13th. May, 1929.

Professor Bhai Balmukand,
Agricultural College,
Lyallpur,
Punjab,
India.

My Dear Balmukand,

It is a great pleasure to hear from you again. The honour I have received this year is of course very gratifying to me, but its chief importance is in showing that in the highest scientific circles in this country the subject of Statistics, straggling and anomalous as it is, is beginning to be appreciated.

I hope that in the fertility enquiry it will be possible to make a rough classification by age, as I am sure that this will be needed in interpreting the results.

I am glad you are using uniformity trials to demonstrate the efficiency of plot arrangement. In Ex. 44, the 19 independent differences between stripe gives a S.S. 289,766 and a mean square 15,251 so the S.D. of a single strip is 123.5. The S.D. of the difference between 2 means of 4 is 87.32 or 2.657% of the general mean 3285.75; in practice with 5 varieties we would have only 15 D./F, so that the 5% point has t 2.131, so that the significant difference between any two varieties is 5.662%

In the diagram I had used an unusual measure of significance by supposing the relative yields of the 5 treatments to be in Arithmetical progression 100-2a, 100-a, 100, 100+a, 100+2a, with S.S. $4 \times (2^2+1^2+1^2+2^2)a^2 = 40a^2$, Mean Square $10a^2$. For z = .5585, $e^{2z} = 3.039$

$$10a^2 = 3.039 \times 14.119 \qquad 2(2.657)^2 = \frac{15,251}{(32.8575)^2}$$
$$a = 2.071$$
$$4a = 8.284 \text{ (instead of 6.82).}$$

at least that is what I should do now, if I were to use the convention of an Arithmetical Progression and evaluate how far apart the best and worst treatment should be to give a detectable difference. I certainly had no table of z when the diagram was made, and may have used an unsatisfactory approximation based on χ^2. Such a convention makes no difference in comparing different field arrangements for the same number of treatments, and is merely a way of interpreting the significant value of the mean square, by calling it $10a^2$.

With the replicated Latin square, using your figures we have:-

	D/F	S.S.	M.S.	S.D.
Rows and columns	64	44772		
Squares	7	19119		
Remainder	128	18925	147.85	12.16
Total	199	82816		

whereas you note that the S.D. from the single square classes chosen out of the field for Ex. 46 is 11.41. This argument is to be expected for the one chosen ~~out of the field~~ was one of the eight for which 12.16 is the average value.

The advantage of using 8 lies wholly in the replication, i.e. in dividing the variance of all comparisons by 8. It has been often stated that increased replication fails to give the full theoretical advantage by reason of the greater heterogeneity of the soil in a larger area. The method of replicated Latin squares enables the whole theoretical advantage to be reaped.

$\frac{147.85}{(3.28575)^2} = 13.695$ 5% $n_1 = 4$, $n_2 = 96$ $z = .4415$ about

40 plots instead of 4. $100a^2 = 33.11$ $e^{2z} = 2.418$

$a = .5755$

$4a = 2.302$ (instead of 1.94),

again bigger than the diagram value. For calculating which I must have used some different convention now forgotten. The comparison, however, seems nearly equivalent.

I am glad the Agricultural Commission reported as you say. Their interview with me may have been some good after all. It should help you in pressing for better experiments.

Yours sincerely,