

STUDIES IN CROP VARIATION. I. AN EXAMINATION
OF THE YIELD OF DRESSED GRAIN FROM BROADBALK

Author's Note (CMS 3.106a)

In the author's early work at Rothamsted much attention was given to the massive records of weather, crop yields, crop analyses, etc., which had been accumulated during the long history of that research station. The material was obviously of unique value for such problems as that of ascertaining to what extent meteorological readings were capable of supplying a prediction of the crop yields to follow.

The present paper was the first of a series devoted to this end, and consists of an examination of the yields of selected plots in Broadbalk wheat field with a view to ascertaining the principal components of their variability as a preliminary to the study of meteorological effects. The use of orthogonal polynomials is extensively developed in this and the following papers as affording the means of isolating groups of components susceptible to separate explanations. An important feature of these wheat yields consists in slow changes too consistent to be ascribable to the immediate effects of annual fluctuations in the weather, and differing in kind from the progressive soil deterioration recognisable on plots with insufficient or unbalanced manuring. Evidence is produced converging to the conclusion that varying weed infestation representing a cumulative integration both of seasonal weather and of cultivation practice was a major factor in these slow changes.

STUDIES IN CROP VARIATION.

I. AN EXAMINATION OF THE YIELD OF DRESSED GRAIN FROM BROADBALK.

By R. A. FISHER, M.A.

(*Statistical Laboratory, Rothamsted Experimental Station, Harpenden.*)

(With Three Figures in Text.)

I. THE VARIATION IN WHEAT YIELD.

1. *Introductory.*

THE crop records available at Rothamsted extend back for over 70 years. In Broadbalk wheat field 13 plots have been continuously under uniform treatment since 1852. The length of this unique series of observations is of special value for statistical purposes. The average yields obtained from the Board of Agriculture's Reports cover only half the period; even the Woburn records date only from 1877. The series of yields of dressed grain from the different plots of Broadbalk was examined from 1852, when uniform treatment commenced, until 1918, the last season available when the examination was begun.

Several minor alterations have been made from time to time in the manurial treatment, even in the selected plots, but in the main it has remained uniform.

* (a) For the first 7 years plots 5, 6, 7, 8, 13, and 17 or 18, received 300 lbs. per acre of sulphate of potash, of which for the subsequent 60 years they received 200 lbs. per acre. At the same time the sulphate of soda was cut down from 200 to 100 lbs. per acre on plots 5, 6, 7, 8, 16, and 17 or 18, and from 550 to 366½ lbs. per acre on plot 12; while the sulphate of magnesia on plot 14 was reduced from 420 to 280 lbs. per acre.

+ (b) For two years, 1862 and 1863, all plots, except half of plot 3 received 400 lbs. per acre of a mixture of silicates of soda and lime, and for 3 subsequent years, one half of plots 5, 6, 7, 8, 17 or 18 received 228 lbs. per acre of this mixture. A comparison of the corresponding

* After 13, insert 16.

† After 8, insert 9 and 16.

sub-plots shows that this dressing had little effect, being, if anything, deleterious to the sub-plots receiving it.

(c) For 12 years, 1868 to 1879, the chopped straw from each sub-plot was ploughed in, on half of plots 5, 6, 7, 8, 11, 12, 13, 14, and 17 or 18. A comparison of the corresponding sub-plots shows that in this case also the effect upon the mean yield may be ignored in the analysis of the majority of the plots. Plots 11, 12, and 14 in which the deficiency of potash is a limiting factor showed a sensible benefit from the straw, and for these series (without straw) has been used from 1868.

(d) For 5 years, 1898 to 1902, a dressing of 400 lbs. per acre of basic slag was substituted for the previous dressing of 392 lbs. of super-phosphate, on all plots receiving phosphate. No statistical evidence can be adduced as to the effect of this change, and it is assumed that the dressings were effectively equivalent.

(e) For 1916 the supply of potassium sulphate was reduced to two-thirds of its previous dressing, the ordinary amount of potash being made up by the use of wood ashes. In 1917 and 1918 potassium sulphate was omitted altogether from all the plots, and owing to the insufficiency of wood-ash the deficiency was not made up. In order to continue the comparative treatment of plots 12, 13 and 14, the sodium sulphate on 12 and the magnesium sulphate on 14 were omitted for the same year. From 1920 the original dressings have been resumed.

(f) From time to time small areas have been incorporated in, or excluded from, the experimental plots. These changes can have had but little influence upon the yield per acre, which has always been reckoned on the actual area of each plot; the changes in size have not been large, and any new portion has received similar treatment to that of the main plot, for some years before incorporation.

Changes of Variety.

The variety of wheat seed employed has not been kept the same. The following varieties are recorded:

1852. Old Red Cluster.

1853-81. Red Rostock, 29 years.

1882-99. Red Club, 18 years.

1900-1916. Squarehead's Master, 13 years.

1917 onwards. Red Standard.

Between 1900 and 1912 the varieties were changed more frequently. Squarehead's Master has been principally used, but intermitted with Giant Red (1905), Browick Red (1910), Little Joss (1911, 1912). Red

Standard was adopted in 1917 and will be used in future. When the varieties are changed infrequently, any effect due to genetic difference of constitution would be included in the slow changes. During the latter period it would appear partly as annual variation. That these genetic differences are not at any rate a principal cause of the slow changes observed, may be seen from the great changes in mean yields which occurred during the use of Red Rostock.

2. *The causes of variation in wheat yields.*

From the series of observations it is possible to distinguish three types of variation in the wheat yield: (1) annual variation, (2) steady¹ diminution due to deterioration of the soil, (3) slow changes other than steady diminution. The annual variations may be ascribed primarily to the weather; including in that term not only the direct effects of meteorological conditions in stimulating plant growth, but also the physical effects wrought upon the soil, such as the washing out of plant nutrients and the indirect effects of light, temperature and moisture in stimulating or retarding the increase of bacteria, protozoa and of the fungal and algal flora of the soil, all of which may be supposed to adjust their activities rapidly to the meteorological conditions. The steady diminution of yield may unhesitatingly be ascribed to deterioration of the soil; either, as in plot 10, to the exhaustion of natural supplies of potash and phosphorus, or, in other cases, perhaps, to that of unknown substances required in small quantities, and not supplied in the artificial manure, or to physical changes as yet but little understood, or, in plot 5, to the gradual exhaustion of the power of the soil of producing nitrates in the soil moisture. The third class of change is unexpected, and it is not easy to assign it entirely to any one cause. To establish the existence of large changes in the mean yield, to show how they may be disentangled from the other types of change, and to suggest their possible cause is the purpose of the present paper.

3. *Mean Yield and Annual Diminution.*

In Table I is shown the mean yield and average annual diminution of mean yield, in bushels per acre, for the 13 plots here considered.

¹ The deterioration must not be assumed to be mathematically linear, although it is here represented by a linear function; on most plots it was probably more rapid in the earlier years than it was later, as is indicated by the parabolic term being on the whole more inclined to positive values on those plots in which the deterioration is more rapid. The true curves of deterioration cannot, however, be disentangled from the slow changes which have taken place owing to other causes.

Table I.

Plot	Manure per acre						Mean (Bushels per acre)	Mean annual diminution (Bushels per acre)	P. for deterioration
	Sulphate of potash lbs.	Sulphate of soda lbs.	Sulphate of magnesia lbs.	Super- phosphate lbs.	Sulphate of ammonia lbs.	Chloride of ammonia lbs.			
2b, dung 14 tons	—	—	—	—	—	—	34·549	·031	·41
3 and 4, no manure	—	—	—	—	—	—	12·269	·097	·000,001,4
5	200	100	100	392	—	—	14·180	·090	·000,000,8
6	200	100	100	392	100	100	22·581	·141	·000,11
7	200	100	100	392	200	200	31·367	·144	·002,1
8	200	100	100	392	300	300	35·694	·092	·056
10	—	—	—	—	200	200	19·504	·157	·000,25
11	—	—	—	392	200	200	22·046	·219	·000,003
12	—	366½	—	392	200	200	28·319	·181	·000,35
13	200	—	—	392	200	200	30·209	·123	·009,1
14	—	—	280	392	200	200	27·765	·231	·000,000,6
17) alter-	200	100	100	392	—	—	14·510	·092	·002,8
18) nate	—	—	—	—	200	200	29·006	·114	·005,6

In the last column, P represents the probability of a larger annual diminution occurring by chance owing to the later seasons happening to be on the average less favourable than the earlier ones. In calculating P it has been assumed that there has been no real deterioration of the average weather, an assumption which is tested in Section 11.

The values show that all but the dunged plot, 2b, have suffered sensible deterioration. Even plot 8, which receives "complete" artificial manures, shows a deterioration which would not be expected more than once in eighteen random trials. It is, therefore, probably real, as are certainly all the others.

4. *Slow changes in mean yield.*

It becomes apparent on inspection of the actual yields that the changes in mean yield are by no means fully expressed as simple deterioration. The mean yield rises up to about 1860, and after a bad period in the seventies reaches a second maximum in the nineties. The probability of such large fluctuations occurring by chance may be calculated by the methods of Section 8, on the assumption, as above, that the incidence of good and bad seasons has not been orderly but fortuitous.

When the variation of any quantity (variate) is produced by the action of two or more independent causes, it is known that the variance produced by all the causes simultaneously in operation is the sum of the values of the variance produced by each cause separately. The variance is defined as the mean square deviation of variate from its mean, and is therefore the square of its standard deviation. The above property of

the variance, by which each independent cause makes its own contribution to the total, enables us to *analyse* the total, and to assign, with more or less of accuracy, the several portions to their appropriate causes, or groups of causes. In Table II is shown the analysis of the total variance for each plot, divided according as it may be ascribed (i) to annual causes, (ii) to slow changes other than deterioration, (iii) to deterioration; the sixth column shows the probability of larger values for the variance due to slow changes occurring fortuitously.

Table II.

Plot	Annual causes	Slow changes	Deterioration	Total	P. for slow changes
2b	33.2	17.6	.4	51.1	.000,002
3 and 4	9.3	2.5	3.5	15.4	.004,3
5	11.7	2.7	3.0	17.5	.007,5
6	30.6	8.0	7.5	46.1	.003,1
7	50.3	13.3	7.8	71.4	.0 0 2 9
8	53.2	15.9	3.2	72.3	.0 0 1 2
10	41.8	3.7	9.2	54.7	.2 6
11	50.2	3.4	1 8.8	72.4	.3 8 5
12	58.4	7.8	12.2	78.4	.086
13	50.7	8.3	5.7	64.6	.040
14	49.2	6.3	20.0	75.4	.10
17) minerals	21.4	4.6	3.1	29.1	.011
18) ammonia	38.2	9.0	4.9	52.1	.0 0 6 6

The majority of the plots could be cited independently as evidence that the slow changes in the mean yield are not fortuitous. The less significant values from plots 10 to 14 are due not so much to the relative absence of slow changes, as to the great sensitiveness of these plots to annual causes of variation.

The great variability of P is itself evidence that the slow change effect is of a different origin from that of the annual causes, for P is calculated from the ratio between these two portions. If, for instance, progressive alterations had been taking place in the average weather to an extent unmistakably distinct from independent annual variations occurring in random order, then we should expect all plots to show very small values of P, whereas, in fact, P is greatest on those plots most subject to meteorological disturbance; on the other hand, if, as appears to be the case, the succession of seasons is not distinguishable from a random order, then the very small values of P observed in some plots must have some explanation other than meteorological.

The actual course and extent of the slow changes in mean yield, other than soil deterioration, is best shown by a smooth curve, a polynomial of the 5th degree, fitted to the series of yields of dressed grain

* Several figures in Table II have been corrected.

from plot 2*b* (Fig. 1*a*). This plot is not only sensibly free from deterioration but, as shown by the value of P, is least affected by accidental fluctuation due to good and bad seasons. In order to obtain data for the study of meteorological effects, it was necessary to eliminate as far as possible

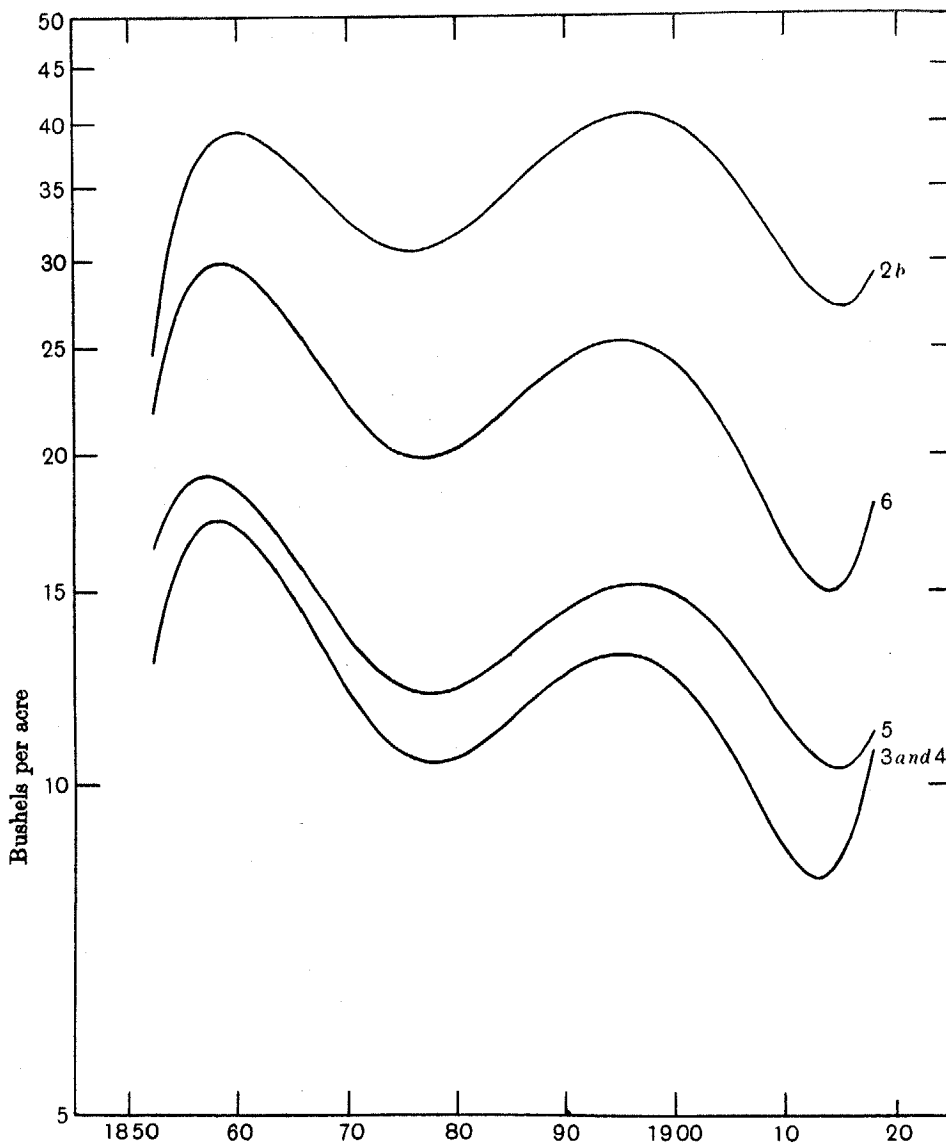


Fig. 1 *a*.

Fig. 1. Showing the course of the changes in mean yield in the continuously manured plots of Broadbalk. The vertical ordinate is plotted on a logarithmic scale in order to shew the proportionality of the slow changes in the several plots, and in order that the relative importance of soil deterioration may be compared.

variations of other than meteorological origin; for this purpose similar curves were fitted to all the plots, and the annual deviation of each plot, from the corresponding curves, will form the basis of a study of the

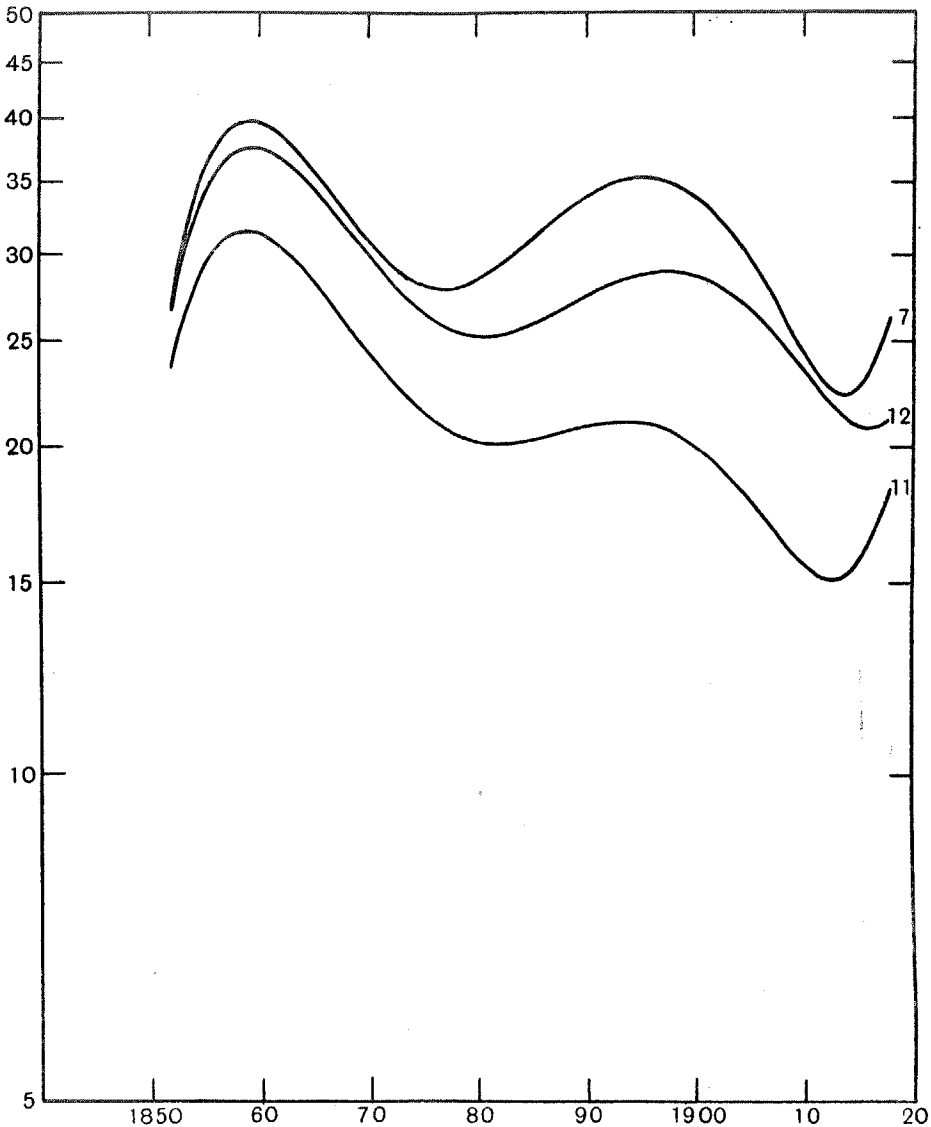


Fig. 1b.

effects of weather upon wheat yield. Fig. 1 shows that during two periods of high yields, centred about 1860 and 1897, the mean yield of plot 2b was about 40 bushels per acre, while in the depression about 1876 it had fallen nearly to 30 bushels per acre, and in 1915 even lower. For special

reasons (Section 9) the extreme values are not so reliable as the others; yet there can be little doubt that the recent minimum is already past, and that a genuine improvement took place in the early years of the experiment.

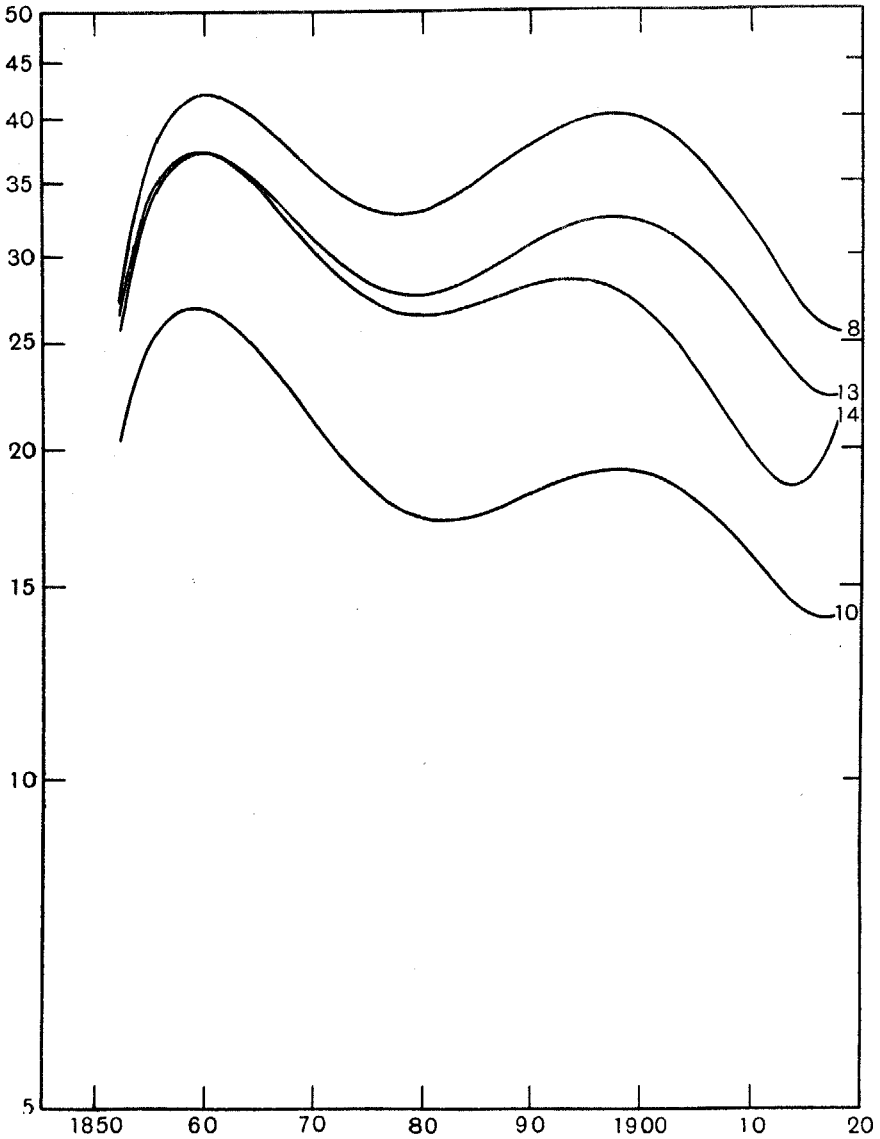


Fig. 1c.

5. *Yield and exhaustion in relation to manurial treatment.*

Before passing to the theory of the significance of these polynomials attention may profitably be drawn to the effect of different manurial treatment upon the mean yields and average annual decrements in the different plots. The probable errors refer in each case to the deviations to be expected from variations of weather occurring in random order.

Tables III, IV, V and VI give the mean yield with the absolute and relative rate of decrement arranged to show best the effect of differences of manurial treatment.

Table III.

Plot	Mean yield (Bushels per acre)	Mean annual decrement (Bushels per acre)	Mean annual decrement %
5, no ammonia	14.18 ± .44	.090	.63 ± .16
6, single ammonia	22.58 ± .71	.141	.62 ± .19
7, double "	31.37 ± .90	.144	.46 ± .15
8, treble "	35.69 ± .93	.092	.26 ± .14

Plots 5, 6, 7 and 8 all receive the same dressing of superphosphate and of sulphates of potash, soda and magnesia, they differ by successive units of ammonium sulphate. The first increment is applied, half in autumn and half in spring, the subsequent increments in spring only. The increases in mean yield due to these increments of ammonium sulphate are 8.40, 8.77 and 4.34 bushels per acre, a series suggestive of "diminishing returns" although not conforming to the geometric series which has been proposed for such cases (Mitscherlich, 1909). The figures for the mean annual decrement, in the second column, show that although the mean of plot 6 exceeds that of plot 5 by 8.4, the latter plot has gained on 6 during the 67 years under discussion, at the rate of .051 per annum. At this rate the difference would disappear in 165 years. Naturally no such result is anticipated from such a prolonged experiment; but this consideration serves to show that not only the mean values, but the rates of decrement observed, are comparable only when taken over the same period of years.

It is to be expected that if the experiment were continued indefinitely, each plot would approach continually to a constant mean yield, but that these yields would differ materially for different manurial treatment. An approximation to the relations which finally exist between the mean yields of different plots may be obtained from the mean of the last nine years of the polynomial fitted to each plot. The values obtained are

10-77, 15-97, 23-57 and 30-16, for these four plots, and the differences produced by the three successive increments are 5-20, 7-60 and 6-59. In these figures there is no longer any evidence of "diminishing return." The comparative yields of these four plots have become the standard example of diminishing returns in agriculture, and in the earlier years of the experiment fully bore out the anticipations of the economists. For the first 9 years of the polynomial, for example, the successive increments were 9-27, 8-15 and 1-43, and we have seen that a similar effect, though less marked, still appears in the means of 67 years. It is only in recent years that the progressive deterioration of the less highly manured plots has gone so far as to make the third increment of the series exceed the first, and so to make apparent the fact that the benefit of the higher dressings was not wholly reaped in the immediate yield, but to some extent is long effective in maintaining the fertility of the soil at a higher level.

The average annual decrements when set out as percentages of the mean yield show a progressive advantage of the nitrogenous manuring. Since all the plots have received soda, potash, magnesia, sulphates and phosphates, equally while greater weights of these ingredients have been removed from the more highly manured plots, the advantage of the latter may be safely ascribed to the supply of nitrogenous plant nutrients. The increasing effectiveness of the nitrogenous manure might be ascribed to the gradual utilisation of slowly available nitrogen compounds in the soil; this view, however, is not borne out by the comparative nitrogen analysis of these plots, which are available since 1865, and it is more probable that the heavier vegetation supported by plot 8 has had some more indirect beneficial effect, equivalent to a more ample supply of nitrogen. The greater root growth on the heavy yielding plots has perhaps increased the effective depth of soil activity, or has supplied more abundantly substances required for bacterial life. If the natural supplies of nitrogen in plots 5 and 6 have become impoverished in these ways we should expect to find the additional dressings of sulphate of ammonia to become, as they have become, an increasingly important factor in maintaining yield.

The four plots (3 and 4), 5, 10 and 7 illustrate the supplementary action of nitrogenous compounds on one side and mineral manures on the other. The mean yield of 7 not only exceeds that of 10 by a greater amount than does 5 exceed (3 and 4), but it exceeds it in a higher ratio. In its absolute rate of decrement 10 exceeds 7 and (3 and 4) exceeds 5, although in both cases giving a lower mean yield. This can only mean

that a part of the decline of (3 and 4) and 10 is to be ascribed to the progressive exhaustion of some of the mineral ingredients supplied to 5 and 7.

Table IV.

Plot	Mean yield	Mean annual decrement	Annual decrement %
3 and 4, no manure	12.27 ± .39	.097	.79 ± .16
5, minerals ...	14.18 ± .44	.090	.63 ± .16
10, ammonia ...	19.50 ± .83	.157	.81 ± .22
7, ammonia and minerals ...	31.35 ± .90	.144	.46 ± .15
17 } minerals ...	14.51 ± .59	.092	.63 ± .21
18 } ammonia ...	29.01 ± .79	.114	.39 ± .14

The comparison with the two series from plots 17 and 18 shows that the mineral series is very little better than 5, and much worse than 7, showing that there is little residual effect of the ammonium salts applied the previous year. That this small advantage is accompanied by greatly increased variability appears from Table II, in which the variance due to annual causes in plot (17 and 18) minerals is nearly double that of plot 5; from the considerations of Section 14 this may be due to the greater effects in variations of weed prevalence. The relative annual decrement also agrees with 5, in being much greater than 7. On the other hand, the ammonia series of (17 and 18) has a mean yield much more close to 7 than to 10, showing that there is relatively little ultimate loss of the mineral ingredients applied in previous years. The relative decrement of (17 and 18) ammonia is even less than that of 7, owing presumably to the smaller crops, averaging only about 21.8 bushels per acre, which have been taken off these plots. The difference in annual decrement, .07 per cent., is very small compared with the difference in mean yield, of 9.6 bushels per acre; and this suggests that, in this case also, if residual minerals could be tested year after year against minerals freshly applied, the higher yielding series would display the less deterioration.

Table V.

Plot	Mean yield in bushels per acre	Mean annual decrement	Annual decrement %
11, no K, Na or Mg ...	22.05 ± .91	.219	.99 ± .21
12, sulphate of soda ...	28.32 ± .98	.181	.64 ± .18
13, sulphate of potash ...	30.21 ± .91	.123	.41 ± .16
14, sulphate of magnesia ...	27.76 ± .90	.231	.83 ± .17
7, All three sulphates ...	31.37 ± .90	.144	.46 ± .15

In comparing plots 7 and 13 it is surprising that the latter, which has slightly the lower mean yield, should have the lower relative rate

of decrement. This fact by itself suggests that the addition of sulphates of soda and magnesia to the sulphate of potash, stimulates production, perhaps by making more potash available, but causes more rapid exhaustion. On the other hand, if we contrast either 12 or 14 with 11, the actual production is much increased by the addition of sulphate of soda or magnesia and the relative annual decrement is reduced.

In these figures it is not easy to distinguish any physical effect of the large quantities of saline matter annually added to the plots. In comparing plot 7 with 13 the former receives a large quantity of sulphates of soda and magnesia, which, in view of the dressings of sulphate of potash which both receive, might be expected to add little to the available plant nutrients while adding much to the salinity. That the average yield of plot 7 exceeds that of plot 13, suggests that even in this case the nutrient advantage outweighs the effect of additional salinity, and it is not clear that the greater deterioration of plot 7 is to be ascribed to the latter cause, since this is also characteristic of plots 12 and 14 in which it must be largely ascribed to the progressive exhaustion of soil potash, which the sulphates of soda and magnesia have naturally facilitated.

Table VI.

Plot	Mean yield in bushels per acre	Mean annual decrement	Mean annual decrement %
2b	34.55 ± .74	.031	.09 ± .11
8	35.69 ± .93	.092	.26 ± .14

The dunged plot 2b can only be well compared with 8. Alone of the plots it shows no significant diminution of yield; plot 8 comes third in this respect, the probability of an increment or decrement as large or larger being the result of a random distribution of favourable and unfavourable years being .056 (or 17 to 1) for 8, against .41 (nearly even odds) for 2b. On the average of 67 years 8 gives the higher mean, but the difference, 1.14, is less than 2b gains on 8 during 20 years, so that during the present century 2b has had the highest mean yield. In addition to having the best sustained, and now the highest yield, 2b is the least variable; the standard error of the annual decrement per cent. in the other plots ranges from .14 to .22, in 2b it is .11. The crop on the dunged plot is therefore better placed to do moderately in a bad year of drought or excessive rain, than to take great advantage of a favourable season.

II. THEORY OF POLYNOMIAL FITTING.

6. *Uncorrelated terms T_r .*

If a quantity x have values $x_1, x_2, \dots x_n$ at a number n of successive times, the general course of its changes may be represented by a polynomial,

$$a + bt + ct^2 + \dots + kt^r,$$

in which t represents the time. In this expression the coefficients $a, b, c, \dots k$, will be altered if more or fewer terms of the series be used; for instance, the value of a will be different according as one takes 2 or 3 terms, for unless c is zero the mean value of ct^2 will not be zero. If the coefficients obtained are to be independent of the number of terms employed, the successive terms must consist of polynomials of degree 0, 1, 2, ... r , which are mutually uncorrelated. Such uncorrelated polynomials may be obtained uniquely in succession, for the term of degree r must fulfil r conditions in order to be uncorrelated with the preceding terms; these conditions specify the term completely with the exception of a numerical factor which is absorbed in the coefficient.

When the values, as do annual numbers, stand at equal intervals of time, t may be conveniently measured from the mid-point of the series, in units equal to the time interval; the series is then

$$A + Bt + C(t^2 - n_2) + D\left(t^3 - \frac{n_4 t}{n_2}\right) + E\left\{t^4 - \frac{n_6 - n_2 n_4}{n_4 - n_2^2} t^2 + \frac{n_2 n_6 - n_4^2}{n_4 - n_2^2}\right\} \\ + F\left\{t^5 - \frac{n_2 n_8 - n_4 n_6}{n_2 n_6 - n_4^2} t^3 + \frac{n_4 n_8 - n_6^2}{n_2 n_6 - n_4^2}\right\}$$

as far as t^5 . n_2, n_4, n_6 , and n_8 represent the mean value of t^2, t^4, t^6 , and t^8 , and for a series of n terms,

$$n_2 = \frac{1}{12}(n^2 - 1), \\ n_4/n_2 = \frac{1}{20}(3n^2 - 7), \\ (n_6 - n_2 n_4)/(n_4 - n_2^2) = \frac{1}{14}(3n^2 - 13), \\ (n_2 n_6 - n_4^2)/(n_4 - n_2^2) = \frac{3}{560}(n^2 - 1)(n^2 - 9), \\ (n_2 n_8 - n_4 n_6)/(n_2 n_6 - n_4^2) = \frac{5}{18}(n^2 - 7), \\ (n_4 n_8 - n_6^2)/(n_2 n_6 - n_4^2) = \frac{1}{1008}(15n^4 - 230n^2 + 407).$$

Variance accounted for by each term.

The coefficients are chosen so as to make the residual variance a minimum; consequently the residual variance is reduced at each stage by an amount equal to the variance of the term itself. If T_r stand for

* The last term within the brackets should be multiplied by t .

the term of degree r , the mean value of $T_p T_q$ is necessarily zero, when p and q are different, while the mean value of T_r^2 is

$$\frac{(|r|)^4}{2r(2r+1)} (n^2 - 1^2) (n^2 - 2^2) \dots (n^2 - r^2).$$

Thus
$$\overline{T_1^2} = \frac{n^2 - 1}{12},$$

$$\overline{T_2^2} = \frac{(n^2 - 1)(n^2 - 4)}{180},$$

$$\overline{T_3^2} = \frac{(n^2 - 1)(n^2 - 4)(n^2 - 9)}{2800},$$

$$\overline{T_4^2} = \frac{(n^2 - 1)(n^2 - 4)(n^2 - 9)(n^2 - 16)}{44,100},$$

$$\overline{T_5^2} = \frac{(n^2 - 1)(n^2 - 4)(n^2 - 9)(n^2 - 16)(n^2 - 25)}{698,544}.$$

The variance contributed by each term, and by which the residual variance is reduced when that term is removed, is therefore of the form $B^2 \overline{T^2}$, $C^2 \overline{T^2}$, and so on.

7. Distribution of variance for unchanging series of independent values.

A series may be said to change when the chance of an observation falling in any given range is a function of the time. When this is not the case the series is unchanging. The main difficulty in the adequate treatment of annual returns of economic and vital statistics lies in the prevalence of profound changes in the population observed. It is probable that the disturbing factors could be largely eliminated by the discriminating use of polynomials as here described.

If A_p is the coefficient of the term of the p th degree, and x is any observed value, then $A_p = \overline{T_p x / T_p^2}$; hence the variance of this term

$$A_p^2 \overline{T_p^2} = \overline{T_p x^2 / T_p^2};$$

if the values of x are independent, each with standard deviation, σ , and the series unchanging

$$A_p^2 \overline{T_p^2} = \frac{\sigma^2}{n}.$$

Hence, if polynomials are fitted to unchanging series of n quantities, the variance contributed by each term is on the average $1/n$ of the total variance. After fitting n terms the fit is necessarily perfect, and the residual variance zero. The first step merely consists in ascertaining the mean, as is ordinarily done before finding the variance, *the subsequent*

curve fitting, applied to an unchanging series divides the variance up into $(n - 1)$ parts, each of which has on the average the same value.

The actual frequency distribution of these fractions may be most simply shown by putting ϵ_p^2 for the variance of the term of degree p ; ϵ_p being supposed +ve or -ve according to the sign of A_p . Thus

$$\epsilon_p = A_p \sqrt{T_p^2} = T_p x / \sqrt{T_p^2}.$$

If x is distributed normally about its mean value, so then is ϵ_p with a standard deviation, which we have already evaluated at σ/\sqrt{n} . In unchanging series every value of ϵ_p , except the first, is distributed about zero as its mean; $\bar{\epsilon}_0 = \bar{x}$, the mean of the series.

In a changing series in which the change consists mainly in an alteration of the mean value, we may represent the change in mean value by a sufficient number of terms of a polynomial, the deviations from the mean being then an approximately unchanging series. The mean value of ϵ_p is then the value corresponding to the series of means; its standard deviation is unaltered, provided that σ is interpreted as the standard deviation of any observation from the changing mean.

8. The significance of an observed term.

If the change in the mean value is sufficiently represented by the terms up to T_r , then the mean value of the residual variance is

$$\frac{n - r - 1}{n} \sigma^2.$$

If the residual variance is σ_r^2 we may therefore take

$$v = \frac{\sigma^2}{n} = \frac{\sigma_r^2}{n - r - 1},$$

as the variance of each ϵ obtained. It may be that σ_r^2 is slightly increased if the terms up to T_r do not fully represent the course of the mean, in this case the significance of the preceding terms may be somewhat underestimated. There is no tendency to overestimate their significance.

If v is the mean variance contributed by each term, the mean value of the total variance contributed by p terms is pv . If t is the variance from an actual sample of p terms, the distribution of t is easily seen to be $df \propto t^{\frac{p-2}{2}} e^{-\frac{t}{2v}} dt$, a curve of type III. The standard deviation of t is therefore

$$\sigma_t = v \sqrt{2p}.$$

The standard error in the determination of the annual variance σ^2 is therefore represented by a coefficient of variation equal to $100\sqrt{\frac{2}{p}}$. In this expression p will be $n - r - 1$. In the application to the Broadbalk wheat yield $n = 67$, $r = 5$, so the standard percentage error

$$100\sqrt{\frac{2}{p}} = 18.11.$$

The combined significance of a group of terms may be derived from Elderton's tables of goodness of fit, taking for χ^2 the ratio of the observed variance to v , and for n' one more than the number of terms.

9. *Magnitude of the Residuals.*

The coefficients of the fitting polynomial are given by the equations

$$A_p = \frac{S(T_p x)}{S(T_p^2)},$$

where S represents summation over the n observations. The polynomial is therefore

$$\sum_0^r \frac{S(T_p x)}{S(T_p^2)} T_r,$$

and the residual is

$$x \left\{ 1 - \sum_0^r \frac{T_p^2}{S(T_p^2)} \right\} - S' \left\{ \sum_0^r \frac{T_p T_p' x'}{S(T_p^2)} \right\},$$

in which S' represents summation over all observations except x .

If x and x' are independently distributed about a mean at zero then for a fixed value of t the mean square residual is

$$\sigma^2 \left\{ 1 - \sum_0^r \frac{T_p^2}{S(T_p^2)} \right\}.$$

$$\begin{aligned} \text{Now } T_0 &= 1, & S(T_0^2) &= n, \\ T_1 &= t, & S(T_1^2) &= \frac{1}{12}n(n^2 - 1), \\ T_2 &= t^2 - n_2, & S(T_2^2) &= \frac{1}{180}n(n^2 - 1)(n^2 - 4), \end{aligned}$$

and so on.

The extreme value for t is $(n - 1)/2$; for this value,

$$\begin{aligned} \frac{T_0^2}{S(T_0^2)} &= \frac{1}{n}, \\ \frac{T_1^2}{S(T_1^2)} &= \frac{3(n-1)}{n(n+1)}, \\ \frac{T_2^2}{S(T_2^2)} &= \frac{5(n-1)(n-2)}{n(n+1)(n+2)}, \end{aligned}$$

and so on.

The average value of the mean square residual for all values of t is

$$\sigma^2 \left(1 - \frac{r+1}{n}\right),$$

as is necessarily the case since each term used removes σ^2/n of the variance; the reduction of variance is not the same for all values of t , and this introduces an element of heterogeneity. The reduction is greatest at the extremes. The polynomials tend to fit the extreme terms more accurately than the others. This effect is only strongly felt in the first and last terms. Fitting 67 terms with a curve of the 5th degree the variance of the first and last residual is reduced to 63.87 per cent. of its average value.

The residual variance for each value of t of a series of 67 terms, fitted to the 5th degree, is shown in the following table; the original variance being taken as 66, and the mean residue therefore 61.

Table VII.

$\pm t$		$\pm t$	
0	63.48	17	62.96
1	63.46	18	62.82
2	63.40	19	62.59
3	63.29	20	62.27
4	63.18	21	61.90
5	63.04	22	61.50
6	62.91	23	61.12
7	62.80	24	60.83
8	62.72	25	60.67
9	62.67	26	60.66
10	62.67	27	60.75
11	62.71	28	60.81
12	62.78	29	60.58
13	62.87	30	59.33
14	62.95	31	56.36
15	63.01	32	50.23
16	63.02	33	38.96

These are shown in Fig. 2.

The introduction of slight heterogeneity is a necessary consequence of the elimination of change by curve fitting; it is a weakness of the polynomial form that the extreme terms should be so much affected. The form of fitting polynomial is therefore unduly affected at its extremes by fortuitous circumstances; its values at the extremes give a less reliable index of mean yield than in the remainder of the range; still less is the rate of change at the extremes to be relied upon as a basis for prediction.

It should be noticed that for each year of the series separately the polynomial value is an approximation to the true mean for that year. The variance of the polynomial value about the true mean is complementary to that of the observation about the polynomial value (the

residual variance). Together they make up σ^2 , the variance of the observation about its true mean. Thus the residual variance of the end values being only 38.96 parts in 67, the variance of the polynomial value must be 28.04; this is four times the average value so that the probable error of the extreme values of the polynomial is double as great as the average value for the whole series.

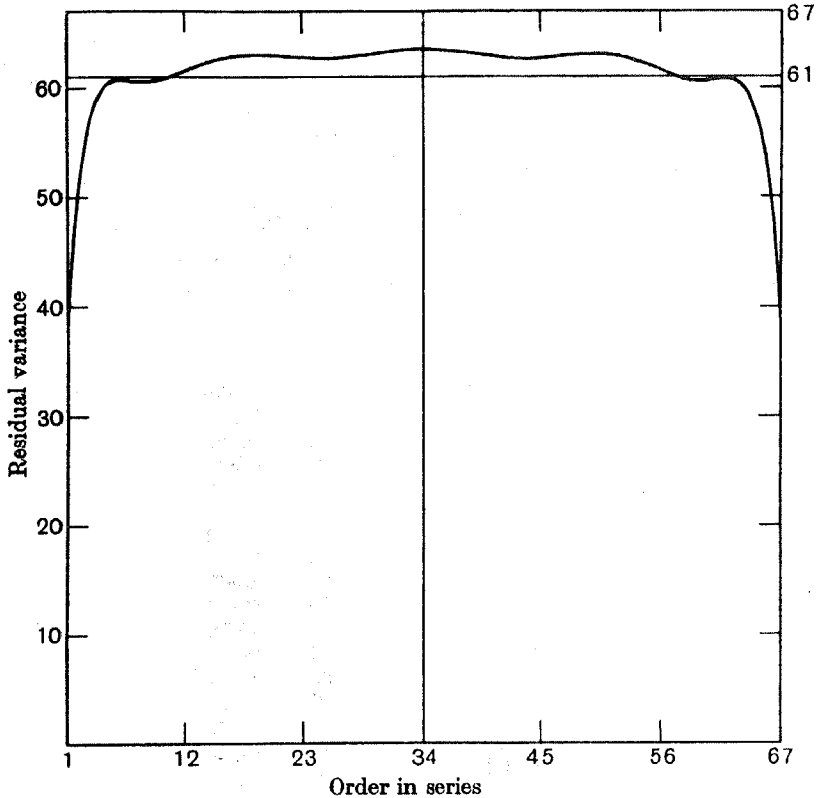


Fig. 2. Residual variance of individual terms, according to their position in the series.

10. *Correlation of residuals.*

The formulae of the last section show that once the changes of the mean have been sufficiently represented, the addition of further terms to the polynomial is disadvantageous (i) in increasing the probable error of all values, (ii) in increasing the heterogeneity of the residuals. On the other hand, if an insufficient number of terms are taken (i) the residual variance will be overestimated, (ii) the residuals will be heterogeneous by the confusion of changes with annual variation.

It is possible to test whether the process of fitting has been carried far enough by means of the correlations between neighbouring

values. The correlation between any two deviations from the mean of n , is, if they are independent,

$$-\frac{1}{n-1};$$

while the correlation between neighbouring values of a changing series will generally be positive. For example, in the yields of dressed grain from the dunged plot, 2*b*, of Broadbalk, neighbouring values are evidently associated. The correlations between harvests 1, 2, 3, up to 6 years apart are:

	r	z	P.E.
1	+·3559	+·3722	±·0850
2	+·1637	+·1652	±·0857
3	+·2767	+·2841	±·0864
4	+·1811	+·1831	±·0871
5	+·2030	+·2058	±·0878
6	+·1488	+·1499	±·0886
Mean	...	+·2429	±·0354
Expectation	...	-·0152	—
Difference	...	+·2581	±·0354

(In calculating averages or differences of correlations, it is advisable to use, as above the function z , connected to r by the equation

$$r = \tanh z;$$

when r is small, as in these cases, z is little different from r ; for large values, however, the inaccuracies involved in using the probable error of r as though errors in r were normally distributed, become considerable.)

Evidently neighbouring values tend to be alike. This correlation between neighbouring values gradually disappears as the changes in mean value are more and more closely represented by the polynomial, at the same time the negative correlation to be expected from an unchanging series is gradually increased. When these two values are no longer significantly different, the subsequent terms of the polynomial cannot be of importance.

The calculation of the correlation to be expected between residuals of k years apart, for an unchanging series or one in which the change is fully represented by the polynomial, is somewhat complex. When k is small compared to n it will be sufficient, as in the present applications of the theory, to ignore terms involving

$$\frac{k}{n-r-1}$$

to the third and higher powers. In these cases the correlation between residuals of the polynomial of degree r taken k years apart is

$$-\frac{r+1}{n-r-1} \left(1 - \frac{rk}{n-r-1} - \frac{rk^2}{(n-r-1)^2} \dots \right).$$

Thus, after fitting the polynomial for 67 values up to the fifth degree, the mean correlations between residuals are:

<i>k</i>	<i>r</i>	<i>z</i>	P.E.
1	-.0902	-.0904	±.0850
2	-.0817	-.0819	±.0857
3	-.0730	-.0731	±.0864
4	-.0640	-.0641	±.0871
5	-.0547	-.0547	±.0878
6	-.0452	-.0452	±.0886
Mean	...	-.0687	±.0354

The figures for plot 2*b* are even more negative than the expectation.

	<i>r</i>	<i>z</i>	
1	+.0034	+.0034	
2	-.1682	-.1698	
3	-.0624	-.0625	
4	-.1490	-.1501	
5	-.0583	-.0584	
6	-.0407	-.0407	
Mean	...	-.0797	
Expectation		-.0687	±.0354
Difference		-.0110	±.0354

Neighbouring residuals of this series are not less unlike than would be expected from an unchanging series; it may be inferred that the polynomial of the fifth degree sufficiently represents the course of the slow changes.

III. POSSIBLE CAUSE OF SLOW CHANGES.

11. *The local character of the slow changes.*

The distinction between the slow changes in mean yield and the annual variations is emphasised by the extremely local character of the former. An examination of the successive yields (*a*) of the experimental wheat at Woburn, (*b*) of the wheat averages for Hertfordshire compiled by the Board of Agriculture, (*c*) of barley, and (*d*) of grass from the experimental plots at Rothamsted, shows that the slow changes of mean yield observed on Broadbalk are not reflected in these series. To make the test objective correlation coefficients were found between the yields on plot 3*a* of Stackyard Field, Woburn, and (1) the actual yields of plot 2*b* of Broadbalk, (2) the polynomial values and (3) the deviations from the polynomial for the same plot.

Table VIII.

Correlations of plot 3*a* of Stackyard Field, Woburn, with

	<i>r</i>	<i>z</i>
Actual yield ...	+.31	+.32 ±.12
Polynomial value	+.02	+.02 ±.12
Deviations ...	+.34	+.35 ±.12

There is no significant association of the polynomial value with the yields on Stackyard Field, Woburn. The latter values are indeed more closely associated with the deviations from the polynomial than with the actual yields. The correlations are all somewhat low, perhaps owing to the very different soil at Woburn. The probable errors refer to deviation in the value of the correlation to be expected in a random sample of 36 years, the three values above are, however, taken for identical years, so that the probable error gives no indication of the significance of their differences.

Table IX.

Correlation of Hertfordshire mean
wheat yield with

	r	z
Actual yield ...	+·31	+·32 ±·13
Polynomial value	-·16	-·16 ±·13
Deviation ...	+·55	+·61 ±·13

As before, the correlation with the polynomial value is not significant. Its actual value is negative, and the effect of including the polynomial terms in the actual yield, is to reduce the strong positive correlation shown by the deviations to a value which by itself would not be clear proof of any association whatever.

The comparison with the dunged barley plot on Hoos Field is interesting, since it might be suspected that the slow changes in mean yield, at any rate on the dunged plots, were partially caused by changes in the manurial value of the farmyard manure employed. That this effect is not of great importance may be seen by comparing the polynomials from the dunged plot with that of plots receiving only artificial manures (Fig. 3). This conclusion is confirmed by the comparison with Hoos barley.

Table X.

Correlation of plot 7 (2) of Hoos
Barley Field with

	r	z
Actual yield ...	+·28	+·29 ±·09
Polynomial value	+·09	+·09 ±·09
Deviation ...	+·23	+·23 ±·09

All the correlations are small, and that with the polynomial is insignificant. The inclusion of these latter has in this case slightly increased the correlation observed from the actual yield.

With the yield of the grass plots the correlations are negative, but are all insignificant.

Studies in Crop Variation

Table XI.

Correlation of plot 9 (first crop) of
Park Grass with

	<i>r</i>	<i>z</i>
Actual yield ...	-.19	-.20 ± .12
Polynomial value	-.12	-.12 ± .12
Residuals ...	-.12	-.12 ± .12

Thus the correlations of the residual yields with neighbouring wheat crops is clearly significant and greater than that of the actual yields, with barley and hay the correlations are less distinct; in no case do the polynomial values give significant correlations. The causes of the changes in mean wheat yield on Broadbalk do not affect the wheat yields of Woburn or Hertfordshire, or the barley and hay of neighbouring fields at Rothamsted; they appear to be limited to Broadbalk field.

The relation between yield and weather cannot be fully treated here; indeed, the present results may be regarded as preliminary to a thorough attack upon that problem. It is here only necessary to justify the assumption that the weather of a succession of years may be treated as an unchanging series (see Section 7). A full investigation of this point cannot be attempted until the whole meteorological data from the Rothamsted Station has been analysed. As an example of the method of investigation appropriate for testing this question in its present application we may cite the case of October rain.

Hooker (1907) found that the rainfall in autumn was more influential than that at other times of the year. The series of October rains was therefore examined to see if any change in mean rainfall could be detected. To do this accurately the series was fitted with a polynomial of the fifth degree, exactly as had been done to the wheat yields, and the chance of obtaining larger coefficients from an unchanging series was calculated. Every term individually proved to be insignificant, the chance of finding greater variance in the first 5 terms of an unchanging series being .88 (compare Table II). There is thus no evidence of more than random changes in the October rain.

One significant, though small, correlation was found between the polynomial values and the weather. On examining the correlation of total rainfall for each harvest year (Sept. to Aug.) with the succeeding crop, the following values were found.

Table XII.

Correlation of total rainfall with

	<i>r</i>	<i>z</i>
Actual yield ...	-.60	-.69 ± .09
Polynomial value	-.29	-.29 ± .09
Deviations ...	-.56	-.64 ± .09

This would indicate that a perceptible proportion, nearly 9 per cent. of the variance of the *mean* yield is due to difference in the year's rainfall. A further point of interest in connection with this result is that the optimum rainfall derived from the deviations is about 21 inches, that from the actual yields 17 inches, while for the polynomial value there is no optimum at all. This suggests that ideal conditions for the wheat plant on this soil requires a rainfall about 21¹ inches, but that lower values though somewhat injurious to the plant, are of permanent value to the field by facilitating the eradication of weeds.

If this be so, it illustrates a point to be borne in mind in considering the effect of weather upon crops, that the ideal weather for the plant is not necessarily ideal for the purpose of farm operations; it is not to be expected that in every case these two classes will be distinguishable as above, the one as an annual effect, the other with relatively permanent consequences.

12. *Possible influence of weeds.*

Of all the organic factors which influence the yield of wheat, it is probable that weeds alone change sufficiently slowly to explain the changes observed on Broadbalk. Farm weeds are notoriously difficult to eradicate; deep-rooted perennials throw up stems year after year from below the cultivated layer, whilst annuals, and on Broadbalk especially the slender Foxtail grass, produce numerous seeds which may lie dormant for several years, so baffling the most thorough attempts to free the field once it is badly infested. The conditions of cultivation of Broadbalk make the control of weeds peculiarly difficult. The wheat, grown year after year, has been sown in autumn in every year save one (1853) leaving but a short space of time for cleaning the land. The mean date of autumn sowing is Nov. 1.4 and that of carting the crop is Aug. 24.9, leaving a mean interval of only 68.5 days. Compared to fields bearing spring-sown corn, or periodical root crops the position of Broadbalk is exceedingly unfavourable. It is therefore not improbable that if for a period the weeds are unchecked either from a relaxation of effort, or from unfavourable seasons, a considerable number of successive wheat yields would be reduced, giving rise to such periods of depressed yield as have been observed in the early fifties, and the seventies of the last century and about 1914.

¹ This refers to rainfall as *actually distributed* throughout the year. To discover the optimum distribution of rainfall will require a far more elaborate analysis.

13. *Records of Weeds.*

(1) An old record exists giving the exact work of every man employed in each day. The record includes parts of the four years 1852 to 1855, and is complete for the year 1853. During this period the wheat yields show that the condition of the field was improving even on the unmanured plots. The amount of hand labour employed on Broadbalk is striking; in 1853 211.5 man-days and 714 boy-days were expended in weeding the field. The whole period of weeding operations little exceeds 100 working days, so that roughly the work done was equivalent to the continuous labour of two men and seven boys, on a field of about 14 acres. The proportion of boy labour is striking and the record is sufficiently detailed to show how it was utilised.

The weeding done by the boys was principally by hand, with some spudding in May or June. Only men used hoes. In April boys were at work "picking twitch," in June they were "pulling garlic" (presumably charlock) "and the larger weeds," and in July they were "pulling wild oats" in the high corn.

If our interpretation of the change in mean yield is correct, this free use of boy labour was eminently successful, for the mean yield mounts rapidly to its maximum about 1860. Unfortunately, no comparable record of the employment of labour exists at other dates.

(2) In the Broadbalk records the first botanical account of weeds is in 1867, a date corresponding to the beginning of the second depression in the mean yields (see Fig. 1). From then to 1889 the field is frequently described as "exceedingly foul"; from 1889 the weeds are not mentioned till 1904, when the weeds were sufficiently dominant to require that the field should be fallowed, as was done by halves in 1904-5, and again in 1914-15. The evidence of the records thus confirms the supposition that the field suffered from weeds at periods about 1877, and 1910, and was relatively free in the high-yielding period in the nineties.

From the observations of the seventies it is clear that the dominant weeds consisted of five perennials, *Sonchus arvensis* (Corn Sowthistle), *Convolvulus arvensis* (Bindweed), *Equisetum arvense* (Horsetail), *Cirsium arvense* (Creeping Thistle), and *Agrostis vulgaris* (Twitch grass), and three annuals, *Polygonum aviculare* (Knotgrass), *Myosotis*¹ *arvensis*

¹ In the opinion of Dr Brechley, whom I have had the privilege of consulting, it is to be doubted that this weed, even when conspicuous, was really present in sufficient quantity sensibly to depress the yield.

(Birdseye), and *Stellaria*¹ *media* (Chickweed). A fourth annual, the Slender Fox-Tail grass, *Alopecurus agrestis*, was certainly regarded as an unimportant weed in 1867 and 1869, and is not mentioned in 1872, 1873 and 1876, but in 1879 and the eighties it has become enormously abundant, and at the present time it is considered to be by far the most troublesome weed. In 1886 this weed had become such a pest that "Sir John decided that pulling up by hand should be resorted to." In connection with this, it may be remembered that the Education Acts of 1876 and 1880 made attendance at school compulsory; the boy labour which had regularly hand-weeded the land in the past had evidently been cut off for some time before 1886. Great efforts were made again in 1887 to eradicate the weed by hand weeding, but the wet summers of 1888 and 1889 prevented this operation and the land again became very foul, and was partially fallowed in 1890 and 1891 by drilling the rows at double widths over half the field. After this the weeds seem to have been kept in check until Sir John Lawes' death in 1901². In 1904 *Alopecurus agrestis* was so thick that the field was given a complete fallow in two halves in 1904, 1905.

14. *Indirect evidence of the influence of weeds.*

Much evidence has already been adduced that the slow changes in the mean yield is due to very different causes from those that produce the annual deviations; in particular, the local character of the former has been mentioned. The variation due to those two causes, *relative to the mean of each plot*, may be conveniently expressed as a coefficient of variation.

The most striking point about these figures is the comparative constancy of the coefficient of variation for slow changes, especially for neighbouring plots, if we exclude plots (17 and 18) for separate discussion. The variation is uniformly greater in the north half (plots 2b to 8) of the field than in the more southern portions (plots 10 to 14), the localisation of the effect thus showing itself even within a single field. The coefficients of variation from annual causes are quite different; they show no influence of locality, the most variable being those plots in which exhaustion has been most rapidly in progress. The proportionality

¹ Possibly *Arenaria serpyllifolia* (Sandwort) was sometimes recorded as *Stellaria media*; the latter weed is, however, the more frequent in recent reports.

² During the nineties parties of schoolgirls were employed at hand-picking, at Easter and Whitsun, on Saturdays and in the evenings. Sir John Lawes took much interest in their work, giving prizes to those who collected the greatest quantity.

of the slow changes is well shown by plotting the polynomials upon a logarithmic scale, as in Fig. 1, which shows the changes in mean yield on plots 2*b* to 14.

Table XIII.

Coefficient of variation due to slow changes and annual deviations

Plot	Slow changes	Annual deviation
2 <i>b</i>	12.13	16.68
3 and 4	12.86	24.90
5	11.56	24.17
6	12.53	24.51
7	11.61	22.61
8	11.16	20.44
10	9.83	33.32
11	9.33	32.13
12	9.88	26.98
13	9.54	23.56
14	9.00	25.26
17 } minerals	14.72	31.88
18 } ammonia	10.35	21.32

This proportionality, while discriminating the slow changes sharply from the meteorological effects, is not an unexpected feature if these changes are due to the varying prevalence of weeds. For in a bad period perhaps 25 per cent. of the area of each plot is unproductive, while after many years of constant attention the percentage lost may be reduced to about 5 per cent.; the constancy of the coefficient of variation in the different plots indicates that the area lost in the same year is roughly proportional in all plots.

An exceptional amount of variation is shown by the mineral series of plots 17 and 18; this plot differs from the others in that while the growth of wheat is limited by shortage of nitrogen, in the previous year it received a nitrogenous dressing. This should be to the advantage of the perennial weeds which should have benefited by the previous year's manure. The mineral plot should therefore suffer particularly when the field is infested by perennial weeds, while the corresponding ammonia plot should show principally the effect of annual weeds.

A comparison of the polynomial for plot 17 and 18 minerals with that for plot 5 which it nearly resembles, shows that (Fig. 3) the former was in fact more seriously depressed in the first two periods of low yields, to such an extent that it has a lower average in the early fifties and about 1880, while in the recent depression the effect is less marked, its minimum being higher than that of plot 5. On the other hand, comparing plot 17 and 18 ammonia with plot 7, which receives

ammonia continuously, the first two depressions of the former are considerably less, while the recent depression is quite as great as that of the latter. These differences are comprehensible if in the two earlier

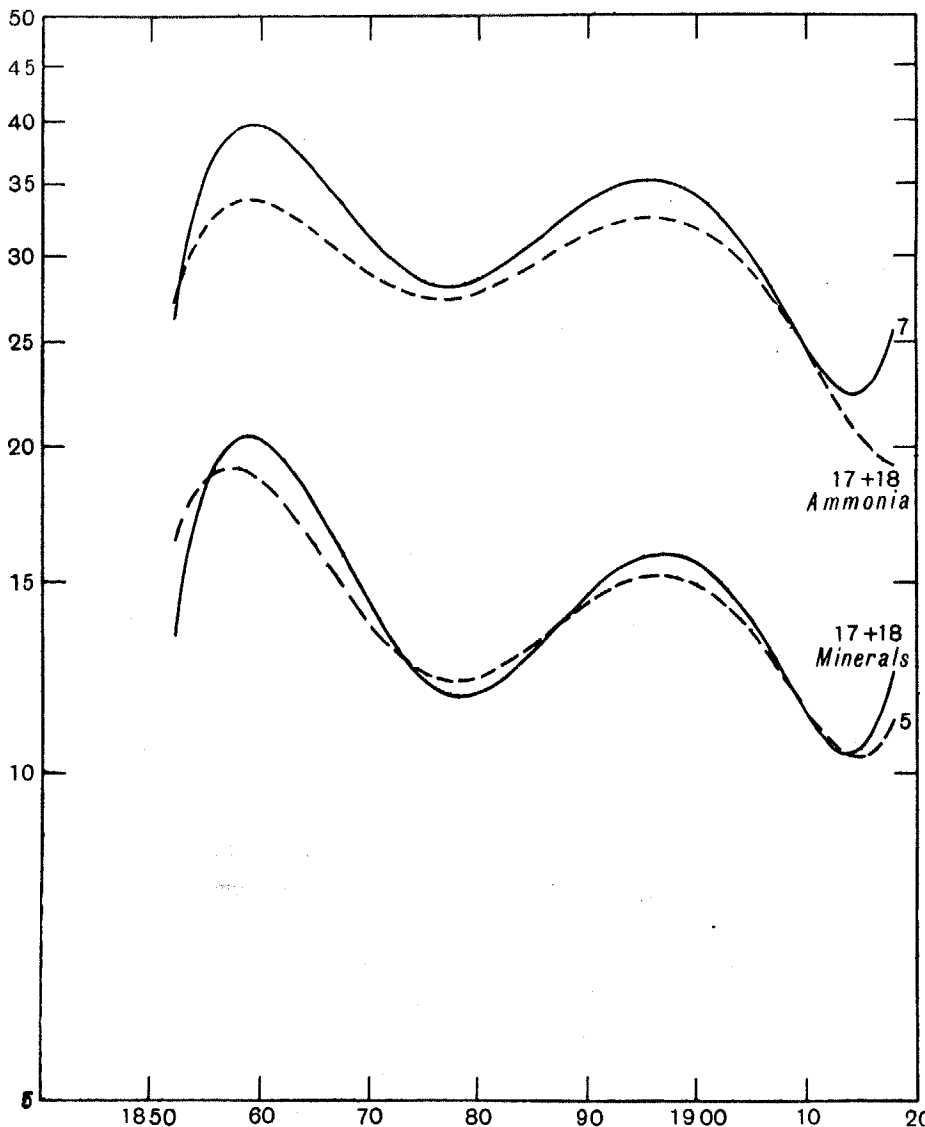


Fig. 3. Yield of alternating plots 17 and 18, compared with plots 5 and 7; showing the exaggeration of the two earlier depressions, where a nitrogenous dressing has been applied in the previous year.

depressions perennial weeds were more influential than has been the case in recent years, while the recent depression is due in larger measure to annual weeds. Similar conclusions have been seen to be indicated by the references to weeds in the records.

SUMMARY.

In Part I is given a survey of the results of a statistical examination of the yield of the plots of Broadbalk Wheat field during 67 years. The main features of the comparison of mean yields are well known; the comparative rates of decrement, shown in Section 5, supply a class of facts well worthy of further study. Particularly striking are the relatively slow rates of decrement of plots 2*b* and 8, compared with plot 7, which would seem to show a permanent advantage in very high nitrogenous dressings, and to emphasise the need for caution in the application of the principle of diminishing returns. The evidence of the influence of potassium sulphate and its substitutes, sodium sulphate and magnesium sulphate, shown in Table V, is also very striking. An unsuspected feature of the changes of mean yields, which precludes the possibility of obtaining from these data true curves of exhaustion has appeared in the slow changes which have taken place in all the plots in a similar manner. In Part II the mathematical methods by which the variation has been analysed has been discussed, partly as a justification of novel procedure, partly to make clear that the three types of variations found have been genuinely distinguished. In Part III such evidence as is available has been presented, in order to throw light upon the possibility that the changes in mean yield have been caused by variations in the prevalence of weeds at different periods.

One point of importance which should be emphasised is that average wheat yields, even over long periods, from different fields or for different seasons cannot approach in accuracy the comparison of plots of the same field in the same seasons. The advantage of the method adopted by Lawes in the permanent experiments which he instituted is very evident. The effects of weather clearly require that the seasons should be identical, unless the series be very long, but the slow changes in mean yield show that even comparatively long series of different years from the same field cannot be accurately compared. Within the same field, however, the slow changes have almost proportional effects, and comparison between the mean yields of neighbouring plots may be made with great accuracy. The only case in which changes in mean yield sensibly affect the comparison of averages is that of plots 17 and 18.

* In comparing these with plots 3 and 4, 5, 7, and 10, it would be more accurate to confine attention to high yielding periods, at which the disturbing causes are at their minimum.

* For 3 and 4, read (3 and 4).

It is believed that the deviations from the smooth curves, which have been freed, for the most part, from the effects of exhaustion and weeds, form statistically homogeneous material for the study of meteorological effects.

REFERENCES

- R. H. HOOKER (1907). "Correlation of the weather and crops," *Journal of the Royal Statistical Society*, **70**, pp. 1-42.
- E. A. MITCHERLICH (1909). "Das Gesetz des Minimums und das Gesetz des abnehmenden Bodenertrages," *Landw. Jahrb.* xxxviii. 537.
- K. PEARSON (1914). *Tables for Statisticians and Biometricians*, Camb. Univ. Press.

(Received November 1st, 1920.)