



Why do we make changes to the long-term experiments at Rothamsted?

Paul R. Poulton^{a,*}, David S. Powlson^a, Margaret J. Glendining^b, Andrew S. Gregory^c

^a Sustainable Soils and Crops, Rothamsted Research, Harpenden, UK

^b Intelligent Data Ecosystems, Rothamsted Research, Harpenden, UK

^c Protecting Crops and the Environment, Rothamsted Research, Harpenden, UK

ARTICLE INFO

Keywords:

Long-term experiments
Essential/planned change
Nutrient/crop management
Acidification
Soil organic matter
Heavy metals
Archived samples

ABSTRACT

The long-term field experiments at Rothamsted in south-east England (UK) are an important resource that has been used extensively to study the effects of land management, atmospheric pollution and climate change on soil fertility and the sustainability of crop yields. However, for these and other long-term experiments around the world to remain useful, changes are sometimes needed. These changes may be required to ensure that the experiment is not compromised by *e.g.* acidification or weeds, but often they are needed to ensure that the experiment remains relevant to current agricultural practice, *e.g.* the introduction of new cultivars and the judicious use of pesticides. However, changes should not be made just for the sake of change or to investigate aspects of management that could be better resolved in a short-term experiment. Rather, modifications should only be made after carefully considered discussion, involving scientists from different disciplines. It must be remembered however that there are limitations to what can be achieved in one experiment. In this paper we give examples of why certain changes were made to the Rothamsted experiments and what the results of those changes have been. We also highlight the value of archiving crop and soil samples for future studies.

1. Introduction

Long-term experiments (LTEs) provide invaluable information, usually not obtainable by other means, for understanding and assessing the sustainability, or otherwise, of cropping systems and agricultural practices. This information includes trends in crop yield and in the associated occurrence of weeds, pests and diseases. Also, importantly, LTEs provide data on changes in soil properties, some of which only become apparent after many years or decades. These include trends in soil organic matter content, pH, nutrient pools, the availability of nutrients to crops, and some aspects of soil physics and biology. In most cases an essential benefit of an LTE is continuity of treatments over many years. However, there are situations where a change in management is either necessary or desirable. Some changes may be necessary to ensure the continuation or integrity of an experiment; examples might include situations where weeds become difficult to control in certain treatments or where soil erosion is serious. Other changes may be desirable to maintain the relevance of an LTE to current agricultural practice in the region or to address emerging scientific questions. The dilemma is how to make such changes without losing the long-term integrity of the experiment.

The LTEs at Rothamsted, and its associated experimental sites,

include the oldest continuous agronomic experiments in the world which were started by J B Lawes and J H Gilbert in the mid-19th Century. Their intention was to study the effects of different combinations of inorganic fertilisers and organic manures on the yield of the major crops grown in England at that time, including *e.g.* winter wheat (*Triticum aestivum*), spring barley (*Hordeum vulgare*), turnips (*Brassica rapa*) and permanent pasture cut for hay. Lawes and Gilbert made frequent changes to the treatments of the Barnfield Turnip and Broadbalk Wheat experiments in the first few years after they were established in 1843 but soon realised that to gain a better understanding of seasonal effects any changes should be kept to a minimum (Lawes, 1847; Lawes and Gilbert, 1864). Nevertheless, they did make further important changes, some of which are discussed later. Following the death of Lawes and Gilbert (in 1900 and 1901, respectively), successive Directors of Rothamsted found it necessary to make significant changes to the LTEs. They also started several new experiments in the mid-20th Century, mainly on the effects on yield of ley-arable cropping and organic amendments, some of which still continue today. But, by the 1960s, it was apparent that, if the experiments started by Lawes and Gilbert and others were to continue, and be of scientific and agricultural relevance, then significant changes were needed.

* Correspondence to: Sustainable Soils and Crops, Rothamsted Research, Harpenden, Herts AL5 2JQ, UK.

E-mail address: paul.poulton@rothamsted.ac.uk (P.R. Poulton).

<https://doi.org/10.1016/j.eja.2023.127062>

Received 3 October 2023; Received in revised form 28 November 2023; Accepted 16 December 2023

Available online 5 January 2024

1161-0301/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The purpose of this paper is to give examples of some of the changes that have been made to the LTEs at Rothamsted, why they were thought necessary and what the consequences of those changes have been. We hope that some of these examples may help inform managers of other LTEs who are perhaps less experienced; a situation that arises where there has been less continuity in management as older staff have retired or where there is a lack of written records.

2. Experiments

A list of the experiments from which examples of significant change have been taken is given in Table 1. The experiments are on contrasting soil types in south-east England in the UK. The soil at Rothamsted (global coordinates: 51.8095, -0.3730) is a flinty silty clay loam over Clay-with flints and is classified as a Chromic Luvisol; at Woburn (global coordinates: 51.9998, -0.6160) the soil is a sandy clay loam and is classified as a Cambic Arenosol (IUSS Working Group WRB, 2022). The long-term average annual rainfall, 1870–2020, at Rothamsted is 706 mm. The long-term average annual air temperature (1878 – 1988) was 9.0 °C, but this has increased markedly since then. The 10-year average for 2013–2022 was 10.6 °C. Average rainfall and temperature were very similar at Woburn – the two sites are 30 km apart.

3. Discussion

Modifications to the Rothamsted LTEs have only been made after lengthy discussions between experienced scientists from different disciplines; we consider that a multidisciplinary approach is essential. Over the years we have avoided making changes to accommodate a line of research that is best answered in a short-term experiment elsewhere, and, as far as possible, we plan that any changes in treatment or management will remain in place for many years. An exception to this is that the cultivars grown on the arable experiments may be changed every

four or five years (see 3.2.2.). Some of the changes made to Rothamsted experiments, and the reasoning leading to them, will be relevant to other LTEs. Changes fall into two broad groups, although there will be overlap between them. We also discuss how LTEs may be utilised to address unexpected issues, the value of archived samples and problems with assessing the impact of climate change. Our discussion is structured therefore on three themes:

1. essential changes that are needed to ensure that the original purpose of the experiment, and/or its continuation and integrity is not compromised
2. planned changes to ensure that the experiment remains relevant to current farming practices in that region and to address new or emerging scientific questions
3. LTEs as a record of environmental and management change

3.1. Essential changes

The most obvious examples of essential changes that have been made to the LTEs at Rothamsted are those concerning acidification, weed control and nutrient management. Below we discuss the issues faced, the approach adopted, and the lessons drawn.

3.1.1. Soil acidification

Although the farm at Rothamsted overlies chalk, the topsoil on many of the fields would be acid if they were not limed. Although there was no clear understanding of acidity in the early-19th Century, it was well known that spreading chalk on some soils improved workability and yields and it was common practice in the area around Rothamsted for many years (Young, 1813; Russell, 1916). It is thought that as much as 200 t ha⁻¹ of chalk was dug-out and spread, probably somewhat unevenly, on many of the arable fields at Rothamsted before the

Table 1
Details of long-term experiments at Rothamsted and Woburn referred to in this paper.

Site and Experiment	Duration	Clay %	Crop ^a	Treatments include	Reference
Rothamsted, Hertfordshire, UK					
Agdell Rotation	1848 - 1990	27 - 35	Various arable crops, grass, legumes	Rates of P, rates of K	Johnston and Penny (1972) Johnston et al. (2013)
Broadbalk Wheat	1843 - ongoing	20 - 38	Winter wheat, potatoes	Rates of N, FYM	Johnston and Garner (1969)
Broadbalk Wilderness	1881 - ongoing	20 - 25	oats, beans, maize	P, K, Mg, S	Poulton et al. (2024)
Geescroft Wilderness	1886 - ongoing	20 - 25	Regenerating woodland	None	Poulton et al. (2003)
Exhaustion Land	1852 - ongoing	18 - 27	Regenerating woodland	None	Poulton et al. (2003)
Hoosfield Barley	1852 - ongoing	18 - 27	Winter wheat, potatoes	Rates of P, K, Mg, FYM	Johnston and Poulton (1977) Poulton et al. (2013)
Park Grass	1856 - ongoing	18 - 27	Spring barley,	Rates of N, FYM, P, K, Mg	Jenkinson and Johnston (1977) Macdonald et al. (2018)
Highfield Ley-arable	1949 - ongoing	18 - 27	Permanent pasture	Forms of N, chalk, FYM, P, K, Mg, Na, Si	Silvertown et al. (2006) Storkey et al. (2015)
Highfield Bare Fallow	1959 - ongoing	18 - 27	Arable crops, leys and permanent grass	All-arable rotations, ley-arable rotations. Grass since 1838	Johnston (1973) Johnston et al. (2009)
Highfield Reversion	2008 - ongoing	18 - 27	Fallow	None	Barre et al., 2010 Macdonald et al. (2022)
Woburn, Bedfordshire, UK					
Ley-arable	1938 - ongoing	11 - 16	Fallow, grass, cereal	Fallow, grass, cereal sited on parts of the Ley-arable and Bare Fallow expts	Redmile-Gordon et al. (2020)
Continuous Wheat & Barley	1876 - 1966	11 - 16	Arable crops and leys	All-arable rotations, ley-arable rotations. FYM, P, K, Mg, Na	Johnston et al. (2017) Poulton et al. (2023) Johnston (1975)
Market Garden	1942 - ongoing	10 - 12	Winter wheat, spring barley	Sewage sludge, composts	Johnston and Wedderburn (1975)
			Market garden crops	FYM, fertiliser	McGrath (1987)

spring barley (*Hordeum vulgare*)

^a Winter wheat (*Triticum aestivum*), Potatoes (*Solanum tuberosum*), oats (*Avena sativa*), beans (*Vicia faba*), maize (*Zea mays*)

experiments started in the 1840s. Consequently, most of the LTEs had considerable reserves of free calcium carbonate (CaCO₃), and it was not until many decades later, that potential problems with acidification became apparent (Johnston, 1997). On the Broadbalk Wheat Experiment, acidifying inputs from the ammonium sulphate and other fertilisers that were applied meant that by 1944 there was little or no free CaCO₃ remaining in the topsoil (0 – 23 cm) of some treatments (Fig. 1) and that pH was < 6 in a few areas (Poulton et al., 2024). Observations of the wheat confirmed that this was beginning to have an effect on the crop and in 1955 differential amounts of chalk were applied to the experiment to increase soil pH to a value such that the yield of wheat would not be affected. Increased rates of fertiliser N in the 1960s and 1980s and increasing acidifying atmospheric inputs until the 1980s (Poulton et al., 2024), has meant that basal and/or differential

applications of chalk have continued to the present (Glendining et al., 2022).

On the Agdell experiment at Rothamsted, growing a 4-course rotation including root crops since 1848, there were also problems with acidification. From the 1920 s, club-root (*Plasmodiophora brassicae*) became more damaging to the root crops, especially on soils receiving nitrogen, phosphorus, potassium, magnesium and sodium (N, P, K, Mg, Na) which were becoming more acid (Warren, 1958). Although soil pH was corrected by liming, *P. brassicae* was still present in the soil which meant that root crops could no longer be grown. The experiment was substantially modified in the 1950 s – 1970 s (eventually creating 384 sub-plots) to test the value of fresh and residual P and K fertiliser (Johnston and Penny, 1972). The experiment was stopped in 1990; but the data is still useful and archived crop and soil samples are available

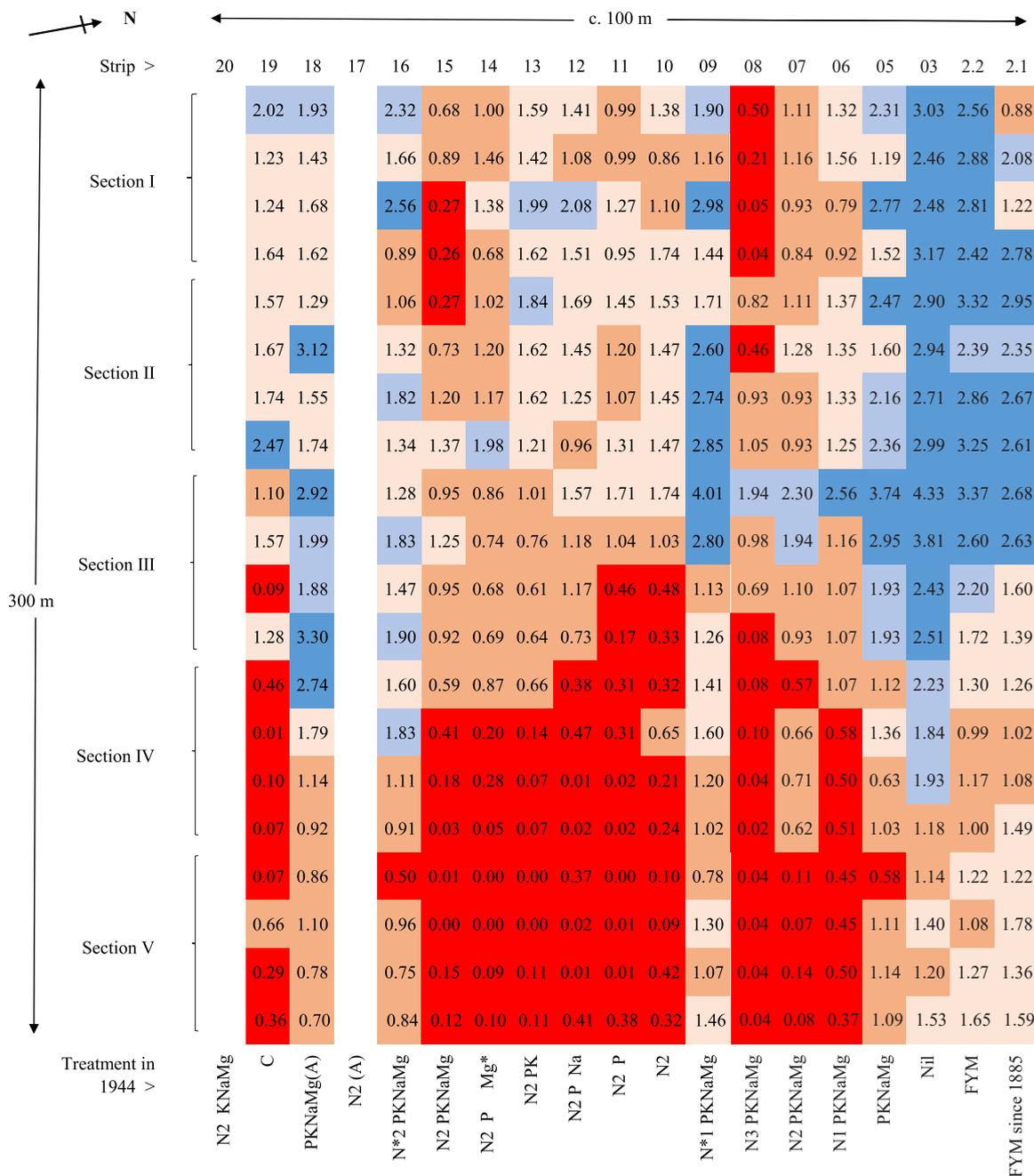


Fig. 1. CaCO₃ in topsoil, 0 – 23 cm, of the Broadbalk wheat Experiment in 1944. N applied as either ammonium sulphate (or mixed ammonium salts) or N* as sodium nitrate; for full details of treatments in 1944 see Poulton et al., 2024. Strips 17 and 20 not analysed. ■, < 0.60%; ■, 0.61 – 1.20%; ■, 1.21 – 1.80%; ■, 1.81 – 2.40%; ■, > 2.41%. For full details of treatments in 1944 see Poulton et al., 2024.

(Johnston et al., 2013; Johnston et al., 2016).

On the sandy soil at Woburn which had little or no reserve of CaCO_3 , problems with acidity became apparent within a few years of the start of the Continuous Wheat and Barley experiments which began in 1876. Liming started in 1897, but yields did not recover (Russell and Voelcker, 1936). In retrospect, it is probable that the yields were also affected by cereal cyst nematodes which can be a problem when continuous cereals are grown on lighter textured soils.

On all the arable experiments at Rothamsted and Woburn, basal or differential dressings of chalk have been applied to correct acidity. It was not thought desirable to incorporate a test of liming within the existing arable LTEs so two separate experiments were established at Rothamsted and Woburn in 1962, specifically to investigate the effects of lime on the yield of arable crops (Bolton, 1977). We have visited LTEs where > 50% of the plots are bare because the soil is too acidic to allow the growth of the intended crop. An often-quoted reason for why this continues year-after-year is that these were the original treatments, and therefore cannot be changed. This does not seem to be a valid reason to us. Once the serious impacts of soil acidification on crop growth have been established, many years of poor or virtually non-existent yields does not provide further useful information, so a change would seem vital. We submit that LTEs should not be regarded as museum exhibits that can never be changed. Indeed, after careful consideration, they should be modified, if necessary, to ensure their continued value and relevance.

However, we do maintain a narrow (3 m drill width \times 200 m in length) strip of land adjacent to the Exhaustion Land Experiment (see 3.2.9.) which, by reason of uneven, historical applications of chalk, has topsoil ranging from pH 8 at one end to pH < 4.5 at the other (Rousk et al., 2009). Termed, the Acid Strip, wheat is usually drilled along the whole length every year but dies out at a pH of c. 4.5–5.0. It is a graphic illustration to visitors of the need to maintain soil pH for arable crops. It has been used to study the effect of pH and the lack of a crop on the microbial community (Rousk et al., 2009).

In contrast to the arable LTEs, liming treatments have become an integral part of the Park Grass experiment at Rothamsted; possibly the foremost ecological experiment in the world (Silvertown et al., 2006). Started in 1856 on long-established pasture, its original purpose was to identify the best means of improving the yield of grass cut for hay. It tests different combinations of organic manures and inorganic fertilisers, with rates of N applied as either sodium nitrate or ammonium sulphate. Within a few years the different treatments had a very noticeable effect on the species composition of the sward and botanical surveys were started (Lawes et al., 1882). A small test of liming was conducted by Lawes and Gilbert in the 1880s but in 1903 the new Director instigated a major change. Most of the plots were divided in 1903 or 1920 and 4 t ha^{-1} CaCO_3 was applied every fourth year to half of each plot (Glen-dining et al., 2021). However, over the next 45–60 years, plots receiving N as ammonium sulphate became progressively more acid even where chalk was applied, while plots receiving N as sodium nitrate did not. Consequently, the effect of the different forms of N on species number and diversity was confounded by acidity. In 1965 most plots were divided again into four sub-plots. On three sub-plots (“a”, “b” and “c”), differential amounts of chalk were applied, where necessary, to raise and maintain soil pH (0–23 cm) at 7, 6 or 5 respectively. The “d” sub-plots do not receive any chalk and are essentially a continuation of the original treatments (Thurston et al., 1976). The effects of the various fertiliser, manure and liming treatments on species number, diversity and the stability of plant communities was discussed by Crawley et al. (2005). The effect on soil pH is shown in Fig. 2 while Fig. 3 shows how species numbers recovered on the ammonium sulphate plots when the soil pH was increased and maintained at pH 7.

To further investigate the interaction between the amount and form of N applied, soil pH and species number, changes were made in 1989, prior to the herbage harvests of 1990. Two plots (each comprising four sub-plots) receiving PKMgNa and 96 kg N ha^{-1} yr^{-1} as either

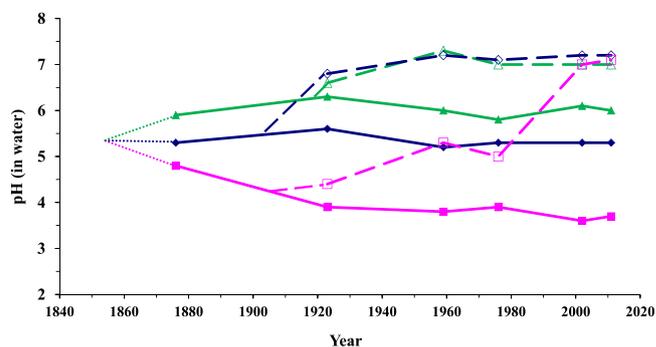


Fig. 2. Soil $\text{pH}_{(\text{in water})}$ of topsoil, 0–23 cm, under selected treatments of the Park Grass Experiment. Solid lines, no chalk; dashed lines, plus chalk: \blacklozenge , No fertiliser; \diamond , No fertiliser + chalk; \blacktriangle , 96 kg N ha^{-1} as sodium nitrate + PKMgNa; \triangle , 96 kg N ha^{-1} as sodium nitrate + PKMgNa + chalk; \blacksquare , 96 kg N ha^{-1} as ammonium sulphate + PKMgNa; \square , 96 kg N ha^{-1} as ammonium sulphate + PKMgNa + chalk. (Macdonald et al., 2018).

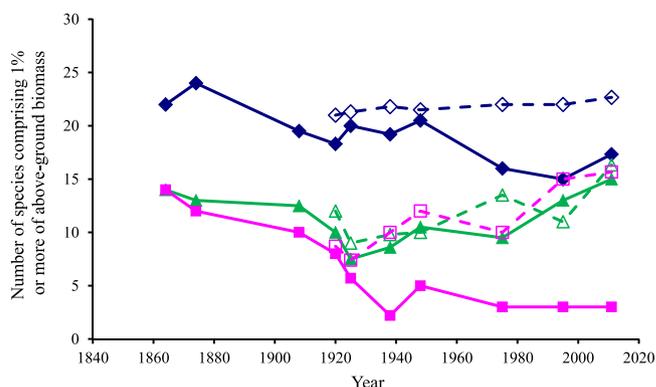


Fig. 3. Changes in the number of species comprising 1% or more of the above-ground biomass under selected treatments of the Park Grass Experiment. Solid lines, no chalk; dashed lines, plus chalk: \blacklozenge , No fertiliser; \diamond , No fertiliser + chalk; \blacktriangle , 96 kg N ha^{-1} as sodium nitrate + PKMgNa; \triangle , 96 kg N ha^{-1} as sodium nitrate + PKMgNa + chalk; \blacksquare , 96 kg N ha^{-1} as ammonium sulphate + PKMgNa; \square , 96 kg N ha^{-1} as ammonium sulphate + PKMgNa + chalk. (Macdonald et al., 2018).

ammonium sulphate since 1856, or sodium nitrate since 1858, were divided and N withheld from one half. Applications of PKMgNa and chalk continued. Table 2 shows the effect on soil pH, yield and the number of species after > 20 years of withholding fertiliser N. Chalk is now applied every third year (if necessary) to maintain the “a”, “b” and “c” sub-plots at pH 7, 6 and 5 respectively; the “d” sub-plots do not receive chalk. Where sodium nitrate was withheld (plot 14/1), soil pH (0–23 cm) declined by about 0.5 unit on the “c” and “d” sub-plots probably reflecting the input of rainfall which had a lower pH of c. 4.9 during that time. Mean pH of the rain was c. 4.6 in 1992–95, c. 5.2 in 2007–10 and c. 5.7 in 2019–22 (Tony Scott, *personal communication*), becoming less acidic as emissions from heavy industry and coal-fired power stations have declined. Yields without added N are similar to those with N (Table 2) almost certainly because of the increase in leguminous species on all subplots (Macdonald et al., 2018). In contrast, where ammonium sulphate has been withheld (plot 9/1), soil pH on the “c” sub-plot has been maintained by liming at 5, but on the unlimed “d” sub-plot the pH has increased slightly. This is presumably a result of the absence of ammonium inputs that lead to acidification during nitrification of the applied N and because of the input of rain with a higher pH than the soil on this sub-plot. Although species numbers have increased on the limed sub-plots where N has been withheld there has not been a significant rise in legumes on the “d” sub-plot because it is too acidic to support most

Table 2

Effect on soil pH, yield and species number on plots receiving PKMgNa and which either last received fertiliser N in 1989 or which continue to receive fertiliser N. Park Grass, Rothamsted.

Sub-plot	Mean dry matter yield ^b		Total Species No. ^c	Mean dry matter yield ^b		Total Species No. ^c
	pH ^a	t ha ⁻¹ yr ⁻¹		pH ^a	t ha ⁻¹ yr ⁻¹	
Ammonium sulphate (plot 9)						
N withheld since 1990 (plot 9/1)						
a	7.1	6.4	30 (6)	7.1	7.1	23 (7)
b	6.4	6.9	34 (6)	6.2	7.0	24 (8)
c	5.2	5.2	31 (10)	5.1	5.7	28 (8)
d	4.1	1.5	11 (2)	3.7	4.6	3 (2)
Sodium nitrate (plot 14)						
N withheld since 1990 (plot 14/1)						
a	6.9	6.5	29 (8)	7.0	6.5	25 (7)
b	6.0	6.9	29 (8)	6.2	7.0	28 (8)
c	5.3	7.2	31 (12)	5.9	6.8	25 (9)
d	5.4	7.0	26 (11)	6.0	6.3	22 (8)

number of species comprising 5% or more of the herbage (Adapted from Macdonald et al., 2018)

^a Soil pH(in H₂O) in 0 - 23 cm soil layer, 2011.

^b Mean total dry matter 2010 - 2012; 2 cuts each year.

^c Total number of species; mean of botanical surveys 2010 - 2012. Figures in brackets are the

leguminous species. Consequently, yields on 9/1d are poor compared with those on 9/2d which, although very acid, continues to receive N. Storkey et al. (2015) showed that, except where there was extreme acidification, the proportion of legumes in the sward, species richness and diversity increased across the whole experiment between 1991 and 2012 as atmospheric N deposition declined by about 50%.

3.1.2. Weeds

Keeping experiments free from weeds can be a major challenge. On the arable experiments at Rothamsted weeds were originally controlled by hand-hoeing and occasional bare fallows. When labour became scarce in the early 20th Century, it became more difficult to control weeds and cereal yields were reduced considerably Fig. 4 (Poulton et al.,

2024). On Broadbalk, the experiment was large enough (most treatment strips are 300 m in length) to allow its division into five sections; regular bare following (one year in five) was introduced sequentially on the sections to control the weeds and wheat yields recovered. Herbicides were introduced in the 1950 s and some sections reverted to continuous wheat. However, half of one section (now Section 8) has never received herbicides. The yields of wheat are much reduced, and it needs to be fallowed regularly in an attempt to control weeds. There are more than 50 arable weeds in this section, some of which are amongst the rarest in the UK. The weeds are a valuable source of material for scientists working on herbicide resistance (Hicks et al., 2018).

3.1.3. Crop nutrients

Nutrient management is essential in any LTE. Many are established to study the effect of different combinations of organic manures and/or inorganic fertilisers on yield. Their purpose might be to test the effect of e.g. applied N fertiliser on yield which would require that other nutrients (or weeds, pests and disease) do not limit yield. Usually this means that those other nutrients (P, K etc.) are applied as basal dressings. However, care must be taken to ensure that differential offtakes of these basal nutrients does not confound results. In a Ley-arable experiment that was established at Woburn in 1938 to compare the yields of arable “test” crops grown in a continuous arable rotation or as part of a ley-arable rotation (Poulton et al., 2023) the arable crops and the 3-year ley crops of grass, grass-clover or lucerne (*Medicago sativa*) received the same basal fertiliser applications. However, the lucerne leys removed far more K than the other crops and later measurements also showed that, on this sandy soil, much K was also leached, particularly on soils in continuous arable cropping. Soil samples taken in the 1950 s showed that, on some treatments, concentrations of exchangeable K in the topsoil were too low and might limit yield (Johnston et al., 2022). To overcome this problem, basal K was increased on all treatments and additional “corrective” K was (and still is) applied every fifth year, mainly to soils in continuous arable rotations which are more prone to losing K by leaching. The aim of the “corrective” K is to raise the concentration of exchangeable K in the topsoil of all plots to 250 mg kg⁻¹ (Johnston et al., 2022). Fertiliser applications are now reviewed and updated regularly if needed (Poulton et al., 2022). Plant-available P (Olsen P) is maintained at > 26 mg kg⁻¹, and exchangeable Mg was

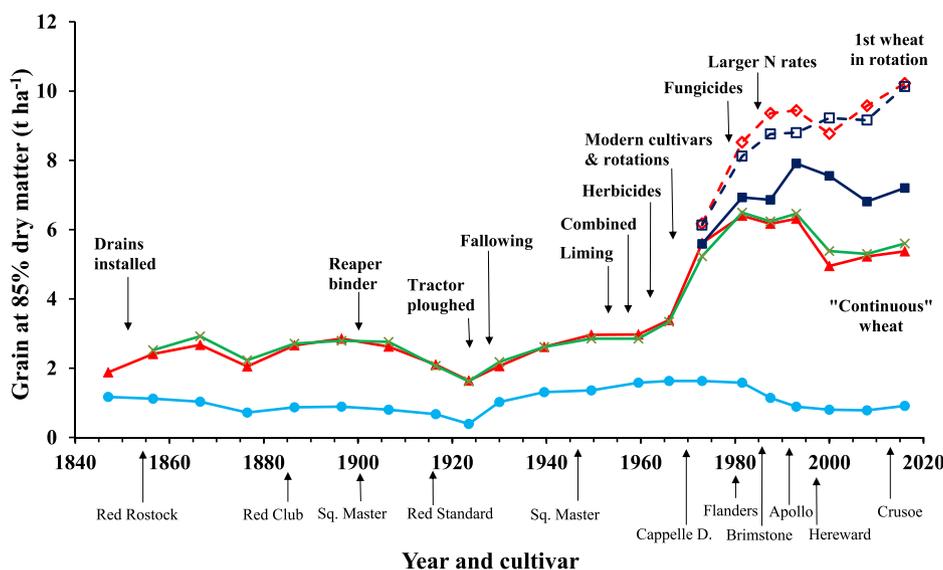


Fig. 4. Mean yields of wheat grain, t ha⁻¹ at 85% dry matter on selected treatments of the Broadbalk Wheat Experiment for 1844–2018, showing changes in cultivar and management. Continuous wheat (mean of Section 1 and 9 after division into 10 sections): ●, Nil; ×, (P)KMg+ 144 kg N ha⁻¹; ■, largest yield with (P)KMg+N (max. 288 kg N ha⁻¹); ▲, FYM. 1st wheat in rotation: □, largest yield with (P)KMg+N (max. 288 kg N ha⁻¹); ◇, FYM+96 kg N ha⁻¹ (+144 kg N ha⁻¹ since 2005).1852–2018. Current treatments given. (Adapted from (Perryman et al., 2023)).

increased by the use of magnesian limestone as a liming agent. Because of the need to maintain adequate nutrients and problems in the 1950 s and 1960 s with soil-borne pests such as stem eelworm (*Ditylenchus dipsaci*) and potato cyst nematodes (*Globodera rostochiensis*, *G. pallida*) which necessitated changes in the crops being grown, it was not until many years after the experiment started that its original purpose, *i.e.* to look at the effects of continuous arable or ley-arable rotations on the yield of two following “test” crops could be realised (Johnston et al., 2022), and incidentally to look at the effect on soil organic matter (Johnston et al., 2017).

3.2. Planned changes

Because the Classical experiments at Rothamsted have been running for so long there have been many changes which we now take for granted but which would have been subject to much discussion at the time. Some of these changes may be relevant to more recently established experiments around the world. Many of the examples cited below reflect the continued development of agriculture from the 19th Century to the present. Not incorporating (or at least acknowledging) such developments into the Rothamsted experiments would have diminished their continued relevance to contemporary agriculture. We again discuss the issues faced, the approach adopted, and the lessons drawn.

3.2.1. Mechanisation

The older arable experiments at Rothamsted were originally ploughed by oxen and horses, probably to a depth of about 10 – 15 cm. Tractors were introduced in the early 20th Century and plough depth increased to about 20 cm. With more powerful tractors, plough depth has still been limited to < 23 cm. This is important because almost all of the topsoil samples have been taken to a depth of 23 cm (9 in.), and soils taken now are directly comparable with soils taken 100 or 150 years ago. Where there is a chance that the depth of ploughing may be increased, great care must be taken to ensure that soils sampled before or after any change are comparable, possibly by taking soils immediately before any change to a depth which exceeds the likely new plough depth.

Another obvious change has been in the harvesting method used. Arable crops were originally cut by hand with a scythe; a reaper-binder was first used on Broadbalk in 1902, but sheaves were still stooked in the field before transfer to a threshing barn. A combine harvester was first used on Broadbalk in 1957.

On Park Grass the plots were originally cut by hand and mowing machines were first used in the 1880 s. The first cut was always made into hay and weighed as such until 1959; the second cut is weighed green. In 1959 a flail-type forage harvester (taking two or four cuts per sub-plot) was compared with the usual cutter-bar machine for the first cut. The cut herbage from the forage harvester was weighed green. A sample for analysis and archiving was then taken but the remainder of the cut herbage *and* the rest of the plot was made into hay in the usual way, *i.e.* continuing what had been standard practice in the preceding 100 + years. Since 1960 a forage harvester has been used for both cuts. In 1992–94 further comparisons were made between yields obtained by forage harvester and those of hay which will be subject to some losses during the hay-making process in the field. These comparisons have allowed us to derive a conversion factor which can be used to compare yields pre- and post-1960. It is important to monitor and record such changes, ideally over several years.

3.2.2. Changes in crop varieties and rotations

For an experiment to remain relevant to farming practice in the region, changes to the cropping may be necessary. In the 1950 s/1960 s a fundamental, world-wide change saw the introduction of short-strawed wheat cultivars as part of the “Green Revolution”. In the UK most farmers were using the newer varieties by the 1960 s and, at Rothamsted, all the cereals grown on the LTEs were changed to short-strawed

cereal cultivars. On the Broadbalk Wheat experiment (Fig. 4) this was accompanied by other major changes (Poulton et al., 2024). The experiment was further divided into 10 sections and while some sections reverted to continuous wheat (after the following regime which had been introduced to control weeds) others grew a rotation of crops; the rotations have changed several times. The short-strawed cultivars allowed the maximum rate of fertiliser N to be increased to 192 kg ha⁻¹ with less chance that the crop would “lodge” (fall over) in heavy rain. The main form of N tested was changed from ammonium sulphate to calcium ammonium nitrate and later to straight ammonium nitrate. The introduction of summer fungicides in 1979 meant that the higher grain yielding potential of the modern cultivars could be exploited by further increasing the maximum rate of N tested to 288 kg ha⁻¹ (Poulton et al., 2024). Lawes and Gilbert changed cultivars occasionally, presumably choosing varieties that were being grown locally. We change cultivars regularly and consider that this is essential to maintain the relevance of the experiment. We choose varieties that are expected to stay on the national recommended lists for at least five years (five of the 10 Sections on Broadbalk being in a 5-course rotation), thus providing a degree of continuity. Bread-making varieties rather than feed wheats are usually chosen. Very occasionally, poor weather in the autumn has meant that a spring sown wheat has been grown instead. Interestingly, this has occurred twice in the last decade.

The inclusion of other crops in rotations with wheat on Broadbalk from 1968 allowed us to compare the yields of continuous wheat with those of the 1st wheat after two break crops, where the effects of take-all (*Gaeumannomyces graminis* var. *tritici*) were minimised. This change has been a major benefit which allows us to compare yields of continuous wheat with the more commercially common growth of wheat in a rotation and providing data on the yield potential of wheat grown under more conducive conditions. Table 3 gives the yield of a modern cultivar, *c.v.* Crusoe, grown either continuously or as the 1st wheat crop after a 2-year break, with either FYM or adequate P, K, Mg, and rates of N. It also gives the yields of the 2nd and 3rd wheat after a 2-year break where take-all limits yield. Yields of the 1st wheat after a 2-yr break were greater than those of continuous wheat or of the 2nd and 3rd wheat crop. For example, for treatments receiving N, yields in the 1st year after a break crop were 43 – 75% greater than from continuous wheat.

Table 3

Yields of winter wheat grain, t ha⁻¹ at 85% dry matter, grown continuously or as the 1st, 2nd or 3rd wheat after a 2-year break from cereals; with combinations of FYM, P, K, Mg and N given as single or split dressings; *c.v.* Crusoe, 2013–2018 (excluding spring wheat in 2015). Broadbalk Wheat Experiment.

	Continuous	1st	2nd	3rd
Treatment ^a	Section 1	wheat	wheat	wheat
FYM	4.89	7.00	5.63	5.28
FYM N3	6.95	10.35	10.03	8.65
Nil	1.01	1.89	0.90	1.01
(P)KMg	1.24	1.72	1.02	1.15
N1 (P)KMg	2.80	4.35	3.12	3.05
N2 (P)KMg	4.38	6.69	4.85	4.37
N3 (P)KMg	5.14	8.61	5.96	5.72
N4 (P)KMg	5.86	9.40	7.16	6.55
N5 (P)KMg	6.53	9.92	7.09	6.56
N6 (P)KMg	6.48	10.25	8.40	7.53
N1 + 2 + 1 PKMg	5.81	9.44	8.11	7.44
N1 + 3 + 1 (P)KMg2	6.71	10.01	8.57	7.50
N1 + 4 + 1 PKMg	6.76	10.54	9.32	8.41

N1, N2, N3, N4, N5, N6: 48, 96, 144, 192, 240, 288 kg N ha⁻¹

Split N applications in mid-March, mid-April, mid-May.

N1 + 2 + 1: 48 + 96 + 48 kg N ha⁻¹

N1 + 3 + 1: 48 + 144 + 48 kg N ha⁻¹

N1 + 4 + 1: 48 + 192 + 48 kg N ha⁻¹

For full details of treatments see Poulton et al. (2024)

^a Single N applications in mid-April.

3.2.3. Timing of N fertiliser applications

In 2000, further changes were made on the Broadbalk experiment as some of the treatments no longer reflected common practice. Since 1968, all spring applied N was given as a single application at the time of stem extension, regardless of the amount being applied. However, by the 1980 s, farmers were recommended to apply spring N as two or three split applications for amounts over 100 kg N ha^{-1} (RB209 (5th Edition), 1988) as this was intended to decrease N losses and increase crop utilisation of N. After much discussion it was therefore decided that four treatments that were considered less useful would be changed so that the four highest rates of N ($144, 192, 240$ and 288 kg ha^{-1}) given as three split dressings could be compared with the same amounts given as single applications. Surprisingly, on this soil type, for continuous wheat and the 1st wheat after a 2-year break there was no significant difference in

yield between single and split applications (Table 3). However, for the 2nd and 3rd wheats after a 2-yr break there were, on average, large benefits in favour of the split applications. Because of take-all these crops had a poorer root system which meant that they could, in some years, make better use of split rather than single applications of N (Poulton et al., 2024). Note that, in Table 3 yields on plots given 144 kg ha^{-1} as three split dressings are not shown as, for this particular treatment, they are now limited by lack of P; Olsen P on this treatment ranged from $16\text{--}30 \text{ mg kg}^{-1}$ in 2000 and some plots are now as low as 11 mg kg^{-1} (the intention is that this will become a low P treatment; see 3.2.4.). Meaned over just three years of the most recent variety (c.v. Zyatt) there has also been a benefit from the split applications for both continuous and 1st wheats, although it has been very variable.

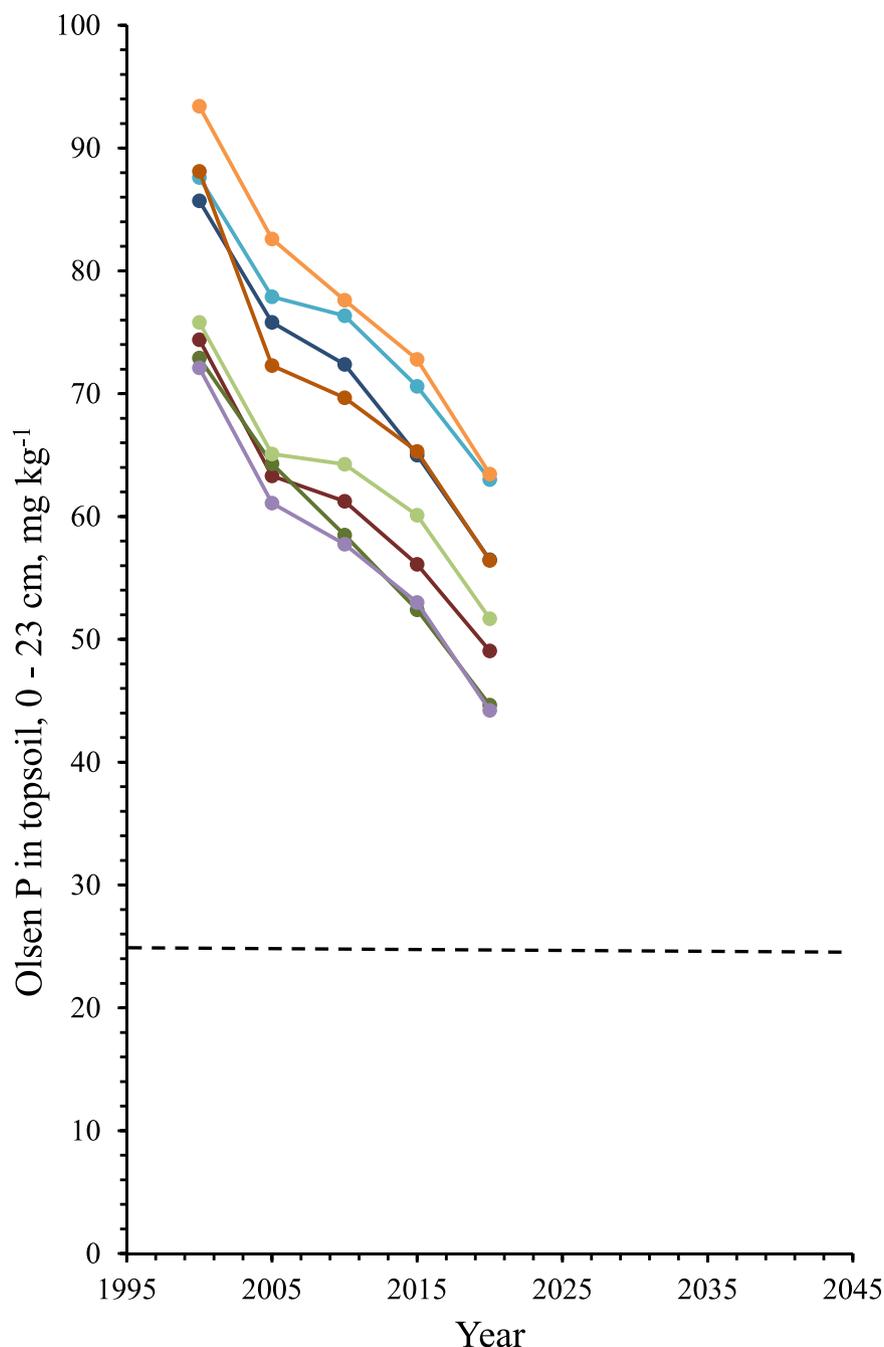


Fig. 5. Decline in Olsen P in topsoil, 0 – 23 cm, on selected treatments of the Broadbalk Wheat Experiment from which fertiliser P has been withheld since 2000. Different symbols/colours represent different treatment strips. The dashed line represents the top of P Index 2 (RB 209 (5th Edition), 1988).

3.2.4. P fertiliser applications

In 2000, after more than 150 years, it was decided that P fertiliser would no longer be applied on some plots of the Broadbalk experiment. The reason was that, in 2000, the concentrations of plant-available P (Olsen et al., 1954) in the topsoil ranged from 65 - 100 mg kg⁻¹, well above the value at which farmers at that time were being advised to maintain their soils, typically P Index 2 (16 - 25 mg kg⁻¹ Olsen P) in the UK advisory system (RB 209 5th edition, 1988). At these high concentrations of Olsen P in the topsoil there is also a risk of phosphate being leached into drainage ditches where it might lead to eutrophication of surface waters (Heckrath et al., 1995). This situation had occurred because, since the start of the experiment, P offtake in the crop was often much lower than the 35 kg P ha⁻¹ that was applied; Lawes had knowingly applied more fertiliser P than the crop removed, presumably to ensure that lack of P did not limit yield (Lawes and Gilbert, 1873). In 2000, it was therefore decided that P applications should be withheld on some plots until Olsen P declined to a more sensible agronomic value, *i.e.* about 30 mg kg⁻¹ Olsen P. Fig. 5 shows the decline in Olsen P over c. 20 years since fertiliser P was withheld on plots that had been receiving P, K and Mg and different amounts of N. Previous work (Johnston et al., 2016) has shown that the rate of decline will slow, and it will take perhaps another 20 years before more acceptable values are reached on those plots in this group which currently contain the least Olsen P, at which point they will all start to receive applications of P fertiliser again, possibly at a lower rate. But, for some years to come, we are confident that withholding fertiliser P applications will not adversely affect crop growth and yield.

3.2.5. Sulphur (S) from air pollution or fertiliser

In the 1980 s, air pollution in the UK from coal-fired power stations, heavy industry and domestic central heating meant that atmospheric inputs of S were as high as 65 kg ha⁻¹; the only positive benefit of this level of pollution was that the amount of S was more than sufficient to meet the needs of arable crops. By 2000, this input had declined to about 5 kg S ha⁻¹ and was no longer enough to fully meet the needs of a higher yielding cereal crop (Zhao et al., 2003). On Broadbalk, S was applied by default as part of the K and Mg fertiliser dressings as these were applied as sulphates. In 2001, in order to enable a comparison of yields with and without S fertiliser, but with other nutrients being non-limiting, the usual form of K fertiliser, K₂SO₄ was replaced by KCl and previous large inputs of MgSO₄ were stopped on one treatment strip. Yields without S (but with N, K and adequate concentrations of Olsen P and exchangeable Mg) can now be compared with plots receiving S and N, K, P and Mg. Since 2000, the benefit to grain yield has been modest at about 0.5 t ha⁻¹ (mean of continuous wheat and 1st wheat after a 2-year break). However, it is known from other work that lack of S can have more significant effects on wheat grain composition with implications for bread making quality (Poulton et al., 2024).

3.2.6. Land-use change: removing land from arable cropping

Major changes were made to parts of the Broadbalk Wheat and Geescroft Bean experiments in the 1880 s. On Broadbalk, a small unfertilised area across the top (western) end of the experiment was fenced off and abandoned. Lawes was interested in seeing how long the unharvested wheat would survive. A few years later a larger area, previously in continuous beans in Geescroft field, was also abandoned (Lawes, 1895). Both are now maturing woodland. Although these two "Wilderness" areas are not experiments in the true sense, they are nonetheless, informative. Botanical surveys have been carried out regularly (Brenchley and Adam, 1915; Harmer et al., 2001) and soils have been sampled occasionally (Hall, 1905; Jenkinson, 1971). Unlike Broadbalk, the Geescroft site had not received large inputs of chalk before the experiment began; the topsoils of both now have a soil pH of 7.7 and 4.4 respectively (Poulton et al., 2003). Both sites accumulated significant amounts of soil organic carbon (SOC), with mean rates of 0.54 and 0.38 t SOC ha⁻¹ yr⁻¹ in Broadbalk and Geescroft respectively

over > 100 years, but much more, *c.* 2.85 and 1.62 t SOC ha⁻¹ yr⁻¹ respectively, has been accumulated in the above-ground biomass (*i.e.* the trees) (Poulton et al., 2003; Poulton et al., 2018). Although these sites were not established as experiments to measure changes in C stock, and they have limitations, especially because Broadbalk Wilderness is a small area, they are nevertheless valuable because there are relatively few sites globally where such data is available. One important aspect of this work is the change in bulk density/soil weight over the 100 + years since the change was made. Prior to the sites being abandoned in the 1880 s, the topsoil, 0 - 23 cm, weighed 2.87 and 2.73 M kg ha⁻¹ on Broadbalk and Geescroft respectively. In 1999, the topsoil weight to the same depth had declined to 2.05 and 2.19 M kg ha⁻¹ respectively. If one wants to calculate the change over time in the *amount* (*i.e.* stock) of SOC or N in the soil (as opposed to their concentration), the same mass of soil needs to be considered both before and after the change in management. If the soil has been sampled to depth on each occasion, and the weight measured, then the SOC in the extra soil that should have been sampled can be calculated. This is illustrated graphically in Fig. 6 which shows, for Broadbalk Wilderness, the cumulative soil weight down to a depth of 69 cm in 1881 and 1999. By referencing the 0 - 23 cm soil weight measured in 1881 to the same weight in 1999 and then referencing across to the depth ("y") axis shows that, to account for the same mass of soil in 1999, an extra 8.5 cm depth of soil needs to be considered. The same correction needs to be made where soils have changed from continuous arable to ley-arable cropping or permanent grass or where large amounts of organic manure are applied. The reverse correction needs to be made where soils have been cleared from woodland or ploughed out of long-term grass and the bulk density has increased (Poulton et al., 2018).

3.2.7. Effects of old and new manure applications on spring barley yields and soil properties

On the Hoosfield Barley experiment which started in 1852, one of the treatments received 35 t FYM ha⁻¹ yr⁻¹. After 20 years, Lawes and Gilbert decided that they would stop the FYM application on one half of the original treatment so that they could measure the residual effect of the manure on the yield of spring-sown barley. They knew that > 4 t N ha⁻¹ had been applied in the manure during the previous 20 years, that about 15% had been removed in the crop, that soil N had increased by about 50% (Dyer, 1902) and that much N would have been lost by leaching. In a publication marking 50 years of experiments at Rothamsted, Lawes and Gilbert (1895) showed that more than 20 years after applications stopped the yields with FYM residues were about half of those achieved where applications continued but were still more than double those on the Nil plot with no inputs. Fig. 7 shows that 150 years after applications stopped the FYM residues plot still yields almost 2 t ha⁻¹ more than the Nil. This is almost certainly because of the P that was applied in the FYM; the topsoil of the FYM residues and the Nil plots contained 9 and 3 kg ha⁻¹ Olsen P respectively in 2018.

For about 30 years after the experiment was modified in the 1960s yields on plots receiving PKMg + sufficient N often matched those receiving FYM. However, since the mid-1990s, this has not been the case; plots receiving PKMg + N have not always given the same yield as those on the FYM treatment. The differences in yield were almost certainly due to the large difference in SOC on the two treatments; 1% and 3.5% SOC on the PKMg and FYM plots respectively. With a short growing season compared to winter wheat, we hypothesise that the spring barley benefitted from the better soil structure, improved water-holding capacity and from the extra N that was mineralised and made available to the growing crop during the growing season, and in parts of the soil profile that were not mimicked from spring applied N fertiliser. The spring barley variety grown has been changed several times since the 1960s (usually every four or eight years to tie-in with a 4-year cycle of rotating N rates) and the newer, higher yielding cultivars were able to exploit these conditions. The question is, which factor is more important, the improved structure or the extra mineralised N? To help answer

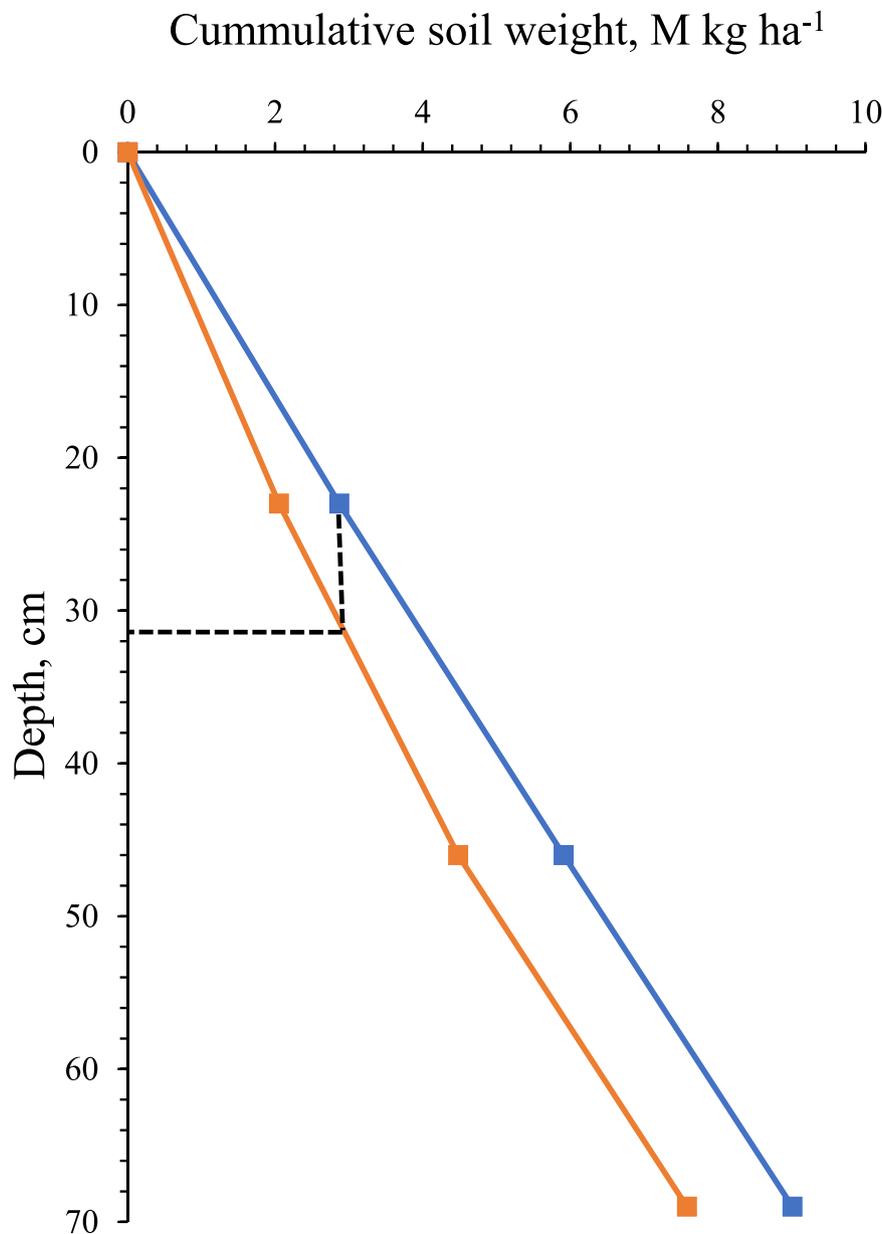


Fig. 6. Changes in cumulative soil weight with depth over time in Broadbalk Wilderness. Assuming that there has been a significant decline in bulk density over time the graph shows how to calculate how much deeper the soil should have been sampled so that the same mass of soil is considered on each occasion. ■, cumulative soil weight in 1881; ■, cumulative soil weight in 1999. (Poulton et al., 2003).

this question a new FYM treatment was started in 2001 (Macdonald et al., 2018). The idea was to see how quickly yields on the new FYM treatment increased to match those of the long-established FYM treatment. If the increase was rapid, say within five years, it could be argued that this was due mainly to the extra N available to the crop; if the increase was slower, say 10 years or more, then it is more likely to be structural. It is also possible that a relatively small increase in SOC within a few years may result in a proportionately larger increase in a fraction within the total soil C important for soil structure. Within 2–3 years there was a clear benefit of the new FYM treatment compared with the NPKMg treatment, which could be assumed to be mainly an N effect. Fig. 7 shows the grain yield with the long continued and the “new” FYM treatments for c.v. Tipple grown between 2008–2015. After 16+ years yields on the new FYM treatments were still about 10% less than on the plots that have received FYM since 1852, which implies that there is also an underlying structural effect which has yet to be fully realised. By 2013 the topsoil on the new FYM treatment contained 0.18% N while

the long continued FYM soil contained 0.34% N. In contrast, the NPK soils contained about 0.10% N. A highly significant downside of the large manure applications is that there will be considerably more nitrate leached from the FYM soils than from those receiving fertiliser N (Powelson et al., 1989). Large nitrate leaching losses where the high rate of FYM has been applied annually have also been observed in the Broadbalk experiment (Goulding et al., 2000).

3.2.8. Reversing long-term cropping treatments: 1) the Highfield Ley-arable Reversion Experiment

The Highfield Ley-arable experiment, started in 1949, compared crop performance and soil properties in different rotations comprising continuous grass, various ley/arable rotations, and continuous arable. The experiment was established on an area that had been in continuous grass since 1838, so, inevitably, ploughing the grass to start the arable treatments, led to a decline in SOC content. In addition to the ley-arable experiment, an adjacent area of the old grassland was ploughed in 1959

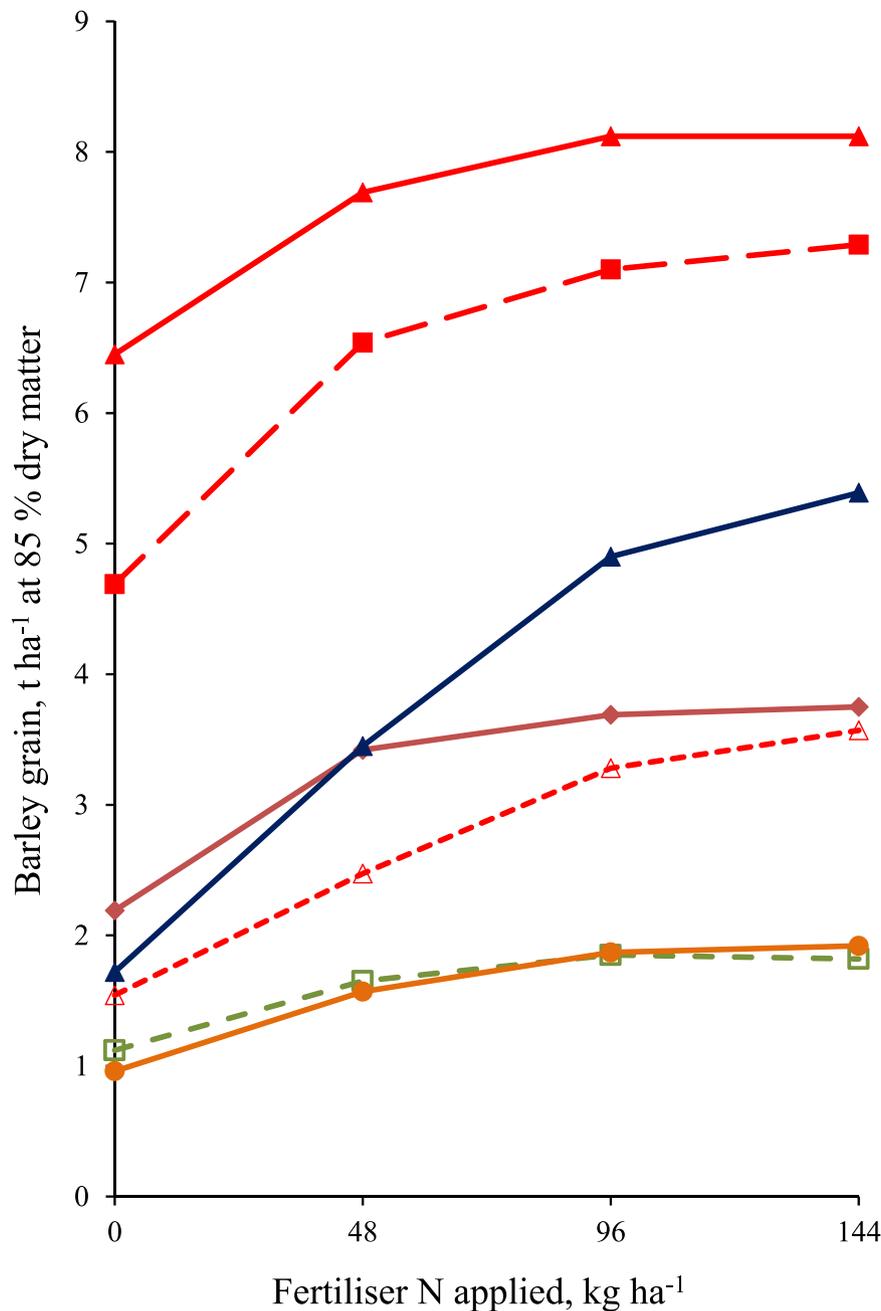


Fig. 7. Mean yields on selected treatments of the Hoosfield Barley Experiment in 2008 – 2015 of spring barley grain (cv. Tipple), given N fertiliser and: □, Nil; ●, K only; ◆, P only; ▲, PKMg; △, FYM, 1852–1871 only; ■, FYM since 2001; ▲, FYM since 1852. (Macdonald et al., 2018).

and has since been kept bare of all vegetation by regular tillage; as expected, the decline in SOC was even greater than in the treatments under arable crops. Results from this experiment, and another established at the same time but on an old arable field (the Fosters Ley-arable experiment), clearly demonstrated SOC changes on a decadal time scale resulting from different management practices and are frequently cited as examples for teaching purposes and used for testing SOM models. (Johnston, 1997; Macdonald et al., 2018). The comparison of the same soil under grass, arable or bare fallow for several decades has also proved valuable as a testbed for many studies on organic matter effects on soil properties. A recent example is a microbial study demonstrating, among many other results, that methods commonly used to assess

microbial diversity in soil do not adequately capture important facets of community adaptation to stresses (Hirsch, 2009; Neal et al., 2020, 2021). Data and soil from the Bare Fallow area has been used, in conjunction with similar bare fallow sites in several other European countries, in studies to quantify the most stable fractions of soil organic matter and to test models that include this fraction (Barré et al., 2010) and to investigate the chemical nature of strongly stabilised fractions and the microbial and chemical factors contributing to stability (Barré et al., 2018).

In 2008, more than 50 years after the Highfield Ley-arable and Bare Fallow experiments were started, it was decided to reverse the treatments within specific plots or areas of the experiments. New arable

(wheat), grass and bare fallow treatments were superimposed on plots or areas previously under long-term arable, grass or fallow (*i.e.* plots or areas were split in three with the continuation of the long-term treatment and the imposition of the two other treatments). After 2.5 years of the altered treatments measurements were made of soil aggregate stability using wet sieving to determine mean weight diameter (MWD) of stable aggregates (Redmile-Gordon et al., 2020). In the fallow-to-grass transition, total SOC increased by 64% during this period but aggregate stability as measured by MWD increased by 125%. The opposite trend was apparent in the grass-to-fallow transition: SOC decreased by 37% but MWD decreased by 78%. The changes in aggregate stability were much more strongly correlated with changes in extracellular polymeric substances (EPS), especially EPS-protein, than with the changes in total SOC. The organic materials measured as EPS are considered to represent microbial metabolites, so the results confirm the importance of microbial activity, and the resulting metabolites, as influential in modifying soil physical properties. By comparison, the “legacy” effect of old organic matter derived from the previous grass was small.

In contrast to the results for aggregate stability, soil structure as revealed by X-ray computed tomography based on > 2 mm aggregates showed that changes in soil porosity mostly required 7–10 years to become apparent after land use change. One exception was that conversion from bare fallow to grass led to a change in the distribution of pore sizes within 2 years (Bacq-Labreuil et al., 2021).

The changed management in this experiment has also allowed other studies on the rate of change of soil physical properties at different spatial scales after land use change. Jensen et al. (2020) estimated pore size distribution using soil cores subjected to different matric potentials under laboratory conditions. Comparing the fallow-to-grass conversion with grass-to-fallow it was concluded that it was faster to degrade than to restore a complex soil structure; a finding of great significance when considering land management practices.

Hirsch et al. (2017) used the site to quantify changes in microbial and mesofauna populations after land use change. It was found that mesofauna populations changed rapidly after conversion, responding to either increased or decreased plant inputs. DNA- and 16sRNA-based methods were used to characterise changes in the population size and diversity of bacteria, archaea and fungi during the 3 years following conversion. With several of the microbial groups studied, recent management was more significant than previous history.

3.2.9. Reversing long-term cropping treatments: 2) the “Exhaustion Land” experiment – several reincarnations to address different questions

Most of the LTEs at Rothamsted have evolved gradually over time, sometimes with minor and occasionally with major modifications. Like the Highfield experiment described above, the “Exhaustion Land” experiment on Hoosfield, has had several distinct phases. Between 1856 – 1875 the influence of various combinations of N, P, K, Mg and Na on the yields of continuous winter wheat were tested (Johnston and Poulton, 1977). In the mid-19th Century devastating potato famines in Europe, particularly in Ireland, resulted in the deaths of millions of people. After another severe attack of potato late blight (caused by *Phytophthora infestans*) in 1873 the Council of the Royal Agricultural Society in the UK initiated investigations on the potato crop (Jenkins, 1874). One aspect that obviously interested Lawes and Gilbert was how fertilisers and manures might influence the severity of the disease. Consequently, in 1876, extra land was incorporated into the experiment, existing plots were divided, and the resulting 10 plots tested Nil, N, P, K, Mg, Na and FYM on the yield of potatoes and on the incidence of disease (Gilbert, 1888). Potatoes were grown continuously until 1901 with best yields obtained with full NPKMgNa applications or FYM (after several years application). Overall, yields declined in the later years (Warren and Johnston, 1960). Of interest is that for the susceptible cultivar grown between 1876 and 1879, Lawes and Gilbert recorded whether the harvested tubers were “good”, “small”, “bad” or “diseased” and dried,

ground, labelled and archived samples in the Rothamsted Sample Archive. In the early 21st Century, Ristaino et al. (2012), analysed archived samples of potato and was able to identify the Ia mtDNA haplotype of *P. infestans* in diseased tubers from the 1870s. Real-time PCR assays were used to detect and quantify *P. infestans* in the tubers. The level of pathogen DNA was greatest in tubers from higher yielding plots receiving NPKMgNa and least in FYM plots.

The potato experiment was stopped in 1901 and from 1902 - 1940, no further treatments were applied, cereals, mainly barley, were grown and the residual effects of the previous fertiliser and manure applications were measured. It is from this period that the Exhaustion Land name is derived because nutrients in the soil declined due to lack of inputs. Crucially, the site was not used for other experiments. In the 1930s, the effects of the previous treatments were still apparent, and, to exaggerate the differences, basal fertiliser N was applied from 1941. Fig. 8 shows that by 1981, plant available P (Olsen P) in the topsoil (0 – 23 cm) had declined markedly from the maximum measured values in 1903, but that there were still differences between soils that had received no P since 1852 and those that had received fertilisers or FYM between 1856 and 1901 (Johnston and Poulton, 1977).

To gain further information on the effect of P and K residues each of the 10 plots were divided to test four rates of N. However, in 1986, after > 80 years measuring the effect of residues in a situation where concentrations of plant-available P and K were declining, it was decided that there was little more to be gained by continuing this course of action and that it would be more useful to see how quickly Olsen P (and later exchangeable K) could be increased such that yields were not restricted. Therefore, from 1986, half of the experiment (20 sub-plots) started to receive basal K and N and tested four rates of fresh P fertiliser. Fig. 9 shows that within three years, concentrations of Olsen P on some sub-plots were such that the yields of spring barley were not further increased by adding more fertiliser P (Poulton et al., 2013). Five of the sub-plots did not receive fresh P and continue to measure the decline in Olsen P where no fertiliser P or FYM had been applied since 1856 or 1901 (the dark blue diamond symbols in Fig. 9).

3.3. LTEs as a record of environmental and management changes

3.3.1. An unexpected use of an LTE to investigate a European-wide issue

The Market Garden Experiment at Woburn was established in 1942 to test the effects on the yield of market garden crops of various organic manures at a time when it was becoming more difficult to import fertilisers (Johnston and Wedderburn, 1975). Compost, sewage sludge, sludge compost and vegetable compost were compared with FYM and fertilisers and were applied for 20 + years. Residual effects on arable

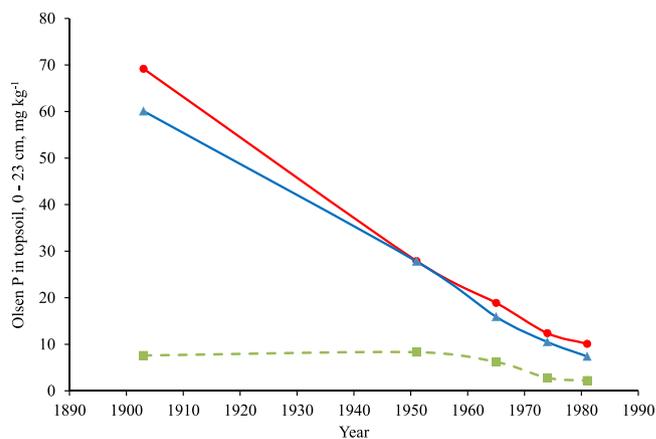


Fig. 8. Decline in Olsen P in topsoil, 0 – 23 cm, of the Exhaustion Land Experiment from 1903 – 1981. ■, No P since 1856; ▲, P fertiliser, 1856 – 1901; ●, 35 t ha⁻¹ FYM, 1876 – 1901. (Updated from Johnston and Poulton, 1977).

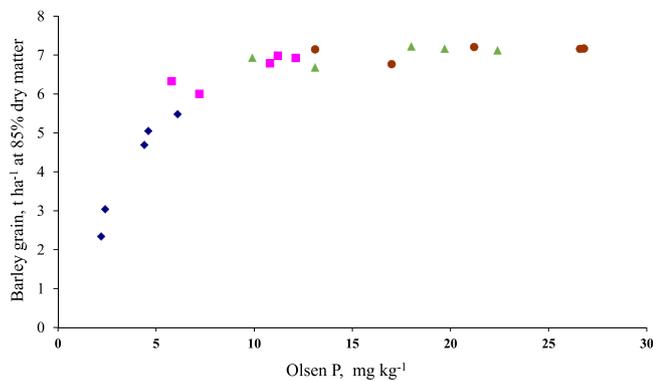


Fig. 9. Yield of spring barley grain ν Olsen P in topsoil, 0 – 23 cm, of the Exhaustion Land Experiment, showing how a plateau had developed after three years of applying 0 (\blacklozenge), 44 (\blacksquare), 87 (\blacktriangle) or 137 (\bullet) kg P ha⁻¹ as triple super-phosphate fertiliser. (Poulton et al., 2013).

crops were measured until 1973 (Johnston and Wedderburn, 1975). From 1974 – 1982 the site was grassed over but not used for other trials. In the 1970s the UK government decided that the wholesale dumping of sewage in the sea was no longer acceptable and that, after treatment, it could be made available to farmers as an organic manure (UK Statutory Instrument, 1989). However, at that time in the UK, wastewater from industrial factories was not kept separate from domestic sewage and there was concern that the treated sewage would contain heavy metal contaminants and that these might be harmful to health particularly if root crops or leafy vegetables were being grown for human consumption. It was realised that the Market Garden experiment, which had received sewage sludge for many years, could be used to gain a better understanding of the accumulation and persistence of metal pollutants applied in the sewage sludge some 20 - 40 years earlier. Samples of the manures that had been applied, of the soil and of the crops that had been grown had been dried and archived. Cropping resumed which, together with the archived samples, were analysed for their heavy metal content to see how much was taken up by the crops and how much remained in the soil (McGrath, 1987). The metal pollutants had a severe effect on the microbial community. McGrath et al. (1988) found that when white clover (*Trifolium repens*) was grown, toxic metals remaining in the soil had a disastrous effect on N fixation which was due to the extinction of *Rhizobium leguminosarum* biovar *trifolii* in the soil. This would have a very damaging effect in cropping systems which relied on the use of biological fixation to supply N to crops. Chaudri et al. (1993) showed that zinc was responsible for this effect at concentrations in soil treated with sewage sludge which were lower than those permitted at that time. After considerable discussion and argument, this led to a decrease in the amount of Zn that was allowed in soils in the UK (DEFRA/EA, 2018). Very little of the heavy metals were taken up by the crops and only 1% was found immediately below plough depth, *i.e.* below 23 cm. Allowing for any dispersion by ploughing and other cultivations, about 80% of the metal load could be accounted for in the topsoil (McGrath and Lane, 1989). It was concluded that most of the metals will remain in the plough layer for a very long time. This information helped draft EU legislation which set limits on the concentrations of heavy metals that were permitted in sewage sludge that was to be spread on farmland. Other long-term sludge trials were established at Woburn as part of a nationwide network of experiments (Chaudri et al., 1993).

An interesting postscript to this work is that in the early-2000s the integrity of the Market Garden Experiment was threatened by a local road widening scheme which would affect two of the four blocks of the experiment. It was decided that soil from the 40 plots (each 5 m x 8 m) would be physically moved. In 2005 an equivalent area on the other side of the experiment was excavated to a depth of 25 cm and soil from each of the 40 plots, to a depth of 25 cm was then moved individually to their new location. More than 800 t of soil was moved (Steve McGrath,

personal communication).

3.3.2. The value of archived samples and long-term datasets

Although not strictly within the remit of a paper dealing with changes which might be needed to maintain the relevance of LTEs, it is worth highlighting the value of archived material. From the very earliest years of the experiments at Rothamsted, Lawes and Gilbert had the foresight to retain crop and soil samples. They anticipated that the rapid advances that were being made in chemistry at that time (the mid-19th Century) would mean that, in the future, archived samples would be a valuable resource. In the 1860s they said "...if our knowledge of the chemistry of soils should progress as rapidly as it has in the last twenty years, the analysis of a soil will ere long become much more significant than it is at present." (Lawes and Gilbert, 1864). Thus, the Rothamsted Sample Archive (RSA) was initiated with grain, straw and herbage samples from every year since the start of the Broadbalk Wheat, Hoosfield Barley, Park Grass and many other experiments together with soils which were sampled less frequently, and samples of the inputs (*e.g.* fertilisers and organic manures). Overall, the RSA now contains > 300, 000 samples. This means that, as analytical techniques change over the course of a very long-term experiment, it is possible to re-analyse archived samples using current methods. Thus, soils on which organic C has previously been measured by Walkley-Black and Tinsley methods and total N which were analysed by Soda Lime and Kjeldahl methods have now been re-analysed by modern combustion techniques. Similarly, plant-available P in current and earlier soils has been analysed by Olsen's method since the 1950s. Archived samples also allow us to look back over > 180 years at many aspects of plant nutrition, soil fertility and atmospheric pollution that could not have been imagined by Lawes and Gilbert. The analysis of archived samples from the Woburn Market Garden and Exhaustion Land experiments discussed above are just two examples; there are many others. These have included work on organic pollutants deposited on land from the atmosphere (Jones et al., 1994) and the analysis of soils for their organic C and ¹⁴C content to construct and validate a computer model, RothC, to simulate the turnover of organic matter in soil (Jenkinson, 1990).

The RSA remains a key resource of the Rothamsted LTE project along with the electronic Rothamsted Archive (e-RA) which provides access to historic datasets, associated experimental details and meteorological records (<http://www.era.rothamsted.ac.uk>). Information about access to archived samples is available through e-RA.

3.3.3. Climate change

By far the most important changes which might affect the sustainability of cropping systems around the world are those caused by climate change. Increasing concentrations of CO₂, and other greenhouse gases in the atmosphere has led to increased temperatures, droughts, wildfires, more extreme rainfall events and resultant flooding which has affected all continents, often with devastating results. All LTEs are being conducted against the background of increased CO₂ and temperature, some might even have been directly affected by fire and flood. At Rothamsted the average annual temperature for 2013–2022 was > 1.5 °C above the long-term average of 9.0 °C for 1878–1988. However, any impact of climate change on the experiments at Rothamsted and Woburn has been confounded by the increase and subsequent decline in atmospheric inputs of N and S and, on the arable experiments by the regular introduction of new cultivars. However, on the Park Grass Experiment, Köhler et al. (2010), using archived herbage samples to examine C isotope discrimination within the plant showed that water-use efficiency has increased since 1857. More recently, Addy et al. (2022) used sophisticated modelling techniques to show that warmer and drier years in the 20th and 21st Centuries reduced hay yield. The model suggested that hay yields might decline by as much as 50% by 2080.

The archived samples and long-term datasets relating to the Rothamsted LTEs exist as a record of both management and environmental change. This allows us to contribute to the global societal challenges of

sustainable agriculture, food security and climate change.

4. Conclusions

Long-term experiments globally continue to provide unique information on the sustainability of cropping systems and agricultural practices. They also act as research platforms, providing the basis for a wide range of scientific investigations connected with soils and crops. In most respects continuity of treatments and management is key to the value of an LTE but for very long-running experiments this is virtually impossible to achieve. On Broadbalk for example, FYM has been applied every year since autumn 1843 (except when sections were fallowed), 48, 96 and 144 kg N ha⁻¹ has been applied since 1852, 192 kg ha⁻¹ since 1968, 240 and 288 kg ha⁻¹ since 1985, but, since the experiment began, ploughing, drilling and harvesting techniques have changed, chalk has been added, and pesticides used, rotations have been included on part of the experiment and up-to-date cultivars have been introduced regularly. Despite all of these changes the experiment is still recognizably the same and still serving its original purpose *i.e.* to identify those factors affecting crop growth. We have concluded from 180 years of experience at Rothamsted that carefully thought-out changes can be extremely beneficial. We have learned that LTEs should not be regarded as static museum exhibits that should never be altered.

As discussed above, some changes are essential to secure the continued value of the experiments. By contrast, some are responses to emerging scientific or agricultural issues that facilitate new research that builds on the foundations provided by the experiments. However, managers of LTEs must be careful to avoid the temptation to include treatments to address every conceivable issue; some questions are better answered by setting up separate shorter-term experiments. Some new uses of LTEs are essentially extensions of the experiments' original aims concerning crop production, efficiency of use of inputs, or environmental impacts but some are totally unexpected and could not have been envisaged by the LTEs' initiators. LTEs frequently provide a platform to test hypotheses or evaluate new techniques. Such uses benefit from the existence of contrasting treatments but with other factors held constant. For example, soils having differing organic matter content or nutrient status but with the same mineral composition or differing soil pH but with other aspects the same. Whilst continuity of treatments is vital for such uses, in some circumstances measuring how soil properties alter in response to a change in management following a long period of continuity can also be extremely informative.

Whilst LTEs are one vital means of investigating agricultural practices and determining their sustainability, or lack of sustainability due to production or environmental reasons, their limitations must also be recognised. The maintenance of crop production in small, well-managed field plots is no guarantee that the practices used will be equally robust when applied by farmers at large scale in a wide range of environments with differing soil types. Some processes such as soil erosion, water movement or pathogen spread operate at landscape scale and are not easily reproduced in small plots. Edge effects and soil movement between plots can occur within LTEs and must be minimised and/or taken into account. From a sustainability perspective, economic, social, and local infrastructure factors are important and cannot be directly addressed within an LTE. But, despite these limitations, LTEs provide a unique source of information on the functioning of agricultural systems. Our experience with the LTEs at Rothamsted provides evidence that this value is greatly enhanced when a degree of flexibility is introduced and some changes in treatments or practices made. But these must be done carefully and be based on interdisciplinary consultations to avoid unintended consequences and to ensure the essential continuity of the experiment. A complementary approach that was initiated at Rothamsted sites 5–6 years ago is to design agricultural systems experiments that are planned from the outset to have more flexible treatments (Li et al., 2023). Such new experiments are easier to adjust to address new agricultural practices but inevitably lack the benefit of continuity within

a traditional LTE. Nevertheless, the intention from the outset is for these new system-based experiments to be long-term platforms and to be incorporated into Rothamsted's general LTE "portfolio". The two styles of experiment each have their own advantages and limitations, and each is strengthened by the existence of the other.

CRediT authorship contribution statement

Glendining Margaret J.: Data curation, Writing – original draft, Writing – review & editing. **Gregory Andrew S.:** Writing – original draft, Writing – review & editing. **Poulton Paul Richard:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Powlson David S.:** Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

We are indebted to J B Lawes, J H Gilbert, and later scientists for establishing the Rothamsted Long-term Experiments and the generations of farm staff who have maintained them. We would also like to thank the many people who have been involved with the sampling, sample processing, chemical, physical and statistical analyses over many years, particularly Suzanne Clark, Chris Hall, Sarah Perryman and Wendy Wilmer, and Tony Scott for the meteorological information. We also thank the Lawes Agricultural Trust and Rothamsted Research. The Rothamsted Long-term Experiments-National Bioscience Research Infrastructure is funded by the UK Research and Innovation- Biotechnology and Biological Sciences Research Council (BBS/E/RH/23NB0007).

References

- Addy, J.W.G., Ellis, R.H., MacLaren, C., Macdonald, A.J., Semenov, M.A., Mead, A., 2022. A heteroskedastic model of Park Grass spring hay yields in response to weather suggests continuing yield decline with climate change in future decades. *J. R. Soc. Interface* 19. <https://doi.org/10.1098/rsif.2022.0361>.
- Bacq-Labreuil, A., Neal, A.L., Crawford, J., Mooney, S.J., Akkari, E., Zhang, X., Clark, I., Ritz, K., 2021. Significant structural evolution of a long-term fallow soil in response to agricultural management practices requires at least 10 years after conversion. *Eur. J. Soil Sci.* 72, 829–841. <https://doi.org/10.1111/ejss.13037>.
- Barré, P., Eglin, T., Christensen, B.T., Ciais, P., Houot, S., Kätterer, T., van Oort, F., Peylin, P., Poulton, P.R., Romanenkov, V., Chenu, C., 2010. Quantifying and isolating stable soil organic carbon using long-term bare fallow experiments. *Biogeosciences* 7, 3839–3850. <https://doi.org/10.5194/bg-7-3839-2010>.
- Barré, P., Quénéa, K., Vidal, A., Cécillon, L., Christensen, B.T., Kätterer, T., Macdonald, A., Léo, P., Plante, A.F., van Oort, F., Chenu, C., 2018. Microbial and plant-derived compounds both contribute to persistent soil organic carbon in temperate soils. *Biogeochemistry* 140, 81–92. <https://link.springer.com/article/10.1007/s10533-018-0475-5>.
- Bolton, J., 1977. Changes in soil pH and exchangeable calcium in two liming experiments on contrasting soils over 12 years. *J. Agric. Sci.* 89, 81–86. <https://doi.org/10.1017/S0021859600027210>.
- Brenchley, W.E., Adam, H., 1915. Recolonisation of cultivated land allowed to revert to wilderness. *J. Ecol.* 56, 759–775.
- Chaudri, A.M., McGrath, S.P., Giller, K.E., Rietz, E., Sauerbeck, D.R., 1993. Enumeration of indigenous *Rhizobium leguminosarum* biovar *trifolii* in soils previously treated with metal-contaminated sewage sludge. *Soil Biol. Biochem.* 25, 301–309. [https://doi.org/10.1016/0038-0717\(93\)90128-X](https://doi.org/10.1016/0038-0717(93)90128-X).
- Crawley, M.J., Johnston, A.E., Silvertown, J., Dodd, M., de Mazancourt, C., Heard, M.S., Edwards, G.R., 2005. Determinants of species richness in the Park Grass experiment. *Am. Nat.* 165, 179–192. <https://doi.org/10.1086/427270>.

- DEFRA/EA (2018) Sewage sludge in agriculture: code of practice for England, Wales and Northern Ireland. (<https://gov.uk/government/publications/sewage-sludge-in-agriculture-code-of-practice-for-england-wales-and-northern-ireland>).
- Dyer, B., 1902. Results of investigations on the Rothamsted soils. USDA Bulletin 106. Government Printing Office, Washington, USA.
- Gilbert, J.H., 1888. The results of experiments at Rothamsted on the growth of potatoes. *Agric. Gaz. N. Ser.* 4, 1–26.
- Glendining, M.J., Poulton, P.R., Macdonald, A.J., Gregory, A.S. 2021. Park Grass Experiment lime treatments, 1856–2021. Electronic Rothamsted Archive, Rothamsted Research, Harpenden, UK. <https://doi.org/10.23637/rpg5-LimeTreats-01>.
- Glendining, M.J., Gregory, A.S., Poulton, P.R., 2022. Broadbalk Wheat Experiment Chalk Applications. Electron. Rothamsted Arch. <https://doi.org/10.23637/rbk1-chalk-01>.
- Goulding, K.W.T., Poulton, P.R., Webster, C.P., Howe, M.T., 2000. Nitrate leaching from the Broadbalk Wheat Experiment, Rothamsted, UK, as influenced by fertilizer and manure inputs and the weather. *Soil Use Manag.* 16, 244–250.
- Hall, A.D., 1905. On the accumulation of fertility by land allowed to run wild. *J. Agric. Sci.* 1, 241–249.
- Harmer, R., Peterken, G., Kerr, G., Poulton, P., 2001. Vegetation changes during 100 years of development of two secondary woodlands on abandoned arable land. *Biol. Cons.* 101, 291–304. [https://doi.org/10.1016/S0006-3207\(01\)00072-6](https://doi.org/10.1016/S0006-3207(01)00072-6).
- Heckrath, G., Brookes, P.C., Poulton, P.R., Goulding, K.W.T., 1995. Phosphorus leaching from soils containing different phosphorus concentrations in the Broadbalk. *Exp. J. Env. Qual.* 24, 904–910.
- Hicks, H.L., Comont, D., Coutts, S.R., Crook, L., Hull, R., Norris, K., Neve, P., Childs, D.Z., Freckleton, R.P., 2018. The factors driving evolved herbicide resistance at a national scale. *Nat. Ecol. Evol.* 2, 529–536. <https://doi.org/10.1038/s41559-018-0470-1>.
- Hirsch, P.R., Jhureeja, D., Williams, J.K., Murray, P.J., Scott, T., Misselbrook, T.H., Goulding, K.W.T., Clark, I.M., 2017. Soil resilience and recovery: rapid community responses to management changes. *Plant Soil* 412, 283–297. (<https://link.springer.com/article/10.1007/s11104-016-3068-x>).
- Hirsch, Penny R., et al., 2009. Starving the soil of plant inputs for 50 years reduces abundance but not diversity of soil bacterial communities. *Soil Biol. Biochem.* 41 (9), 2021–2024. <https://doi.org/10.1016/j.soilbio.2009.07.011>.
- IUSS Working Group WRB, 2022. World Reference Base for Soil Resources 2014. International soil classification system for naming soils and creating legends for soil maps, 4th edition. International Union of Soil Sciences (IUSS), Vienna, Austria. (https://wrb.isric.org/files/WRB_fourth_edition_2022-12-18.pdf).
- Jenkins, H.M., 1874. Report on the cultivation of potatoes with special reference to the potato-disease. *J. R. Agric. Soc.* 10, 475–477 (2nd Series).
- Jenkinson, D.S. 1971. The accumulation of organic matter in soil left uncultivated. Rothamsted Experimental Station Report for 1970, Part 2, 113–137. <https://doi.org/10.23637/ERADOC-1-34803>.
- Jenkinson, D.S., 1990. The turnover of organic carbon and nitrogen in soil. *Philos. Trans. Roy. Soc. B* 329, 361–368.
- Jenkinson, D.S., Johnston, A.E. 1977. Soil organic matter in the Hoosfield Continuous Barley Experiment. Rothamsted Experimental Station Report for 1976, Part 2, 87–101. <https://doi.org/10.23637/ERADOC-1-34448>.
- Jensen, J.L., Schjonning, P., Watts, C.W., Christensen, B.T., Munkholm, L.J., 2020. Soil Tillage Res. 199, 104597 <https://doi.org/10.1016/j.still.2020.104597>.
- Johnston, A.E. 1973. The effects of ley and arable cropping systems on the amount of organic matter in the Rothamsted and Woburn Ley-arable experiments. Rothamsted Experimental Report for 1972, Part 2, 131–159. Lawes Agricultural Trust, Harpenden, UK. <https://doi.org/10.23637/ERADOC-1-34692>.
- Johnston, A.E., 1975. Experiments made on stackyard field, Woburn, 1876–1974. I. history of the field, details of the cropping and manuring and the yields in the Continuous Wheat and Barley Experiments. Rothamsted Exp. Station Rep. 1975 (Part 2), 29–77. <https://doi.org/10.23637/ERADOC-1-33158>.
- Johnston, A.E., 1997. The value of long-term field experiments in agricultural, ecological and environmental research. *Adv. Agron.* 59, 291–333.
- Johnston, A.E., Garner, H.V. 1969. Broadbalk: Historical Introduction. Rothamsted Experimental Station Report for 1968, Part 2, 12–25. <https://doi.org/10.23637/ERADOC-1-34916>.
- Johnston, A.E., Penny, A., 1972. The Agdell Experiment, 1848–1970. Estimates of the P and K accumulated from fertiliser dressings given between 1848 and 1951, their recovery by grass between 1958 and 1970, and their effect on the response by grass to new dressings of P and K. Rothamsted Exp. Station Rep. 1971 (Part 2), 38–68. <https://doi.org/10.23637/ERADOC-1-37287>.
- Johnston, A.E., Poulton, P.R., 1977. Yields on the Exhaustion Land and changes in the NPK content of the soils due to cropping and manuring, 1852–1975. Rothamsted Exp. Station Rep. 1976 (Part 2), 53–85. <https://doi.org/10.23637/ERADOC-1-34447>.
- Johnston, A.E., Wedderburn, R.W.M., 1975. The Woburn Market Garden experiment 1942–1969 I. A history of the experiment, details of the treatments and yields of the crops. Rothamsted Exp. Station Rep. 1974 (Part 2), 79–101. <https://doi.org/10.23637/ERADOC-1-33161>.
- Johnston, A.E., Poulton, P.R., Coleman, K., 2009. Soil Organic Matter: It's Importance in Sustainable Agriculture and Carbon Dioxide Fluxes. *Adv. Agron.* 101, 1–57.
- Johnston, A.E., Poulton, P.R., White, R.P., 2013. Plant-available soil phosphorus. Part II: the response of arable crops to Olsen P on a sandy clay loam and a silty clay loam. *Soil Use Manag.* 29, 12–21 <https://doi.org/10.1111/j.1475-2743.2012.00449.x>.
- Johnston, A.E., Poulton, P.R., White, R.P., Macdonald, A.J., 2016. Determining the longer-term decline in plant-available soil phosphorus from short-term measured values. *Soil Use Manag.* 32, 151–161. <https://doi.org/10.1111/sum.12253>.
- Johnston, A.E., Poulton, P.R., Coleman, K., Macdonald, A.J., White, R.P., 2017. Changes in soil organic matter over 70 years in continuous arable and ley-arable rotations on a sandy loam soil in England. *Eur. J. Soil Sci.* 68, 305–316. <https://doi.org/10.1111/ejss.12415>.
- Johnston, A.E., Poulton, P.R., Macdonald, A.J.; Glendining, M.J. 2022. The Woburn Ley-arable experiment, 1938–2020: its management, the issues and problems that arose, and their resolution to maintain the continuity and relevance of the experiment. Rothamsted Research, Harpenden, UK. 28pp. <https://doi.org/10.23637/wrn3-manag1938-2020-01>.
- Jones, K.C., Johnston, A.E., McGrath, S.P., 1994. Historical monitoring of organic contaminants in soils. In: Leigh, R.A., Johnston, A.E. (Eds.), Long-term Experiments in Agricultural and Ecological Sciences. CAB International, Wallingford, UK, pp. 147–164.
- Köhler, I.H., Poulton, P.R., Auerswald, K.A., Schnyder, H., 2010. Intrinsic water-use efficiency of temperate seminatural grassland has increased since 1857: an analysis of carbon isotope discrimination of herbage from the Park Grass Experiment. *Glob. Change Biol.* 16, 1531–1541. <https://doi.org/10.1111/j.1365-2486.2009.02067.x>.
- Lawes, J.B. 1847. On Agricultural Chemistry. Turnip Culture. *J. Roy. Ag. Soc. England*, VIII, Part II, 226–260.
- Lawes, J.B. 1895. Upon some properties of soils which have grown a cereal crop and a leguminous crop for many years in succession. *Agricultural Students Gazette*, New Series, Vol VII, April 1895, 65–72.
- Lawes, J.B., Gilbert, J.H., 1864. Report of experiments on the growth of wheat for twenty years in succession on the same land. *J. Roy. Ag. Soc. Engl.* XXV, Parts I & II, 93–185 and 449–501.
- Lawes, J.B., Gilbert, J.H., 1873. Report of experiments on the growth of barley for twenty years in succession on the same land. *J. Roy. Ag. Soc. Engl.* IX – S.S. Parts I & II, 5–79 and 79–178.
- Lawes, J.B., Gilbert, J.H. 1895. The Rothamsted Experiments: being an account of some of the results of the agricultural investigations conducted at Rothamsted in the field, the feeding shed, and the laboratory, over a period of fifty years. *Transactions of the Highland and agricultural Society of Scotland*, Fifth Series, VII, pp 354.
- Lawes, J.B., Gilbert, J.H., Masters, M.T., 1882. Agricultural, botanical and chemical results of experiments on the mixed herbage of permanent meadow, conducted for more than twenty years in succession on the same land. Part II: The botanical results. *Philos. Trans. R. Soc. (A B)* 173, 1181–1413.
- Li, X., Storkey, J., Mead, A., Clark, I., Ostler, R., Roberts, B., Dobermann, A. 2023. *Agronomy for Sustainable Development* 43, 60. 023) 43:60 <https://doi.org/10.1007/s13593-023-00914-8>.
- Macdonald, A., Poulton, P., Clark, I., Scott, T., Glendining, M., Perryman, S., Storkey, J., Bell, J., Shield, I., McMillan, V., Hawkins, J. 2018. Guide to the Classical and Other Long-term experiments, Datasets and Sample Archive. Rothamsted Research, 57 pp. <https://doi.org/10.23637/ROTHAMSTED-LONG-TERM-EXPERIMENTS-GUIDE-2018>.
- Macdonald, A.J., Poulton, P.R., Gregory, A.S., Glendining, M.J., Ostler, R.J. 2022. Highfield Bare Fallow soil chemical properties, 1959–2014. Electronic Rothamsted Archive, Rothamsted Research, Harpenden, UK. <https://doi.org/10.23637/rrs1-SOILCN1959-2014-01>.
- McGrath, S.P., 1987. Long term studies of metal transfers following applications of sewage sludge. *Special Publication No.6 of the British Ecological Society*. In: Coughtrey, P.J., Martin, M.H., Unsworth, M.H. (Eds.), *Pollutant Transport and Fate in Ecosystems*. Blackwell Scientific, Oxford, pp. 301–317. *Special Publication No.6 of the British Ecological Society*.
- McGrath, S.P., Lane, P.W., 1989. An explanation for the apparent losses of metals in a long-term field experiment with sewage-sludge. *Environ. Pollut.* 60, 235–256. [https://doi.org/10.1016/0269-7491\(89\)90107-3](https://doi.org/10.1016/0269-7491(89)90107-3).
- McGrath, S.P., Brookes, P.C., Giller, K.E., 1988. Effects of Potentially Toxic Metals in Soil Derived from Past Applications of Sewage-Sludge on Nitrogen-Fixation by *Trifolium repens* L. *Soil Biol. Biochem.* 20, 415–424. [https://doi.org/10.1016/0038-0717\(88\)90052-1](https://doi.org/10.1016/0038-0717(88)90052-1).
- Neal, A.L., Bacq-Labreuil, A., Zhang, X., et al., 2020. Soil as an extended composite phenotype of the microbial metagenome. *Sci. Rep.* 10, 10649 <https://doi.org/10.1038/s41598-020-67631-0>.
- Neal, A.L., Hughes, D.J., Clark, I.M., Jansson, J.K., Hirsch, P.R., 2021. Microbiome Aggregated Traits and Assembly are More Sensitive to Soil Management than Diversity. *mSystems* 6 (3), e01056-20. <https://doi.org/10.1128/mSystems.01056-20>.
- Olsen, S.R., Cole, C.V., Watanabe, F.S., Dean, L.A. 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. *USDA Circular* 939.
- Perryman, S., Glendining, M., Gregory, A., Poulton, P., 2023. Dataset: Broadbalk mean long-term yields of winter wheat 1852–2018. Electronic Rothamsted Archive, Rothamsted Research, Harpenden, UK. <https://doi.org/10.23637/rbk1/meanWYields1852-2018-02>.
- Poulton, P., Johnston, J., Macdonald, A., White, R., Powlson, D., 2018. Major limitations to achieving “4 per 1000” increases in soil organic carbon stock in temperate regions: Evidence from long-term experiments at Rothamsted Research, UK. *Glob. Change Biol.* 2563–2584 <https://doi.org/10.1111/gcb.14066>.
- Poulton, P.R., Pye, E., Hargreaves, P.R., Jenkinson, D.S., 2003. Accumulation of carbon and nitrogen by old arable land reverting to woodland. *Glob. Change Biol.* 9, 942–955. <https://doi.org/10.1046/j.1365-2486.2003.00633.x>.
- Poulton, P.R., Johnston, A.E., White, R.P., 2013. Plant-available phosphorus. Part I: the response of winter wheat and spring barley to Olsen P on silty clay loam. *Soil Use Manag.* 29, 4–11. <https://doi.org/10.1111/j.1475-2743.2012.00450.x>.
- Poulton, P.R., Johnston, A.E., Macdonald, A.J., Glendining, M.J. 2022. Woburn Ley-arable experiment fertilizer and lime applications 1938–2020. Electronic Rothamsted Archive, Rothamsted Research, Harpenden UK. <https://doi.org/10.23637/wrn3-fert1938-2020-01>.

- Poulton, P.R., Johnston, A.E., White, R.P., 2023. Response of three cereal crops in continuous arable or ley-arable rotations to fertiliser nitrogen and soil nitrogen at Rothamsted's Woburn Ley-arable experiment. *Soil Use Manag.* 39, 771–784. <https://doi.org/10.1111/sum.12872>.
- Poulton, P.R., Johnston, A.E., Glendinning, M.J., White, R.P., Gregory, A.S., Clark, S.J., Wilmer, W.S., Macdonald, A.J., Powelson, D.S. 2024. The Broadbalk Wheat Experiment, Rothamsted, UK: crop yields and soil changes during the last 50 years. *Advances in Agronomy*, Volume 184 (in press).
- Powelson, D.S., Poulton, P.R., Addiscott, T.M., McCann, D.S., 1989. Leaching of nitrate from soils receiving organic or inorganic fertilizers continuously for 135 years. In: Hansen, J.A., Henriksen, K. (Eds.), *Nitrogen in Organic Wastes Applied to Soils*. Academic Press, London, pp. 334–345.
- RB 209 (5th edition), 1988. Fertiliser recommendations for agricultural and horticultural crops. HMSO, London.
- Redmile-Gordon, M., Gregory, A.S., White, R.P., Watts, C.W., 2020. Soil organic carbon, extracellular polymeric substances (EPS), and soil structural stability as affected by previous and current land-use. *Geoderma* 363, 114143. <https://doi.org/10.1016/j.geoderma.2019.114143>.
- Ristaino, J.B., Hu, C.H., Fitt, B.D.L., 2012. Evidence for the presence of the founder Ia mtDNA haplotype of *Phytophthora infestans* in 19th century potato tubers from the Rothamsted archives. *Plant Pathol.* 62, 492–500. <https://doi.org/10.1111/j.1365-3059.2012.02680.x>.
- Rousk, K., Brookes, P.C., Bååth, E., 2009. Contrasting soil pH effects on fungal and bacterial growth suggest functional redundancy in carbon mineralization. *Mar 2009 Appl. Environ. Microbiol.* 1589–1596. <https://doi.org/10.1128/aem.02775-08>.
- Russell, E.J., 1916. Chalking: a useful improvement for clays overlying the chalk. *J. Bd. Agric. Fish.* 23, 625–633.
- Russell, E.J. Voelcker, J.A. 1936. Fifty years of field experiments at the Woburn Experimental Station. The Rothamsted Monographs on Agricultural Science; Ed. E.J. Russell. Longmans, Green and Co. London. pp. 392.
- Silvertown, J., Poulton, P., Johnston, E., Edwards, G., Heard, M., Biss, P.M., 2006. The Park Grass Experiment 1856 – 2006: its contribution to ecology. *J. Ecol.* 94, 801–814. <https://doi.org/10.1111/j.1365-2745.2006.01145.x>.
- Storkey, J., Macdonald, A.J., Poulton, P.R., Scott, T., Köhler, I.H., Schnyder, H., Goulding, K.W.T., Crawley, M.J., 2015. Grassland biodiversity bounces back from long-term nitrogen addition. *Nature* 528, 401–404. <https://doi.org/10.1038/nature16444>.
- Thurston, J.M., Williams, E.D., Johnston, A.E., 1976. Modern developments in an experiment on permanent grassland started in 1856: effects of fertilisers and lime on botanical composition and crop and soil analyses. *Ann. Agron.* 27, 1043–1082.
- UK Statutory Instrument. 1989. The Sludge (Use in Agriculture) Regulations 1989, no. 1263. (<http://www.legislation.gov.uk/ukSI/1989/1263/made>) (accessed 13 Feb. 2016).
- Warren, R.G., 1958. The residual effects of manurial and cropping treatments in the Agdell rotation experiment. Rothamsted Experimental Station. Report 1957, 252–260. <https://doi.org/10.23637/ERADOC-1-35007>.
- Warren, R.G., Johnston, A.E., 1960. The Exhaustion Land site. An account of experiments made before and during the period of soil exhaustion. Rothamsted Experimental Station. Report 1959, 230–239. <https://doi.org/10.23637/ERADOC-1-39375>.
- Young, A. 1813. General view of the agriculture in Hertfordshire. London: pp. 236.
- Zhao, F.J., Knights, J.S., Hu, Z.Y., McGrath, S.P., 2003. Stable sulfur isotope ratio indicates long-term changes in sulfur deposition in the Broadbalk experiment since 1845. *J. Env. Qual.* 32, 33–39.