THE PATH TO PROCESS DATA

F D. Gault, R. B. Hoffman, B. C. McInnis

Analysis of transformation processes is a potent tool for the understanding of materials utilization, energy consumption, environmental pollution and the productivity of capital and labour. This paper introduces process analysis with emphasis on the need for physical rather than monetary data. Some uncoordinated data gathering activities relevant to process analysis are examined. Two prototype process databases are described to illustrate ways of approaching process analysis and it is noted that a process database is necessary for the next generation of economic models. Although there is as yet little progress in assembling the organization to coordinate the building of such a database, some strategies for the future are proposed.

Keywords: industrial production; process analysis; future studies

PRODUCTS DO NOT HAPPEN. They are made in a transformation of labour, materials and energy which uses the capital stock of the production plant. The transformation yields the product, and other outputs which might be by-products of some commercial value, or pollutants like sulphur dioxide or acoustic energy. Information about transformation processes is important to the understanding of the utilization of materials, the consumption of energy, the pollution of the environment, and the productivity of labour and capital. Process analysis, which is the application of this information to production analysis, is a powerful tool complementary to the highly aggregated approach of macroeconomic analysis.

This article starts at the macroeconomic level and looks underneath it at the structure of the economy revealed by input–output models, and then at the greater understanding possible if materials and energy are required to obey physical conservation laws in these models. As a progressively more disaggregated view of the economy is taken, the measurement of physical quantities becomes more important and money less important. Money, which is a natural measure of aggregated quantities, is seen to connect to the underlying

The authors are with the Structural Analysis Division, Statistics Canada, 24th Floor, R. H. Coats Building, Tunney's Pasture, Ottawa, Ontario, Canada K1A OT6. The authors gratefully acknowledge the comments of R. U. Ayres, G. O. Barney and colleagues at Statistics Canada.

0016-3287/85/050509-17$03.00

physical flow of materials and energy through the price mechanism which provides an interface between economics and the physical economy.

The measure of the physical flows of material and energy into and out of a transformation process which involves capital stock and labour, requires the gathering and storage of large amounts of information. The size and complexity of this problem has been one of the inhibiting factors in the evolution of process analysis.[1] However, the present state of information technology and the work of prototype projects,[2] suggest that there are no insuperable barriers remaining to hinder progress.

A theme which parallels the move from money and aggregation to physical processes, is the growth and availability of computing power. The spread of microcomputers, easy-to-use software packages and networks presents the economist with powerful tools and access to computer-searchable databases on larger computers. While these tools support conventional economic analysis there is a gap in the information necessary to support process analysis. As there are no software or hardware difficulties with storing and retrieving physical flow and capital stock data, this gap may now be filled. How the gap is to be filled is an important and current issue.

The economic debate on externalities[3] is an area of application for process analysis. The issues include the opportunity cost of using a scarce resource, the capital cost involved in pollution control and abatement and the substitution cost of processes to satisfy government regulations. Process analysis is a natural framework for such questions as process descriptions capture information on the factors necessary to production (skilled labour, energy, minerals, topsoil, etc), the capital stock needed to effect the transformation to products and on the residual outputs (by-products, pollutants, energy, etc).

The description of a transformation process requires a model. The model allows statements to be made about changes in outputs resulting from changes in inputs. With the help of computers, process models can be connected together into networks representing an entire industry, its consumption, production and residual output. Such models could also ensure that the physical requirement of material and energy conservation is maintained for each constituent process and it could account for the capital stock involved at each stage. This approach of combining processes to represent an industry makes technology substitution analysis straightforward in that it permits the building of an industry model using several technologies, the combination of which may change with time, depending on the life of the capital stock.

The utility of process-based analysis raises an obvious question: where are the computer-searchable process databases to provide the support for this work? The answer is that there is no complete and universally available process database in existence although studies nave shown that such a database is both possible and useful. This article explores the reasons for this state and suggests ways of changing the situation.

History of process analysis and the problem of disaggregation

The macroeconomic models which dominate modern economics predict changes in employment and national income resulting from monetary and fiscal adjustments. While this approach is useful for broad planning, it does not provide the detail necessary to account for changes in individual sectors of the economy. For example, a decision on the part of a consumer to pay for riding lessons rather than to spend the same money on a personal computer is not noticed in macroeconomics as the aggregate consumption is unchanged, and so is the contribution to the GNP.

The decisions of consumers, however, do change the underlying economy, as the people involved in building, shipping, selling and servicing personal computers will differ in number and skill from those involved with keeping horses and providing riding lessons. The choice made by the consumer in this example will result in a different consumption pattern. Of course, it is not just individual consumers that change consumption patterns but also governments as a result of economic planning.

Wassily Leontief was the first to develop the input–output techniques that reveal the patterns obscured by the aggregation of macroeconomic models.[4] Leontief divided the economy into sectors or industries and recorded the flows of goods and services between sectors, including the flow of intermediate products. The resulting table showed which inputs in what quantities were required by a sector to make its product or output. The success of input–output models was considerable as they allowed, for example, for impact studies on output changes if final demand changed, and they described the effect on prices which resulted from a change in factor prices. They were also able to represent the allocation of fixed resources. Since 1936, when the first input–output table depicting the American economy of 1919 was published, a considerable body of empirical knowledge has grown up, and the use of these tables in planning has become widespread.

Part of the success of the input–output model is due to the simplicity of the underlying assumptions. For example, the model is linear, which means that outputs are directly proportional to inputs and there is no mechanism for changing this relationship as the scale or technology of production changes. Once the input–output table is constructed, the technology embodied in the production processes is implicitly assumed to be fixed, and all subsequent projections or impact studies are based on this implicit assumption. Production, however, is not in reality a linear function of inputs and as the rate of technological change increases the relaxation of the fixed technology assumption becomes progressively more important.[5]

With the energy supply shock of the 1970s and the environmental concern of that decade, attempts were made to represent energy, limited resources and unwanted products in the context of input–output models. Part of the motivation for attempting this was to provide uniform data for large-scale simulation models. A major exponent of this approach was R. Ayres who developed the integrated materials/energy balance statistical system (MEBSS).[6]

Extending input–output models to include environmental information poses a number of problems, as Ayres points out. An environmental question tends to be physical: how much material/energy/pollutant is used or generated directly or indirectly in a process, in the productiuon of a commodity by or a sector of the economy. The input–output model does not impose physical constraints like energy and material conservation, and unless these constraints are added, absurd predictions can result.

A further problem of extending input–output models is the implicit assumption that the supply and demand of goods and services is assumed to be in equilibrium and that the specification of final demands should reflect a Pareto-optimal consumption pattern. As the environment provides unpriced goods and services which are subject to physical laws but are not in the market place, it is difficult to define an economic equilibrium. If anything, these goods and services are subject only to clumsy regulatory control.

Once physical constraints are accepted, however, linear programming can be used, to relax the fixed technology assumption of the input–output model. This gives rise to the static models of both Koopmans and Kantorovich[7] and also to dynamic models. The important observation is the significance in these models of physical data subject to physical laws in contrast with constant dollars and economic assumptions.

The concern with physical constraints has led to the MEBSS developed by Ayres under the auspices of the United Nations Statistical Office. It required[8] that all material and energy inputs to the world economic system, as well as to individual countries, be accounted for either as final outputs or as changes in accumulated stocks, including durable goods in service as well as inventories. It required two balance principles, a gross (volume) balance applied to production, consumption and trade of major resources and commodities, and a more refined materials and energy balance by process to elucidate the relationship between resource/commodity production and consumption and the generation of waste flows.

While MEBSS in principle means that resource consumption and residual production can be monitored, the construction of such a framework requires data of standard classification and uniform quality. The source of suitable data is seen as some combination of engineering studies and of survey and monitoring data compiled by monitoring and regulatory agencies. The lack of such data, noted by Ayres in 1978, poses a fundamental problem for the MEBSS approach.

Before physically-based models of the economy can be constructed, there is a need for data on industrial transformation processes. Here, an industrial transformation process means any process or operation which converts materials and energy to a product. It covers the obvious manufacturing and chemical processes as well as, for example, the conversion of natural resources to extracted resources. The question arises as to where the boundary is drawn to define a process.

A process could be an entire industry, a plant or a process within the plant. Ayres[9] introduces the concept of 'economic process' to provide a lower limit for the definition and he requires two criteria. First, the product must be transportable and storable and have an ascertainable market value. Second, the activity must be carried out in the context of a cost-minimizing or a profit-maximizing enterprise. Ayres later introduces the concept of 'significant' process[10] and argues that although one could rank processes according to some value-added criterion to establish those most economically significant, there will be those low on this scale which are important by other criteria. For example, an economically insignificant process producing a heavy metal as a by-product represents an environmental hazard which is significant to the running of a sewage treatment plant if the metal kills the bacteria in the anaerobic digesters.

The definition of a transformation process has to evolve over time, from interactions between users of the data and those who compile and evaluate the information.[11] While the definition is evolutionary, the concept is fundamental to the representation of socioeconomic-resource systems. It has its roots in activity analysis[12] and has been further elaborated by Geogescu-Roegen.[13] Process analysis, however, cannot be done without process data.

The problem of the data is not limited to process data as Leontief points out: "An unreasonably high proportion of material and intellectual resources devoted to statistical work is now spent not on the collection of primary information but on a frustrating and wasteful struggle with incongruous definitions and irreconcilable class fixations."[14] Leontief is talking about economic statistics in 1971 but the same problem exists for much of the physical data necessary to industrial processes. To illustrate this point, data programmes in material properties and energy are considered, as are the problems of gathering data on residuals. The data are the problem.

A materials database: an example problem

Any database, especially one which is computer-searchable, improves the use of the data it contains and, because of ease of use, broadens the user community. While one can understand the lack of easily available data on industrial transformation processes because of the complexity of the information and its commercial sensitivity, it is less easy to see why a universally available materials database does not exist. This problem was addressed in 1982 at a workshop of interested parties convened in Tennessee.[15]

The proceedings of the workshop make interesting reading as they deal with the same issues that must be confronted by a process database. First, the database does not exist, second, it is a desirable input for computer-aided design (CAD), manufacturing (CAM), and testing and evaluation (CATE), programs and, third, if action is not taken soon, independent programmes will evolve with incompatible standards and duplication of effort. Implicit, of course, is the need for action to get the project going.

It is encouraging that since 1982 action has been taken, in that an autonomous US corporation, The National Materials Property Data Network, Inc[16] has been set up to plan and implement a network to provide information on materials. This follows a feasibility study[17] conducted by the Metal Properties Council, Inc which indicated that it was possible to combine existing databanks into a network with common access.

This initiative recognizes that in the USA alone, tens of billions of dollars are lost, due to avoidable materials failure and inefficient use of materials through improper selection or incorrect design parameters. The likely users of the new network are people who own or operate equipment, those involved in engineering design, manufacturing, field service, failure analysis, remaining life assessment, R and D, codes and standards activities, writing specifications and selecting materials.

Once this network is established the expected benefits will be considerable,[18] not just for the reasons cited, but also because of the more rapid diffusion of new materials and technologies which will result. New materials and new technologies again raise the problems of processing and use of materials.[19] Synthetic materials like composites, ceramics and plastics are replacing metals in manufacturing and their processing is different. At the same time new processes are emerging for forming and working metals and new materials which reduce the need for traditional machinery.[20]

A materials database network is an important beginning in support of all aspects of computer-aided design and manufacture. However, materials cannot be separated from the machines, the chemicals, and the energy used to process them, not if a complete picture of the manufacturing process is to be available for design purposes.

Including energy in the analysis

All materials applications and production processes require energy but it was not until the supply shock of the 1970s that energy became a major issue. The analytical issue was twofold: could the process be modified to use energy more efficiently and could different sources of energy be substituted for those that are costly or of strategic significance. Once the question was raised of how energy was used, the entire production process had to be reconsidered, including the possibility of substituting a totally different production process which still yielded the same product but with reduced energy demand.

Two representative examples of process-oriented energy data projects are outlined. These are major projects which have been in place for some time. The strategic importance conferred on energy, and consequently on energy information programmes, by the supply shock of the 1970s means that these programmes are more advanced than materials programmes and considerably more advanced than databases of residuals production. Nonetheless, it should be noted that, conceptually, the energy database problem is much simpler than that for materials because there is only one type of energy which can be classified by a small number of sources and by the services provided to the transformation process.

The first example is the Industrial Energy Productivity Project,[21] which produced the ISTUM-2 Industrial Sector Technology Use Model. This model provided projections of energy use, efficiency and technology penetration for 27 industries at the 2-, 3- and 4-digit Standard Industrial Classification (SIC) levels in ten regions for the period from 1985 to 2000 in five-year intervals. The model included over 450 proven and emerging processing, energy conversion and heat recovery technologies to represent industrial technology choices in 52 energy services (eg, steam generation, machine drive, iron making, distillation etc). For each of the 27 industries, energy use was projected by total demand and by fuel type (18 primary and by-product fuels).

The Industrial Energy Productivity Project used the process approach to promote the idea of providing energy services (heat, light, mobility etc) at least cost rather than addressing the narrower problem of reducing the US consumption of imported oil. Process data were gathered and the energy consumption of the 27 industries was represented. In doing this, database problems were encountered because of the volume of the data and the paucity of systematic surveys at the high level of disaggregation embodied in the model's database. ISTUM-2 also tried to model the factors influencing industrial decision makers to delay the adoption of new technologies. This was ambitious, and it required the introduction of financial as well as technological constraints. Now that the project has reported, the model is available for use along with extensive documentation. The need to update the database from 1976 to 1980 has been identified by the project organizers and interested organizations have been invited to participate.[22] This data gathering is not part of an on-going programme as it is in the next example.

The next example of a process-oriented organization is le centre d'études et de recherches économiques sur l'énergie (CEREN) in France. Since the mid-1960s, CEREN has maintained a comprehensive database on French industrial, commercial and domestic energy consumption. It surveys over 4000 plants every four years and has a high level of participation as a private, non-profit organization cooperating with the support of the French energy supply industry.[22]

In the CEREN approach, an industrial process is characterized by a set of ten elementary energy operations (EEOs), each of which is qualified by five parameters (equipment characteristics, material inputs, energy inputs, intermediate products and waste energy). Product–process combinations are specifically identified and the approach allows energy data to be clustered at the plant level by energy type and activity and enables data on production and energy to be linked. The CEREN organization is successful and on-going and studies have been made on transferring the CEREN techniques to other countries in the European Economic Community and to the USA.[24]

Process information is more prominent in the two energy examples just considered than in the case of materials. Materials can be viewed as a list of properties independent of the process in which they are used and this information is quite valuable. The corresponding information for energy is its source, whether it is derived from coal, or oil or whatever and while this information is useful for decisions on substitution, the principal means of classifying energy is by the service it supplies to a production process. Viewed another way, it is useful to know the energy consumption of a particular process and that requires knowledge of the energy services necessary to the process.

Residuals

The problem of incorporating residuals information in a process database is more subtle than for materials or energy. Materials data are clearly important for substitution and determining the capital stock appropriate for the material processing. The importance of energy data is equally apparent in determining the energy cost of operations or processes and in deciding on substitution of energy sources. The decisions influenced by material and energy data are economic and can be seen in terms of maximizing the profit of the firm. For residuals production this connection is not as clear and involves a wider, social perspective.

A plant which consumes natural resources, releases sulphur dioxide into the atmosphere and chemicals and waste heat into the river, and is a source of noise, suffers no economic penalty until society, through government, imposes a cost for these activities. The determination of this cost is a matter of on-going social debate,[25] and it has been the subject of economic studies incorporating environmental factors for some years.[26]

Ayres[27] has reviewed various models incorporating residual production, waste management and pollution abatement and he made the point, in 1978, that "the necessary understanding and data for a satisfactory, comprehensive 'macro-economic-ecologic' model do not yet exist". The situation is essentially unchanged and is likely to remain so until the relevant data are gathered, evaluated and disseminated within a framework of manifest benefit to industry.

As regulations on the use of common resources like pure air and clean water have proliferated, the cost of pollution abatement has become part of the cost of production. Production decisions now have to be made to maximize profits in a way which takes these regulations into account. As a result, designers of new plant and processes may select different production processes which fulfil the requirements of the new constraints. Thus the production process provides a paradigm for classifying information on material, energy and labour used, capital stock necessary for the process, and on products and residuals produced. Tying this information together facilitates decision making which takes all aspects of production into account. It also allows firms to be classified by their principal activities.[28] This approach has been studied for some years by Statistics Canada and the International Institute for Applied Systems Analysis (IIASA) in Austria.

The Process Encyclopedia at Statistics Canada

As a result of economic modelling[29] in the 1970s, which included energy demand and residal production, Statistics Canada became aware of the need for a detailed physical description of industrial processes. The need also arose in the representation of technological change in dynamic economic models using an underlying input–output structure.[30] To assess the problems and costs of meeting this need, an exploratory programme was carried out between 1978 and 1983 with the encouragement of the United Nations Statistical Office and R. U. Ayres, who acted as a consultant to the project in its early stages.

The Process Encyclopedia was the principal product of the programme. Enough process data and software for data storage and retrieval were assembled to test the usefulness of the database. The problems of gathering, evaluating and retrieving the data were also assessed. The programme was reviewed in 1981 with a workshop[31] which identified the strengths and weaknesses of the prototype implementation and made suggestions for future development.

The first problem encountered by anyone wishing to store information in a database is the construction of a logical model, or schema, which describes the structure of the data. *The Process Encyclopedia* was no exception and its designers chose to divide the description into four parts, a topography, a 'generic' model, parameters and observations. In addition, there was a bibliographic file indicating where the data came from. It should be emphasized that in *The Process*

Encyclopedia the word 'process' is used to describe either a chemical or biological transformation or an 'operation' or physical transformation. This differs from engineering literature where the two are distinguished.

The logical modelling of process data differs from that of some scientific data as there is, from the start, a theory of the machine in the form of an engineering design. The design specifies that for a particular capital stock, certain inputs of labour, materials and energy, yield specified amounts of product, waste energy and residues. It might also specify an optimum level of operation to produce the most product for least cost or to use least energy, or to produce the least amount of residue. The design relates the outputs of the process to the inputs.

The logical model of the data starts with a picture of the stocks and flows, their names and how they are connected. This is the topography, a simplified example of which is given in Figure 1. The storage of such connectivity informa-

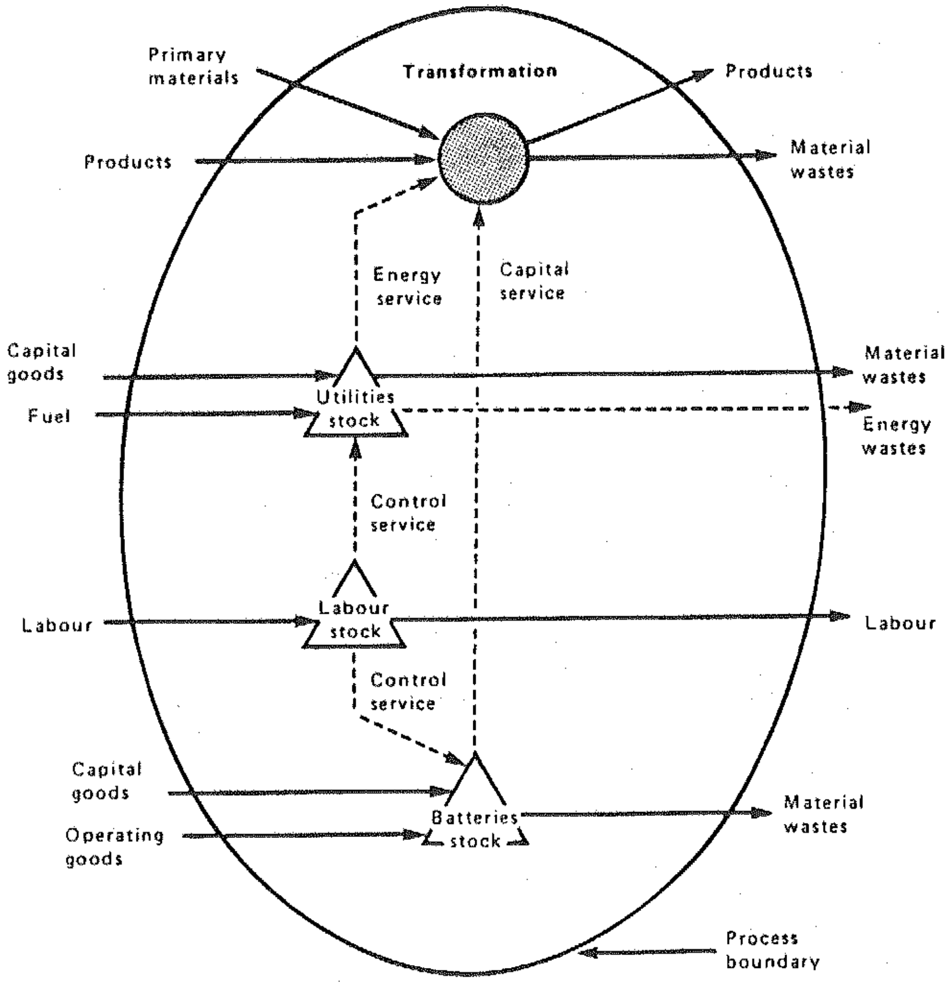

Figure 1. Simplified representation of a process topography.

tion, which can be of great complexity, is common to crystallographic databases[32] and presents no fundamental problem in storage, retrieval or graphical display. The description of a crystal ends when the constituents and links are named and the connectivity is specified. With a process, the topography is only part of the description and the same topography can represent a number of processes.

The next requirement is for the functional relationship to be specified, relating inputs to outputs, and here the design of the process is helpful. In the case of input–output models the relationship had to be linear, but for a process model this is not so. any dependence on the inputs of energy, material, labour and capital stock consistent with physical laws is possible.

The model need not be specific to any one process and it is in this sense that it is called 'generic' in *The Process Encyclopedia*. To specify a generic model it is sufficient to define the functional form, as in the example of a linear dependence. To say that output is linearly dependent on input is not to say what the linear dependence is. To specify that, requires a set of parameters.

The parameter specification is not general but it is necessary before a generic model can describe a particular process. If the model is linear the fact that ten units of an input yields five units of an output is specified by a parameter.

The coming together of a generic model and a topography generates a container, or mini database, to store a set of parameters for the process representation. What is usually called a 'model' of a process is the combination of the generic model, the topography and the set of parameters. The combination describes how all of the stocks and flows combine and how they change in response to adjustments of control settings. It is called a 'realized' generic model and constitutes a complete description of the transformation process.

The final set of data which can be stored for a process consists of observations. These are actual measurements of flows through the process which can be compared with the predictions of the process model to see if it is performing to specification. Observations can also be used to construct a mathematical model of the process if there is no design model available. The absence of a design model can occur in the case of a large process of great complexity which has just been put together over time, or in the case of a commercially secret design.

As a feasibility test, data on several hundred processes were gathered[33] and stored as topographies, generic models, parameters and observations as well as bibliographic references. This information was available in computer-searchable form for *The Process Encyclopedia* workshop in March 1981.

Part of *The Process Encyclopedia* programme was a study of data sources and methods of data capture. Secondary sources, published documents and unpublished contractor reports were found to be rich in data and were sufficient to get a process database started before it was found necessary to collect data from engineering firms and by performing surveys of establishments engaged in process activities.

Data capture presented a problem. It could not be done by technical support staff but required a minimum level of training and experience equivalent to that of a third year engineering undergraduate at a Canadian university. As the data capture studies were short, no benefits of learning were derived and no data evaluation took place.

The dissemination of the captured data was principally to the workshop and limited experience was gained of working with users of process data over time or of direct feedback from expert users on the quality of the data, the software or the ease of use of the system.

The workshop[34] recognized the usefulness of bringing technical information to bear on economic analysis and it supported the development of a production version of a process database after one or two applications had been developed with the existing prototype. The applications were felt desirable so that the utility of the prototype could be convincingly demonstrated before embarking on long-term data gathering.

Interest was also shown in the next stage of the prototype *Process Encyclopedia*, which was the specification of software to permit process description to be linked together to represent more complex processes, plants or industries.[35] The linking of processes to represent industries provides a way of representing competing technologies, processes and products in one framework. It naturally supports dynamic modelling and substitution analysis, and by bringing process data together, it provides a way of identifying and costing of various desired processes which may not yet exist. However, the software specified to make this possible suggested that it would be a costly undertaking both in implementation and in the eventual demand on computing resources.

The computing cost of linked process models, however, should not become a serious problem because economic modelling is moving towards the use of large computers, a practice already well established in the physical sciences. In this computing environment a process database and associated software should be seen as a natural tool in support of modelling, as natural as a language compiler or statistical package. The cost of developing such a process tool should be seen in terms of its return to the modelling community. Its demand on computing resources would not be much greater than that of other support packages.

The Process Encyclopedia study ended in 1983 with the specification of the linking software. Statistics Canada are now studying potential applications of process databases with a view to building up a community of users and evaluating techniques of data gathering, evaluation and dissemination.

The WELMM project at IIASA

At about the same time as *The Process Encyclopedia* was being developed at Statistics Canada, IIASA in Austria was working on the Water, Energy, Land, Materials and Manpower (WELMM) project. There are close parallels between the two:[36] both are process oriented and both projects have ended.

While WELMM was concerned with primary resource consumption[37] from a process perspective, it consisted of databases of primary resources and of resource requirements for industrial processes to produce commodities. The process information was held in a Facilities Data Base (FDB) with the process boundary drawn to correspond to an industrial unit or facility. Not only were data stored on the process but also on the construction process of the plant containing the transformation process. This meant that there was a complete picture of process inputs and outputs augmented by a picture of the resource flows necessary to build and maintain the plant.

The WELMM data were collected from literature, existing databases or questionnaires and they were analysed before they were entered into the database. This data analysis and judgement was stored in the FDB along with bibliographical and numerical information.

The process orientation of WELMM made it suitable for a variety of applications including comparisons of alternative technologies for the production of specific products or services[38] and for the comparison of whole resource processing systems.[39] The project was brought to an end after nearly seven years of operation and like *The Process Encyclopedia* it demonstrated the potential value of process data.

The examples considered in this paper have concentrated on either a specific part of the transformation process—materials, energy and residuals—or process studies necessarily limited in time and scope. Process-based industries like oil, chemicals and pharmaceuticals, maintain plant-specific models for control and optimization purposes which take all aspects of processes into account, but these models are not readily available. Of the examples considered, only CEREN is on-going. Nowhere, however, is there a current general process database which captures, stores and disseminates data on all aspects of transformation processes.

Policy, futures and process analysis

Why is it important to have an on-going commitment to the collection and dissemination of process data? There are two components to the response. The first is the proven usefulness of process information in policy analyses. The second is the potential usefulness of a process database in modelling complex systems in support of futures studies.

Two examples of how process analysis is used now in support of policy analysis have already been given: ISTUM and CEREN. The Industrial Sector Technology Use Model came out of a two-year project in which US government and industry cooperated to develop a tool for analysing the impacts of public policy and business decisions on industrial energy use. CEREN, which differs in approach, also provides support for analysis of contingency plans in the event of energy supply disruption, and for the planning of transitions from threatened energy supplies to alternative sources and technologies.

An example of process analysis applied to a series of linked processes and to technology substitution is a study of plastic and glass bottle production.[40] The resulting model was used to forecast future market behaviour in cases where the two products were competing and effluent taxes were taken into account. The comments of the reviewers of the study, given in the foreword, are illuminating as they show an awareness in 1974 of both the need for, and the cost of, a database to support this type of work.

There is a wide range of applications of process analysis in studies of liquid and solid waste production and of gaseous emissions.[41] These studies concern themselves with environmental damage, or economic externalities. An externality is an effect of production and consumption on other producers and consumers without corresponding payments.[42] It can be positive, as in the case of inventions, or negative as in the case of environmental damage. To understand the root cause of such damage, it helps to know which industrial transformation processes are present in the area being studied and the rate at which they consume energy and material.

The advantage of process analysis in the evaluation of externality problems is that it allows the connection of residuals production to other variables which might be better reported or more easily measured, like rate of production of the principal product, or fuel consumption. If it is known that the fuel used contains sulphur, then it is possible to ask how the production process takes this into account. As process analysis identifies all possible material and energy inputs to a process, and as material and energy cannot be destroyed in the macroscopic world, questions can also be raised about how used materials are dealt with. An example is a policy decision on siting a firm which manufactures silicon chips and printed circuit boards. Trichloroethylene and trichloroethane are widely used in cleaning chips and trichloroethylene is suspected of causing cancer in humans. Copper from circuit boards kills the bacteria in sewage treatment plants if it is washed down the sewers in sufficient quantity. A land use policy analysis which did not take these factors into account could lead to costly problems for the community.

A disadvantage to the application of process analysis to a limited area is the temptation to deal only with the process information relevant to the particular policy issue, like water, land or energy use. While this approach can be justified in terms of available time and budget, it can also lead to problems. A simple example is a study of zinc extraction with a view to deciding how best to dispose of the solid waste. This could be analysed just in terms of zinc and waste production. However, a more thorough analysis would show that zinc ore frequently contains cadmium which would form part of the solid waste and might influence waste disposal policy.[43] Here, there is a case for an easily accessible database of complete process descriptions.

The importance of process analysis to policy analysis is well established, however, it has a broader application in long-term or futures analysis. Dahmen[44] points out that the problem of externalities occurs in both planned and free market economies and he attributes this to the process of industrialization common to them both. Industrialization, he argues, disaggregates economic activities into discrete stages and as a consequence, different stages of decision making have arisen to control different parts of the production process. This disaggregation has possible, and far-reaching, consequences for the public which has little influence over these decisions.

This raises the question of how to model a complex organization and also how to represent decision-making processes and feedback mechanisms. There are various schools of thought on how this should be done[45] and the debate is well covered elsewhere.[46] The important point is that linked process descriptions have the potential to represent the physical transformation aspects of plants, industries or entire economies. The addition of material and energy balance requirements can keep track of physical stocks and flows within the transformation processes. Once such a physical model is constructed, it will ensure that its predictions are consistent with physical reality, whatever the additional control, demographic or economic mechanisms used to represent the complex organization.

This stratified approach, with the organization divided into various layers is close to the hierarchical, multilevel, systems method of Mesarovic[47] and applied in the World Integrated Model (WIM) of Mesarovic and Pestel.[48] In WIM the world is divided into ten regions and each region is stratified into layers representing the individual, the group, demographics and economics, technology, and the environment. Dynamic feedback loops connect the various strata and equilibrium micro and macroeconomic models are used to account for a Gross Regional Product.[49] The regions interact through a Trade Matrix[50] and the physical transformations are not made explicit.

Statistics Canada, also using a hierarchical systems approach, are developing a long-term socioeconomic and resource policy analysis tool.[51] The current implementation, called the Socio-Economic-Resource Framework, or SERF,[52] separates decision making, or control variables, from the underlying demographic, resource and production models which deal principally in physical quantities. An input–output model is used for resource allocation but the long-term intention is to replace this with linked process representations to facilitate technology substitution studies. The approach, which combines control theory and hierarchical systems theory is applied to the Canadian Economy but it is equally applicable to larger- or smaller-scale studies.

To support the modelling of systems in general, using a hierarchical approach with one level devoted to physical transformation processes, requires the collection and availability of high-quality process data. While the gathering together of such data will benefit futures studies in the long term, it will also have a direct benefit for more limited and immediate policy studies.

The next step?

Process-based models can supplement macroeconomic models, show why policy decisions have the consequences they do, and can help evaluate proposals on technology substitution, labour and capital productivity and on the use of scarce and public resources.[53] They can do this because estimated changes resulting from policy decisions of governments, consumers, or firms are not lost in macroeconomic aggregation but are explicit in the model and are constrained by basic physical laws.

Large models involving social parameters, natural resources, population data, capital investment and pollution and their projections have made a significant contribution to the awareness of future problems. Examples are the work of Forrester,[54] Meadows,[55] Mesarovic and Pestel,[56] and the *Global 2000 Report* of Barney[57] where other such models are reviewed. These models involve physical data in their calculations and projections, but out of this an organization with an on-going commitment to compile the physical data so important to the modelling of the future has failed to materialize. This is due partly to the complex nature of the data, the difficulty of data capture and the diversity of skills required to capture, evaluate, store and disseminate such information.

The problems of the industrialized world—energy, the environment, natural resources—are partly economic, partly technical and partly social.[58] To address them calls for economic models with a firm physical base and as the data needs of economics shift in this direction, the conventional distinction between economics, engineering, geology and even biology are disappearing. This is a point made by Wassily Leontief[59] who goes on to call for a large organization to oversee the construction of a large economic model and the collection and evaluation of relevant data.

Leontief uses the Standord Linear Accelerator Center as a model for this organization but he could equally well have chosen any large national or international experimental centre for physical science. The economic model would replace the accelerator, and would serve as a laboratory for experiments in economics in the same way that large accelerator laboratories serve experimental physicists. The importance of this approach is that the physical data and the decision-making process in the economic 'experiment' would necessarily be separate. In mounting experiments on the Leontief 'machine', users would have to provide a set of economic and social decisions as their 'experiment' and the 'machine' would provide consequences of these decisions consistent with physical conservation laws. In this respect the 'machine' differs from econometric models which attempt to project historical time series without regard for the physical consequences of their predictions. The design and maintenance of the model would require a considerable and disciplined team effort.

The analogy with an accelerator is illuminating, as the technology to build an accelerator, although advanced, is based on established physical principles while it is the experiment run on the accelerator which probes new physics. The proposed Leontief 'machine' would embody physical and social science principles and could then be used to run social science experiments to probe new phenomena.

Data gathering for this project poses a fundamental problem as the collection would require close cooperation between engineers, demographers, labour specialists, sociologists and economists.[60] An integral but separable part of such a project would be a database of industrial transformation processes. Such a database, while useful in its own right, would provide the physical foundation on which models that integrate prices, labour and demography could be built. The collection, evaluation and dissemination of process data is an important next step in the progress of economics.

Data collection

In the physical sciences, funds are forthcoming to support data collection and evaluation for a variety of reasons. Some data have economic and strategic importance, such as atomic data for fusion and nuclear data for reactor design. In both cases there are networks of data collectors and evaluators coordinated by international agencies.[61] In some subject areas the participants are sufficiently aware of the worth of a central database to fund it by subscription. This is the case in the Mossbauer community[62] and to some extent in crystallography.[63] In others where the data have less strategic or economic importance, networks of compilers are coordinated by groups supported by national research funds as in the case of elementary particle physics.[64] What then is the appropriate paradigm for physical process data compilation in economics?

The choice seems to lie between a national or international programme of coordination or a smaller effort conducted by interested groups as part of the support for their own research and coordinated informally. Either way, the problem lies in the coordination. As both Leontief and Ayres observed[65] that data should be compiled by expert agencies with the standards maintained by the central coordinating agency. In the physical sciences the coordinators and evaluators are physicists, while in economics the establishment of a large- or small-scale data organization would require a broadening of the discipline of economics to encompass physical data.

Conclusion

Process analysis has been introduced and various programmes of data compilation relevant to process analysis have been examined. What has emerged is the clear absence of an on-going programme of collection, evaluation and dissemination of process data. As process data are a prerequisite to process analysis, the establishment of such a programme is an important step in the progress of economics.

The importance of the physical process data which underlies the economy has been emphasized by Leontief in his call for a large computer-based representation of the economy. The Statistics Canada Socio-Economic-Resource Framework (SERF) has practised the separation of choice from the physical consequence of choice for some years, although not on the scale envisaged by Leontief. The success of the approach at Statistics Canada has been sufficient to suggest that a large-scale representation of the economy would be beneficial, if not essential, to the advancement of understanding of economic problems.

Process data are useful in more immediate applications of process analysis such as investment decision analysis, residual management analysis, resource extraction analysis, and process design and control. These applications[66] are particularly important in engineering design and it is no exaggeration to say that a computer-searchable database of process data could save $ millions as a result of more effective computer-aided design and control.

In spite of the manifest utility of a universally available and general process database, it is not clear that present programmes, with separate interests and constituents, will come together into one network of process data without a central agency to provide standard data description, standards of data evaluation and of data exchange. In this respect the work of the US National Materials Property Data Network will provide an instructive model in one important area.

Even with a central agency to coordinate process data there is the problem of the commercial sensitivity of process information. A recent technology transfer study[67] makes the point that companies would rather transfer their product technology than their process technology as the diffusion of process technology is harder to control once it is transferred. A process database organization would have to take these problems into account and strike a balance between the secrecy of new process technology and the need for its availability.

The future of process analysis lies with the establishment of a coordinating agency to set and maintain standards and to disseminate data. As a first step, prototype process database organizations are called for, to work out the problems of this approach with users and suppliers of the data.

Notes and references

1. T. C. Koopmans, comment on "Energy and economic growth", by E. A. Hudson and D. W. Jorgenson, *American Economic Review*, Papers and Proceedings, May 1978.
2. R. B. Hoffman, B. C. McInnis and W. S. Page, "Process analysis", *Encyclopedia of Materials Science and Engineering* (New York, Pergamon Press, in press); and A. Guebler, R. Hoffman and B. McInnis, *Process Information Systems: A Synthesis of Two Independent Approaches*, WP 81-12-01 (Ottawa, Statistics Canada, Structural Analysis Division, 1981).
3. R. U. Ayres and A. V. Kneese, "Production, consumption and externalities", *American Economic Review, 59*, 1969, page 282.
4. W. W. Leontief, *The Structure of the American Economy 1919–1939, Second Edition* (New York, Oxford University Press, 1951).
5. T. C. Koopmans, "Analysis of production as an efficient combination of activities", *Activity Analysis of Production and Allocation*, Cowles Commission Monograph, *13*, New York, J. Wiley and Sons, 1951; and L. V. Kantorovich, "On the calculation of production input", translated in *Problems of Economics*, 1960.
6. R. U. Ayres, *Resources, Environment and Economics: Applications of the Material/Energy Balance Principle* (New York, J. Wiley and Sons, 1978).
7. Koopmans, *op cit*, reference 5; Kantorovich, *op cit*, reference 5.
8. R. U. Ayres, *Preliminary Description of Materials/Energy Balance Statistical Sub-system (MEBSS), IRT-358-R, Draft Working Paper for United Nations Statistics Office* (New York, 1976); and R. U. Ayres, "The need for an Integrated Materials/Energy Balance Statistical System (MEBBS), in *op cit*, reference 6.
9. R. U. Ayres, J. Cummings-Saxton, M. O. Stern and R. W. Roig, *Optimizing Materials/Energy Process Models, op cit*, reference 6, chapter 5.
10. Ayres, *op cit*, reference 8.
11. Ayres, *op cit*, reference 8.
12. Koopmans, *op cit*, reference 5.
13. N. Georgescu-Roegen, "The economics of production", *American Economics Association, LX* (1), 1970.
14. W. W. Leontief, "Theoretical assumption and unobserved facts", *American Economic Review*, 1971.
15. J. H. Westbrook and J. R. Rumble, Jr, *Computerized Materials Data Systems* (Washington, DC 20234, Standard Reference Data, National Bureau of Standards, USA).
16. The National Materials Property Data Network, Inc, c/o The Metal Properties Council Inc, United Engineering Centre, 345 East 47th Street, NY 10017, USA.
17. *Feasibility Study of an Inter-Society Computer-Based Materials Property System* (New York, Materials Properties Council, 1982).
18. J. R. Rumble, Jr, "Progress towards a computerized materials data system", *Proceedings Computerized Materials Workshop for the Ground Vehicle Industry* (Columbus, Ohio, April 1984).
19. L. J. Lau, "The measurement of raw material inputs", in V. K. Smith, J. V. Krutilla, editors, *Explorations in Natural Resource Economics* (Baltimore, Johns Hopkins University Press, 1982).
20. R. J. Kopp and V. K. Smith, "The perceived role of materials in neoclassical models of production technology", *ibid*, and *The US Machine Tool Industry and the Defence Industrial Base* (Washington, DC, National Academy Press, 1983).
21. *Industrial Energy Productivity Project—Final Report*, (1–9), DOE/CS/40151-1 (Washington, DC, US Department of Energy, 1983).
22. *The Industrial Sector Technology Use Model, Project Summaries* DOE/CE-0038 (Washington, DC, US Department of Energy, 1982).
23. CEREN, 89, rue de Miromesmil, 75008 Paris, France.
24. *The CEREN Industrial Energy Data System: A Transferability Study* (Washington, DC, US Department of Energy, 1979).
25. G. Hardin and J. Baden, editors, *Managing the Commons* (San Francisco, W. H. Freeman and Co, 1977); and V. K. Smith and J. V. Krutilla, editors, *Explorations in Natural Resource Economics* (Baltimore, Johns Hopkins University Press, 1982).
26. O. Herfindahl and A. V. Kneese, *The Quality of the Environment* (Baltimore, Johns Hopkins University Press, 1965); and G. O. G. Lof and A. V. Kneese, *The Economics of Water Utiliza-*

tion in the Sugar Beet Industry (Baltimore, Johns Hopkins University Press, 1968); and P. Bohm and A. V. Kneese, eds, *The Economics of Environment* (London, The Macmillan Press, 1971); and A. V. Kneese, S. E. Rolte and J. W. Harned, eds, *Managing the Environment* (New York, Praeger Publishers, 1971); and C. S. Russell, *Residual Management in Industry* (Baltimore, Johns Hopkins Press, 1973).

27. Ayres, *op cit*, reference 6.
28. K. Pavitt, *Patterns of Technical Change—Evidence, Theory and Policy Implications* (Brighton, Science Policy Research Unit, University of Sussex, UK, 1982).
29. B. C. McInnis and K. Hamilton, *Economic Environmental Modelling—Discussion Paper for a Proposed Residuals Tracking System*, WP 74-10-25 (Ottawa, Structural Analysis Division, Statistics Canada, 1974).
30. T. Gigantes, "The representation of technology in input–output systems", in A. P. Carter and A. Brady, eds, *Contributions to Input–Output Analysis* (Amsterdam, 1970); and R. Hoffman, *The Adjustment of Input–Output Coefficients in Time Structural Economic Models*, 76-03-15 (Ottawa, Structural Analysis Division, Statistics Canada, 1976).
31. B. C. McInnis, *The Process Encyclopedia Workshop, 26, 27 March 1981*, WP 81-08-25 (Ottawa, Structural Analysis Division, Statistics Canada, 1981).
32. P. A. Machin *et al*, "Crystal structure search retrieval", *Daresbury Laboratory Report*(Daresbury, UK, 1977).
33. B. C. McInnis and D. Tobalt, *Sample Process Data from the Prototype MINISIS Implementation of the Process Encyclopedia*, WP 81-03-15 (Ottawa, Structural Analysis Division, Statistics Canada, 1981).
34. Workshop, *op cit*, reference 31.
35. F. Daneliuk, *Design Specification for an Interactive Modelling System Based on the Concepts of the Process Encyclopedia*, WP 82-09-30 (Ottawa, Structural Analysis Division, Statistics Canada, 1982); and F. Daneliuk, "The Process Encyclopedia—a software tool for modelling", in C. Keren and L. Perimutter, eds, *The Application of Mini- and Microcomputers in Information, Documentation and Libraries*, Amsterdam, 1983; and *op cit*, reference 2.
36. Geubler, *op cit*, reference 2.
37. J. M. Merzeau, *Resources Group Progress Report on Unconventional Oil Studies*, WP 80-28 (Laxenburg, IIASA, Austria, 1980); and A. Greubler, *Resource Requirements for Industrial Processes: A WELMM Comparison of Energy Chains*, WP 80-50 (Laxenburg, IIASA, Austria, 1980).
38. *Ibid.*
39. Resources Group, IIASA, "Comparaison WELMM de scénarios énergetiques", *Revue de L'Energie, 316*, Paris, France, 1979; and A. Gruebler, "Natural resource commitments for long-term energy supply options", in B. Kursonoglu, A. Perimutter, M. Bauer and L. Scott, eds, *International Scientific Forum on Changes in Energy*, Lexington, MA, D. C. Heath and Co, 1981.
40. R. Ayres, J. Saxton and M. Stern, *Materials–Process–Product Model: A Feasibility Demonstration Based on the Bottle Manufacturing Industry, Report, IRT-305-FR* (Virginia, International Research and Technology Corporation, 1974).
41. Herfindahl, *op cit*, reference 26.
42. Dahmen, *op cit*, reference 26.
43. H. Inskip, V. Beral and M. McDowall, "Mortality and Shipham residents: 40-year follow-up", *Lancet*, 1982, page 896.
44. Dahmen, *op cit*, reference 42.
45. R. U. Ayres, "Limits and possibilities of large-scale long-range societal models", *Technological Forecasting and Social Change, 25*, 1984, page 297; and D. H. Meadows, "Charting the way the world works", *Technology Review, 54*, February/March 1985; and J. M. Richardson, "A decade of global modelling", *Futures, 14*, 1982, page 136.
46. "Status of futures research", *Futures, 16*, 1984, page 382.
47. M. D. Mesarovic, D. Macko and Y. Takahara, *Theory of Hierarchical, Multilevel, Systems* (New York, Academic Press, 1970).
48. M. Mesarovic and E. Pestel, *Mankind at the Turning Point* (New York, Dutton, 1974).
49. M. Mesarovic *et al, Construction of a Dynamic Model of the Regionalized World Economic System* (unpublished report, August 1972).
50. Mesarovic, *op cit*, reference 48.

51. F. D. Gault, K. E. Hamilton, R. B. Hoffman and B. C. McInnis, *A Design Approach to Socio-Economic Modelling* (Ottawa, Structural Analysis Division, Statistics Canada, to be published).
52. G. J. Cameron *et al*, *The Socio-Economic-Resource Framework*, WP 83-10-28 (Ottawa, Structural Analysis Division, Statistics Canada, 1983).
53. Cameron, *ibid*, reference 52.
54. J. W. Forrester, *World Dynamics* (Cambridge, MA, Wright-Allen Press, 1971); and J. W. Forrester, "Global modelling revisited", *Futures*, *14*, 1982, page 95.
55. D. H. Meadows, J. Richardson and G. Bruckmann, *Groping in the Dark: The First Decade in Global Modelling* (Chichester, John Wiley, 1982).
56. Mesarovic, *op cit*, reference 48.
57. G. O. Barney, *The Global 2000 Report to the President*, *I–III* (Washington, DC, Government Printing Office, 1980).
58. W. W. Leontief, "An information system for policy decisions in a modern economy", in M. Dutta, J. C. Hartline and P. D. Loeb, eds, *Essays in Regional Economic Studies* (Durham, North Carolina, The Acorn Press, 1983).
59. Leontief, *ibid*.
60. W. W. Leontief, *The Long-Term Impact of Technology on Employment and Unemployment* (US National Academy of Engineering Symposium, 30 June 1983).
61. F. D. Gault, "Physics, databases and their uses", *Computer Physics Communications*, *22*, 1981, page 125; and D. R. Lide, "Critical data for critical needs", *Science*, *212*, 1981, page 1343.
62. J. G. Stevens, "The operation of a small data centre", *Computer Physics Communications*, *31*, 1984.
63. O. Kennard *et al*, *Pure and Applied Chemistry*, *49*, 1977, page 1807.
64. F. D. Gault and M. R. Whalley, "Compilation programmes in elementary particle physics", *Proceedings of the 9th International CODATA Conference, Jerusalem, 1984*.
65. Ayres, *op cit*, reference 6; Leontief, *op cit*, reference 58.
66. Hoffman, *op cit*, reference 2.
67. E. Mansfield and R. Wagner, "Technology transfer: foreign trade and R and D", in E. Mansfield *et al*, *Technology Transfer, Productivity, and Economic Policy* (New York, W. W. Norton and Co, 1982).