Control
and Cybernetics

VOL. 3 (1974) No. 1/2

On diagonal models for bank assets management*)

by

G. P. SZEGÖ

University of Venice
Faculty of Economics and Commerce

This paper gives a presentation of the most recent results in the construction of bank asset magagement models. In particular it discusses a few analytical results which are available trough a very simplified model, the various hypotheses about the structure of the variables to be taken into account as well as a few results about dynamical models.

1. Introduction

The area of assets gathering (financing) and assets allocation (investment) is the area in which the use of mathematical models is becoming more and more popular.

These mathematical models together with their counterpart, the simulation models, provide a decision tool which is becoming indispensable for financial managers.

Im most industrial situations the two problems are correlated, i.e. the film will choose the "best" investment and the "best" financing corresponding to it in order to achieve its own gains.

This simultaneous presence of both problems which is already apparent for instance in the case of *insurance companies*, is evident in the case of *banks* which as financial intermediaries indeed are the most tipical examples of institutions which indeed present both the financing as well as the investment problems.

In this paper we shall give a brief account of the various types of bank asset management models and present the solution of some related mathematical problems.

---

*) Presented at Polish-Italian Meeting on Modern Application of Mathematical Systems and Control Theory in Particular to Economic and Production Systems, Cracow (Poland), Sept. 1972.

2. A simple one-period linear deterministic model

Such a model can be represented by the following equations:

$$D = R_o + I + P, \qquad (2.1)$$

$$R_o = q_o D, \qquad (2.2)$$

$$D = A_2 + a_1 r_d - a_2 \tilde{r}_i, \qquad (2.3)$$

$$P = A_2 - a_3 r_p + a_4 \tilde{r}_i, \qquad (2.4)$$

where D, R_o, I and P are respectively the deposit, the statutary reserve, the investment and the loan levels at the end of a unitary period of operation which may be considered as the first period of operation of the bank.

The constant q_o is fixed. The parameters r_d, r_p and $\tilde{r}_i$ describe the deposit, loan and investment interest rates available to the public respectively.

The period in which this model operates must be so small that the linear market equations (2.3) and (2.4) are an acceptable approximation of the reality. It will be assumed that $a_i > 0$, $i = 1, \ldots, 4$, and that it must be

$$D \geqslant 0, P \geqslant 0, I \geqslant 0. \qquad (2.5)$$

Now the first two positivity conditions, under the realistic assumption that

$$r_d \leqslant \tilde{r}_i \leqslant r_p \qquad (2.6)$$

imply that

$$a_1 \geqslant a_2, a_4 \geqslant a_3, \qquad (2.7)$$

since equations (2.3) and (2.4) must hold (given $\tilde{r}_i$) for all values of the decision variables r_d and r_p.

Now we have the simple mathematical problem of finding r_d and r_p as to maximize a suitable goal function. For instance, let the goal be the maximization of the following profit function

$$G = r_o R_o + r_i I + r_p P - r_d D, \qquad (2.8)$$

where r_i is the interest rate of the investments available to the bank with $r_i \geqslant \tilde{r}_i$.

Under the simplifying assumption $r_i = \tilde{r}_i$ the values of r_p and r_d will maximize G as follows:

$$r_d^* = a_1 r_o q_o - A_1 + r_i [a_1 (1 - q_o) + a_2]/2a_1 \qquad (2.9)$$

and

$$r_p^* = [A_2 + r_i (a_3 + a_4)]/2a_3 \qquad (2.10)$$

and the corresponding optimal levels are

$$D^* = (A_1/2) + a_1 q_o + (r_i/2) [a_1 (1 - q_o) - a_2] \qquad (2.11)$$

and

$$P^* = (A_2/2) + (r_i/2) (a_4 - a_3). \qquad (2.12)$$

Notice now that under the assumption (2.7), indeed we have that

$$r_d^* \leqslant r_i \leqslant r_p^* \tag{2.13}$$

and that all the levels D^*, P^*, $R_o^* = q_o\, D^*$ and the corresponding I^* which can be derived from (2.1), are all nonnegative.

Indeed the decision rule is that the bank fixes (for all values of the exogeneous variable r_i) the values r_d^* and r_p^*, obtains the deposit D^* and gives the loans P^* and then invest the residual

$$D^* - P^* - q_o\, D^* = I^* \tag{2.14}$$

in the investments (bonds) at the rate $r_i \leqslant r_p^*$.

This model, which is extremely simple, shows however already the complexities of the problem. In most of the literature on the subject the constraints (2.5) are not taken into account, the assumption (2.7) is also not considered and therefore the inequality (2.13) is not satisfied, thus leading to absurd results.

These fundamental ideas can of course be expanded as to maximize other bank goals like the profit under some suitable inequality constraints on r_p, r_d and D, consider the case $r_i \neq \tilde{r}_i$ and more realistic equations.

3. Problems for the construction of a dynamic model

The first idea is to replace the model of Sec. 2 with a multiperiodal model and that the equations (2.1)—(2.4) represent one period.

The variables in this case will not be levels, but variations of levels.

The model of Sec. 2 must however be changed in the sense that the inequalities (2.5) do not hold for the variations, but they must hold for the total levels.

The real dynamic model which indeed describes the time behaviour of the relationships among variables is much more involved since the cause-and-effect relationships (2.3) and (2.4) are not instantaneous, but, for instance, the deposit at the time $t{:}D(t)$ depends on the history of the interest rates $r_i(t)$ and $r_d(t)$ on a whole interval from a given $t-x$ to t.

The structure of the functional relationship is very likely rather involved, and there is no reasonable hope to obtain the numerical values of this functional relationship from the available data.

4. Problems arising from uncertainty

The static deterministic models presented in Sec. 2 may lead to quite erroneous results, since they do not allow for the presence of uncertainty in the estimates defined by the market equations.

Due to the presence of those random effects on the estimates it is however quite possible that, for instance, the deposits accumulated on the given time interval in which the linear approximation is defined, are different from what is predicted. These variations around the expected value, which we may define as the value predicted by the linear approximation, can be also taken into account. In particular for the bank, an unexpected contraction of deposits must be faced, since, for the bank to meet this withdrawal, it must sell its assets, namely bonds, and use if possible the loans as collaterals fro obtaining credit from the central bank or other banks, the second alternative bearing obviously a very heavy psychological cost for the management and it must be avoided if at all possible.

A rather rough technique is to consider under some assumption on the probability distribution of withdrawals the expected values of the decrement of deposits: the so-called deposit low. Then the problem again becomes a static problem: one feeds into the model the cost of selling bonds, the cost of using loans as collateral and the deposit low. Then again one can solve the problem of maximization of the profit function.

This approach, even if it does extend the range of validity of the linear model, does not provide a complete description of the situation: in particular the variance of the deposit withdrawals is not shown, as well as other probabilistic aspects like the risk involved in each loan, the different transaction costs of each type of bonds, etc.

In order to provide a good description of these phenomena, one can however apply a suitable modification of the *Markowitz portfolio selection* model, namely a variation of it which allows for random transaction costs as well as for random total supply to be invested (deposits).

The Markowitz portfolio selection model is an application of the Von Neumann-Morgenstern utility theory for uncertain situations.

The problem can be stated in the following way: assume that we have a unitary *fixed capital* to be invested among n investments which are characterized by random returns. The random return of the i-th investment is assumed to be such that it can be characterized through its standard variation and its expected return. Let the standard deviations of the distribution of the i-th investment be denoted by σ_i, the expected value with R_i, the correlation coefficient between the i-th and j-th investment be ρ_{ij} and finally let the x_i the percentage of the capital which will be invested in the i-th investment.

Then

$$\sum_{i=1}^{n} x_i = 1,\ x_i \geqslant 0 . \tag{4.1}$$

Let

$$\sum_{i=1}^{n} x_i R_i = R(P) \tag{4.2}$$

and

$$\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_i \sigma_j \rho_{ij} = V(P) \tag{4.3}$$

be the expected return and the variance of a portfolio, respectively.

Then the portfolio selection problem can be stated as the search for n real numb ers x_i subject to the constraint (4.1), such that a certain utility function

$$u=u(R(P), V(P)) \tag{4.4}$$

which describes the preference of the investor with respect to expected return versus expected risk (variance).

This fixed investment problem can be translated into a banking asset management problem by assuming random liabilities and random transaction costs to meet them.

The theory, seven if it makes a nice pace of mathematical system theory, does not have practical application due to the extreme complexity of the model.

In fact for each asset one has to measure R_i, σ_i, ρ_{ij} and the transaction cost and for each group of liabilities (deposits etc.) one has also to consider its withdraval probabilities characterized again in a simplified model by their expacted values, variances and covariances.

The only hope to use this methodology r_i to real problems may lay in the construction of index-related models along the same line as done in the modern portfolio selection theory; and in particular in the so-called diagonal models which will be presented in the following section.

5. Index related models and optimal index construction

Let r_i be the random return from the i-th investment and let I be a "index of return of the market" (for instance the arithmetic mean of the returns of all shares in the market, etc.).

The return r_i can be represented in the following form

$$r_i=a_i+b_i\, I+c_i \tag{5.1}$$

where the representation

$$r_i=a_i+b_i\, I \tag{5.2}$$

is obtained by linear regression of the random return r_i. Thus the variance of a_i is zero and the expected value of c_i is zero.

Now we make the simplifying assumption that the covariance between c_i and c_j is zero for all investmenzs i, j. Preceeding exactly in the same way as in the previous section it is then possible to develop a portfolio selection technique which is based upon only $2\,(4+1)$ data instead of $n\,(n+1)$ data as in the complete model presented in the previous section. This diagonal model which has been developed by W. Sharpe is now in wide use in investment managing.

The problem is however shifted in the *proper selection of the index I* since, clearly, the error inherent in the diagonal model with respect to the "complete" model will depend upon I. Thus the crucial problem for the applicability of the diagonal model is the choice of I.

G. P. Szegö, M. Gaviano and C. Sutti made a preliminary investigation of the problem related with the construction of an index I, which is a linear combination of the returns of the n investments with the coefficients chosen in such a form as to minimize the total residual variance

$$F=\sum_{i=1}^{n} c_i^2 \tag{5.3}$$

defined by the linear regression (5.1). For this operation it is required to have a set of data of the returns of the n investments.

Assume a set of returns r_{ij} of each asset i at time instants t_j, $i=1, \ldots, n, j=1, \ldots, m$, with $m>n$. The desired index at the time instant t_j has the form

$$I(t_j)=I_j=\sum_{i=1}^{n} d_i\, r_{ij}=\langle d, r_j\rangle . \tag{5.4}$$

The problem to be solved is the identification of the real numbers d_i to minimize (5.3) in which from the definition (5.3) each c_i^2 can be expressed in the form

$$\sum_{j=1}^{m} (a_i+b_i\, I_j-r_{ij})^2=c_i^2 . \tag{5.5}$$

By differentiating (5.3) with respect to a_i and b_i, respectively, and equating the result to zero, we obtain

$$ma_i+b_i \sum_{j=1}^{m} (I_j-r_{ij})=0, \tag{5.6}$$

$$a_i \sum_{j=1}^{m} I_j+b_i \sum_{j=1}^{m} (I_j^2-I_j\, r_{ij})=0 . \tag{5.7}$$

Let next be

$$\bar{R}_i=\frac{1}{m} \sum_{j=1}^{m} r_{ij} \tag{5.8}$$

and

$$\bar{I}=\frac{1}{m} \sum_{j=1}^{m} I_j . \tag{5.9}$$

Then instead of the actual returns F_{ij} and index return I_j we shall consider the following deviations from the income $\bar{R}_i$, $\bar{I}$ as follows

$$q_{ij}=r_{ij}-\bar{R}_i \tag{5.10}$$

and

$$s_j=I_j-\bar{I} . \tag{5.11}$$

Then from (5.6) and (5.7) we get

$$a_i=\bar{R}_i-b_i\, I \tag{5.12}$$

$$b_i=\Big(\sum_j q_{ij}\, s_j\Big)\Big/\sum_j s_j^2 . \tag{5.13}$$

If we substitute these relationships in the form c_i^2 (eg. (5.5)) after some purely algebraic computations, we obtain

$$\sum_{j=1}^{m} q_{ij}^2 - \left(\sum_{j=1}^{m} q_{ij}\, s_j\right)^2 \Big/ \sum_{j=1}^{m} s_j^2 = c_i^2 . \tag{5.14}$$

The problem is now to compute the weights d_i in the index (5.4) so that the following sum of all squared residual (5.14)

$$F = \sum_{i=1}^{m} \left[\sum_{j=1}^{m} q_{ij}^2 - \left(\sum_{j=1}^{m} q_{ij}\, s_j\right)^2 \Big/ \sum_{j=1}^{m} s_j^2 \right] \tag{5.15}$$

is a minimum (notice that through (5.11), (5.9) and (5.4) s_{ij} is a function of d_i).

Define next the following vectors and matrix

$$d = \{d_i\};\; s = \{s_i\};\; Q = \{q_{ij}\}. \tag{5.16}$$

Then again after some purely algebraic computation the index (5.16) taken the form

$$F = \sum_{i=1}^{n} \sum_{j=1}^{m} q_{ij}^2 - \frac{d^T (QQ^T)^2\, d}{d^T (QQ^T)\, d} . \tag{5.17}$$

Now the first term of (5.17) does not depend upon d. Thus the minimization of F is achieved through its second term which reaches its minimum value when the ratio of two quadratic forms in the variables d_i is a maximum. Now if λ is an eigenvalue of QQ^T, λ^2 is an eigenvalue of $(QQ^T)^2$. Thus this ratio assumes its maximum value when the vector d is the eigenvector corresponding to the largest eigenvalue of the matrix QQ^T.

Consider now that d is indeed the eigenvector corresponding to the largest eigenvalue and in addition that such eigenvector has been normalized so that

$$\sum_{i=1}^{n} d_i^2 = 1 \tag{5.18}$$

hen the optimal value of the index F becomes

$$F = \sum_{i=1}^{n} \left(1 - \sqrt{\lambda}\, \frac{w_i}{\sqrt{\sum_j q_{ij}^2}}\right) \sum_{j=1}^{n} q_{ij}^2 \tag{5.19}$$

in which the correlation coefficient inside the paranthesis is a measure of the fitness of the best index.

In spite of the success in the solution of the problem, we can still argue if this is the "best" possible index for all circumstances or not. Matter of discussion is the choice of the function F (eqs. (5.3) and (5.15)) as the aim of the squares of the residuals, it would have been possible to use also some different function like the sum of the squares of c_i where each c_i is first premultiplied by $\sum_j r_{ij}$ or by the price of the i-th investment. It is our opinion however that these additional performance indexes (which do not considerably alter the structure of the mathematical solution) could be used only in very special cases in which the structure of the problem is already well known, but in general the F of the form (5.3) since it is not based with respect to any particular investment is the most suitable. The weights d_i will indeed themselves at the end of the computation express the "weight" that each particular investment has one the index as a whole.

6. Conclusions

This exercise in model making clearly does not present a complete model of bank asset management, but a technique to be followed for its construction.

Our conclusion is that the only possibility for obtaining mathematical models for the description of the asset allocation problem in presence of uncertainty is to construct first suitable aggregations by means of the theory of the "best" index which we have developed in collaboration with Dr. Hersom of the Numerical Optimization Centre of Hatfield Polytechnic.

Numerical experimnets are now on the way to compare this index to the indexes currently being used.

References

1. Fried J., Extension of Markowitz portfolio selection model to bank asset management. Ph. D. Thesis. Northwestern University 1970.
2. Markowitz H., Portfolio selection. New York 1956.
3. Sharpe W., Portfolio selection. New York 1970.
4. Szegö G. P., Shell K., Mathematical methods in investment and finance. Amsterdam 1972.
5. Szegö G. P., Gaviano M., Sutti C., A remark on the stock market index to which the share prices are most linearly correlated. In: G. P. Szegö, K. Shell — ref. above.

Modele diagonalne zarządzania aktywami bankowymi

Przedstawiono najnowsze wyniki w zakresie konstrukcji modeli zarządzania aktywami bankowymi. Omówiono zwłaszcza pewne wyniki analityczne możliwe do otrzymania za pomocą bardzo uproszczonych modeli, założenia dotyczące struktury zmiennych wchodzących w skład modelu, a także pewne wyniki dotyczące modeli dynamicznych.

Диагональные модели управления банковскими активами

Представлены последние результаты в области разработки моделей управления банковскими активами. В особенности рассмотрены некоторые аналитические результаты, которые возможно получить с помощью очень упрощенных моделей, предпосылки касающиеся структуры переменных входящих в состав модели, а также некоторые результаты касающиеся динамических моделей.

Control
and Cybernetics
VOL. 3 (1974) No. 1/2

Combined technological and economic control process*)

by

JANUSZ GOŚCIŃSKI

University of Lodz
Institute for Light Industry

In the paper some general statements based upon the theory of adaptive control processes and the theory of reservation are given.

The role and function of an initial and working information in common and adaptive systems are discussed and the characteristics of both types of systems and some operational conclusions are given.

The problem of reliability of a combined technological management system is presented together with a treatment of reliability of a system as related to the characteristics of a manufacturing process.

Two practical examples of an adaptive control system are included.

The last paragraph contains some final remarks reflecting the gap in our knowledge in the field discussed which causes that the systems design process is not enough effective.

1. Introduction

Efficient management depends in general upon two basic factors, namely on:

(a) the reliability of the manufacturing process;

(b) the reliability of the decision-making as well as the data processing systems.

The reliability of both the decision-making process and the information systems are of the greatest importance for the effectiveness and the efficiency of management. These processes include the following relevant questions:

— the scope of managerial responsibilities (strategic and tactical activities as well as all activities related to the policy of resources aquisition, allocation, and use);

— the structure of an organization (number of management levels, span of control, job descriptions, procedures etc.);

— the system of communication within the organization and between the organization and its environment.

---

*) Presented at Polish-Italian Meeting on Modern Application of Mathematical Systems and Control Theory in Particular to Economic and Production Systems, Cracow (Poland), Sept. 1972.

Is the manufacturing process in fact a highly unreliable one, the most perfectly designed decision-making and information systems would be not able to fulfill their missions to run the business successfully.

There exists a feedback between manufacturing and management process. Each of these processes affects directly the other. They differ, however, from each other.

Is the reconstruction or modification of one of those systems decided, the other system requires because of the mutual interactions, an immediate re-evaluation and usually some modifications. In one of the next paragraphs the relationships between those systems will be considered in some details.

2. Initial and working information

By an initial information I mean such a set of messages describing the subject of a control process, the elements of which are necessary to design a control subsystem and to ensure that this subsystem will operate. A system requiring complete initial information is called a COMMON system. We need full initial information to describe the system and to design the formal model of that system which will make the system controllable.

Common systems are divided into closed and open systems. A system is closed when it has at least one feedback. An open system is, in contrary, a system without feedback. In this case monitoring information describes disturbances arriving from the environment.

On this basis we can conclude that manufacturing systems are closed common systems. The problem how to affect the environment with the purpose to decrease the number and the power of environmental disturbances belongs to the management system.

The manufacturing systems belong to the class of common systems because they do not have the capability to be adaptive against the arriving changes. The management systems, what means the decision-making process and the data processing system are able to follow all changes and to adapt the organization to them. As a result, decisions are taken to modify the programme of production.

An open system can be closed through establishing of buffer or safety stocks which separate the system from the environment and protect it on the one hand from the input disturbances and from the random arrival of orders on the other hand. It is, however, obvious that different types of manufacturing processes can be protected to different extent from the disturbances generated by the environment.

In general there are three types of control processes which are usually designed and applied in the field of production. They can be listed as follows:

(a) programmed control which occurs when parameters vary in a period of time on a basis of a given programme (computer controlled equipment);

(b) follow-up control which occuıs when variables change in an unlimited way within a given dispersion of their values;

(c) stabilization process with fix values of variables in a given period of time.

Management systems are, in contrary, adaptive control systems. An ADAPTIVE control system does not require complete initial (preliminary) information. In this case the process of control is performed partly on a basis of initial information but mostly through the use of working information gathered and processed during the monitoring process.

Working information means a set of meassages describing the state as well as the course of the controlled process which are used in the process of control.

From the point of view of the subject being discussed here the management system plays the function of a control subsystem. It should be, therefore, helpful to achieve two most important goals:

(1) to increase, through correct decisions, the reliability of the manufacturing process using for that purpose the general principle for constructing highly reliable systems containing unreliable elements;

(2) to increase the reliabiltiy of the management system itself through objective setting-up, proper job and responsibilities descriptions, effective structure designing and implementation as well as maintaining of smooth and useful information flows.

An effective management system must be able to meet three fundamental conditions. On the background of the theory of regulation these conditions can be defined in the following way.

The effectiveness by which I mean the reliability of management system in a broad sense depends upon:

A. The SPEED of reaction of the subsystem of control. It is measured by the time to be used by the control system between the moment a deviation occurs and the moment the deviation is reduced after it is monitored and recorded. It is the total time being used by the control subsystem to measure and record any deviation, to transmit adequate information, and to act. It is, more generally speaking, the ability of a control system to regain the static or dynamic equilibrium of the entire system by corrective actions.

B. The ACCURACY of the control process in reducing all possible deviations which can occur in a given period of time.

C. The RELIABILITY of the control system by which I mean that decision-making process and information flows work in a given period of time without failure.

3. Reliability of manufacturing systems

Each process of production is a sequence of a given and limited number of elements in series. Within each plant as well as among them there exists always a chain of cooperation. The more sophisticated are manufacturing techniques the more complex are final products and more diversified is the equipment the greater is the importance of the reliability of a production process. A system becomes more unreliable with the increasing number of components and couplings in series.

The reliability problem contains the following aspects:

(a) the level of reliability of each particular element of a system;

(b) the choice of conditions under which the manufacturing process will be performed;

(c) the implementation of the principle of highly reliable system built up of unreliable subsystems;

(d) the size and allocation of stocks (reserves) which will increase the reliability of the system as a whole;

(e) the optimization approach which will prevent the failure of system elements.

The extention of series (the increasing length of the chain of operations) causes that the probability of process failure increases rapidly. The system becomes in this case less reliable. Moreover, this kind of extension leads to the loss of ability of the system to keep the same level of quality of the manufacturing process performed under the same conditions.

A system characterized by elements in series where each component has a given probability to be reliable signed by p_n does have the following probability of the entire system:

$$P(t)=\prod_{j=1}^{n} p_j(t) \tag{1}$$

for $j=1, 2, ..., n$.

This formula shows that the reliability coefficient is equal to the (1) product of probabilities of all elements in series.

This fact leads to the conclusion that attempts to increase the reliability of a system through the increase of reliability of particular components of the system are not effective. The reliability of such systems decreases in a geometrical progression with the number of elements in series increasing in an arithmetic progression. Systems become in this case unreliable both when failures are dependent or independent events.

Let me use τ to mark the time of life of an element. The probability of failure of a system in a given moment t being written by $Q(t)$ and being a continuous function, if there exists a continuous density of the probability a failure occurs, can be written symbolically as

$$Q(t)=P\{\tau<t\} \tag{2}$$

and the probability of being reliable as

$$P(t)=1-Q(t)=P\{\tau>t\}. \tag{3}$$

The average time to failure or, in other words, the expected value of a random variable is equal to

$$T_0=E(\tau)=\int_0^{\infty} tq(t)\,dt \tag{4}$$

where $q(t)$ is the density of probability the break-down will occur before the moment t.

4. Constructing of highly reliable systems

Now the problem arises how to achieve a very high level of reliability of a complex system consisting of unreliable components.

Two strategic approaches could be applied to solve this problem. The first approach is based on the assumption that the achievement of a high and explicitly defined level of reliability is required despite outlays. This kind of approach can be accepted in this case only if the required reliability is based on the fact that people being involved must survive. Space expeditions, jet flights with passengers are examples for this kind of requirements. The other strategy relates to the case that we are searching for a solution being optimal at a given level of expenses we have decided to accept.

Different types of production processes and management systems require different approach and techniques to achieve high level of reliability. Despite several possible ways of problem solving there exists a common point in the area of demanded reliability achievement. This common and universal approach is called the principle of constructing of highly reliable systems consisting of unreliable components (subsystems).

The problem looks at the first glance like a paradox. The problem solving techniques are, however, based on a simple assumption one is able to achieve each demanded level of reliability by applying of the proper reservation policy which will increase the reliability of the system.

Assume a system consists of n elements in series. Assume in addition each j-th element is supported by m alternative elements. Note, $i=1, 2, \ldots, m$. The master element and m its alternative elements create a group of elements which is called a RESERVATION GROUP of $(m+1)$ components. Each element on line (a master element) represents a piece of a sequence. The system works in each part of the whole sequence as far as even one element from a reservation group is not defected and works.

A system consisting of n elements in series where each of these master elements can be replaced by one of m alternative elements from reservation groups does have certain reliability. The reliability of j-th master element is a function of the probability of unfailure of that element. It can be denoted as follows:

$$P_j = 1 - \prod_{i=1}^{m} (1 - p_{ij}). \tag{5}$$

The reliability of the entire system with n elements in series and m alternative reserved elements can be presented using the notation

$$P = \prod_{j=1}^{n} \left[1 - \prod_{i=1}^{m} (1 - p_{ij}) \right]. \tag{6}$$

The reliability of a system increases rapidly with the growth of the number of alternative elements in reservation groups.

5. Types and methods of reservation

The policy of reservation is the most successful way to increase the reliability of systems. The policy can be formulated by answering a set of questions; the logic of them is shown below:

(a) define the types of reserves to be used within the system;
(b) design the allocation of these reserves;
(c) determine the value and size of allocated reserves;
(d) calculate costs of these reserves;
(e) analyse costs of system defecting if reservation policy not implemented;
(f) compare costs of reservation and system defecting and take the final decision for your reservation policy.

Applying this procedure it will be possible to optimize the size of reserves, the allocation of them as well as acceptable level of expenses for a given level of reliability of any system of production.

Three kinds of reservation can be distinguished. Is a reserve in a state which is identical with the state of the master element we call this kind of reserve a LOAD RESERVE. The reliability of a load reserve does not, of course, depend on the moment in which it replaces an other element.

If a reserve is in a state of expectation to be used by replacing another element and if it does not work at the moment, we call it an UNLOADED RESERVE. It cannot defect before it replaces an load element which will get out of order and stop to work.

Is a reservation element loaded partly only until it replaces another element which was working in the master sequence of the manufacturing process it is called a PARTLY LOADED RESERVE. In this case the element can defect before it replaces a working master element. The probability of defecting before replacement is, however, smaller than in a case of load reserve.

An average time of working without failure can be computed for load reserves as follows:

$$T_m = P_m(t)\, dt. \tag{7}$$

The average time to failure for a reservation group is

$$T_m^{avr} = \sum_{i=1}^{m} t_i^{avr}. \tag{8}$$

Problems are considered here in respect of reliability of manufacturing process if renewal policy is not applied. The renewal approach based on a procedure of restoring or replacement of parts of a defected element generates another type of strategy. This strategy can be used if elements can be repaired and used again. The implementation of the renewal policy generates conditions to apply queuing models to find the probability of failure, the intensity of defecting and many other characteristics of the problem of systems reliability.

6. Types of stocks used to increase the reliability of mixed parallel-serial systems

Several types of stocks can be listed as follows:

(a) raw materials and semi-products on hand;
(b) reserves of unfinished production;
(c) tools and instruments on hand;
(d) reserve machines available (load, partly loaded and unloaded reserves);
(e) machine parts and components on hand;
(f) finished products on hand.

In my opinion all these kinds of stock-on-hand have to be divided into two basic groups of reserves because of two reasons:

(1) they protect in different way the system from disturbances;
(2) the location of reserves is different from the point of view of the manufacturing process.

To the first class of reserves belong all those reserves which are located parallely to the master process. This class includes, there-fore, all reserved machines. If a machine is working full-time and operates as a load reserve in a lateral production line it will be treated as alternate coupling.

The role played by a part on hand with respect to a given machine is identical with the function played by the machine with regard to the process of production. The analogy can only take place if the strategy of renewal was applied. Parts on hand are alternate reserves.

This class of reserves could be called.

ALTERNATE RESERVATION STOCK-ON-HAND.

Reserves of the second class are located, in contrary, in the chain of a master production process. Places where reserve elements are stored divide the sequence of the manufacturing process into subsequences. Because of this position they prevent proceeding of disturbances from the subsequence they occur in toward next subsequences. Providing buffers we create conditions under which the production system can work despite the defecting of some components. Stocks of raw materials and semi products belong to this class of reserves. They are located between the environment and the manufacturing process. They protect therefore the process from supply disturbances. The same role is played through finished products between the demand expressed by random arrival of orders and the manufacturing process. In these terms, unfinished production is an element of reservation policy. The stocks are in this case reserves which are located among production departments, work places, and subsequences of the process.

The level of this stock-on-hand is determined through the probability of defecting of reservation groups, random arrival of orders and supply disturbances.

On the background of these considerations I suggest to call this class of reserves:

SERIAL STABILIZATION STOCK-ON-HAND.

7. Replacement policies

Is the reservation with replacement applied then arises the problem of choice of the proper strategy of replacement. The replacement of components of a system can be in general done in two different ways.

The first approach is based on an assumption that each defected element is replaced immediately.

The next approach is founded upon the idea of prevention against defecting of elements and, as a result, a failure of the whole system. In accordance with this assumption replacement occurs before the failure of elements.

In addition to this classification of replacement approaches we distinguish another one, namely, the INDIVIDUAL and GROUP replacement. Individual replacement is often identified with replacement which takes place just after failure of an element. It is, however, a misleading identification. Group replacement cannot be identified with preventing replacement because group replacement means usually that both defected and working elements are replaced at the same moment.

Before the choice and implementation of one of these approaches is made one has to calculate outlays generated by each of them and to compare results.

There are costs caused by failure which stops elements to work. Is the last element of a reserve group out of order a break-down of the manufacturing process is the consequence. On the other hand there are some costs of replacement activities such as cost of parts, wages of maintain and repair gang workers.

There are many strategies of replacement which can be applied depending on actual conditions.

The characteristic of the first approach is that just after failure of an element it is individually replaced. The next strategy is based on the assumption that periodical group replacement is only undertaken. This strategy of replacement requires the determining of intervals. At the end of an interval group replacement is undertaken. There are two possibilities now. We replace all defected elements or all elements of a given sequence of the manufacturing process. It would be reasonable to use the second strategy if:

— the number of reserve elements is great,

— the elements are of low price,

— the elements are short life components,

— the probability of process failure is rather low because the number of available reserve elements is great.

The third strategy is a combination of both strategies being mentioned before.

Let us assume that there are intervals t, $t+1$, $t+2$, ..., $t+k$. Between each pair of successive moments, for example t and $t+1$, each defected item is replaced immediately. After an interval is reached group replacement is implemented. All in the meantime not defected elements are now replaced or we replace all working elements with the greatest probability to defect in the next period of time. More sophisticated is the next strategy we can use to formulate our replacement policy. In this case all defected elements as non-defected items are replaced individually and in groups.

The significant difference between these two variations of strategy is that in the second case the period of time $(t, t+1)$ is divided by a moment t' in two subperiods. If a break-down of an element occurs before the moment t' what means it occurs in the subperiod (t, t') the element will be replaced individually. All failures occuring in the next subperiod $(t', t+1)$ will in contrary, not cause an immediate action. All elements being defected during this subperiod will wait to the end of this period of time an then be replaced at the same time.

The choice of the best replacement policy depends upon:

A. The characteristic of the manufacturing process,

B. The number and time of life of elements in particular reservation groups,

C. The outlay of replacement policy compared with costs of opportunities being lost because of the production breakdown.

8. Reliability of different types of a manufacturing process

There are many kinds of production classification available. According to the subject of our consideration I will use the classification of technological types of a manufacturing process. The list of these types is as follows:

1. Production of units to customers individual orders.
2. Production of single but technologically compound products.
3. Fabrication of large equipment in stages.
4. Small batch production.
5. Production of elements in large batches which are combined or assembled into batches of varying size.
6. Production of large batches.
7. Production of large batches on assembly lines.
8. Mass production.
9. Intermittent process of production of chemicals.
10. Production process of substances prepared for sale diversely.
11. Continuous process of liquids, gases and crystalline substances.

All eleven kinds of production can be grouped because of some similarities. Thus we distinguish three basic types of production. They can be described in the following way:

Type I	Unit and small batch production (containing numbers 1 to 4).
Type II	Large batch and mass production (numbers from 5 to 8).
Type III	Production process (includes numbers 9 to 11).

The mutability of the production process is the greatest in Type I. The nature of operations, operations tooling, the succession of operations, the time of machine setting-up and the variety of materials and semiproducts used depend upon the technological characteristic of customers' orders and their additional requirements.

The poorer is the variety of ordered products, the more smoothed and stabilized is the manufacturing process and the easier it is to group ordered items in batches. The more extorted is the order of operations in the process, the smaller is the freedom of choice. The manufacturing process becomes a system with extorted serial couplings.

The feature of production affects the number and types of equipment to be installed. Machines can be in majority general-purpose (universal) machines. This fact allows the performance of several operations using the same machine. Are, the machines to a great extent unique facilities it would be not more possible to performe given operations using each type of machines. The manufacturing process will be in this case much more sensitive to any kind of possible disturbances. Large batch production, mass production and process production are usually sequential in character with relatively long chain of co-operation. Because of this reasons they are the most unreliable manufacturing processes.

The point is that unit production needs very high number of components being bought from the environment. Many large industrial organizations manufacturing final products produce not more than 20 per cent of all components they use on their assembly lines. For these types of production is, therefore, highly important to secure smoothing deliveries of semiproducts, components, and subassemblies accordingly to an exact time schedule. Short-range planning system and shop-floor control look like networks.

CASE STUDIES

9. Examples of production of type II and III

9.1. Paper pulp production (typ III, point 11)

The paper pulp plan is divided into departments which are adequate to technological phases (sequences) of a continuous process of substances. Technological phases and adequate departments can be listed in the following way:

(1) timber preparation;

(2) timber boiling;

(3) washing of the pulp mass;

(4) bleachery;

(5) second screening phase;

(6) dehydration or drying process.

Interoperational stock on hand (unfinished production) is for this type of process production of very limited importance. Technological conditions make impossible to store the pulp mass just after two most important process stages, namely after boiling (safety stock container can be used after the washing stage only) and bleaching because the buffer stock can be hold after the next stage what means after the screening stage.

Facilities used in two most important stages of the whole process work three shifts a day. The interruption of a continuous process means that the semiproduct will be destroyed and is, therefore, unfeasible.

Let me take the example of the Boiling Department to ilustrate the combined technological and economic process of control. Timber is the input prepared during the first stage of the manufacturing process. It must be dimensioned and keep the proper humidity. As an outcome we get alkalified mass which is washed passing through filters which are in the meantime diluted.

The automatically controlled sulphit pulp process requires the following measurements:

(1) The efficiency of the paper pulp digest:
  (a) pulp flow (flowmeter);
  (b) pulp mass concentration (sensing device).
(2) Feedings:
  (a) compressed air (measurement sensor);
  (b) process steam (readings);
  (c) water consumption (measurements; device);
  (d) power supply (counter).
(3) Wood saturation (white and black paper digesting liquours):
  (a) measurements of percentage of active alkali (NaOH) per unit of dry pulp mass;
  (b) flows of black lye through screens.
(4) Digestion zone — process control by "permanganate number" (alkali batchering).
(5) Mass flows — measurements of mass blow-out from the digest bottom to the mass washing section.

Technological readings are used both to control the manufacturing process and management. Measurement instruments readings allow to compute:

— the consumption of water, process steam, power and compressed air;

— the loss of alkali per weight of the pulp mass using the pulp digest efficiency and NaOH contents data.

The most important elements of Boiling Department equipment are:

— white alkali container,
— paper pulp digest (boiler),
— black alkali expander,
— filters,
— alkali tanks,
— blow-out container,
— water container,
— turpentine oil separator,
— turpentine oil container.

The paper pulp digester is very expensive. Because of this reason it is rather impossible to keep another digester empty as an unloaded reserve.

On the other hand each failure generates disturbances which move forward to the successive sequences of the process. An effective solution requires, therefore that two conditions should be met:

(a) the continuous readiness of maintain and repair gangs,

(b) the proper level of parts-on-hand necessary to ensure the immediate replacement of elements which will probably defect in the next period.

There are several stages of implementation of combined technological and economic control processes based on automatic control systems.

The simpliest computer assisted control system is able to gather data, to classify and process these data and to reflect the state and the flow (or performance) of the process only. Decisions are in this case made by people.

Allows the software programming of decisions with varying parameters of the process and taking into account significant limitations (e.g. machines or job working places capacities) we can speak about the higher second level of computer implementation. This stage is usually called: man — machine system.

Is the computer used on-line the optimal automatic control system can be applied. Decisions are made in this case by the control system. Detaching of machines, deviation recording, and reducing, operations planning and other decisions will be made by the system itself.

Finished product stock-on-hand (pulp) as well as raw materials (timber, alkali and other substances) consumed fully or used many times after regeneration e.g. NaOH regeneration they all play the some role of stabilizers of a serial process. They block disturbances and avoid movement of them along the process line. Unifinished production (inter-operational stock-on-hand of details and semi-products) are of small importance because they can be stored after few subsequences of the process only and cannot be kept longer than a couple of hours because they loose after short time the quality required.

There are not auxiliary operations parallel to the master process of production where machines can be used as partly unloaded or load reserves which could be interrupted at any time if a given master machine defects.

Keeping of unloaded reserve machines is beyond discussion because of economic reasons. The alternative reservation is, therefore, limited and can be applied in the field of parts, storage and maintaining policy only. This sort of reservation is of second-range importance from the viewpoint of systems reliability.

Because of these reasons all unique machines should be highly reliable items and the strategy should accept relatively high expenses to keep the reliability of each master machine on a given and high level.

9.2. Production of grinders (type II, point 4 and 5)

The largest division of a given enterprise manufacturing grinders is divided into four main sections. They are as follows:

(1) rough mechanical working section,

(2) fine mechanical working section,

(3) heat treatment section,

(4) assembly section.

We omit the description of some auxiliary sections such as toolmakers shop (because it is not of any importance for the problem being discussed here).

Sections are divided into groups of work places and working places. There are the following group work places. In the fine mechanical working section: shafts, toothed wheels, spindels, lead screws, and nuts. The rough mechanical working section is divided into the following groups of places: scrapers, bodies and machine tool, beds, tables and saddles, and finally headstocks.

In this section the following types of machines are installed: grinders, horizontal boring machines, slotters and drillers. Particular types of machines differ from other types because of such parameters like speed, accuracy etc. The number of installed machines is 54, however, there are 33 dimentionally different types of machines only.

Analyzing these data we can see that in average less than 50 per cent of the total number of machines can be used alternatively to perform particular sorts of operations being used as load reserves. The reason for this unprofitable situation is that the majority of equipment used in this section are specialized machines which cannot be replaced by other available machines. As a consequence defecting of those machines causes an immediate process break-down.

The annual production of all types of machines, the structure of manufactured goods and the applied technology of which is discussed here can be considered using data from table.

Position	Group of products	Number of types	Annual production
1	Surface grinders	15	1410
2	Cylindrical grinders	12	460
3	Internal grinders	6	657
4	Centerless grinders	3	195
5	Special purpose grinders	2	60

The production is dominated by surface grinders. The average number of pieces per one type of them is annually 94 items. In this case we can speak about small batch production.

The technological structure of production can be presented as follows:

number of groups of products	5
number of types (sorts) of products	38
number of details	about 23.000

but about 11 thousand of them are standarized details which are manufactured in large batches.

The surface grinders are finally dominated by one type of them, namely by type SPD X. The three structure of this product shows five levels of elements being assembled during the manufacturing and assembly process. They are presented below:

level		number of elements
level 0 (final product)		1
level 1 (assemblies)		5
level 2 (sub-assemblies)		30
level 3 (details)		1840
level 4 (operations)	about	8000

As a conclusion of the analysis technological tree — two critical problems of the manufacturing process can be distinguished:

(1) parallel performance of details and related operations is possible to some extend only in a frame of the apllied technology;

(2) collecting and storing of elements (details, subassemblies) before the assembly process starts is of the greatest importance for the whole process and its final results.

The reliability of the manufacturing process depends, because of the characteristics of the process, on:

(a) continuous flow of elements through all work places,

(b) proper stock-on-hand of standarized details,

(c) relatively high level of reliability of single machines or very small reservation groups of them because they are in majority specialized units which cannot be replaced or substituted by other unique facilities.

That implicates that the stabilization can be achievad successfully through the following kinds of reserves:

— by stock-on-hand of details both made and bought

— by stock-on-hand of parts allowing relatively fast repair of defected machines used in the master process.

To tell the truth some alternative reserves are undervaluated by top-management despite the fact that this type of production is very suitable for that approach. In this situation very useful are.

(a) unloaded machines (grinders, drillers etc.) which could be used in the master process after a break-down of working items;

(b) reserves of partly unloaded machines working during one shift (the plant works in two shifts) but unloaded during the second shift and used from time to time to produce outstanding details to replenish the stock for the next shift;

(c) moving reserves of multi-operational transfer machines able to replace or substitute a number of specialized master machines and used along a seqvence if somewhere defects occur.

Two types of manufacturing process were reflected and discussed here namely paper pulp process production and production of several types of grinders. The second process differs from the first one because of the number of degrees of freedom which is in this case greater. Many details can be manufactured in varrying order or parallel by different work places. The extorted serial system is in the second case not as vexing as in a case of a continuous process of production. The advantage of a large number of standarized details produced in large batches independently on the arriving customers orders in the context of reliability is quite obvious.

Types of reserves, allocation and optimal magnitude of them require different approach for different types of production. The problem requires an extension of research. Results could be very useful to increase the effectiveness of management. Optimal and adaptive control is urgently more and more needed.

Sterowanie złożonym procesem techniczno-ekonomicznym

Przedstawiono ogólne uwagi wywodzące się z teorii adaptacyjnych procesów sterowania i teorii rezerwacji.

Rozważono rolę i funkcję informacji początkowej i bieżącej w systemach zwyczajnych i adaptacyjnych oraz przedstawiono charakterystyki systemów obu typów wraz z pewnymi wnioskami na temat ich działania.

Przedstawiono zagadnienia niezawodności złożonego systemu zarządzania obiektem przemysłowym oraz rozważono zagadnienie niezawodności systemu w odniesieniu do charakterystyk procesu wytwarzania.

Podano dwa przykłady praktyczne systemów sterowania adaptacyjnego.

Ostatni punkt zawiera pewne uwagi końcowe na temat niedostatku naszej wiedzy w zakresie omawianych tu problemów, będącego przyczyną niezadowalającej efektywności procesu projektowania systemów.

Управление сложным технико-экономическим процессом

Представлены общие замечания берущие начало из теории адаптивных процессов управления и теории резервирования.

Рассмотрена роль и функция начальной и текущей информации в обычных и адаптивных системах а также представлены характеристики систем обоих типов вместе и некоторыми выводами в отношении их действия.

Представлены вопросы надежности сложной системы управления промышленным объектом, а также рассмотрен вопрос надежности системы по отношению к характеристикам процесса производства.

Даны два практических примера систем адаптивного управления.

Последняя часть содержит некоторые заключительные замечания, касающиеся недостаточности наших знаний в области рассматриваемых здесь вопрос, являющейся причиной неудовлетворительной эффективности процесса проектирования систем.