Journal of Industrial Microbiology, 12 (1993) 295–300

Published by The Macmillan Press Ltd

Design of experiments for predictive microbial modeling

Kenneth W. Davies

Unilever Research Laboratory, Colworth House, Sharnbrook, Bedfordshire, MK44 1LQ, UK

Key words: Experimental design; Predictive modeling

SUMMARY

Good predictive microbial models can be built with appropriate data from well-designed experiments. Anyone setting up an experiment should consider the sources of variability, possible screening experiments, optimum spacing between points on a continuous scale, and the most appropriate type of design, e.g. factorial, screening, or central composite.

INTRODUCTION

In order to use time and resources to the best advantage, it is essential that all experiments have clear aims. The two most common aims are (i) understanding and (ii) optimization, which can have conflicting experimental requirements.

Sometimes the concern is to know which out of many factors affect the growth of certain microbes. At other times, interest may center in a small number of factors, but with the desire to know more precisely how these factors influence growth either singly or in combination with each other. In both these cases the aim is obviously to achieve understanding.

Other situations can arise where it is necessary to find an optimum or near optimum set of conditions for preservation or growth. For these situations it is not essential to have complete understanding over a wide region, but just enough information at or around the optimum to be able to predict it well.

Variability

To avoid misleading conclusions and to get the best out of any experiment it is essential to design it well, and this includes controlling as much as possible. It is therefore necessary to know what sources of variability will influence the final result other than the factors which are of specific interest.

Variation can occur between different strains of the same bacterium. It can also be due to previous bacterial history, such as time spent in stationary phase, or time spent at a certain pH, temperature, or salt level. Bacteria can grow in slightly different ways on different batches of media.

It is generally desirable to make conservative predictions (i.e. ones that err on the side of safety), and for this reason it is best to perform experiments using a cocktail of strains. Alternatively if understanding of microbial behavior is desired it is best to use a single strain.

Variation can occur not only between different batches of raw material, but between different samples from the same batch. A sample may not be homogeneous. For example, dried material that has been shaken will generally have larger particles at the top, and smaller ones at the bottom.

The process of dilution, involving as it does, a pattern of mixing and sampling several times, can increase the variability at each stage, and is a factor that contributes to the lognormal distribution of results that are obtained from plate counts.

Last, but not least, it has often been observed that different experimenters performing essentially the same experiment can get different results, particularly where operations such as pipetting, spreading and counting are done manually. Consistent results within an experiment or across a series of experiments can therefore be achieved more easily if a given task is performed by the same person each time.

Factorial designs

A common aim is to find out how microbes grow when there are several influential factors. One way of approaching the problem is to take each factor in turn and see how the growth pattern changes as that factor is varied. An alternative is to consider all the factors simultaneously, looking at appropriate combinations of them. The latter approach makes it possible to find out whether the factors interact, and by how much. The following figures illustrate where these interactions are important.

Consider the situation where two factors X and Y influence the growth of bacteria in a given time. Figure 1 illustrates the sort of results that could be obtained if 25 experiments were performed. It will be seen that the results

Correspondence to: K.W. Davies, Unilever Research Laboratory, Colworth House, Sharnbrook, Bedfordshire, MK44 1LQ, UK.

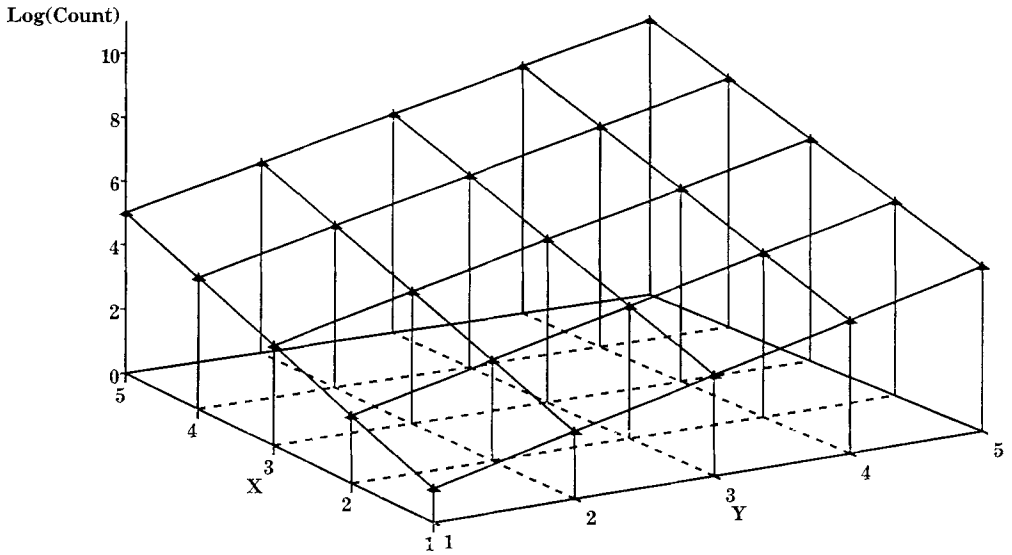

Fig. 1. Factors X and Y do not interact.

all lie on a plane, i.e. a flat but sloping surface. These sort of results can be explained by an additive model in which X and Y do not interact. If it were known there was no interaction it would have been sufficient to take each factor in turn and keep it constant while varying the other one (e.g. take X = 1 and vary Y from 1 to 5, then take Y = 1 and vary X from 1 to 5). The results from these experiments would be sufficient to predict all the others.

Figures 2 and 3 show examples of interactions that could occur. Figure 2 illustrates 25 results that may arise from combinations of factors X and Y. If an experiment were performed fixing X = 1 and varying Y from 1 to 5, a fairly shallow response would be seen. A similar picture would be seen by fixing Y = 1 and varying X. What would not be found from this one-factor-at-a-time approach is that X and Y act synergistically and give a much higher response than expected if increased together.

Adopting the one-factor-at-a-time approach in Fig. 3 would lead to an overestimate of the effect of increasing X and Y simultaneously. Both of these examples show clearly the benefit of varying the factors simultaneously.

Confounded designs

A confounded design is one where two effects cannot be separated, e.g. two samples containing bacteria are to be compared but the first is assessed by one person, and the second by another. If there is an apparent difference it is impossible to deduce from the figures whether the samples or the two people's techniques are different. For this reason, it is important to ensure that conditions are consistent

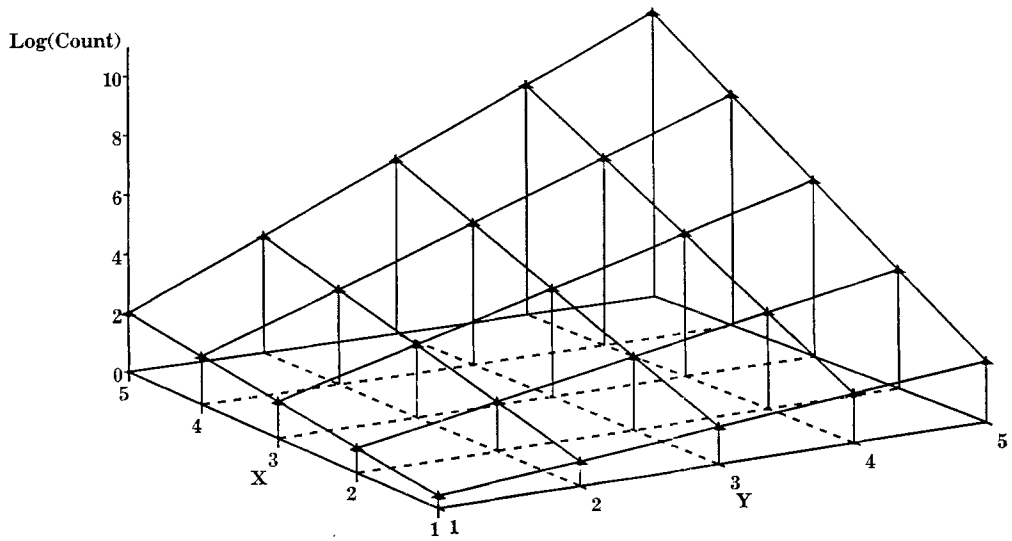

Fig. 2. Factors X and Y interact, example 1.

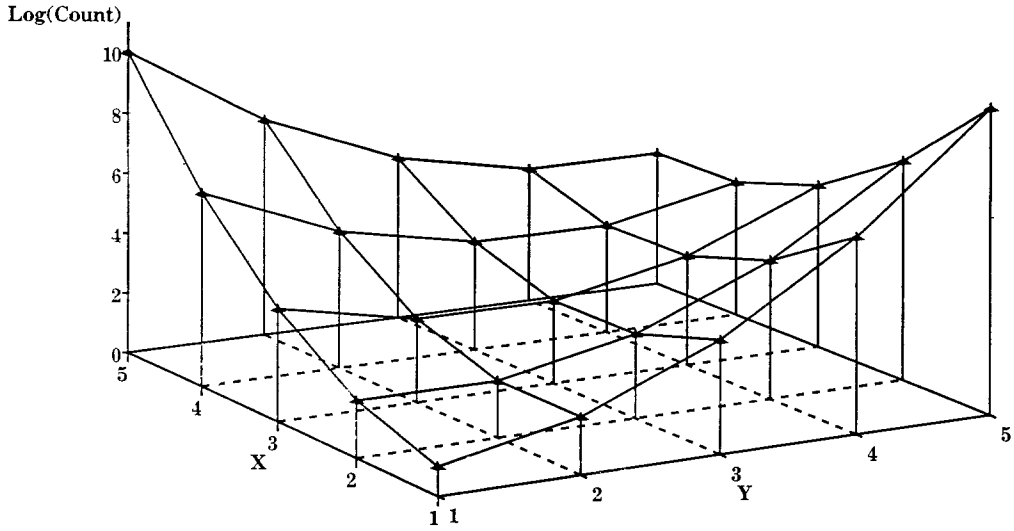

Fig. 3. Factors X and Y interact, example 2.

throughout each experiment. Results can then be interpreted more easily and with more confidence.

Screening designs

When there are a lot of factors that might influence our results, and it is not apparent which are important, it is advisable to do an experiment in order to find this out. A large factorial trial would yield this information, but it would take a lot of resources. Instead it would be better to do a small subset of the complete factorial, and initially as small a subset as possible. There is a class of designs called fractional factorial designs which are admirably suited to this. Consider the situation where there are three factors which each have three levels of interest. Using all the combinations in an experiment would entail using $3^3 = 27$ combinations. Instead, just a third of them could be used, from which could be derived information on the main effect of each of the factors.

Figure 4 shows how to select a subset of 9 out of 27 combinations in the situation where there are three pHs (A, B,C), three different temperatures, and three levels of salt. This type of a fractional factorial is often called a latin square. It can be seen that each pH occurs in every row and every column.

		Temperature		
		5	10	15
Salt %	1	A	B	C
	2	B	C	A
	3	C	A	B

Fig. 4. Example of latin square.

Whenever any design is used it should be randomized so that an experiment is not going to be conducted in exactly the same order as on a previous occasion. In the case of latin squares, a suitable one should be taken at random. The rows should then be randomized, followed by a randomization of the columns. The result will still be a latin square, so the essential nature of the design will be preserved. It must be noted that fractional factorial designs will not allow estimation of all the interactions. The smaller the design, the fewer interactions that can be estimated.

Figure 5 shows an example of a design in which there are seven factors A–G, each at two levels, which for convenience are called low and high. The design consists of just eight out of the $2^7 = 128$ possible combinations which could have been studied. Each combination in the design is carried out with some of the factors at the low level and others at the high one. The diagram contains a block to indicate that factor is at the high level, and is blank where it is low. Thus it can be seen that the first experimental combination will be done with factors A, D, E, and F at high levels and the others at low levels.

This particular design has the property that the overall effect of each factor can be found by taking the average result obtained from experiments where the factor is at a high level, and subtracting the average result where the factor is at a low level.

Figure 6 shows some results from an experiment where this design was used, together with the computed effects of each factor (shown as an absolute value). It will be seen that factor E appears to have the most effect on the final results, followed by B and F. Only by doing significance tests can we tell whether any of these effects are statistically significant. In this case only factor E had a significant effect.

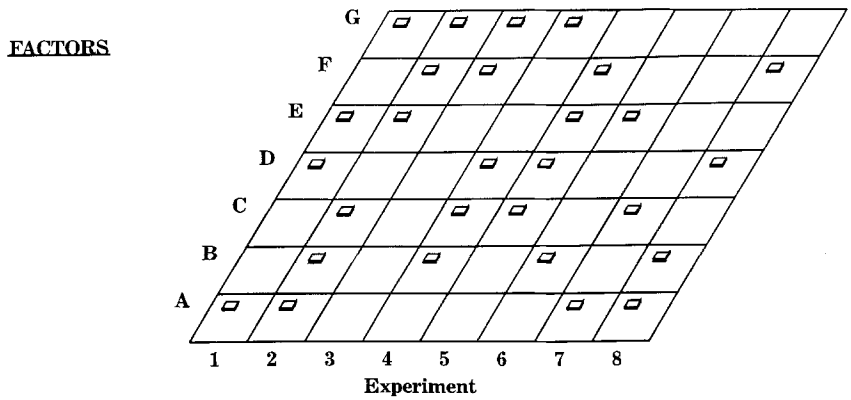

Fig. 5. Screening design with seven factors each at two levels.

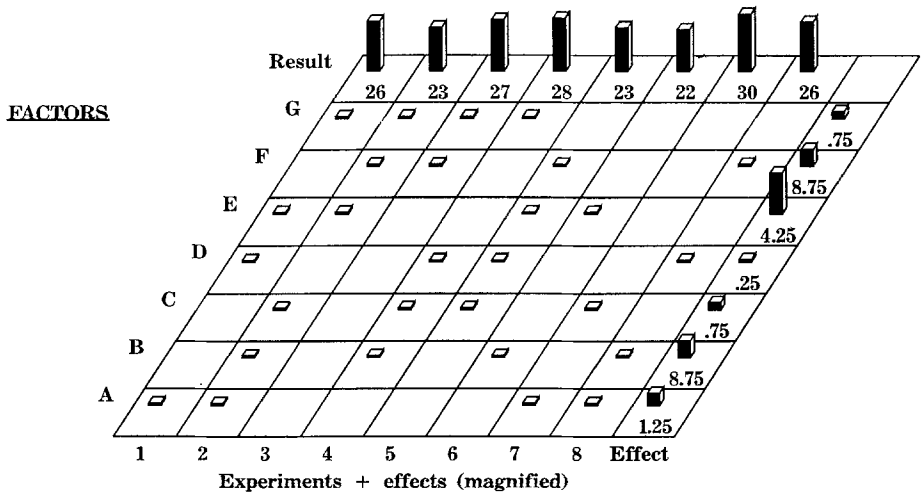

Fig. 6. Design, results and effects.

Screening experiments

If a factorial design is being contemplated, but there is prior information on which combinations will give no information, these should be left out of the design and effort concentrated on the ones that are worthwhile. If the main experiment is concerned with bacterial growth as measured by plate counts, a screening experiment could be done looking at the simpler measurement of time to turbidity in order to determine which factor combinations are worthwhile in the main experiment. The main experiment may then become a partial factorial, such as is shown in Fig. 7. This example presumes that bacteria are known not to grow at low salt and temperatures when there is a high $[H^+]$. These have therefore been excluded from the main experiment.

Scale for design (explanatory variables)

When choosing levels for design variables it is essential that they encompass the full range over which predictions are to be made, because extrapolation of results to any points outside could be dangerous. If an experiment is needed to investigate the effect of salt on the growth of microbes it is generally best to choose equal intervals of salt or water activity, whichever is considered most appropriate, e.g. salt levels of 1%, 2%, 3%, 4%. In the absence of any other information, this is a very reasonable thing to do, but if a greater change took place between salt levels of 2% and 3% than between any other levels, it could be of interest to look also at salt of 2.5%.

This idea can be generalized by saying that where changes are taking place fastest, it is generally useful to have more levels (or observations). It is often considered that microbes react to pH, but it has been the experience at Colworth that they are actually reacting to the concentration of hydrogen ions. Therefore if it is desirable to look at the effect of pH, more useful information will be obtained from spacing the pH at intervals of equal hydrogen ion concentration rather than equal pH intervals.

Similarly if population growth is measured with time, more information will be obtained for the same effort if intervals between cell counts are short for a rapidly growing population, and longer for a slower growing one. Also if the end of lag phase is to be estimated more precisely, it is

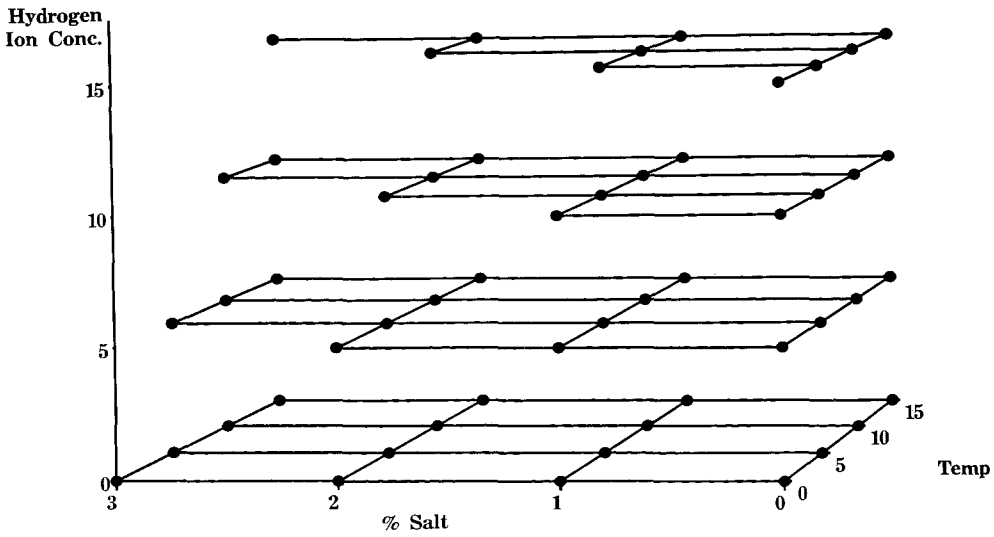

Fig. 7. Partial factorial screening out lethal combinations.

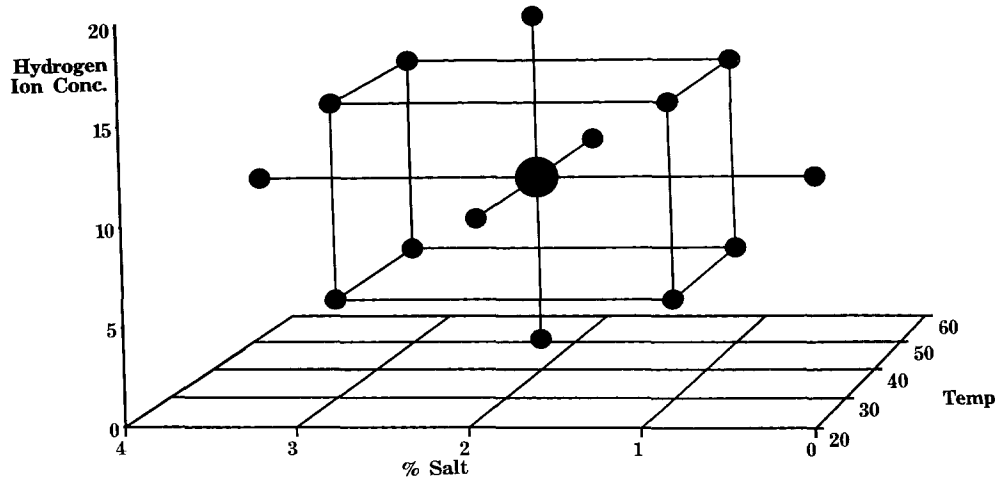

Fig. 8. Central composite design for estimating optimum.

strongly advisable to take more counts around the time when this is expected to occur. Much of this is just common sense.

Central composite designs

It has been said that factorial designs are good for obtaining a broad understanding over a region of interest. If this is not desired, but an optimum of some sort is sought instead, the best design in this situation is a central composite design. An example of a central composite design in three dimensions is shown in Fig. 8.

The design should be centered around the main region of interest, and the larger symbol in the center indicates that greater replication is needed there. The length of a unit step should be chosen to correspond to an equal expected change in response in each of the three directions. It can be seen that the exterior points lie roughly on a sphere, and the positions are ideally suited to the fitting of a generalized quadratic, or response surface equation, as it is sometimes called. It should be noted that the spherical nature of this design means that there are no points in the 'corners' of the region. More information is being gained about the center of this region at the expense of the extremes, and as a result predictions have the greatest error farthest from the center. If however, there is an optimum as expected, it can be estimated quite well from the generalized quadratic equation.

This type of design is ideal for the study of fermentation processes and vaccine production, where interest centers on obtaining optimal growth of bacteria. It is a poor design to use when trying to prevent proliferation of pathogenic microorganisms.

GENERAL STRATEGY

A general strategy for experimental design with several factors should include the following:

(1) Define experimental objective.

(2) List all the possible factors.
(3) Design a screening experiment to determine which factors are the most important.
(4) Follow up, if necessary, with a screening experiment to determine the most appropriate range for these factors, e.g. measure time to turbidity.
(5) Choose a central composite design (to find an optimum) or factorial design (to improve overall understanding).

CONCLUSION

A brief introduction has been given to the subject of experimental design in the context of predictive microbial modeling. It has only scratched the surface of a very large subject on which there are many excellent books. Those who wish to pursue this interest would benefit from reading Box et al. [1], who cover basic concepts and methods of design with an emphasis on building models.

REFERENCE

1 Box, G.E.P., W.G. Hunter and J.S. Hunter. 1978. Statistics for Experimenters. Wiley, New York.