EVOLUTIONARY OPERATION:*
A METHOD FOR INCREASING INDUSTRIAL PRODUCTIVITY

GEORGE E. P. BOX†
Imperial Chemical Industries Limited
Dyestuffs Division Headquarters, Manchester

The rate at which industrial processes are improved is limited by the present shortage of technical personnel. Dr Box describes a method of process improvement which supplements the more orthodox studies and is run in the normal course of production by plant personnel themselves. The basic philosophy is introduced that industrial processes should be run so as to generate not only product, but also information on how the product can be improved.

Introduction

Much scientific effort in industry is directed on the one hand to the discovery of new products and processes and on the other to their development and improvement. This paper is concerned with a particular aspect of the problem of *improving* industrial processes.

Industrial organisations usually have specialist groups of scientific workers in research, development, and experimental departments, permanently occupied with improving manufacturing processes, who employ a wide variety of techniques, ranging from the fundamental study of reaction mechanisms to the purely empirical assessment of the effects of changes in variables. Associated experimentation may be conducted in the laboratory, on the pilot plant, and on the full scale process, and in particular may involve the use of statistical techniques having a fairly high degree of sophistication.[1–7] As a result of the application of this variety of specialised effort a steady rate of increase in productivity is usually attained.

Ultimately the rate of improvement is limited by the shortage of technical personnel. This shortage can be expected to become more rather than less severe, and in searching for further ways of attaining greater process efficiency one must look for methods which are sparing in their use of scientific manpower. The object of this paper is to outline one such device, which has been applied with considerable success over the past few years.

This is called 'Evolutionary Operation'. It is a method of process operation which has a 'built-in' procedure to increase productivity. It uses some simple statistical ideas and is run during normal routine

* Based on a paper given at the International Conference on Statistical Quality Control organised by the European Productivity Agency of OEEC in Paris in July 1955.

† Now Director of the Statistical Techniques Group at Princeton University.

production largely by plant personnel themselves. Its basic philosophy is that it is nearly always inefficient to run an industrial process to produce product alone. A process should be run so as to generate product *plus information on how to improve the product.*

The technique is in no sense a substitute for the more fundamental investigations referred to above. On the contrary, the effects discovered by the application of evolutionary operation, particularly those which are of an unexpected kind, help to indicate new areas where fundamental research might be rewarding. Although the method has been specifically developed as a production technique for the chemical industry, it is believed that it has more general applicability.

Plant-scale and Small-scale Experiments

In the chemical industry the plant process will usually have been arrived at after considerable experimentation on the small scale. Now the optimum conditions of operation on the small scale usually provide no more than a good first approximation to the full-scale optimum. Because of this it is commonly found that considerable modification of the conditions arrived at from small-scale work is necessary before a comparable result can be obtained on the plant itself.

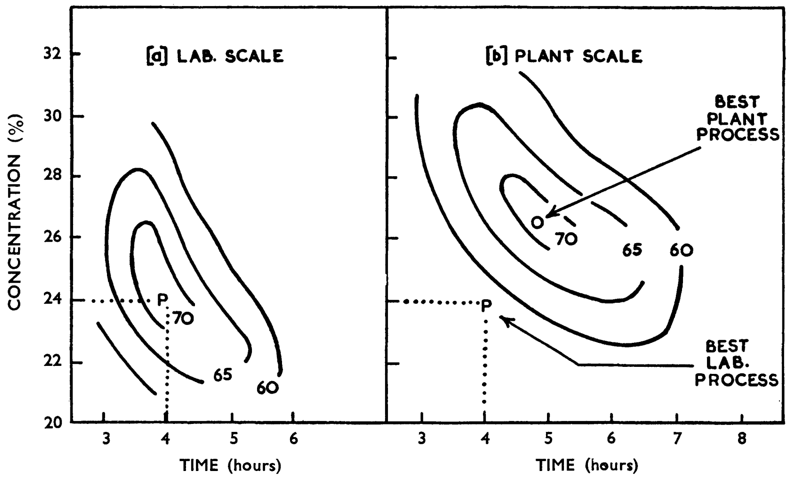


FIG. 1. Possible appearance of yield surfaces, showing contours of percentage yield, for a process conducted (*a*) on the laboratory scale and (*b*) on the plant scale.

Fig. 1 shows how, in translating the process from the laboratory to the plant, the yield surface can become distorted and displaced due to scale-up effects. It will be seen that when this happens one might expect the best laboratory conditions to give disappointingly low yields on the full scale. The effort to move from the laboratory maximum-yield conditions at P to the plant maximum-yield conditions at O must evidently be exerted *on the plant itself* and not in the laboratory—investigation in the laboratory can only lead back to P.

For those unfamiliar with the representation adopted in Fig. 1 it should be explained that the relationship between the response ('yield') and the process variables ('time of reaction' and 'concentration of one of the reactants') is imagined to be represented by a solid graph or 'response surface'. In the neighbourhood of maximum yield such a surface may have the appearance of a mound. The height of the mound at any particular point represents the yield at some set of reaction conditions. To allow representation in two dimensions the yield is shown by contours in the same way that the height of land is represented by contours on a map. Although for simplicity the above discussion is conducted in terms of yield, it should be understood that conditions giving highest yield would often not represent the optimum process. There would, for example, be no advantage in obtaining a higher yield if to do so involved the use of a disproportionate amount of some expensive starting material. The principal response usually considered, therefore, is 'the cost of producing unit quantity of product under the specified manufacturing conditions', or some other measure of productivity which takes account of the cost of running the process.

In addition to improvements made possible by adjustment of process conditions already studied on the small scale, further progress is usually possible by the introduction of new modifications not considered—and often not capable of being studied—at the small-scale stage of development.

Adjustments are made when the plant is first installed, but these seldom result in the location of the ultimate plant optimum, and as a result of special experimental campaigns, chance discoveries, and new ideas, improvement usually continues over many years. The object of evolutionary operation is to speed up this process.

Analogy with Evolutionary Process

The method used to speed the improvement is illustrated by the following analogy. Living things advance by means of two mechanisms:

(i) Genetic variability due to various agencies such as mutation.

(ii) Natural selection.

Chemical processes advance in a similar manner. Discovery of a new route for manufacture corresponds to a mutation. Adjustment of the process variables to their best levels, once the route is agreed, involves a process of natural selection in which unpromising combinations of the levels of process variables are neglected in favour of promising ones.

Fig. 2 illustrates diagrammatically the possible evolution of a species of lobster. It is supposed that a particular mutation produces a type of lobster with 'length of claws' and 'pressure attainable between claws' corresponding to the point P on the diagram and that in a given environment the contours of 'percentage surviving long enough to reproduce' are like those shown in the figure.

The dots around P indicate offspring produced by the initial type of lobster. Since those in the direction of the arrow have the greatest chance of survival, over a period of time the scatter of points representing succeeding generations of lobsters will automatically move up the survival surface. This automatic process of natural selection ensures, without any special effort on the part of the lobsters, that optimum-type lobsters exist. It also ensures that if the environment changes so that the survival surface is altered, the lobsters will change correspondingly to the new point of maximum survival.

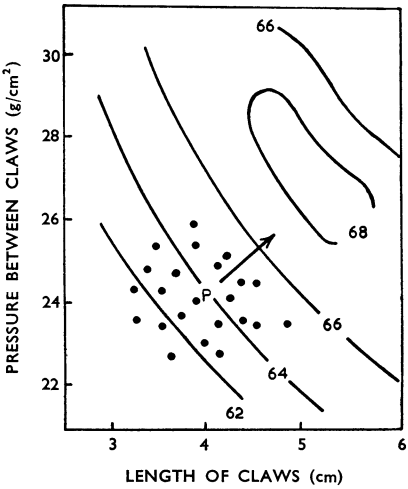


Fig. 2. Evolution of a species of lobster. Contours show the percentage surviving long enough to reproduce in a given environment.

What we have to do is to imitate this process. That is to say, we have to institute a set of rules for *normal plant operation* so that (without serious danger of loss through manufacture of unsatisfactory material) an evolutionary force is at work which steadily and automatically moves the process towards its optimum conditions if it is not operating there already. Such a technique will gradually nudge the operating procedure into the form that is ideally suited to the particular piece of equipment which happens to be available. The two essential features of the evolutionary process are:

(i) *Variation.*

(ii) *Selection* of 'favourable' variants.

Static and Evolutionary Operation

Routine production is normally conducted by running the plant at rigidly defined operating conditions called the 'works process'. The works process embodies the best conditions of operation known at the time. The manufacturing procedure, in which the plant operator aims always to reproduce exactly this same set of conditions, will be called the method of *Static Operation.* Although this method of operation, if strictly adhered to, clearly precludes the possibility of evolutionary development, yet the *objectives* which it sets out to achieve are nevertheless essential to successful manufacture, for in practice we are interested not only in the productivity of the process, but also in the physical properties of the product which is manufactured. These physical properties might fall outside specification limits if arbitrary deviations from the works process were allowed. Our modified method of operation must therefore include safeguards which will ensure that the risk of producing appreciable amounts of material of unsatisfactory quality is acceptably small.

In the method of Evolutionary Operation a carefully planned cycle of minor variants on the works process is agreed. The routine of plant operation then consists of running each of the variants in turn and continually repeating the cycle. The cycle of variants follows a simple pattern, the persistent repetition of which allows evidence concerning the yield and physical properties of the product in the immediate vicinity of the works process to accumulate during routine manufacture. In this way we use routine manufacture to generate not only the product we require but also the information we need to improve it.

Controlled variation having thus been introduced into the manufacture, the effect of selection is introduced by arranging that the results are continuously presented to the plant manager in a way which is easily comprehended. This allows him to see what changes ought to be made to improve manufacture. The stream of information concerning the products from the various manufacturing conditions is summarised on an *Information Board* prominently displayed in the plant manager's office. This is continuously brought up to date by a clerk to whom the duty is specifically assigned. The information is set out in such a way that the plant manager can at any time see what weight of evidence exists for moving the centre of the scheme of variants to some new point, what types of change are undesirable from the standpoint of producing material of inferior quality, how much the scheme is costing to run, and so on.

In making a permanent change in the routine of plant operation the situation is very different from that which we meet in running specialised experiments on the plant. The latter will last a limited time, during which special facilities can be made available. Furthermore some manufacture of substandard material is to be expected and will be budgeted for. Evolutionary operation, however, is virtually a *permanent* method of running the plant and cannot therefore demand special facilities and concessions. For this reason only small changes in the levels of the variables can be permitted, and only techniques simple enough to be run continuously by works personnel themselves under actual conditions of manufacture can be employed.

The effects of the deliberate changes in the variables will usually be masked by large errors customarily found on the full scale. *However, since production will continue anyway, a cycle of variants which does not significantly effect production can be run almost indefinitely, and because of constant repetition the effect of small changes can be detected.*

An Example

To illustrate the procedure we consider one phase of evolutionary operation for a particular batch process. At this stage of development two process factors—the percentage concentration of one of the reactants, and the temperature at which the reaction was conducted—

A*

were being studied following the scheme of variants shown in Fig. 3. The works process is labelled (1) and the four variants are labelled (2), (3), (4), and (5). One batch of product was made at each set of conditions, which were run successively in the order 1, 2, 3, 4, 5; 1, 2, 3, 4, 5; and so on. Three responses were recorded:

(i) The cost of manufacturing unit weight of product. This was obtained by dividing 'the cost of running at the specified conditions' by 'the observed weight yield at those conditions'. It was desired to bring this cost to the smallest value possible, subject to certain restrictions listed in (ii) and (iii) below.

(ii) The percentage of a certain impurity. It was desired that this should not exceed 0·5.

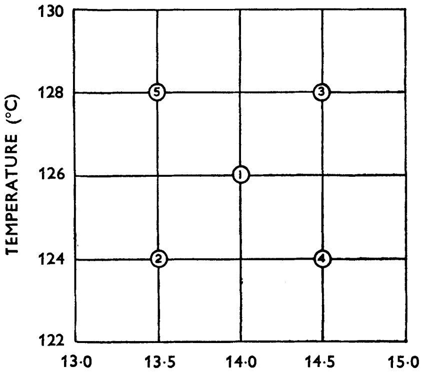


FIG. 3. A cycle of variants about the works process.

(iii) A measure of fluidity. It was preferred that this should lie between the limits 55 and 80.

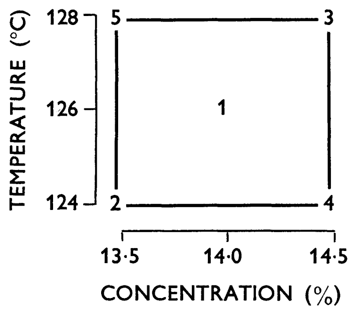
The information coming from the experiment was recorded by writing in chalk on an ordinary blackboard. Alternatively, wax pencil on a white plastic board or magnetic letters and numbers on a steel board could have been used. The essential thing is that it should be a simple matter to erase or remove one number and replace it by another. The scheme set out in Fig. 4 is not the only one which could have been adopted, but is intended to show the sort of calculations and layout of the results which have been found useful.

The phase number at the top left-hand corner of the board indicates that two previous phases of evolutionary operation have already been completed on this process. In general these might have involved other variables or the same variables at other levels. In order that the new results may be considered in proper relation to those obtained previously, the final average values recorded in previous phases should also be available (for example on sheets of paper pinned to the board). The cycle number in the top right-hand corner indicates that 16 cycles of this third phase of operation have been completed. There follows a plan of the cycle of variants being run.

The table below this summarises the current situation. First are shown the requirements which it is desired to satisfy. These are followed by the running (i.e. up-to-date) averages at the various manufacturing conditions set out so as to follow the plan of the cycle of variants. This arrangement makes it easy to appreciate the general implications of the results. A measure of the reliability of the individual running averages is supplied by the '95% error limits'. These are simply the quantities

$\pm ts/\sqrt{n}$ appropriate to the calculation of fixed sample-size confidence limits for individual means. In this expression s is the standard deviation, n the number of cycles, and t the appropriate significance point of Student's t distribution. The scheme can be run until these limits for the means of the principal response have been reduced to an acceptably small preassigned width.[8]

PHASE 3 LAST CYCLE COMPLETED 16



		Cost	Impurity (%)	Fluidity
Requirement		Minimum	Less than 0·50	Between 55 and 80
Running Averages		32·6 33·9 32·8 32·3 33·4	0·29 0·35 0·27 0·17 0·19	73·2 76·2 71·3 60·2 67·6
95% Error Limits		±0·7	±0·03	±1·1
Effects with 95% Error Limits	Conc.	1·2 ± 0·7	0·04 ± 0·03	5·2 ± 1·1
	Temp.	0·4 ± 0·7	0·14 ± 0·03	10·8 ± 1·1
	C × T	0·1 ± 0·7	0·02 ± 0·03	−2·2 ± 1·1
	Change in Mean	0·2 ± 0·6	−0·02 ± 0·03	−1·6 ± 1·0
Standard Deviation 95% Error Limits		1·44 1·22 1·76	0·059 0·050 0·072	2·12 1·80 2·59
Prior Estimate		2·71	0·054	3·22

FIG. 4. Appearance of information board at the end of cycle 16.

Below this are shown the 'effects' of the variables and their 95% error limits. The concentration and temperature effects are each calculated in the usual way as the difference in the average values of the response at the higher and the lower level of the variable. The value at the centre conditions does not enter the calculations except in computing the 'change in mean' effect. If y_1, y_2, y_3, y_4, and y_5 are the running averages of one of the responses after n cycles of operation, the effects and their limits of error for this particular example are as follows:

Effect		*Value*	*Limits of Error*
Concentration	..	$(y_3 + y_4 - y_2 - y_5)/2$	$\pm ts/\sqrt{n}$
Temperature	..	$(y_3 + y_5 - y_2 - y_4)/2$	$\pm ts/\sqrt{n}$
Interaction	..	$(y_2 + y_3 - y_4 - y_5)/2$	$\pm ts/\sqrt{n}$
Change in mean	..	$(y_2 + y_3 + y_4 + y_5 - 4y_1)/5$	$\pm 2ts/\sqrt{(5n)}$

The effect referred to as the 'change in mean' is simply the grand mean for all the runs $(y_1 + y_2 + y_3 + y_4 + y_5)/5$ less the mean for the 'standard' conditions y_1. It is therefore an estimate of the difference in average response resulting from the use of the evolutionary scheme. In the present example the effect of introducing the evolutionary scheme is thus:

(i) To increase the cost per batch by $0{\cdot}2 \pm 0{\cdot}6$ units.

(ii) To reduce the average impurity by $0{\cdot}02 \pm 0{\cdot}03\%$.

(iii) To reduce the average fluidity by $1{\cdot}6 \pm 1{\cdot}0$ units.

(It will be seen that if the effect of blending batches of slightly different qualities was to average the physical properties, the 'change in mean' would measure the difference between the product of the works process and a blend of the products from evolutionary operation. In many actual examples partial or complete blending of the products does naturally occur at the later stages of manufacture, so that where physical properties behave approximately additively there may in fact be remarkably little overall change in the manufactured product due to evolutionary operation. In some cases, especially if the effect of introducing small variations in the levels of the variables was unexpectedly large, blending could be deliberately introduced to produce an acceptable product.)

If the variants cover a region of the response surface which is a sloping plane, then it is not difficult to see that the true 'change in mean' will be zero. For a convex surface such as that near a maximum it will be negative, whereas for a concave surface such as that near a minimum it will be positive. It can be shown that, on certain plausible assumptions, the 'change in mean' is proportional to the sum of the quadratic constants which measure curvature in the directions of the variables. When the interaction effect and the change in mean effect are not small compared with the single effect of the variables, this indicates that a maximum or minimum is being approached and for exact location and exploration a technique fully set out elsewhere[4–6] is adopted.

For the cost response the change in mean supplies a continuous measure of the cost of obtaining information by the process of evolutionary operation. In the present example this is estimated at $0{\cdot}2 \pm 0{\cdot}6$ units of cost per batch. If the cost surface were locally planar, evolutionary operation would cost nothing. In practice, concavity of the cost surface is to be expected, since a minimum is usually being approached, so that there will usually be a small cost associated with running the evolutionary process. Except when the process has been brought very close to its ultimate optimum this cost will be redeemed many times over by the value of the permanent process improvements that occur from time to time as a result of the information generated by the evolutionary scheme.

After the calculated effects, the experimental error standard deviations calculated from the observations themselves are shown. Except in the initial stages of the scheme these are used in computing the limits of error for the running means and for the 'effects' of the variables. The normal theory fixed sample-size 95% confidence limits for these estimates of the standard deviations are also shown. The final items are the estimates obtained from prior data used in initiating the scheme.

In general, by inspecting the results set out on the information board in the light of the requirements it is desired to satisfy and from expert knowledge of other factors which affect plant operation, the plant manager decides at any particular stage whether

(*a*) to wait for further information;

(*b*) to modify operation.

Under (*b*) some of the alternatives open to him are:

(i) To adopt one of the variants as the new 'works process' and to recommence the cycle about this new centre point.

(ii) To explore an indicated favourable direction of advance and recommence the cycle around the best conditions found. (This exploration may be done, for example, by making a series of tentative advances in the indicated direction, at each stage running the new conditions and the previous best conditions alternately, until sufficient evidence has been gathered.)

(iii) To change the pattern of variants to one in which the levels are more widely spaced.

(iv) To substitute new variables for one or more of the old variables.

In the example actually discussed it is seen that a decrease in concentration would be expected to result in reduced cost, reduced impurity, and reduced fluidity. The effect on cost of a decrease in temperature is uncertain but is more likely to be favourable than not, and would almost certainly result in marked reductions in impurity and fluidity. These facts have to be considered, bearing in mind that a fluidity of less than 55 is undesirable and that although a further large reduction in impurity is welcome it is not necessary in order to meet the specification. It was decided in the event to explore the effects of reducing concentration alone. Phase 3 was terminated, and in the next phase the three processes (13%, 126°), (13·5%, 126°), (14%, 126°) were compared. The first of these gave a mean cost of 32·1 with an impurity of 0·25 and a fluidity of 60·7, and was adopted as a base for further development.

A geometrical display of the results on the information board may also be used when three variables are jointly considered. In this case perspective drawings of the sort illustrated in Fig. 5 are used. The

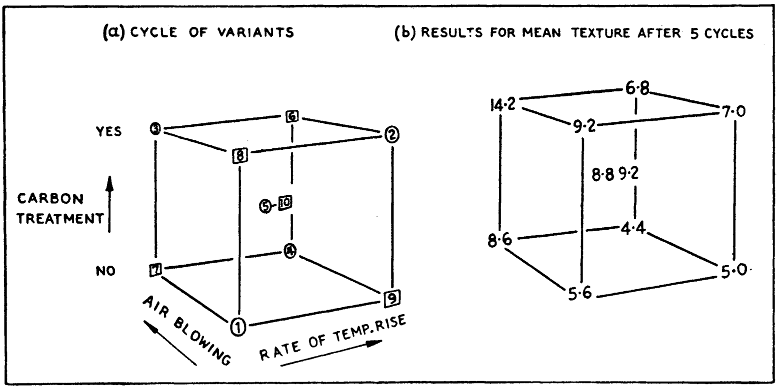


FIG. 5. Display of pattern of variants and results for three variables.

cycle of variants for a three-variable scheme is shown in Fig. 5(*a*), while 5(*b*) shows the results from one of the responses after five cycles of operation. This method may, of course, be applied to responses which cannot be measured directly. In the case illustrated in Fig. 5(*b*) an important property of the product was a somewhat esoteric quality referred to as 'texture'. A set of artificial standard samples were prepared which were judged by experts to have a range of 'textures' in approximately uniform steps, and these were arbitrarily scored. The texture of a sample of each manufactured material was then matched against the standards and an appropriate score given to it. In a similar way a scheme to evolve conditions which would give a product less inclined to 'cake' has been run, in which caking was judged by visual inspection and scored by comparison with a verbally defined scale.

Selection of the Variants

The technique outlined differs from the natural evolutionary process in one vital respect. In nature the variants occur spontaneously, but in our artificial evolutionary process we have to introduce them. Variants involving the levels of temperature, concentration, pressure, etc., are natural choices, but there are usually an almost unlimited number of less obvious ways in which manufacturing procedure can be tentatively modified. Frequent instances of marked improvement due to some innovation never previously considered in a process which has been running for many years testify to the existence of valuable modifications waiting to be thought of.

To make our artificial evolutionary process really effective, therefore, one more circumstance is needed—we must set up a situation in which useful ideas are continually forthcoming. An atmosphere for the generation of such ideas is perhaps best induced by bringing together at suitable intervals a group of people with special, but different, technical backgrounds. In addition to plant personnel themselves, obvious candidates for such a group are, for example, a research man with an intimate knowledge of the chemistry of the similar reactions and a chemical engineer with special knowledge of the type of plant in question. The intention should be to have complementary rather than common disciplines represented.

These people should form the nucleus of a small evolutionary operation committee, meeting perhaps once a month, whose duty it is to help and advise the plant manager in the performance of evolutionary operation. The major task of such a group is to discuss the implications of current results and make suggestions for future phases of operation. Their deliberations will frequently lead to the formulation of theories which in turn suggest new modifications that can be tried with profit.

Since questions of modification of certain physical properties of the manufactured product may arise, a representative of the department responsible for the quality of manufacture should also be on the evolutionary operation committee. Rather more may be got from the results and more ambitious techniques adopted if a statistician is also present at the meetings.

With the establishment of this committee all the requirements for an efficient evolutionary method of production are satisfied and the 'closed loop' illustrated in Fig. 6 is obtained. We are thus provided with a practical method of 'automatic optimisation' which requires no special equipment and which can be applied to almost any manufacturing process, whether the plant concerned is simple or elaborate.

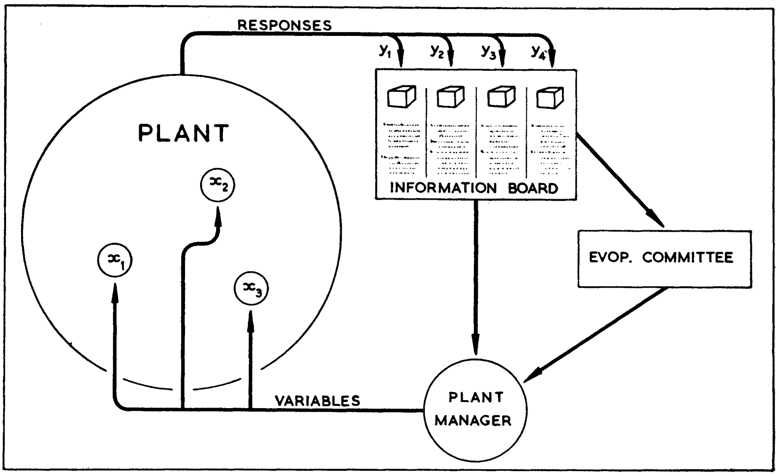


FIG. 6. Diagrammatic representation of the 'closed loop' provided by Evolutionary Operation.

At the beginning of this article I spoke of evolutionary operation as being run largely by plant personnel themselves rather than by specialists. The use of some specialists as advisers on the evolutionary operation committee does not seriously vitiate this principle. In practice, the time spent by the specialists is perhaps one afternoon a month, and the ultimate responsibility for running the scheme still rests with the plant manager and not with the specialists.

When not to Stop

With an alert team of workers new ideas should be continually forthcoming and the evolutionary method becomes virtually a permanent mode of operation and should be so regarded. Only if it seemed that more would be lost than gained from the evolutionary procedure would the reintroduction of static operation be justified. In practice it is found that even very small gains will justify the continual operation of the evolutionary method. The situation at any given time can be appraised by the use of a pictorial log like that shown in Fig. 7.

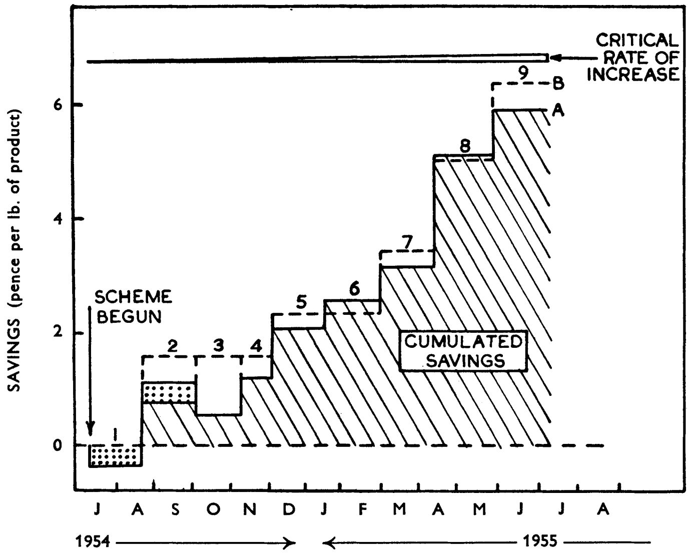


FIG. 7. Progress of an evolutionary scheme, showing critical rate of increase.

The full line A in this diagram shows the savings in pence per lb. which have been achieved in the various phases of operation. The dotted line B shows the savings that would have resulted if the centre or 'works process conditions', appropriate for each particular phase, had been run. On the assumption of constant output the shaded area is proportional to the accumulated savings resulting from the scheme, while each of the rectangular areas between the dotted and full lines shows the accumulated 'expenses' of running the scheme during that phase. The speckled area in phase 1 represents the cumulated expenses for the scheme during this phase and is debited from the cumulated savings in phase 2.

Whereas each phase of evolutionary operation, and consequently the expenses associated with it, lasts for only a limited time, any improvements which result go on for as long as the process is used. Suppose it is assumed that process improvements will go on earning money for p years after they are discovered and that the running of the evolutionary scheme adds c pence per lb. to the cost of the product. Then the question of whether or not, at any instant of time, evolutionary operation should be continued may be resolved by comparing the rate of improvement r which it is expected may be produced by the evolutionary process (measured in pence per lb. per year) with the critical rate of improvement r_0 given by $r_0 = c/p$. For if it is expected that the evolutionary scheme will need to be run for time t years to produce an improvement at the end of that period of rt pence per lb., and if k lb. of product is made per year, then the total saving during the p years for which the discovery is used will be $rtkp$ pence. During this time, kt lb. of product will be made and the loss due to running the evolutionary scheme will be ckt pence. Thus the scheme will pay off if $rtkp$ is greater than ckt, that is if r is greater than $r_0 = c/p$.

As an example consider the situation in Fig. 7. Should the scheme there shown be continued or not? Let us suppose (very conservatively) that improvements on this process are expected to go on earning money for 3 years after their discovery, so that p is put equal to 3. Suppose also that c is taken to be the average of the values experienced in the 9 previous phases. This gives the value $c = 0{\cdot}3$ pence per lb. We then find for the critical rate $r_0 = 0{\cdot}3/3 = 0{\cdot}1$ pence per lb. per year. Thus so long as the rate of improvement due to the evolutionary process is expected to be at least as great as 0·1 pence per lb. per year the evolutionary scheme should be continued. This critical rate of increase is shown diagrammatically at the top of Fig. 7. It will be seen that the actual rate of improvement which had been experienced over the previous year was about 6 pence per lb. per year (about 60 times the critical rate), and there is no evidence as yet of any flagging in this rate of improvement. There is therefore no doubt whatever that this scheme should be continued.

The example given is by no means atypical, which explains my insistence that the evolutionary method should be regarded as virtually a permanent mode of operation. It is psychologically wrong to talk of production under such a scheme as 'experimental manufacture', since an experiment is something which is done for a limited period and is not part of the normal run of things.

Some Questions and Tentative Answers

Like statistical quality control, evolutionary operation is designed for application in the factory itself. Its aims are different and more ambitious than those of quality control, however, since it is directed to ensure not a more uniform product but a cheaper and better product. I believe that this technique, if applied sufficiently widely, can have a marked affect in achieving greater productivity in industry by ensuring that the plant that is available, whether old or new, is operated in the best possible manner. The outline above is intentionally general because it is the general attitude and philosophy that is important here, and not the particular details of application. A full account is nevertheless being prepared in which a number of technical questions are discussed. In the present paper I shall do no more than indicate some of these questions and some tentative replies, which it is hoped to amplify and to justify in the later discussion.

Q. How many variables should be included at one time in an evolutionary scheme?

A. Usually two or three variables can be handled satisfactorily under the normal production conditions with which I am familiar. It should perhaps be emphasised once again that what is being discussed is the normal production situation in which evolutionary operation is applied. For specialist short-period investigations (which, as has been explained in the introduction, are *not* the subject of this account but which nevertheless play an extremely important part in the general scheme of process development to which evolutionary operation is a supplement) the situation is entirely different. In these specialist investigations where, *for a limited period*, it is permissible to interfere with production, and where special supervision and other facilities are made available, the object should be to saturate (or following Satterthwaite[9] possibly 'super-saturate') the experiment with as many factors as possible.

By studying the variables in groups of three or so at a time we forgo the possibility of detecting dependence (interaction) between variables not in the same group. The effect of this limitation should be minimised as much as possible by examining in the same group sets of variables which are expected to be interrelated. Periodically variables whose effects have been found to be important in different phases of operation should be tested together.

Q. What patterns of variants are of most value?

A. A variety of patterns of variants are useful for particular purposes. Among the most valuable for initial use are those based on two-level factorial designs with one or more added points at the centre conditions. These are the arrangements shown for two and three variables in Figs. 3 and 5. They have the advantages that:

(i) They are simple to comprehend, perform and analyse.

(ii) The added centre points allow continual reference to the 'standard' process and permit the 'cost' of the evolutionary scheme to be assessed.

(iii) Complexity of the surface is easily detected by considering the relative magnitudes of the simple 'main effects' on the one hand and the 'change in mean' and interaction effects on the other.

(iv) They can be made the nucleus of more elaborate designs (in particular of composite second-order rotatable designs[10]) by which complexity of the surface may be elucidated.

(v) They lend themselves conveniently to 'blocking arrangements' whereby extraneous disturbances due to such uncontrolled factors as time trends may be reduced. In Fig. 5(*a*), for example, the circles and squares indicate two sub-cycles into which the complete cycle may be divided. A general change in mean occurring between sub-cycles will not bias the estimation of effects.

Q. How should past plant-records be used in planning the evolutionary scheme?

A. When, as is often the case, past plant-records covering long periods of normal operation are available, the planning of an evolutionary scheme should always begin with a careful study of these records. In particular they may be used to determine the approximate magnitude and nature of the uncontrolled variation in the various responses, and consequently the number of repetitions of the cycle likely to be needed to detect effects of a given size.

From these records we can find out whether the errors in the principal response can be regarded as effectively independent and, if not, we can determine the nature of the dependence that exists. This is of some importance in choosing the period for which each variant should be run before changing to the next variant. In practice, of course, this period depends partly on convenience of operation and, for continuous processes, on the time it takes for the plant to settle down after a change in reaction conditions. The time for running each variant which gives the maximum amount of information for a given total period of production can be shown to depend on the nature of the dependence between the observations, and may be determined by a fairly simple use of the correlogram along the lines considered by Jowett[11] or, equivalently, by considering the spectral properties of the record.

Q. Should the variants be run in random order?

A. Faced with the possibility of serial correlation between successive observations, the statistician would normally wish to perform the variants in random order within each cycle, thus guaranteeing the validity of the simple type of analysis used in the example of Fig. 4. However, in some cases, particularly where the time for running each cycle is short, it is much simpler to run a systematic routine of variants on the plant than a random one. In these circumstances randomisation is usually abandoned (as was in fact done in the example considered). The observations after n cycles of k variants can be written in a table having n rows and k columns, and we are only concerned with comparison of *column* means. Now the major part of the dependence occurs *within* rows, and in this situation, as was pointed out by R. A. Fisher,[14,15] the simple analysis of the type we have considered above will usually not be seriously invalidated. Where the correlogram is available it is possible to determine how far dependence between observations will affect the simple analysis and what remedial measures, if any, need be taken, although this refinement would seldom be worth while.

Q. How should multiple responses be considered?

A. Although it is theoretically possible to equate all responses to a single criterion such as profitability, this usually presents great practical difficulties. As a general rule it is best to represent the problem as one of improving a *principal* response (for example the cost per lb. of product) subject to satisfying certain conditions on a number of *auxiliary* responses. These auxiliary responses usually measure the quality and important physical properties of the product.

Very careful thought in the selection of the principal response is essential. The vital question to ask is: 'If this response is improved will it mean *necessarily* that the process is improved?'

In the example of Fig. 4 the reconciliation of the requirements for the various responses in the light of the experimental results was done intuitively and led to the decision to reduce concentration alone. This intuitive approach has the virtue of simplicity and allows background information not emanating from the experimental results themselves to be taken into account. It is fairly satisfactory in the situation specifically dealt with here when there are only two or three variables to consider. However, as has been pointed out elsewhere,[12] the problem is really one of programming, in the sense of linear programming, with the added complications that the problems are not always approximately linear and that the restraints are not known exactly but must be estimated. In the fuller account we show how certain calculations can help with the more difficult cases.

Q. How best can the stream of information coming from the plant during the evolutionary process be presented to those responsible for deciding what to do?

A. Two things are necessary: first, to show how much weight ought to be attached to the results, and second, to present them in such a way that their interpretation is facilitated as much as possible.

To convey a sense of the degree of reliability which the plant manager should associate with the results, a number of ideas have been tried. In the original schemes various types of sequential charts and significance tests were used. It is now felt, however, that the problem is not one of significance testing and that what is needed is a presentation of the information contained in the data unweighted by external features subsequently injected into the situation. For example, the particular choice of the risks α and β and of the hypotheses 'which it is desired to test' (subtleties not readily comprehended by plant managers) can completely alter the apparent implications of a set of data when these are plotted on a sequential chart. If the observations are roughly normally distributed, are independent, and have constant variance, then all the information they contain is included in the mean, the standard deviation, and the number of observations. These statistics seem best comprehended in the form of a mean with its 95% confidence limits. In appraising the results prior information about the importance of different sorts of effect must be used. It seems best, however, to separate this from the presentation of the results, which then refer to information supplied by the observations and to nothing else. This problem is regarded as being far from solved. It involves many intricacies which cannot be discussed in the present account and is probably best considered in terms of stochastic approximation[16] and servo-mechanism theory. All that would be claimed for the present method is that it does allow a satisfactory evolutionary process to go on.

To show the implications of the mean results once it becomes apparent that these are determined sufficiently accurately, there is no doubt that for two or three variables geometrical representation, such as that shown in Figs. 4 and 5, is ideal. It allows the general trend in the responses and their relationship to each other to be appreciated in a manner not possible in any other way.

The plant manager should run evolutionary operation in much the same way as he would play a card game. The information board shows him his 'hand' at any given time, and depending on that hand there are a number of actions he can take (including drawing a further card and deciding what to do then).

Where one or more of the variants is clearly better than the works process or where clear-cut trends in the results exist, the plant manager will have no difficulty in following the indications of the information board. Where the results indicate that complexity exists he will be able to obtain the help of the statistician on the evolutionary operation committee in elucidating the results and, where necessary, in augmenting and modifying[4-6] the cycle of variants in the next phase of operation so as to resolve the complexity.

A duplicate information board may be kept on the plant itself and its significance explained to process workers. This provides added interest and is an incentive to accurate operation, which itself can result in general improvement in productivity.

Q. In what way can the results from small-scale experimental studies be used in planning the evolutionary scheme, and how should this affect the way in which these small studies are conducted?

A. The complexity mentioned above arises principally because the variables studied fail to behave independently in their effects on the response; that is, they interact. The plant process will usually have been arrived at as the result of a small-scale investigation of at least some of the variables. This investigation should have culminated in a study of the local 'geography' of the response surfaces in the neighbourhood of the proposed operating conditions. The principal features of the laboratory response surfaces will normally be preserved on the plant scale even though some distortion occurs. If the characteristics of the laboratory surfaces have been determined in the manner mentioned above, it is frequently possible to discover transforms of the variables originally considered which act approximately independently, at least for the principal response. By working in terms of these new variables, difficulties due to complexity of the surface can be greatly reduced in the plant-scale investigation.

Q. In practice it is impossible to attain truly static operation. Small variations in the process conditions are bound to occur from one run to the next. In cases where these changes are recorded, why should one bother to carry out a special pattern of variants? Why not use the 'pattern of variants' supplied by the natural variation of the process to supply information on which evolutionary improvement can be based?

A. It is true, of course, that for data generated by natural variation the simple type of analysis of the results which has been used above would no longer be applicable. However, this itself is no reason why the natural pattern of variants should not be used. Suppose that the level of response is denoted by y and that there are k variables whose levels are denoted by $x_1, x_2, \ldots, x_k$; suppose that the works process is defined by the particular set of conditions $x_{10}, x_{20}, \ldots, x_{k0}$ and that, owing to imperfect control, fluctuations about those levels occur and are recorded. Then we can, for example, assume a local relationship of the form

$$y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k$$

and estimate the coefficients continuously by the method of least squares (multiple regression). If our assumptions were correct, these coefficients would measure the individual effects of the variables. The calculations required to fit the equations are laborious but, as has been suggested by Professor Goodman,[13] could in principle be done mechanically (e.g. by an electronic computer).

At first sight such a method appears attractive, for here we seem to have an evolutionary scheme in which we do not need to bother about introducing variants deliberately. On closer examination, however, its value seems much more doubtful. Many investigations have been made by statisticians over the years in which plant records have been analysed by multiple regression in an attempt to determine the 'effects' of the variables and so to improve the process. In my experience the results of such investigations are nearly always disappointing. The reasons are not far to seek:

1. Many of the factors that may vitally affect the efficiency of the process are not in the normal course of events altered at all.

2. Those factors which vary naturally do so, not over the ranges we should like, but over ranges dictated by the degree of control which happens to exist. The more control is improved, the less information we get.

3. The fluctuations that naturally occur in the variables are often heavily correlated. This results in poor precision of the estimates when we try to disentangle the effects of the variables one from the other.

4. Accidental modifications often tend to happen in 'phases' and so become spuriously correlated with causally unrelated time-trends in response. Such effects can lead to completely wrong conclusions. Attempts to eliminate time-trends computationally usually eliminate the effects of the factors at the same time.

What this all amounts to is that a naturally occurring scheme of variants is not very likely to provide a good, or even passable, 'design' and consequently that the amount of information generated by natural variation may be scarcely worth salvaging.

Q. Can evolutionary operation be made automatic?

A. The procedure of evolutionary operation so far described is a 'manual' one. It requires no special facilities and can be immediately applied in one form or another to a very large proportion of industrial processes. This is so whether the available plant is of the crudest kind or whether it includes such refinements as automatic controllers and recorders. The plant manager is himself a part of the 'closed loop', thus ensuring that sensible action will be taken even in unforeseen circumstances.

With a sufficiently instrumented plant the evolutionary procedure is, of course, capable of being made completely automatic. Thus variables whose levels are regulated by a controller can be automatically changed at regular intervals so as to follow a cycle of variants, and a response such as cost per lb. can be automatically computed from the readings of instruments which measure the properties of the product. The cumulated differences in response at the various process conditions can be used to trigger off adjustments in the location of the pattern of variants, so completing the evolutionary process.

In continuous processes (where there is a continuous input of starting materials and a continuous output of product) the 'pattern of variants', instead of being a discrete set of points as in Fig. 3, can consist of a continuous locus. The problem of detecting the effects of the variables is then precisely that which arises in communication theory, of detecting a signal of known form in a noisy channel.

The introduction of automatic evolutionary operation would usually be worth while only if the response surface itself was changing in some way and it was desirable to attempt to follow that change. For many chemical processes the response surfaces are reasonably stable. In some, however, unpredictable but steady changes can occur owing to slow changes in raw material (such as crude oil) or in catalyst activity. Here unpredictable differences in the position of the optimum conditions may occur between batches of catalyst and also within the life of a single catalyst batch. In these cases automatic evolutionary operation may be effective in keeping the plant operating near its best performance, but only if the rate at which information is generated is sufficiently large compared with the rate at which the optimum conditions are changing. This is essentially a problem in the theory of servomechanisms.

Discussion

The device I have described is of course only a more powerful and concentrated form of the naturally occurring evolutionary process which goes on during all manufacture. In the ordinary course of events once the favourable effect of a deliberate or accidental modification is *recognised*, that modification is included in the works process. Unfortunately, because of a high level of variation, which usually obscures all but very large effects, favourable modifications frequently go unrecognised unless they are forced to reveal themselves by the device used here of constant repetition and consequent averaging-out of errors.

That many of the problems touched on in the later part of this paper are still the subject of active investigation should not obscure the fact that evolutionary operation, as set out in earlier sections, is a practical and immediately available method which ought to be more widely applied.

Both practical experience and theoretical consideration show that very little can be lost and a great deal can be gained by application of the technique, and for this reason evolutionary operation should be adopted as a *normal production method*. Static operation should be tolerated only if good reasons for not using the evolutionary procedure can be advanced.

REFERENCES

[1] YOUDEN, W. J. (1954 onwards). Bimonthly articles on 'Statistical Design', *Industrial and Engineering Chemistry*.

[2] DANIEL, C. and RIBLETT, E. W. (1954). 'A multifactor experiment', *Industrial and Engineering Chemistry*, **46**, 1465.

[3] VAURIO, V. W. and DANIEL, C. (1954). 'Evaluation of several sets of constants and several sources of variability', *Chemical Engineering Progress*, **50**, 81.

[4] BOX, G. E. P., CONNOR, L. R., COUSINS, W. R., DAVIES, O. L. (Editor), HIMSWORTH, F. R., and SILLITO, G. P. (1954). *The Design and Analysis of Industrial Experiments*. Oliver and Boyd, Edinburgh and London.

[5] BOX, G. E. P. and WILSON, K. B. (1951). 'On the experimental attainment of optimum conditions', *J. R. Statist. Soc.*, B, **13**, 1.

[6] BOX, G. E. P. (1954). 'The exploration and exploitation of response surfaces: some general considerations and examples', *Biometrics*, **10**, 16.

[7] BOX, G. E. P. and YOULE, P. V. (1955). 'The exploration and exploitation of response surfaces: an example of the link between the fitted surface and the basic mechanism of the system', *Biometrics*, **11**, 287.

[8] ANSCOMBE, F. J. (1954). 'Fixed-sample-size analysis of sequential observations', *Biometrics*, **10**, 89.

[9] SATTERTHWAITE, F. E. (1956). (Unpublished communication.)
[10] BOX, G. E. P. and HUNTER, J. S. (1956). 'Multifactor experimental designs for exploring response surfaces', *Ann. Math. Statist.* (In the press.)
[11] JOWETT, G. H. (1955). 'The comparison of means of industrial time series', *Applied Statistics*, **4**, 32.
[12] BOX, G. E. P. (1955). Discussion at Symposium on Linear Programming, *J. R. Statist. Soc.*, B, **17**, 198.
[13] WILKES, M. V. (1956). Discussion of paper 'Application of digital computers in the exploration of functional relationships' by G. E. P. Box and G. A. COUTIE at the Convention on Digital Computer Techniques. *Proceedings of the Institution of Electrical Engineers*, **103**, Part B, Supplement No. 1, 108.
[14] FISHER, R. A. (1941). *Statistical Methods for Research Workers*, 8th edition, p. 226. Oliver and Boyd, Edinburgh and London.
[15] BOX, G. E. P. (1954). 'Some theorems on quadratic forms applied in the study of analysis of variance problems. II: Effects of inequality of variance and of correlation between errors in the two-way classification'. *Ann. Math. Statist.*, **25**, 484.
[16] ROBBINS, H. and MONRO, S. (1951). 'A stochastic approximation method', *Ann. Math. Statist.*, **22**, 400.

B*