SSSU 132
ISSN 0140 3818

University
of Southampton

School of Engineering Sciences

Ship Science

**AUTOSUB Propulsion System Investigation:
Theory of Experiment Design**

CD Fallows

Ship Science Report No 132

January 2005

SSSU 132
ISSN 0140 3818

AUTOSUB Propulsion System Investigation: Theory of Experiment Design

CD Fallows

Ship Science Report No 132

January 2005

AUTOSUB Propulsion System Investigation

Theory of Experiment Design

University of Southampton

School of Engineering Science

Ship Science

C.D.Fallows

Report Number : 132 **January 2005**

Contents

THEORY OF EXPERIMENT DESIGN

1 Introduction

As a result of analysis of the propulsion system, it was concluded that a series of laboratory experiments is required to determine the drag characteristics of the hull and its appendages. However, it was demonstrated that the potential size of the experimental space was such, that careful design is required to determine a series of measurements that will produce useable results within a realisable level of resource. It was shown that the experimental space to be explored is far too large to enable investigation of all combinations of factors at all levels.

It has also been shown that for any experiment there will be factors that will affect the results, but which will neither be under control, nor be specifically measured. The effect of these will be to appear as errors in measurement of the principal effects. These errors will need to be allowed for in any subsequent analysis. The primary issue is, therefore, how best to sample the experimental space such that the maximum amount of information can be extracted from a realistic number of measurements. The aim is to produce the minimum sample of measurements that will describe the response surface with sufficient confidence. The levels for each factor need to be chosen such that there is a high probability of detecting critical inflections in the response surface.

2 Design Approach

Because of this complexity, a primarily statistical and descriptive approach, rather than that based purely on physics, is proposed, (although, as always, the statistical approach must be informed by knowledge of the underlying physics of the problem). Because of the complexity of the real vehicle, and the limited number of runs feasible for any realistic experiment, careful selection of the combinations of levels for each factor to be set for each experiment is required. As a result of the intention to reflect complexity, the experiment may well expose unexpected interactions between factors under certain conditions. The experiments are, therefore, designed based on the experience of those fields of science that habitually deal with complexity and uncertainty in their experiments: those of the life sciences (Cox, 1992), (Underwood, 1997), and those associated with quality assurance (QA) (Davis and Grove, 1992) (Taguchi, 1988). Those who practice in these fields have developed a body of knowledge that enables the relationships between significant numbers of factors to be established in the presence of unknown, and possibly many, uncontrolled or uncontrollable factors, together with unknown interactions between the controlled factors.

The experiment under consideration here has similar characteristics to those designed for QA and life sciences in that it involves many factors at many potential levels. However, it differs in some key aspects. The characteristics of the experiments for which designs are proposed in the literature were developed for two main classes of work:

- Biological or agricultural experiments, where the aim is to quantify repeatability. The number of controlled factors is chosen to be low, but the number of uncontrolled factors may be high. Interaction between factors is unknown and may be significant. The time required to produce data can be long, but large numbers of data samples may be generated relatively cheaply.
- Experiments to improve the quality of mass-produced components where the aim is to identify the causes of critical differences in performance between nominally identical items. Again the number of uncontrolled factors can be high, but the number of samples taken can be large.

The experiment considered here, by contrast, is to be devised as a means of exploring how the performance of a nominally consistent design changes as the detail of the design varies, both with time, and between missions. To capture the full extent of possible changes to detail, the number of factors is necessarily large and the number of potential levels is high. The number of samples compared with the extent of the interaction surface being explored is, therefore, likely to be significantly smaller than would be expected in the other disciplines. On the other hand it should be possible, by careful design of the apparatus, to ensure that the statistical variability between measurements is comparatively small. Additionally it should be possible to characterise the residual noise so that it can be identified as such.

3 One at a time or all at once?

The instinctive reaction to the need for any scientific exploration is to maintain maximum control by changing, so far as possible, only one parameter at a time. This is done on the assumption that the accuracy of the measurement may otherwise be degraded by unforeseen interference between factors. However, this approach is very expensive in terms of the number of measurements required and enables only a limited amount of information to be extracted from the data. For example, if in an imaginary experiment, three non-interacting factors, A, B and C, may exist in either of two states, 1 or 2, then the minimum set of experiment that measures the effect of all combinations of levels and factors by varying one at a time is illustrated in Table 1.

Experiment	Factor		
	A	B	C
1	1	1	1
2	1	2	1
3	1	2	2
4	1	1	2
5	2	1	1
6	2	2	1
7	2	2	2
8	2	1	2

Table 1 Varying one factor at a time

Thus, the minimum number of experiments, n_{min}, for f factors each of which can take l levels is:

$$n_{min} = l^f$$

Clearly the number of experiments increases very rapidly as the number of factors and levels grow. Thus, 5 factors, each of which can take 3 levels, would require 3^5 or 243 experiments. And this number assumes that the measurements made are perfect. If measurement error is suspected then a number of measurements would be required for each set of circumstances so that some form of averaging could take place. Under these circumstances the number of experiments required would increase pro rata.

However, if we make two bold assumptions then considerable economies in terms of the number of measurements required becomes possible. The assumptions are:

- That none of the factors selected for the experiment interact i.e. their effects are additive.
- That measurements are error free.

All that is necessary now is to make one measurement for each factor at each level, with the other factors held constant. The number of experiments required then reduces from that shown in Table 1 to that shown in Table 2. Notional results have been added to Table 2 to illustrate calculation of the effects.

		Factor		Results
Experiment	A	B	C	R
1	1	1	1	10
2	1	2	1	12
3	1	2	2	15
5	2	1	1	19

Table 2 Minimum number of experiments

The minimum number of experiments now only exceeds the number of factors by one. The effect of changing factor A from level 1 to level 2 is the difference between the results of those two experiments where factor A changes but both of the other two factors remain constant. In this case the difference is provided from the results for experiments 1 and 5. Similarly the effect of changing factor B is the difference between the results for experiments 1 and 2, and that for C is the difference between results 2 and 3.

4 Measurement Error or Noise

Now re-introduce the possibility of measurement error. The error will have some form of distribution. For the sake of argument, assume that the distribution is normal, (although the logic is equally valid for other forms of distribution). Then the distribution of the total population, assuming that an infinite number of measurements are made, will be defined by the population mean and the variance. The greater the number of sample measurements we obtain, the greater the confidence we will have that the mean and variance of the sample represents the actual population. But in the previous example (Table 2) we already have 3 measurements for factor A at level 1, albeit with the other two factors at varying levels. Ignoring for the moment that the other two factors vary, we could subtract the mean figure obtained for the 3 measurements at level 1 from the single measurement at level 2. However, because we have fewer measurements for level 2 we would have less confidence in that figure and it would be difficult to ascribe an overall confidence to the magnitude of the net effect. This could be corrected by introducing two more measurements at level 2. We could then compare the mean of the measurements with the level set at 1, with the mean of the measurements set at level 2, and quantify the confidence with which we express the result.

Now let us address the little difficulty of having the other factors varying between levels. Provided that we retain our assumption that there are no interactions between the factors, then, if we ensure that the other factors change between their levels an equal number of times, their effects on the factor in which we are interested will cancel out. An example is given in Table 3.

This time the experiments have been designed to meet our two conditions. Thus, factors A, B and C each occur twice at each level and, for example, when factor A is at level 1, factors B and C each change between levels 1 and 2.

The effect of each factor is the difference between the mean results obtained when they are set at each level. Thus, for experiment X the effect of changing factor B from level 1 to 2 is:

$$E_B = \frac{R3 + R4}{2} - \frac{R1 + R2}{2} = \frac{9 + 14}{2} - \frac{6 + 8}{2} = 4.5 \cdot$$

Experiment X	Factor A	Factor B	Factor C	Results R
1	1	1	2	6
2	2	1	1	8
3	1	2	1	9
4	2	2	2	14
Mean level 1	7.5	7	8.5	
mean level 2	11	11.5	10	
Effect	3.5	4.5	1.5	

Experiment Y	Factor A	Factor B	Factor C	Results R
1	1	1	2	6
2	2	1	1	18
3	1	2	1	9
4	2	2	2	24
Mean level 1	7.5	12	13.5	
mean level 2	21	16.5	15	
Effect	13.5	4.5	1.5	

Table 3 Effect of factors varying between levels

Now let us suppose that the effect of factor A is increased by 10. Thus, the results of factor A at level 2 in experiment Y are increased to 18 and 24. However, provided the experiment is balanced, the effect of the change in effect of factor A does not change the measured effect of factor B. Thus, for experiment Y:

$$E_B = \frac{9+24}{2} - \frac{6+18}{2} = 4.5 \ .$$

Hence, provided that the experiment is consistent with the assumption, E_B is unchanged. Equally, the effect of factor C is unchanged.

The independence of factors is analogous to functions in mathematics being independent. Under such circumstances the functions are described as being orthogonal. This analogy leads to the practice of describing the experimental factor arrays as orthogonal.

5 Balanced experiments and orthogonal arrays

The illustration just provided indicates that a balanced experiment may be defined as one that satisfies the following two conditions:

- Each level for each factor is measured an equal number of times.
- For each measurement at one level the other factors are represented an equal number of times.

Arranging the factors and levels in the form of an orthogonal array produces by definition a balanced experimental programme. Arrays for different combinations of factors and levels that satisfy these conditions have been published by a number of authors (Davis and Grove, 1992, Taguchi, 1988, Underwood, 1997, Cox, 1992). They have been formulated to enable balanced factorial experiments to be designed. Thus, assuming that any noise (measurement error) is randomly (normally) distributed, the standard statistical descriptors of mean, variance, etc, can be applied to provide a sound representation of the signal level. Standard regression techniques may then be applied to the results to fit a response surface to the levels and factors.

Orthogonal arrays are numbered according to the number of experiments to be performed. Each row of the matrix defines a combination of levels for the factors. The orthogonal array for the balanced experiment in this case is designated an $L_8 2^7$ array (Taguchi, 1988) (Table 4). It allows for 8 experiments to identify a maximum of 7 effects, each with a maximum of 2 levels.

Factor / Experiment	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	2	1	2	1	2	1	2
3	1	2	2	1	1	2	2
4	2	2	1	1	2	2	1
5	1	1	1	2	2	2	2
6	2	1	2	2	1	2	1
7	1	2	2	2	2	1	1
8	2	2	1	2	1	1	2

Table 4 $L_8 2^7$ array

6 Interactions between factors

The argument to date has assumed that there are no interactions between factors and that their effects are additive. However, in the experiment under consideration it is quite likely that items of detail on the hull will interact, especially when they are placed in close proximity. Any interaction will reveal itself as an apparent measurement error. It will, therefore, be important to be able to eliminate, or where this is not possible, quantify as many of the noise sources as possible so that unanticipated interactions can be identified and explored. The subject of noise characterisation is addressed in (Fallows, 2005).

It should be possible to quantify the effect of interactions a priori provided that: the interaction is anticipated; meets certain conditions; and is allowed for in the experiment from the beginning. In particular, if it is assumed that the effect of the interaction is itself additive, the interaction can be treated as if it is an independent factor. In this case, the method outlined previously is used, with the interaction allocated to another column in the orthogonal array. The net effect is then the sum of the individual effects and the interactions. The disadvantage is that to identify interactions the number of measurements must increase. In the case under consideration we will, therefore, adopt the strategy of placing considerable emphasis on reducing and quantifying noise, and using any increase in the overall noise level to indicate significant interactions. Once interactions have been identified, the method for assessing the effects of interactions is as follows.

Consider the $L_8 2^7$ array at Table 4. When originally presented, one factor was allocated to each column subject to there being a maximum of 7 factors, each of which can occupy 2 levels, from a minimum of 8 runs. But as observed above, not all effects need be due to independent factors. Columns can be allocated to interactions.

To illustrate how interactions are included in an orthogonal array, suppose that we revert to 3 factors and allocate the first two factors, A and B, to columns 1 and 2. On the assumption that the interaction is reversible, then two levels of interaction are possible. Moving from both factors at level 1 to both at level 2 will be the same as moving from both at level 2 to both at level 1. Call this interaction level 1. Similarly holding A constant and changing the level of B will produce an interaction identical to that of holding B constant and changing A. Call this interaction level 2. We can now apply this convention to the array to produce the interaction column A x B, as shown in Table 5.

Factor / Experiment	A	B	AxB
1	1	1	1
2	2	1	2
3	1	2	2
4	2	2	1
5	1	1	1
6	2	1	2
7	1	2	2
8	2	2	1

Table 5 Single interaction

This just happens to generate a column identical to the third column of the $L_8 2^7$ array at Table 4.

If we now allocate factor C to column 4 of the $L_8 2^7$ array we find that columns 5 and 6 provide the interactions between A & C and B & C as shown in Table 6. Applying the same rule to any of the factors, A, B or C, with the interaction columns, A x B, B x C or A x C, gives the third order interaction A x B x C in column 7. The notation A x B, A x C, B x C and A x B x C is quite deliberate and reflects the fact that column A x B equates to the product of the contents of column A and column B.

Factor / Experiment	A	B	AxB	C	AxC	BxC	AxBxC
1	1	1	1	1	1	1	1
2	2	1	2	1	2	1	2
3	1	2	2	1	1	2	2
4	2	2	1	1	2	2	1
5	1	1	1	2	2	2	2
6	2	1	2	2	1	2	1
7	1	2	2	2	2	1	1
8	2	2	1	2	1	1	2

Table 6 Multiple interactions

It should be noted that if we wish to measure all 4 possible interactions as well as the principal effects of 3 factors, then the number of runs required increases from 4 to 8 (compare Tables 3 and 6). The total experiment then becomes as shown in Table 7, with the analysis being performed as in Table 6.

Factor / Experiment	A	B	C
1	1	1	1
2	2	1	1
3	1	2	1
4	2	2	1
5	1	1	2
6	2	1	2
7	1	2	2
8	2	2	2

Table 7 Experiment design enabling interaction measurement

From Chapter 1.7 it is clear that the number of factors and levels required for this investigation are such that it is highly desirable to limit the number of interactions to be investigated, and hence the number of additional runs required, to a minimum. Reliance will be made on detecting unsuspected interactions by detecting increase in signal noise. For this to be possible the background noise will need to be kept very low. Thus, the experiment will need to be designed such that the effect of uncontrolled factors is far less significant than is the norm for those designed for QA and Life Sciences. This will necessitate, so far as possible, control of the environment under which the experiment is conducted. Where this is not possible the relevant factors will need to be monitored. Additionally the experimental facilities will need to be characterised in detail, so that the remaining sources of extraneous signals may be identified and quantified in advance. Finally likely interactions between the factors will need to be identified in advance so that they can be catered for in the design and not appear as additional noise.

7 Fractional Experiments

Fortunately, it is not necessary to explore all interactions. A mixture of purely main effect and interactions can be designed into an experiment or, alternatively, an experiment may be designed to enable the extraction of the maximum amount of information from a limited number of measurements. However, the number of independent measurements made does limit the number of contrasts that can be derived. Thus, if only n independent measurements can be taken, then only a maximum of (n-1) contrasts can be measured. Any interactions not allowed for in the experiment will appear as an addition to a main effect.

Consider the plan of Table 6. An additional factor D can be introduced in lieu of the second order interaction A x B x C. We now have an increase in the number of possible interactions, with A x D, B x D, etc, as well as additional 3-way interactions and now a 4-way interaction. But the maximum number of contrasts that can be measured remains at 7. These contrasts are said to be 'confounded' by the unmeasured interactions. The new plan is shown at Table 8.

Factor / Experiment	A	B	C	D
1	1	1	1	1
2	2	1	1	2
3	1	2	1	2
4	2	2	1	1
5	1	1	2	2
6	2	1	2	1
7	1	2	2	1
8	2	2	2	2

Table 8 Fractional Experiment

Using the formula already established for identifying the magnitude of interactions it can be seen that, for experiment 1, the interaction of C x D (both the same and, therefore, level 1) and B with C x D (both level 1) results in a net interaction B x C x D at level 1. This happens to equate to the level in the column for factor A. This observation is true for every row of the factor A column and indicates that column 1 not only measures the main effect of factor A, but also the B x C x D interaction. Applying this process to the remainder of the array results in the so called confounding pattern shown in Table 9 (Davis and Grove, 1992).

Factor / Experiment	A +BxCxD	B +AxCxD	AxB +CxD	C +AxBxC	AxC +BxD	BxC +AxD	D +AxBxD
1	1	1	1	1	1	1	1
2	2	1	2	1	2	1	2
3	1	2	2	1	1	2	2
4	2	2	1	1	2	2	1
5	1	1	1	2	2	2	2
6	2	1	2	2	1	2	1
7	1	2	2	2	2	1	1
8	2	2	1	2	1	1	2

Table 9 Confounding pattern

This, of course, does not matter if the effects of the confounding interactions are small.

8 Higher level experiments

The argument to this point has been limited to experiments where each of the factors varies between only 2 levels. Clearly this can only give an accurate indication of the relationship between factors when that relationship is linear. If the relationship is expected to be quadratic, then a minimum of three levels is required. Where greater confidence in the nature of the curve is required, or where higher order relationships are suspected, then still more levels are required. Using similar arguments to those outlined above, orthogonal arrays can be constructed to produce balanced experiments at higher levels, as demonstrated in the example of a 4 factor, 3-level array in Table 10.

Factor / Experiment	A	B	C	D
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Table 10 $L_9 3^4$ orthogonal array

9 Mixed level Experiments

For many experiments it is necessary to explore some factors at higher levels, but not all. In this case an orthogonal array needs to be selected such that it is designed for the appropriate number of factors and the predominant number of levels to be used. Thus, for a 4 factor experiment where only one factor, A, is to be set at 2 levels and 3 factors, B ,C and D, at 3 levels, then the $L_9 3^4$ orthogonal array shown in Table 10 may be chosen. Factor A would then be allocated to column 1 with levels 1 and 2 corresponding to that in the array, but where level 3 occurs in the array, either level 1 or level 2 should be allocated. In effect we are treating factor A as if it was a 3 level factor with the levels being either 1,1,2 or 1,2,2. This mixed level experiment would then appear as in Table 11.

Factor / Experiment	A	B	C	D
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	1	1	3	2
8	1	2	1	3
9	1	3	2	1

Table 11 The $L_9 2^1.3^3$ mixed level experiment

Factor / Experiment	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	2	1	2	1	2	1	2
3	1	2	2	1	1	2	2
4	2	2	1	1	2	2	1
5	1	1	1	2	2	2	2
6	2	1	2	2	1	2	1
7	1	2	2	2	2	1	1
8	2	2	1	2	1	1	2

A	B		4 level factor		3 level factor
1	1		1		1
2	1		2		2
1	2	=	3	=	3
2	2		4		3
1	1		1		1
2	1		2		2
1	2		3		3
2	2		4		3

Factor / Experiment	A	C	D	E	F	G
1	1	1	1	1	1	1
2	2	2	1	2	1	2
3	3	2	1	1	2	2
4	3	1	1	2	2	1
5	1	1	2	2	2	2
6	2	2	2	1	2	1
7	3	2	2	2	1	1
8	3	1	2	1	1	2

Table 12 Mixed level experiment

If, on the other hand, the predominant number of factors is to be set at the level 2, then the 3 level factor can be allocated to a 2 level array by combining 2 columns to produce a 4 level factor and then reducing this to a 3 level column as described above. This is illustrated in Table 12 (Davis and Grove, 1992).

10 Partitioning

Even allowing for the reduction in number of runs possible by rigorously reducing the effect of noise, reducing the number of levels to the minimum and judicious selection of the array to minimise the confounding effects of interactions, the size of the space to be explored may remain unaffordable. Under these circumstances it is necessary to break the space down into discrete blocks. This enables the factors and levels to be tailored to a particular region of the experimental space and allows the number of levels and factors to be kept within manageable bounds.

11 Summary of the design process

The minimum number of factors is chosen to describe the response sought. Each is measured for the minimum number of levels consistent with describing the response surface to the accuracy required. This is determined by the need to detect the shape of the surface and identify the inflection points. Where interactions between factors are suspected to be significant these are treated in the analysis as if they were separate factors. An orthogonal array is selected for the number of contrasts required and the number of levels chosen. Analysis of the arrays assumes orthogonality, i.e. that there is no interaction between the factors selected and their effects may, therefore, be added. The effect on the overall response of a change in level of any factor is then the difference in the mean of all of the responses to that factor as it changes between levels. The variance in response to any factor provides an indication of the error, or noise, in the measurement. Analysis of variance (ANOVA) will then enable quantification of the significance of the response. The underlying statistics is explained in (Fallows, 2005).

There are two potential sources of error: that due to noise in the measurement system; and that due to interactions between factors. The noise due to interactions can be detected by treating each suspected interaction as an independent factor within the array. This is done by means of allocating an interaction to the column that comprises the product of the factors that interact. Unanticipated interactions may not be easily distinguished from measurement noise: hence the emphasis in these experiments on identifying and quantifying the performance of the measurement system (Chapter 2.4).

When designing experiments it is generally better to establish the lower order interactions first, i.e. those that involve only 2 interactions, since higher order interactions are likely to be less significant. It is also desirable to limit the number of factors investigated in any one experiment since, as the number of factors grow, so do the effects of confounding. For an exploration of a phenomenon dependent on a large number of factors it may be necessary to partition the experiment to keep the number of factors within bounds.

The number of levels chosen for any factor should reflect the form of the anticipated response to it. Thus, if a linear response is expected, two levels may be sufficient. However, if a higher order response is expected then a greater number of levels will be needed if the inflection points are to be detected. For example, in the case of AUTOSUB, analysis of data collected during missions provides the probability of the vehicle adopting a particular stance. The experiment can then be designed to ensure that the most accurate description of the response curve occurs in the region most likely to occur in practice.

12 Summary of the model implied by the method.

The following model underlies the design and analysis:

- Measurement error, or noise, is normally distributed.
- Contrasts are additive
- Factors are orthogonal or, where not, then the interaction between them may be treated as a separate orthogonal factor.
- Each measurement of the effect of a combination of factors and levels provides an independent sample of the response surface.

Conclusions

This Report has addressed how complexity may be handled in an affordable programme of experiments. It concludes that, for the full complexity to be explored in a realistic timescale, three conditions must be met. Firstly the experiments require careful design so that the maximum amount of information may be derived from a realisable number of runs. These are achieved by using balanced fractional designs based on orthogonal arrays. Experiments to determine the drag effects of the hull and appendages of the AUV, needs to be carefully designed. To enable detection of the effects of unanticipated interactions between factors, the apparatus needs to be carefully designed so as to minimise the signal noise. The residual noise in the measurement system must be carefully characterised so that it can be distinguished from any anomalies caused by unexpected interactions between apparently controlled factors.

References

Cox, D. R. (1992), '*Planning of experiments*', John Wiley and Sons Inc.

Davis, T. P. and Grove, D. M. (1992), '*Engineering quality and experimental design*', Longman Scientific & Technical, Harlow.

Fallows, C.D. (2005), '*AUTOSUB Propulsion system investigation - Characterisation of the force measuremen systemt*', University of Southampton, University of Southampton, School of Engineering Sciences, Ship Science, Report 133, pp8.

Fallows, C.D. (2005), '*AUTOSUB propulsion system investigation – Supplementary information*', University of Southampton, School of Engineering Science, Ship Science, Report 135, pp 8.

Taguchi, G. (1988), '*System of experimental design - Volume 1&2*', Kraus International Publications, New York.

Underwood, A. J. (1997), '*Experiments in ecology*', Cambridge University Press.