

NOT FOR QUOTATION  
WITHOUT PERMISSION  
OF THE AUTHOR

PROCESS INFORMATION SYSTEMS:  
A SYNTHESIS OF TWO INDEPENDENT  
APPROACHES

A. Grübler  
R. Hoffman  
B. McInnis

May 1982  
CP-82-24

*Collaborative Papers* report work which has not been performed solely at the International Institute for Applied Systems Analysis and which has received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS  
A-2361 Laxenburg, Austria

## THE AUTHORS

A. GRUEBLER is with the Energy and Mineral Resources Task of the Resources and Environment Area at the International Institute for Applied Systems Analysis, Schloss Laxenburg, A-2361 Laxenburg, Austria.

R. HOFFMAN and B. MCINNIS are from the Structural Analysis Division of Statistics Canada, 24th Floor, R.H. Coats Building, Tunney's Pasture, Ottawa, Ontario, Canada K1A 0T6.

## PREFACE

The development of process information systems to support analysis of energy (or mineral) resource development strategies has been a continuing concern within the WELMM (Water, Energy Land, Materials and Manpower) Project since 1977. The resulting process information system, the Facility Data Base, has demonstrated its viability in a number of application studies, both in-house and outside IIASA.

Motivated by its experience in socio-economic resource modeling, the Structural Analysis Division at Statistics Canada has initiated a feasibility study on a data base of industrial process descriptions (completed in 1981): the Process Encyclopedia.

The independent development and the joint interest in process information systems has led to collaboration between the Structural Analysis Division of Statistics Canada and the Energy and Mineral Resources Task of the Resources and Environment Area of IIASA. This paper (also published by Statistics Canada: Working Paper No. 81-12-01) documents the first results of this collaboration: a synthesis of the conclusions which led to the development of the two systems; of their main features and of the lessons learnt from their design and implementation. The conceptual mapping developed between the Facility Data Base and the Process Encyclopedia and documented as an Appendix of this paper provides a basis for future exchange of information between the two systems and should be seen as a first step in promoting interdisciplinary and international process information exchange and system building.

Janusz Kindler  
Chairman  
Resources & Environment Area

## CONTENTS

1. INTRODUCTION	1
2. PROCESS INFORMATION	1
3. THE FACILITY DATA BASE OF THE WELMM APPROACH	5
4. PROCESS ENCYCLOPEDIA OF STATISTICS CANADA	8
5. CONCLUSIONS	12

### APPENDIX

Conceptual Mapping between the Facilities Data Base and the Process Encyclopedia

## 1. INTRODUCTION

The immediate objective of this paper is to compare the Facility Data Base of IIASA's WELMM system and the Process Encyclopedia of Statistics Canada in order to establish a conceptual mapping between the two systems. This conceptual mapping when transformed into appropriate computer code will permit the exchange of 'process information' in both directions. This objective has been accomplished and is documented in the Appendix to this paper.

The other objective of this paper is to draw from the experiences of both IIASA and Statistics Canada a number of conclusions with respect to process information systems. The second section of the paper introduces the concept of a process and its application in input-output and sector specific process models. It then discusses the sources of process information for these applications. Sections three and four summarize the Facility Data Base and the Process Encyclopedia respectively. Section five presents the conclusions from these two experiences.

## 2. PROCESS INFORMATION

The concept of a process is fundamental in the representation of socio-economic-resource systems. It has its theoretical roots in activity analysis (2-1) and has been further elaborated by Georgescu-Roegen (2-2).

A process is simply a system separated from its environment by an imaginary boundary. It is connected to the environment through input and output flows. Processes may have internal structure which determines the relationships among the flows. The flows may be mass, energy, service, or information. When mass and energy flows relate to a process, they are of course subject to physical laws of mass and energy conservation. Processes may be naturally occurring, for example biological processes, or they may be initiated and controlled by man.

In recent years, problems of resource scarcity and environmental degradation have led to the extension of economic systems to include natural resources, externalities and socio-demographic factors. Energy analysis, environmental impact analysis, and technology assessment have become significant new fields for applied research. The feature common to these areas of interest is the focus on the processes that transform resources and energy into the goods that meet human needs.

Over the last 25 years, there have been two relatively independent lines of development of applied process models in economics: input-output models at the macro-economic level (2-3), (2-4) and (2-5), and, sector specific process models (2-6), (2-7), (2-8), (2-9), (2-10), (2-11) and (2-12). The importance of process models has also been recognized in the field of ecology (2-13).

Process models tend to be information rich. Limitation in the availability of process information imposed by the cost of obtaining such information has restricted the development of process models (2-7).

Input-output process information is readily available in most countries as it is compiled and published by national statistical offices, however, for most applications, the

input-output representation of a process is not adequate for the following reasons: the process boundaries are not well defined; flows are measured in currency units; there are no stocks; the information is retrospective in the sense that the descriptions are derived from measurements of past - usually 3 to 5 years - flows; processes are represented at a relatively aggregate level, as a result of the use of currency units; the degree of aggregation and the use of currency units make the process descriptions time and nation specific in that each description represents the specific mix of processes at that time and place; finally the relationships between input flows and the output flow for each process are linear and proportional.

Process information compiled for sector specific models is normally compiled by the model developers. In fact, the compilation of process data usually requires the preponderance of research funds in the development of process models. Information from this source is usually richer than input-output data in that it is measured in physical quantity units; it is much less aggregated and process boundaries better defined. However, process data collected for sector specific process models is not readily accessible because it is dispersed over a large number of research institutions, none of which have a mandate for information dissemination at that level. Furthermore, process descriptions compiled for sector specific models are often incomplete - for example process descriptions compiled for energy modeling may not include non-energy flows. Also the information is in the specific form required by the mathematical structure of the model and its computer system. For example the process descriptions take the form of constraints in an optimization framework.

Both the WELMM project of IIASA and the Process Encyclopedia project of Statistics Canada represent attempts to compile data bases of process information. Initially these data bases were intended to support particular modeling applications. In both cases it became clear that process descriptions could be compiled in such a way that they could support a variety of different

modeling applications. Indeed processes could be defined in such a way that they are neither site, nor plant nor nation specific - at least for a large number of processes.

### 3. THE FACILITY DATA BASE OF THE WELMM APPROACH

In order to assess the natural resource requirements of resource development strategies (in particular energy strategies) an analytical approach called WELMM has been developed at IIASA (3-1). The WELMM approach involves an assessment of the requirements and the availability of Water, Energy, Land, Materials and Manpower resources. For quantitative analysis, the WELMM approach is based on computerized data bases of primary resource availability at the global, national or regional level (3-2), (3-3), (3-4), and (3-5), and on data bases of resource requirements for industrial processes deployed in processing primary (energy) resources to the commodities required by the final consumer. At each of the transformation steps of such a resource processing system, corresponding industrial processes can be defined. Within the Facility Data Base (FDB), the boundaries of the process analyzed are drawn in such a way that a process corresponds to an industrial unit or facility. The technological characteristics of any particular facility are independent from their location; in addition there is an increasing trend towards standard size classes, particularly for energy facilities (e.g. Pressurized water reactors of 1000 or 1300 MWe, crude oil tankers of 250000 or 300000 DWT, etc.). Both factors mentioned above make WELMM-type analyses (i.e. at a "typical" industrial facility level) easier for a variety of applications including comparisons of alternative technologies for the production of specific products or services (3-6), (3-7), and (3-8) or for the comparison of whole resource processing systems (3-9), and (3-10). However, the process boundary defined for the FDB is flexible - at a conceptual level and from the point of view of the actual data base structure. Aggregation and disaggregation of processes stored in the FDB have been tested and have proved feasible (3-11). The FDB is not entirely model-independent despite a wide range in possible applications. It should be noted



that the concept of a "typical" facility with respect to technology and size cannot be applied to primary resource extraction processes because the individual deposit characteristics determine the choice of the appropriate technology as well as the proper resource flows for the construction and operation of a mine. The FDB has therefore been complemented by a substructure on coal mines, taking the individual deposit characteristics into account. The data of the Coal Mines Data Base (CMDDB) is further analyzed by statistical analysis techniques to obtain relational data on the influence of deposit parameters on technology choice and WELMM resource requirements (3-12) and (3-13).

Another characteristic of the FDB is that it covers both the actual transformation process, as well as the construction process of the plant (the fund), in which the transformation takes place. Within the FDB the resource-flows are accounted for either as a total of the construction period or per year of full stream operation. The resource flows accounted for are direct resource requirements, e.g. the resource flows consumed on-site for the construction and operation of a facility (concrete, structural steel, water, chemicals, etc.) and those resources (e.g. metals) embodied in the capital goods of a facility which are physically on site.

After the collection of raw data (through literature, existing data bases or questionnaires) the data is analyzed before computerization. The information on a particular process in the FDB is paradigmatic in the sense that the data stored are "hard" data. The data analysis and judgement is documented within the FDB which contains qualitative and bibliographical information in addition to numerical information.

The data in the FDB is divided into four blocks: process identification (name, location, capacity, etc., including two text files with information on the particular facility as well as on the process(es) it deploys), process characterization (list of primary and secondary inputs and net outputs), WELMM requirements for construction and, finally, WELMM requirements for operation. In the last two, data is accompanied by a quality indicator (ranging from one to five and indicating possible ranges of uncertainty) and a footnote (giving details

about the origin of the data, conversion factors employed and/or alternative data estimates). For the CMDB the data organization is similar, however the particular mine and its deposit is described in more detail in 14 blocks followed by the WELMM requirements for construction and operation (stored as specific values to facilitate further analysis and generation of relational information) and a text footnote.

All data are stored in a relational data base management system called INGRES (developed at the University of Berkeley) (3-14) and (3-15), which operates on top of the UNIX system on a PDP 11/70. Additional interactive programs developed within the WELMM Project for data entry and retrieval, statistical analysis and linking to other data bases on primary resource availability, have been implemented.

Additional documentation on the FDB and the CMDB is contained in the following references: (3-1), (3-6), (3-12), (3-13), (3-16), (3-17), (3-18) and (3-19).

#### 4. THE PROCESS ENCYCLOPEDIA PROJECT OF STATISTICS CANADA

The Structural Analysis Division of Statistics Canada is a research group concerned with the development and operation of 'structural' economic models of the Canadian economy. The first models were comparative static input-output models which were extended in a variety of ways to include energy flows in physical units, employment, interactions among provinces, and prices (4-1). In the last five or six years, the focus of the development work shifted to time structured socio-economic-resource modeling. The first version of such a model, called the Long Term Simulation Model, became operational in 1977 (4-2). A second version of the model, called the Socio-Economic-Resource model will be operational in 1982.

From this experience it became clear that input-output representations of 'production' processes were inadequate for the reasons outlined in section 2 above. As a result it was decided to establish a project to determine the feasibility of compiling a data base of industrial process descriptions; this data base became known as the Process Encyclopedia.

The Process Encyclopedia feasibility study was initiated in early 1979. The United Nations Statistical Office encouraged Statistics Canada to undertake this project as a pilot project related to UNSO's conceptual work on a Materials/Energy Balance Statistical System. Robert U. Ayres was retained as a consultant. The feasibility study consisted of the development of concepts and methodology, the implementation of a prototype computer system using MINISIS on an HP 3000 mini computer, and the compilation of approximately 300 process descriptions of selected energy transformation, chemical, metallurgical, and mining processes. The results of the feasibility study were the subject of an international workshop held in Ottawa on March 26 - 28, 1981. Since then work has proceeded on the design and implementation of a 'production' version of the computer system to be implemented on a VAX 11/780.

In the methodology of the Process Encyclopedia, a process is described by means of three basic sets of information: *Definitional Information*, which consists of a process name, the names of the flows that cross the process boundaries, and the names of the transformation nodes and funds within the boundaries. This definitional information is in fact a directed graph which in the language of the Process Encyclopedia is called topography.

*Relational Information*, which describes the form of the relationships among the flows of a process. This relational information defines the parameters of the functional forms of the generic model.

*Quantitative Information*, which is simply the values of the parameters defined by the generic model associated with the topography of the process.

There can be more than one generic model associated with each topography and in turn there can be more than one set of parameter values associated with each generic model.

The explicit recognition of relational information or generic models permits the representation of non-linear relationships between input flows and output flows, the definition of control variables, and the introduction of time lags between input flows and output flows.

The use of definitional information permits the representation of structure within a process. Funds or stocks can be distinguished from transformation nodes - thus allowing for dynamic modeling applications.

In addition to the three basic sets of information, the Process Encyclopedia contains indexing information in order to access the data base and also 'observations' of processes. 'Observations' are measured values of the input and output flows. A set of observations is associated with a topography which defines the flows. Parametric information may be obtained through the analysis of observations. Provision is also made for including bibliographic information on data sources and quality.

The Process Encyclopedia is a tool for refining process descriptions as well as a data base of 'good' process descriptions. Raw data in the state in which it has been found can be entered into the Process Encyclopedia without transformation. Such data may be incomplete and inaccurate in the sense that it has not been subject to mass and energy balances or other filter edits. This raw data can be analyzed and refined within the Process Encyclopedia and the resulting 'good' or paradigm process can be saved in the data base. The Process Encyclopedia has been extensively documented (4-3).

## 5. CONCLUSIONS

Process descriptions, which encompass definitional, relational, and quantitative information are an appropriate set of building blocks from which to build process models (which can be overlaid by behavioural models). Process units can be defined and quantified in such a way that they are neither site nor nation specific.

Process descriptions that are complete in the sense that they encompass all input and output flows including those associated with stocks can serve a wide variety of modeling applications. The mathematical form of the description of each process should be 'natural' and independent of the mathematical form of the model(s) which may use the process description. The process descriptions can be transformed by the model builder as required.

Process information is a higher level (more detailed) of information to support modeling activities than any other type of information (e.g. Input-output tables (I-O tables)). If the process information system is adequately designed it is model- (or application-) independent. This in turn implies that the user of process information has to generate his own "paradigmatic" process data out of the process information system. The process information system therefore has to support this "clean" data generation with appropriate tools and/or should store flagged "hard" data, which has been analyzed already for a certain field of application (e.g. the WELMM Facility Data Base data).

Process information is not necessarily country or site specific if collected at the appropriate level (e.g. the engineering level). Engineering information of this type as well as the use of *physical indicators* result in data validity over long periods (unlike economic data or I-O tables). Also within the concept of process analysis, the dynamics of a given system (e.g. introduction of new technologies, changing economic or resource intensiveness of a particular process (e.g. exploration or mining)) in the long-term can be dealt with relatively easily (again unlike I-O coefficients).

Process data is information-rich, i.e. the building of process information systems is a data-intensive, long-term activity that requires continuity. Because process information has long-term validity it is possible to build up process information systems (provided they are carefully and flexibly

designed) over long periods, starting first with specific application-orientated process information and enlarging the system later on until sufficient information becomes available for the process information system to be model - or application - non-specific. This goal can only be achieved if the information system is built up through an interdisciplinary and international effort. Because process information is not country or site-specific, problems should not occur in the exchange of process information on the engineering level. However, because of the long-term nature of the exercise of process information collection, the information system should be hosted within an environment assuring continuity, and adequate (computer and manpower) resources to support a certain level of permanent effort. This environment would best be provided by a governmental statistical office.

It is obvious that IIASA cannot be 'collector' of process information. It lacks the computer and staff facilities to support such a program. In principle, IIASA could be a significant user of process information for the development of a wider range of application specific models. With appropriate computer networks, IIASA could be a significant disseminator of process information. As a user of process information, IIASA would inevitably contribute process descriptions to such a data base. But the main role of IIASA would be to promote and coordinate interdisciplinary and international process information exchange. The task of process information system build-up and maintenance should, however, remain with the institution that can assure the continuity and adequate resources required for this task.

Once such a process information data base is developed and built up, it can easily be enriched and enlarged whenever it is accessed for specific applications which in return will generate additional information to be included into the system. Again in order to promote this access or straight link to other process information systems (such as those described in this paper) it is necessary to ensure a certain minimum, permanent effort in process information systems construction.

REFERENCES

- 2-1 Koopmans, T.C. ed. (1951). Activity Analysis of Production and Allocation. Cowles Foundation Monograph 13. New Haven. Yale University Press.
- 2-2 Georgescu-Roegen (1970). "The Economics of Production". American Economic Association, Volume LX, May 1970, No. 2.
- 2-3 Leontieff, W. (1951). The Structure of the American Economy 1919-1939. Second Edition. New York. Oxford University Press.
- 2-4 Matuszewski, T. (1972). "Partly Disaggregated Rectangular Input-Output Models and their Use for the Purposes of a Large Corporation." Pages 301-318, Input-Output Techniques, A.P. Carter and A. Brody eds. Amsterdam. North Holland.
- 2-5 Gigantes, T. (1970). "The Representation of Technology in Input-Output Systems." Pages 270-290, Contributions to Input-Output Analysis. A.P. Carter and A. Brody eds. Amsterdam. North Holland.

- 2-6 Bower, B.T. (1964). Industrial Water Utilization: Substitution Possibilities and Regional Water Resources Development. Proceedings of the Regional Science Association, Western Section, Urban Systems Report No. 1, Tempe, Arizona State University.
- 2-7 Hudson, E.A. and P.W. Jorgenson. (1978). "Energy Policy and U.S. Economic Growth." American Economic Review, Papers and Proceedings, May, Vol. 68, No. 2, Page 118-130.
- 2-8 Russell, C.S. and W.O. Spofford Jr. (1972). "A Quantitative Framework for Residuals Management Decisions." Pages 115-179, Environment Quality Analysis. A.V. Kneesse and B.T. Bower eds. Baltimore and London. The John Hopkins University Press.
- 2-9 Russell, C.S. (1973). Residuals Management in Industry Baltimore and London. The John Hopkins University Press.
- 2-10 Kydes, S. and J. Rabinowitz. (1981). "Overview and Special Features of the Time - Stepped Energy System Optimization Model (TESOM)." Resources and Energy, 3; 65-92. North-Holland Publishing Co.
- 2-11 Pilati, D.A. and F.T. Sparrow. (1980). "The Brookhaven Process Optimization Models." Energy Vol. 5 pp. 417-428. Pergamon Press Ltd.
- 2-12 Carasso, M. et. al. (1975). The Energy Supply Planning Model Volume 1: Model Structure and Use. NTIS PB-245 382.
- 2-13 Clark, N.C. and C.S. Holling. (1979). "Process Models, Equilibrium Structures, and Population Dynamics: On the Formulation and Testing of Realistic Theory in Ecology." Fortschr Zool. 25, 23, pp. 29-52. Gustav Fisher, Stuttgart, New York.



- 3-1 Grenon, M., and B. Lapillonne. (1976). "The WELMM Approach to Energy Strategies and Options." IIASA RR-76-19. Laxenburg, Austria.
- 3-2 Medow, S. (1979). "ENERTREE Data Base System." Fourth IIASA Resources Conference. 1979. IIASA-RSI Conference on Systems of Energy and Mineral Resources. Systems Aspects of Energy and/or Mineral Resources Management. July 9-14, Laxenburg, Austria. Preprints Vol. IV.
- 3-3 Grenon, M., and S. Medow. (1979). "Some Questions on World Oil Resource Data Bases." (3-2)
- 3-4 Merzeau, J.M. (1980). "Resources Group Progress Report on Unconventional Oil Studies." IIASA WP-80-28. Laxenburg, Austria.
- 3-5 Gourmelon, D., and M. Grenon. (1979). "IIASA's Resources Data Bases." (3-2)
- 3-6 Gruebler, A. (1980). "Resource Requirements for Industrial Processes: A WELMM Comparison of Energy Chains." IIASA WP-80-50. Laxenburg, Austria.
- 3-7 Grenon, M., A. Gruebler, and J.M. Merzeau. (1979). "Systems Aspects of Development of New Liquid Fuel Sources." (3-2).
- 3-8 Emard-Katsonis, F., and D. Gourmelon. (1981). "Resources and Economic Assessment of Centralized and Decentralized Solar Electric Systems." Draft of Final Report for the Electric Power Research Center, Palo Alto, U.S.A.
- 3-9 Resources Group, IIASA. (1979). "Comparaison WELMM de scénarios énergétiques. Revue de l'Energie No. 316. Paris, France.
- 3-10 Gruebler, A. (Forthcoming). "Natural Resource Commitments for Long Term Energy Supply Options" in: International Scientific Forum on Changes in Energy. B. Kursunoglu, A. Perlmutter, M. Bauer, L. Scott (eds.). D.C. Heath and Co. Lexington, Massachusetts, U.S.A.

- 3-11 Kopytowski, J.A., J. Wojtania, M. Zebrowski. (1981).  
"Fossils as Key Resources of Hydrocarbons for the  
Chemical Industry." IIASA CP-81-20, Laxenburg,  
Austria.
- 3-12 Astakhov, A., M. Grenon, and A. Gruebler. (1979).  
"A Data Base for Coal Mines." (3-2).
- 3-13 Astakhov, A., M. Grenon, and A. Gruebler. (Forthcoming).  
"Resource Requirements and Economics of Coal Mining  
in the U.S.A. and the U.S.S.R. and Selected Countries."  
IIASA Research Report.
- 3-14 Held, G.D., et al. "INGRES - A Relational Data Base  
System, AFIPS - Conference Proceedings." Vol. 44,  
Montreal, N.J., U.S.A.
- 3-15 Woodfill, J., et al. (1979). "INGRES Version 6.2  
Reference Manual." Memorandum UCB/ERL M79/43.  
University of California, Berkeley, U.S.A.
- 3-16 Cellerier, M. (1979). "The WELMM Facility Data Base."  
(3-2).
- 3-17 Energy Systems Program IIASA, W. Haefele, Program Leader,  
Energy in a Finite World. Vol. II. Paths to a  
Sustainable Future. Chapter 5. pp. 279-306.
- 3-18 Resources Group IIASA. Abstract and Examples of the  
WELMM Resources Data Bases. IIASA, Laxenburg, Austria.
- 3-19 Gruebler, A. Facility Data Base Manual. IIASA Memorandum.  
Laxenburg, Austria.
- 4-1 Structural Analysis Division, Statistics Canada. (1980).  
Users Guide to Statistics Canada Structural Economic  
Models. W.P. No. 80-01-01.
- 4-2 Structural Analysis Division, Statistics Canada. (1977).  
Users Guide to the Statistics Canada Long Term  
Simulation Model. W.P. No. 77-02-28.
- 4-3 Structural Analysis Division, Statistics Canada. (1980).  
"Documentation Index for the Process Encyclopedia."  
W.P. No. 80-02-08.

- 4-4 Page, W.S. (1977). "The Dynamic Representation of an Economic Process - A Systems Approach to Production Function Theory." W.P. No. 77-05-01.
- 4-5 Page, W.S. (1977). "The Process Dictionary - An Application of A Systems Approach to Production Function Theory." W.P. No. 77-09-02.
- 4-6 McInnis, B.C., W.S. Page, and R. Hoffman. (1978). "The Process Encyclopedia, Introduction and Feasibility." W.P. No. 78-10-05.
- 4-7 McInnis, B.C., and W.S. Page. (1979). "The Process Encyclopedia - A Systems Approach to Technology Assessment." W.P. No. 79-05-23.
- 4-8 Ayres, R.U., R.B. Hoffman, B.C. McInnis, and W.S. Page. (1979). "The Process Encyclopedia - A Conceptual Framework." W.P. No. 79-10-01.
- 4-9 McInnis, B.C. and W.S. Page. (1979). "Pictorial Languages for the Process Encyclopedia." W.P. No. 79-10-30.
- 4-10 McInnis, B.C., and W.S. Page. (1980). "Design and Prototype Implementation of the Process Encyclopedia Information System PRECISE, User Level Documentation." W.P. No. 80-02-01.
- 4-11 McInnis, B.C. (1980). "Generic Models for the Process Encyclopedia." W.P. No. 80-05-01.
- 4-12 McInnis, B.C. (1981). "The Process Encyclopedia - Sample Process Data from the Prototype MINISIS Implementation of the Process Encyclopedia." W.P. No. 81-03-15.
- 4-13 McInnis, B.C. (1980). "The Process Encyclopedia - Its Role, Structure and Organizational Environment." W.P. No. 80-04-01.
- 4-14 McInnis, B.C. Editor. (1981). "The Process Encyclopedia Workshop; March 26, 27, 1981." W.P. No. 81-12-20.
- 4-15 McInnis, B.C. (1981). "Indices for the Prototype Process Encyclopedia." W.P. No. 81-03-06.

APPENDIX 1: WELMM PROCESS ENCYCLOPEDIA MAPPING

1.1) FACILITY DATA BASE

1.1.1) GENERAL

In order to describe the data in the Facility Data Base we use the three node topography shown in figure A1. The node T1 represents facility construction. The node S1 represents facility existence and operation and the node T2 represents the actual transformation being accomplished. The generic model that is implicit in the WELMM data base is shown in figure A2 and is represented as three procedures. The first procedure, P1, models facility construction and is driven by the exogenous triplet  $(N, C, t_R)$  being the number,  $N$ , of facilities of capacity,  $C$ , commissioned at time  $t_R$ . The second procedure, P2, represents the ability to control plant operation and we show a switch to let the plant run at constant load factor to represent average conditions over long periods. The third procedure represents the actual transformation of materials and energy.

It may be useful to give a time graph representation of this generic model in the case of one facility being deployed at  $t_R$ . The model represents the time extension of this process in all its flows by a set of simple

'square waves' as shown in figure A3. The generic model itself also allows the simulation of a 'deployment plan',  $N(C^k, t_R)$ , to give the flow structure at the boundary of the system.

1.1.2) FACILITY DATA BASE  $\longrightarrow$  PROCESS ENCYCLOPEDIA

In this section we indicate the mapping from the Facility Data Base, FDB, to the Process Encyclopedia, PE. The FDB is organized by six files called, NEWFAC, NEWINOUT, WELMMCON, WELMMOP, FOOTNOTES and RESOURCE FILE plus a set of conventions we may call a classification system. (See figure 4A) Each FDB file is composed of fields some of which may be subfielded. The PE structure consists of five files or modules each also consisting of fields that may be subfielded. In the FDB there is implicit information about a field because of the file it is stored in and that information is important for the PE definition of record structure thus we also require 'mask' and 'dummy' input to the PE. The specification language we use is in general as follows:

FDBFILE.FIELD.SUBFIELD  $\longrightarrow$  PESUBFIELD.PEFIELD.PEMODULE  
MASK; STRING1.XX.STRING2  
DUMMY; STRING3  $\longrightarrow$  PESUBFIELD.(PEFIELD)

where the convention is that the data in the field or

subfield of the FDB file is to be embedded in the mask repeating through the subfield then field. Dummy data strings are to be entered repeating and if PEFIELD is left out it refers to the PEFIELD in the first line. We may also have some non FDBFILE in the first line. We may also have some non FDBFILE related dummy entries which are indicated as:

DUMMY; STRING→PEFIELD.PESUBFIELD

Each file in the FDB is documented separately so FDBFILE is left out of the next sections.

1.1.2.1) FDB FILE NEWFAC: This file contains 17 fields some of which are subfielded.

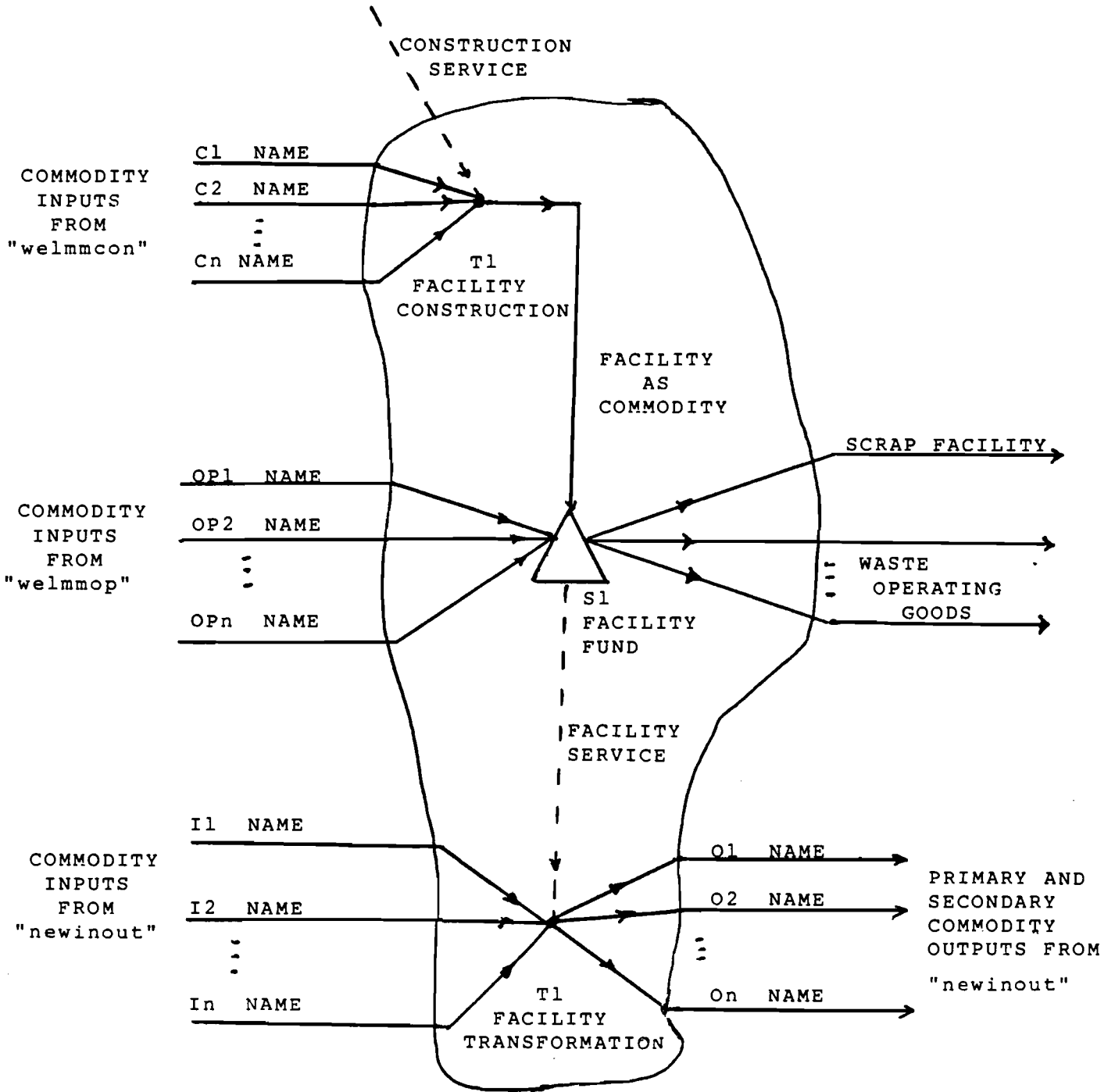
1) CODE OF FACILITY: 8 character field to identify facility

(1.1) CODE OF FACILITY→NOTES GENERAL (A2.1).TOPOGRAPHY  
MASK; WELMM FACILITY DATA BASE CODE.XX

2) NAME: 40 character field used to name process

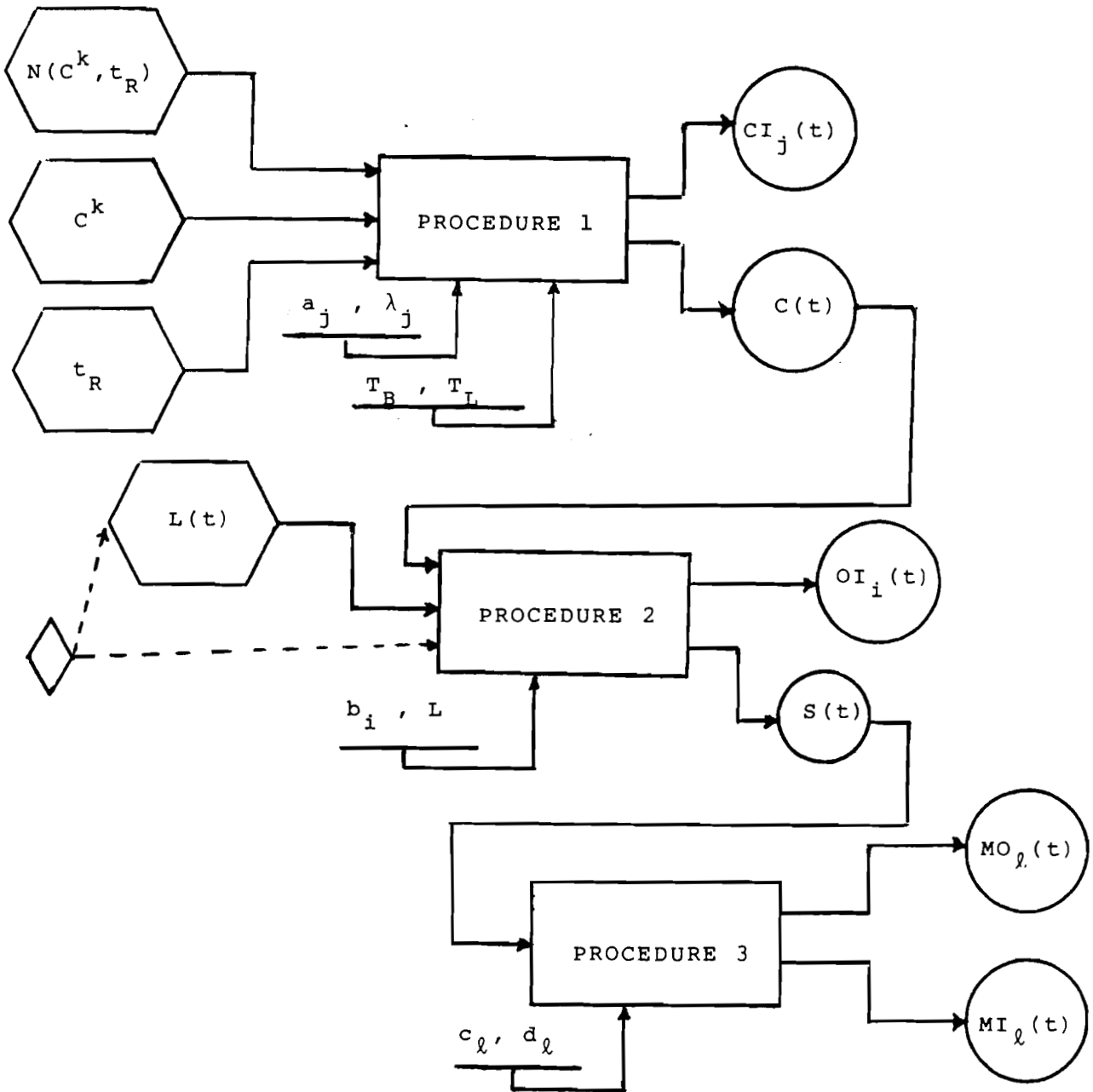
(2.1) NAME→PROCESS NAME (A1).TOPOGRAPHY

3) LOCATION:



TOPOGRAPHY "WELMM FACILITY"

FIGURE A1: TOPOGRAPHY FOR WELMM FACILITY DATA BASE



PROCEDURE 1: 
$$C(t) = \sum_{VC^k} \sum_{t-T_L \leq t_R < t} N(C^k, t_R) * C^k$$

$$CI_j(t) = \sum_{VC^k} \sum_{t \leq t_R < t+T_B} a_j (C^k / C^0) * e^{-\lambda_j (C^k - C^0)} * N(C^k, t_R) / T_B$$

PROCEDURE 2: 
$$S(t) = L(t) * C(t) \quad \text{or} \quad L * C(t)$$

$$OI_i(t) = b_i * S(t)$$

PROCEDURE 3: 
$$MI_l(t) = c_l * S(t)$$

$$MO_l(t) = d_l * S(t)$$

FIGURE A2: GENERIC MODEL FOR WELMM FACILITY DATA BASE



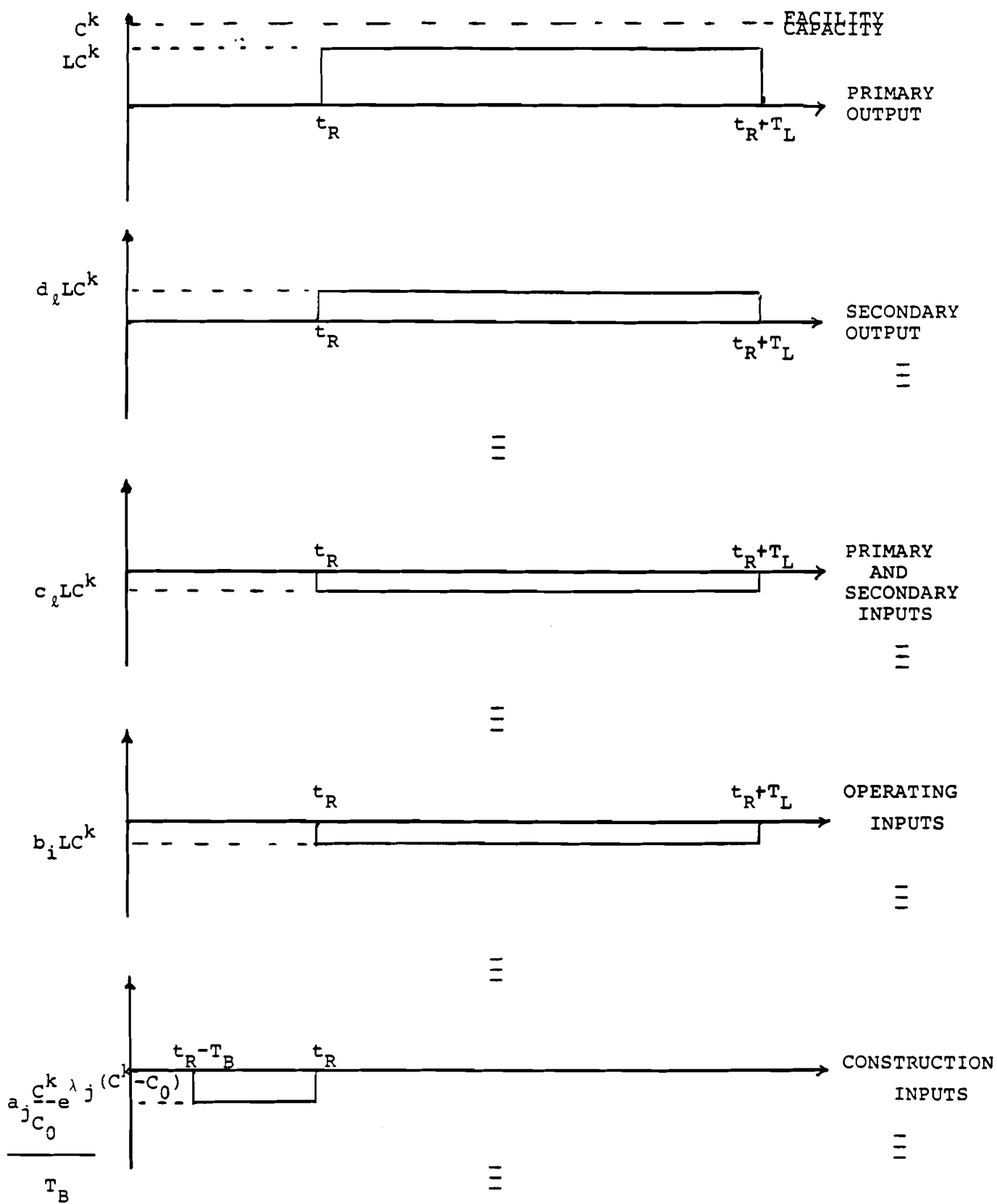


FIGURE A3: TIME GRAPH REPRESENTATION OF "HISTORY" OF A TYPICAL FACILITY IN WELMM:  $N(C^k, t_R) = 1$

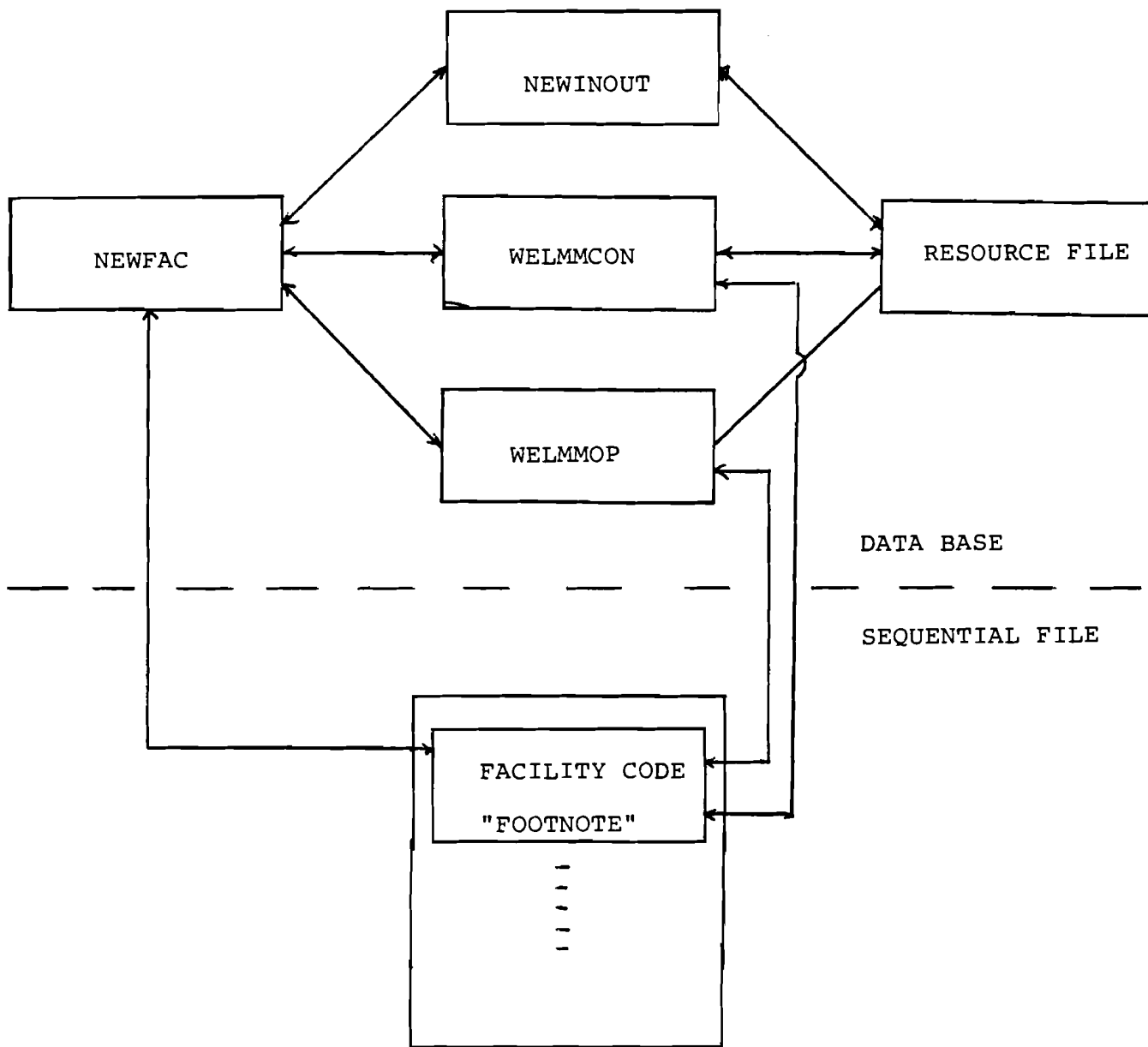


FIGURE A4: FILE STRUCTURE FOR FDB OF WELMM

(3.1) LOCATION → NOTES GENERAL (A2.1). TOPOGRAPHY  
MASK; THIS DESCRIPTION IS FOR THE GEOGRAPHICAL AREA  
DESIGNATED. XX

(4.1) LOCATION → PLANT LOCATION (A6). OBSERVATION

4) PRIMARY CAPACITY:

(5.1) PRIMARY CAPACITY → LEVEL. B1. OBSERVATION

DUMMY; S1 → LABEL

FUND → TYPE

Ø1 → REF. FLOW

5) PRIMARY CAPACITY RESOURCE NUMBER:

(NUMBER, NAME, UNIT, COMMENT) = RESOURCE FILE

(6.1) NUMBER → NOTE NO.. B1. OBSERVATION

(7.1) NAME → TRANS OR FUND NAME. B1. OBSERVATION

(8.1) UNIT → UNITS. B1. OBSERVATION

(9.1) COMMENT → NOTE. NOTES SPECIFIC (A8.2). OBSERVATION

(10.1) NUMBER → NOTE NO.. NOTES SPECIFIC (A8.2). OBSERVATION

6) PRIMARY CAPACITY EQUIVALENT :

- (11.1) PRIMARY CAPACITY EQUIVALENT → NOTES GENERAL (A8.1)  
.OBSERVATION  
MASK; FOR COMPARISON THE PRIMARY CAPACITY EQUIVALENT IS  
.XX
- 7) PRIMARY CAPACITY EQUIVALENT:  
RESOURCE NUMBER :  
(NUMBER, NAME, UNIT, COMMENT) = RESOURCE FILE
- (12.1) UNITS → NOTES GENERAL (A8.1).OBSERVATION  
MASK; IN UNITS OF .XX
- (13.1) NAME → NOTES GENERAL (A8.1).OBSERVATION  
MASK; FOR .XX
- 8) SECONDARY CAPACITY AND SECONDARY CAPACITY:  
RESOURCE NUMBER - no mapping
- 9) EFFICIENCY :
- (14.1) EFFICIENCY → NOTES GENERAL (A8.1).OBSERVATION  
MASK; THE EFFICIENCY OF THE TRANSFORMATION NODE IS.XX
- 10) PLANNING DURATION:  
PLANNING DURATION - no mapping
- 11) CONSTRUCTION DURATION

(15.1) CONSTRUCTION DURATION → VALUE.C.PARAMETER  
DUMMY; TB → GENERIC MODEL SYMBOL  
CONSTRUCTION DURATION → NAME  
YEAR → UNIT OF MEASURE  
C1 → NOTE NO.  
C1 → NOTE NO..NOTES SPECIFIC (A2.2)  
THIS IS TIME BETWEEN CONSTRUCTION START AND  
COMMERCIAL OPERATION START → NOTE.NOTES  
.SPECIFIC (A2.2)

12) LOAD FACTOR:

(16.1) LOAD FACTOR → VALUE.C.PARAMETER  
DUMMY; L → GENERIC MODEL SYMBOL  
LOAD FACTOR → NAME  
UNITLESS → UNIT OF MEASURE  
C2 → NOTE NO.  
C2 → NOTE NO..NOTES SPECIFIC (A2.2)  
(NUMBER OF HOURS (DAYS) FACILITY IS OPERATING IN  
ONE YEAR)/(8160(365)) → NOTE.NOTES  
SPECIFIC(A2.2)

13) LIFETIME:

(17.1) LIFETIME → VALUE.C.PARAMETER  
DUMMY; TL → GENERIC MODEL SYMBOL  
ENGINEERING LIFE TIME → NAME

YEARS → UNIT OF MEASURE

14) REFERENCE YEAR:

(18.1) REFERENCE YEAR → DATE OF OBSERVATION.A4.OBSERVATION

15) FOOTNOTE CONSTRUCTION:

(NUMBER, FOOTNOTE)

(19.1) FOOTNOTE → NOTES GENERAL (A8.1).OBSERVATION

MASK; CONSTRUCTION RELATED NOTES .XX

16) FOOTNOTE OPERATION:

(NUMBER, FOOTNOTE)

(20.1) FOOTNOTE → NOTES GENERAL (A8.1).OBSERVATION

MASK; OPERATION RELATED NOTES .XX

1.1.2.2) FDB FILE NEWINOUT: This file contains 4 fields

1) CODE OF FACILITY: used now to link this file to right  
topography

2) KIND OF INPUT/OUTPUT, VALUE, RESOURCE NO.

(1.2) IF 'KIND OF INPUT/OUTPUT' = XX .. E THEN

KIND OF INPUT/OUTPUT → CLASS.FLOWS (B2.1).TOPOGRAPHY

```
IF 'KIND OF INPUT/OUTPUT' = XIXX THEN
DUMMY; INPT' → ORIGIN.FLOWS (B2.1).TOPOGRAPHY
DUMMY; T2 → DESTINATION.FLOWS (B2.1).TOPOGRAPHY
DUMMY; OUTP → DESTINATION.FLOWS (B2.1).TOPOGRAPHY
DUMMY; ICOUNT1 → LABEL.FLOWS (B2.1).TOPOGRAPHY
DUMMY; ICOUNT1 → LABEL.B2.OBSERVATION
```

```
//VALUE OF INPUT/OUTPUT//
//RESOURCE NUMBER INPUT/OUTPUT//
//(NO., NAME, UNIT, COMMENT)//
```

- (2.2) VALUE OF INPUT/OUTPUT → AMOUNT.B2.OBSERVATION
- (3.2) RESOURCE NO. I/O.NAME → NAME OF FLOW.B2.OBSERVATION
- (4.2) RESOURCE NO. I/Ø.NAME → NAME.B2.TOPOGRAPHY
- (5.2) RESOURCE NO. I/O.UNIT → UNITS.B2.OBSERVATION
- ELSE
- ```
DUMMY; T2 → ORIGIN.FLOWS (B2.1).TOPOGRAPHY
DUMMY; OUTP → DESTINATION.FLOWS (B2.1).TOPOGRAPHY
DUMMY; OCOUNT2 → LABEL.FLOWS (B2.1).TOPOGRAPHY
DUMMY; OCOUNT2 → LABEL.B2.OBSERVATION
```
- (6.2) VALUE OF INPUT/OUTPUT → AMOUNT.B2.OBSERVATION

(7.2) RESOURCE NO. I/O.NAME → NAME OF FLOW.B2.OBSERVATION

(8.2) RESOURCE NO. I/O.NAME → NAME.B2.OBSERVATION

(9.2) RESOURCE NO. I/O UNIT → UNITS.B2.OBSERVATION

1.1.2.3) FDB FILE WELMMCON:

1) CODE OF FACILITY: used to link this file to the right  
topography.

2) RESOURCES NO. (NO., NAME, UNIT, COMMENT):

(1.3) RESOURCE NO..NAME → NAME.B2.TOPOGRAPHY  
RESOURCE NO..NAME → NAME.B2.OBSERVATION  
DUMMY; CCOUNT → LABEL.B2.TOPOGRAPHY  
DUMMY; CCOUNT → LABEL.B2.OBSERVATION  
DUMMY; INPT → ORIGIN  
DUMMY; T1 → DESTINATION

(2.3) RESOURCE NO..NO. → CLASS.B2.TOPOGRAPHY  
MASK; WR. XX

(3.3) FOOTNOTE NO. → NOTE NO..B2.OBSERVATION

(4.3) FOOTNOTE NO. → NOTE NO..NOTES SPECIFIC  
(A.8.2).OBSERVATION



- (5.3) QUALITY OF DATA → NOTE.NOTES SPECIFIC (A8.2)  
.OBSERVATION  
MASK; QUALITY INDICATOR IS. XX
- (6.3) FOOTNOTE NO. → NOTE.NOTES SPECIFIC (A8.2).OBSERVATION
- (7.3) RESOURCE QUANTITY → AMOUNT.B2.OBSERVATION
- (8.3) RESOURCE NO..UNIT → UNITS.B2.OBSERVATION
- (9.3) RESOURCE NO..NO. → NOTE NO..B2.TOPOGRAPHY
- (10.3) RESOURCE NO..NO. → NOTE NO..A2.2.TOPOGRAPHY
- (11.3) RESOURCES NO..COMMENT → NOTE.NOTES SPECIFIC (A2.2)  
.TOPOGRAPHY

1.1.2.4) FDB FILE WELMMOP: This file has the same structure as WELMMCON and the above section (1.1.2.2) is repeated with the following substitutes;

Replace;      CCOUNT      by      OPCOUNT  
                  TI            by      SI  
and            (xx.3)    by      (xx.4)  
for equation numbers.

1.1.2.5) MISC.: We need to fill in the fund and transformation

of the topography. This is done as follows;

For a fixed facility code we have,

(1.5) RESOURCE FILE.NAME → NAME OF FUND OR TRANSFORMATION  
.B1.TOPOGRAPHY

MASK; XX.FUND

DUMMY; SI → LABEL

DUMMY; FUND → TYPE

(2.5) RESOURCE FILE.NAME → NAME OF FUND OR TRANSFORMATION  
.B1.TOPOGRAPHY

MASK; XX. TRANSFORMATION

DUMMY; T2 → LABEL

DUMMY; TRANS → TYPE

(3.5) RESOURCE FILE.NAME → NAME OF FUND OR TRANSFORMATION  
.B1.TOPOGRAPHY

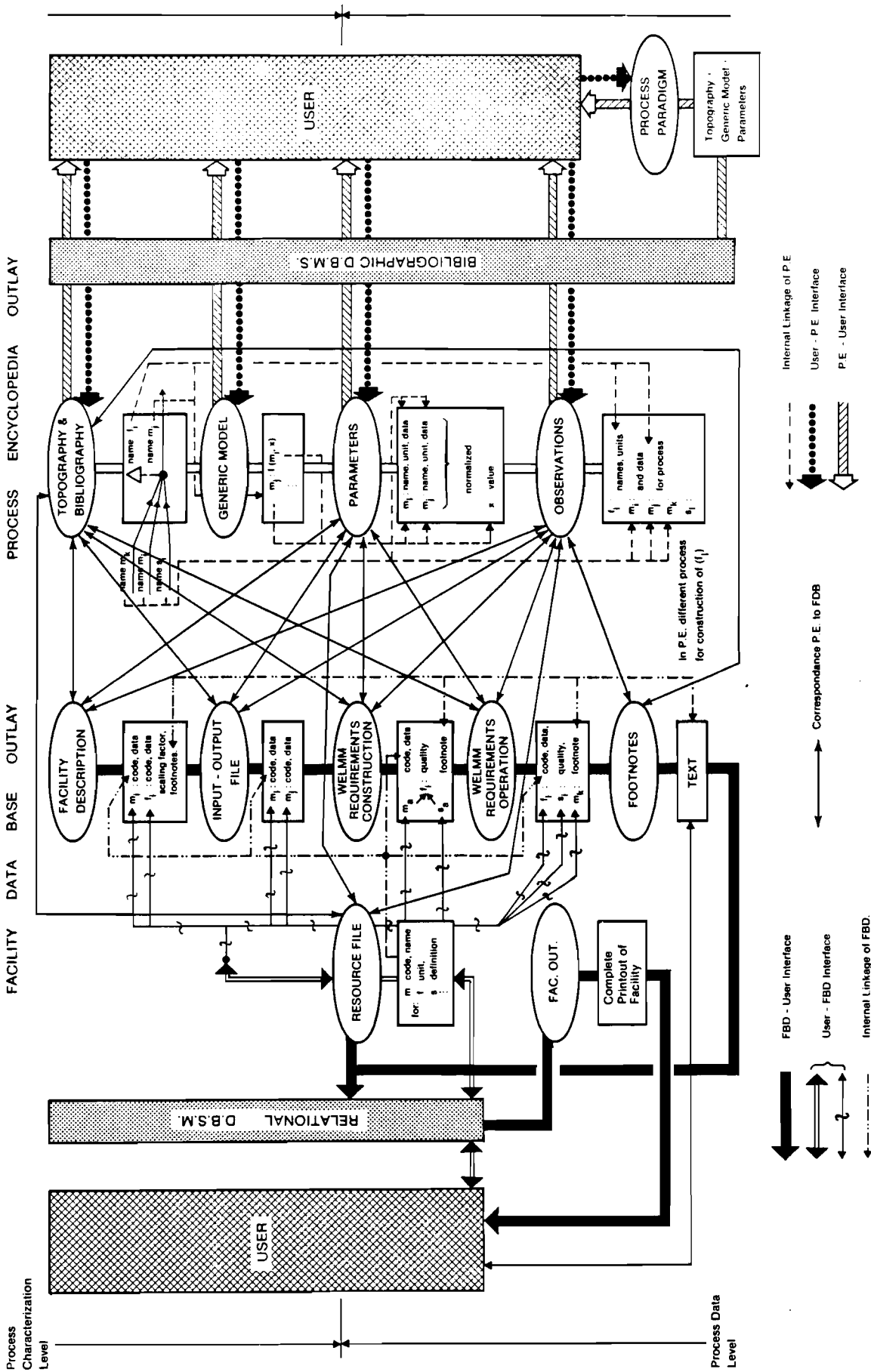
MASK; XX. CONSTRUCTION

DUMMY; T1 → LABEL

DUMMY; TRANS → TYPE

Inputs and outputs will be handled by editing files.  
Furthermore the generic model description will be put in  
by hand to give the sets for the parameters.

A high level diagramatic representation of the PE → FDB  
mapping detailed above is shown on the following page.



High Level Representation of PE → FDB Mapping