., Hillebrand, H., De Kroon, H., Lange, M., Leimer, S., Le Roux, X., Milcu, A., Mommer, L., Niklaus, P.A., Oelmann, Y., Proulx, R., Roy, J., Scherber, C., Scherer-Lorenzen, M., Scheu, S., Tschamtké, T., Wachenford, M., Wagg, C., Weigelt, A., Wilcke, W., Wirth, C.,

- Schulze, E.-D., Schmid, B., Eisenhauer, N., 2017. Biodiversity effects on ecosystem functioning in a 15-year grassland experiment: patterns, mechanisms, and open questions. *Basic Appl. Ecol.* 23, 1–73.
- Whitehead, D., 2020. Management of grazed landscapes to increase soil carbon stocks in temperate, Dryland Grasslands. *Front. Sustain. Food Syst.* 4, 585913.
- Whitehead, D., Schipper, L.A., Pronger, J., Moinet, G.Y.K., Mudge, P.L., Calvelo Pereira, R., Kirschbaum, M.U.F., McNally, S.R., Beare, M.H., Camps-Arbestain, M., 2018. Management practices to reduce losses or increase soil carbon stocks in temperate grazed grasslands: New Zealand as a case study. *Agric., Ecosyst. Environ.* 265, 432–443.
- Wilkins, R.J., Jones, R.J.A., 2000. Alternative home-grown protein sources for ruminants in the United Kingdom. *Anim. Feed Sci. Technol.* 85, 23–32.
- Zaralis, K. & Padel, S. Effects of High Stocking Grazing Density of Diverse Swards on Forage Production, Animal Performance and Soil Organic Matter: A Case Study. *In: Theodoridis, A., Ragkos, A. & Salampasis, M., eds. Innovative Approaches and Applications for Sustainable Rural Development*, 2019 Chania, Crete, Greece. 131–146.