Decomposition of a large-scale
energy model

Evgeni Nurminski
Todor Balabanov

October 1982
WP-82-95

Decomposition of a large-scale energy model

Evgeni Nurminski and Todor Balabanov

International Institute for Applied Systems Analysis,

Laxenburg, Austria

ABSTRACT

A large-scale linear programming energy systems model is decomposed and analyzed using new decomposition algorithms, thus revealing some characteristic features of the model. This approach suggests a possible method of constructing a formally integrated system of linked models.

1. BACKGROUND AND OBJECTIVES

In recent years a number of projects concerned with the analysis of international energy systems have been initiated within various institutions, including the International Institute for Applied Systems Analysis (IIASA) [1], the International Energy Agency (IEA) [2], and the European Community [3]. Their primary objective has been to assess the options for long-term development of energy supply systems and the interaction between general economic development and the use of energy at a national or regional level.

The above studies have each been built around a large-scale multiperiod linear programming model of energy supply, which at IIASA is called MESSAGE, and at the IEA is called MARKAL. Current investigations of the relationship between energy use and the introduction of energy conservation measures [4] seem to be leading to the development of additional linear programming models of, for example, industrial or transportation sectors, thus increasing the dimension of the problem.

The overall problem can be considered as one of investment and production planning in an economy divided into an energy sector supplying a specified set of energy forms and several energy-intensive economic sectors to which energy is supplied at a given price. The overall objective is to minimize the total investment and operating cost of the system over some time horizon, which is usually of the order of 65 years.

The scale of the problem and the variety of sectors considered are such that, in a perfect world, the various submodels should be developed by different groups of specialists working in the appropriate fields, but this lies more in the realms of fantasy than reality. One of the main reasons why this approach is not used more in practice is the absence of efficient methods for linking or bringing into agreement the heterogeneous submodels based on different time-scales and different degrees of aggregation, operating with different policy variables and involving different economic agents.

Another possible reason for the infrequent use of this modular principle in the development of large-scale models may lie in the many reported failures (for purely computational reasons) of attempts to implement decomposition ideas in large-scale optimization. This can be overcome by the development of more advanced decomposition techniques and the accumulation of relevant practical experience.

All of the above considerations, combined with the practical necessity of squeezing a large-scale problem into a small computer, motivated the authors to undertake the decomposition of the large-scale IIASA energy supply model MES-SAGE. The theoretical foundation for the particular decomposition algorithms used here is actually a by-product of a study on nondifferentiable optimization currently underway at IIASA. This work has resulted in the development of a family of decomposition algorithms [5, 6] which have already been successfully

applied to medium-scale problems.

2. THE MODEL MESSAGE II

The energy supply model considered in this paper is MESSAGE II, which is currently under development in the Energy Systems Group at IIASA. It is an extended version of the model MESSAGE (Model for Energy Supply Systems Alternatives and their General Environmental impact) [7] also developed at IIASA .

MESSAGE II is a detailed energy supply model in a dynamic linear programming formulation which is designed to compare alternative existing and prospective technologies for the extraction, production, secondary conversion, storage, distribution, and end-use of energy, capable of meeting the set of demands for useful energy specified outside the model. Constraints are imposed by the availability of primary energy resources, the market penetration rates of new technologies and the rates of decline of existing ones, as well as by limitations on capital and materials. The objective is to minimize the total cost, discounted over some time interval.

The model allows a large number of technologies to be included in each of the energy transformation chains, starting from resources and proceeding via central conversion, transmission and decentralized conversion to end use. A detailed description of different types of energy storage and the load distribution of the intermediate energy forms is also included. More detailed information about the model is given in [8].

The comprehensiveness of this approach is achieved at the cost of increasing the size of the resulting linear programming problem (up to 3000 rows and 4000 columns), leading to some difficulties both in obtaining and interpreting the solution. It also hinders the wider application of the model by restricting its

effective implementation and use to computers above a certain size.

The test problem reported here is part of a study undertaken as a follow-up to the global energy systems analysis conducted at IIASA and reported in full in [1]. This later study concentrates on specific energy options for certain Latin American countries, including Mexico and Brazil. In this study the model MES-SAGE II was used to investigate the feasibility and timing of the introduction of a menu of new technologies in order to explore the possibilities of existing and potential energy resources, to examine import/export strategies, as well as to determine the impact of the rapid introduction of nuclear or hydropower plants on the development of the energy supply system. The analysis was constrained throughout by a set of exogenous energy demands.

The case study for Mexico was chosen to test the decomposition algorithm. This example contains a detailed description of the energy forms and technological options considered (see Tables 1 and 2) ,and thus represents a good off-the-shelf problem including all the features characteristic of an energy system study at the national level .

Table 1. Forms of energy considered in the Mexican case study.

Primary	Secondary and final	End-use
Hard coal (indigenous)	Coke	Process heat (high temp.)
Hard coal (imported)	Light fractions of oil	Process heat (medium and low temp.)
Crude oil (onshore)	Heavy fractions of oil	Coal, specific uses
Crude oil (offshore)	Domestic gas	Liquid fuels
Natural gas	Industrial gas	Electricity (industrial)
Uranium	Electricity	Electricity (domestic)
	District heat	Space and water heating
		Cooking

The matrix generating program of the MESSAGE II model is very versatile, and this made it possible for the adaptations required by the decomposition algorithm to be made relatively easily.

Table 2. Technologies considered in the Mexican case study.

Extraction/production	Conversion	Transportation	End-use

Oil:
• offshore
• onshore
 	
Coal:
• indigenous
• imported
 	
Coking coal
 	
Uranium
 	
Natural gas
	Power plants:
	• liquid fuels
	• gaseous fuels
	• diesel
	• hard coal
	• municipal waste
	• hydropower
	• pumped hydropower
	• solar electric
	• light water reactor
	Liquefaction of coal
	Basic refining
	Cracking of heavy distillates
	Methanol from natural gas
	Gasification of coal
	Cogeneration of heat
	  and electricity
	Production of heat from
	  municipal waste
		Pipeline:
		• light liquids
		• gases
		Transport of fuel oil for industry
		Trucks (coal)
		Electricity transmission network
		District heating network
			High-temperature process heat derived from:
			• coking coal
			• gas
			• fuel oil
			• electricity
			Low-temperature process heat derived from:
			• soft solar
			• district heat
			• heat pump
			• electricity
			• coal
			• gas
			• fuel oil
			Industrial use of electricity
			Cogeneration of heat and electricity from:
			• gas
			• fuel oil
			• coal
			Cooking:
			• natural gas
			• electricity
			• kerosene
			Space and water heating:
			• solar
			• electricity
			• oil
			• gas

It can be seen from Table 2 that there is only a small number of intermediate final energy flows, reflecting the exchange between the energy sector and decentralized users, and this gives us the opportunity to divide the model into two parts :

S1. The first submodel (called CENTR) describes the production chains for a given set of final energy forms from sources such as fossil and nuclear fuels, solar energy and hydropower (see Tables 1 and 2). The final energy forms are electricity, district heat, light and heavy fractions (residual) of oil, coal, gaseous fuels, and metallurgical coke (Table 1).

S2. The second submodel (called END) is concerned with the transformation of final energy into useful energy forms. It describes the flows of final energy through the different stages of transformation, distribution, and on-site conversion to meet the demands of end-users.

A schematic representation of the model structure is given in Figure 1.

The matrix statistics of the submodels are compared with those of the undecomposed model (COMBINED or COMB) in Table 3.

Table 3. Matrix statistics of the original problem (COMBINED) and the two subproblems (CENTR, END).

Problem	Total	Normal	Free	Fixed	Bounded	Matrix elements	Density
COMBINED						11276	0.306
Rows	2132	1851	201	80	0		
Columns	1729	1555	0	67	63		
CENTR (19-th cycle)						10546	0.435
Rows	1581	1243	201	137	0		
Columns	1532	1416	0	53	63		
END (19-th cycle)						5181	1.234
Rows	724	626	33	65	0		
Columns	580	519	2	15	44		

Table 3

The linking variables (LINKS) are the flows of final energy between submo-

dels, and could be interpreted as the energy supplied by the energy sector to the consumers. These linking variables are listed in Table 4.

Table 4. Linking variables.

| Link	Energy form	Load region
1	Electricity	1
2	Electricity	2
3	Electricity	3
4	Electricity	4
5	Electricity	5
6	District heat	1
7	District heat	2
8	District heat	3
9	District heat	4
10	District heat	5
11	Direct utilization of coal	
12	Light oil fractions	
13	Heavy oil fractions	
14	Gaseous fuels	
15	Metallurgical coke	

Table 4

This table gives the linking variables for one time period. The model was set up for eight such periods, bringing the number of linking variables to 120. We assume a time horizon of 65 years, subdivided into three steps of five years and five steps of ten years. This horizon was chosen to allow time for the depreciation of existing capital stock and the penetration of new energy technologies .
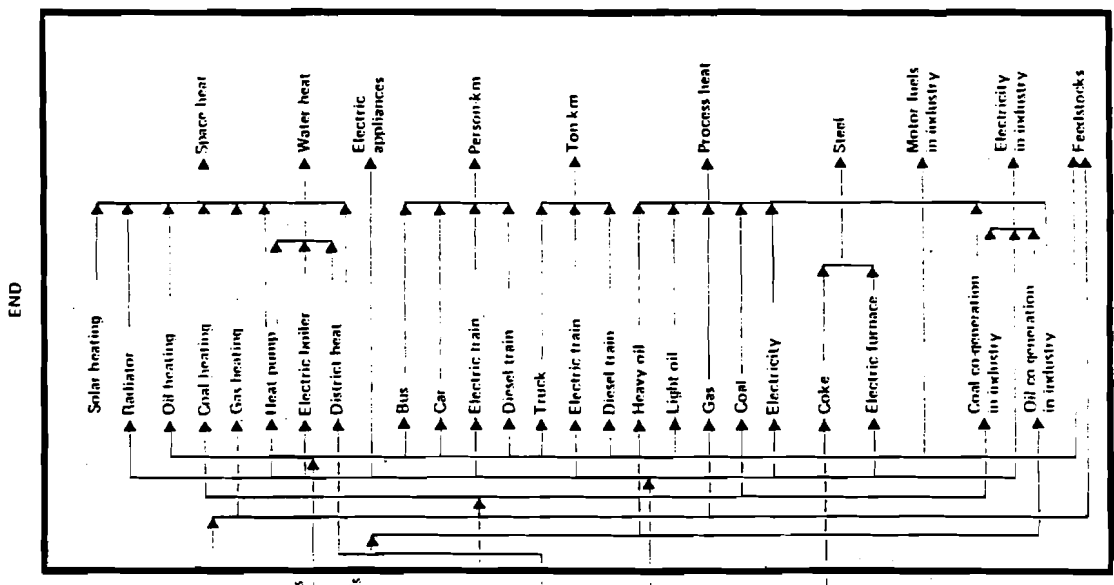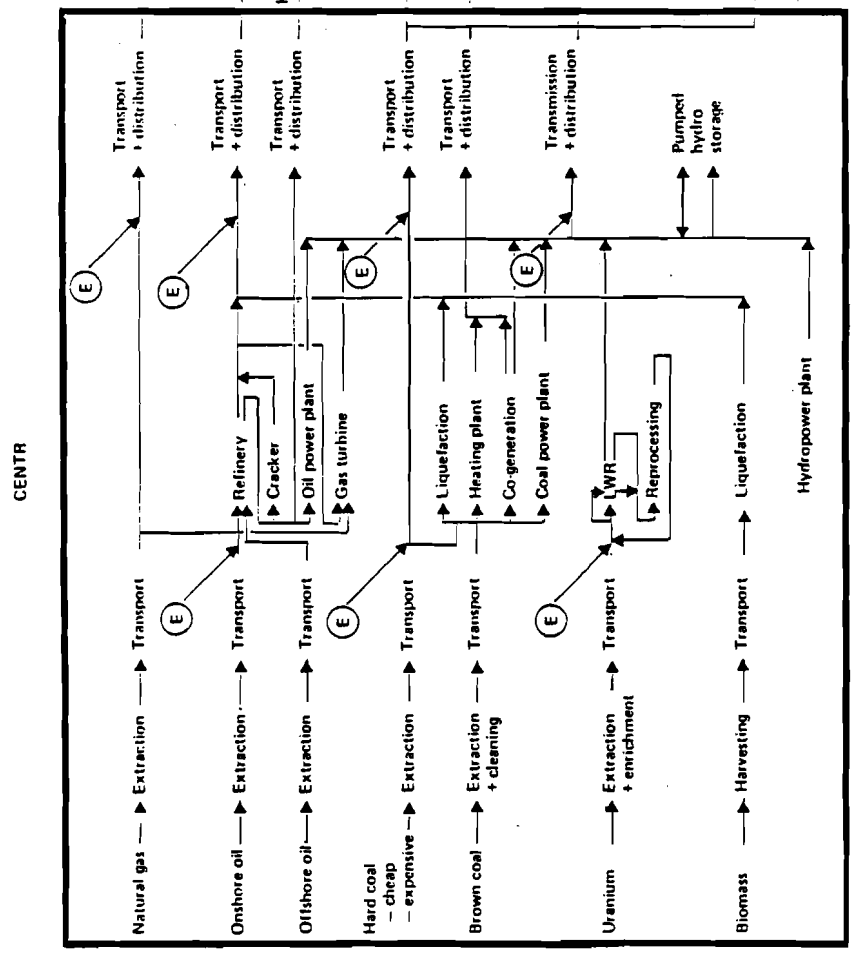
3. DECOMPOSITION ALGORITHMS

The problem discussed above may be described formally by the following two-block linear programming problem with linking variables:

$$\min \{ \, c_A z_A + c_B z_B \, \} \qquad (1)$$

$$A_A z_A \quad + \quad B_A x \leq b_A \qquad \cdot$$

$$A_B z_B \; + B_B x \leq b_B$$

By defining subproblems

$$f_A(x) = \min c_A z_A \qquad (A)$$

$$A_A z_A \leq b_A - B_A x$$

$$f_B(x) = \min c_B z_B \qquad (B)$$

$$A_B z_B \leq b_B - B_B x$$

problem (1) can be restated as the problem of finding the optimum value x^{*} of linking variables

$$v^{*} = \min_{x} \{ \, f_A(x) + f_B(x) \, \} \; = f_A(x^{*}) + f_B(x^{*}) \qquad (2)$$

Variables z_A and z_B can be viewed as internal variables of subproblems (A) and (B), respectively, with values which become known after solution of subproblems (A) and (B) with fixed optimal linking variables. Functions $f_A(x)$ and $f_B(x)$ are piecewise linear functions attaining possibly infinite values for those x which result in empty feasible sets for (A) or (B). We shall refer to (2) as the primal form of problem (1).

The application of standard convex duality theory to problem (2) leads to the following equality:

$$\min_{x} \{ f_A(x) + f_B(x) \} = -\min_{p} \{ h_A(-p) + h_B(p) \} \qquad (3)$$

where $h_A(-p)$, denotes the conjugate of a convex function $f_A(x)$:

$$h_A(-p) = \sup_{x} \{ -px - f_A(x) \} = -x_p p - f_A(x_p)$$

$$= -\min \{ c_A(z_A) + p\,x \}$$

$$A_A\, z_A + B_A\, x \leq b_A$$

The conjugate function may be interpreted as the pay-off obtained in the subproblem for given prices p of linking variables x.

Two different decomposition algorithms were applied to problem (2). The first algorithm, proposed in [5], is based on the idea of replacing (2) by the sequence of problems

$$\min_{x} \{ f_A(x) + f_B^k(x) \} = v_k \leq v^* \qquad (4)$$

where $f_B^k(x)$ is the approximation of the function $f_B(x)$ obtained on the k-th iteration.

It was suggested that this approximation should be derived by constructing a piecewise linear support function for $f_B(x)$ based on the values of this function and its subgradient computed at the solutions of auxiliary problem (4). This approximation is gradually refined, directing the sequence of solutions of the auxiliary problem (4) toward the solution of the problem (2).

The resulting algorithm performs quite satisfactorily for small and medium-sized problems [5]. On applying it to the dual part of (3) it is possible to use the optimal solution obtained in one major iteration as a starting basis for the next cycle, with the result that the number of auxiliary simplex iterations decreases rapidly as the algorithm progresses.

However, this algorithm does not make full use of the information available during the optimization process. Another drawback is that it does not produce both upper and lower estimates of the optimum, which makes it difficult to determine the rate of convergence.

Also, in many practical cases, information on the solution of the dual form of problem (3), which may be interpreted as a set of shadow prices for linking variables, can provide additional insight into the qualitative properties of problem (1). This information is not readily available even if the solution of the primal form is known; substantial further analysis of the problem is required to reveal it. Similarly, if the algorithm is applied to the dual form of problem (1) then the primal solution cannot be found immediately.

These considerations stimulated the development of another algorithm, which is based on the simultaneous use of approximation in the primal and dual formulations of problem (3). This second algorithm provides both primal and dual solutions of problem (3), supplies upper and lower estimates of the optimum during solution, and, as the numerical experiments show, converges more rapidly than the first.

The theoretical basis of this algorithm is described in some detail in [6]; here we simply explain the underlying ideas.

The main cycle of the algorithm involves the solution of two auxiliary problems:

$$\min_{x} \; \{ f_A(x) + f_B^k(x) \} \; = f_A(x^k) + f_B^k(x^k) = v_k \tag{P}$$

and

$$\min_{p} \; \{ h_A^k(-p) + h_B(p) \} \; = h_A^k(-p^{k+1}) + h_B(p^{k+1}) = -w_k \tag{D}$$

where the solution of problem (P) is used to update the approximation of the

function $h_A(-p)$:

$$h_A^k(-p) = \max \{ h_A^{k-1}(-p) , -x^k p - f_A(x^k) \}$$

and the solution of problem (D) is used to update the approximation of the function $f_B(x)$:

$$f_B^k(x) = \max \{ f_B^{k-1}(x) , xp^k - h_B(p^k) \}$$

When solving the auxiliary optimization problem (D) it is again possible to use the preceding optimal solution as a starting point for each new cycle; for the problem (P), however, the previous optimal solution is not feasible but it can still be used as an advanced starting basis for the next iteration. Both strategies lead to a rapid decrease in the number of simplex iterations performed in solving subproblems (P) and (D).

Both algorithms were implemented on a VAX-11/780 computer under the UNIX [9] operating system, using the code MINOS [10] to solve the auxiliary linear problems. For simplicity the auxiliary subproblems were formulated and updated through modification of the input files.

This is clearly not the most efficient way to implement the algorithm, but at this stage we are more concerned with the number of major iterations required than with computational efficiency as a whole. One advantage of this approach was the small amount of additional programming needed to supply codes for generating updated input files: UNIX functions proved very useful in this respect.

4. SOLUTION OF THE PROBLEM

It is clear from the theoretical description of the algorithms that they are unsymmetrical with respect to the subproblems into which the original problem (1) is divided. In the primal decomposition algorithm, subproblem A is considered in its full form while subproblem B is approximated. The primal-dual

algorithm again considers A in its full form but in this case subproblem B is also represented in full, though in dual form.

The computational performance of the algorithms can depend quite strongly on which of the subproblems is considered in full, and this may partly explain the different computational experiences with the decomposition approach reported in the literature or passed along the scientific grapevine. We will call the subproblem which is considered in full in the primal problem the *primal master problem*, and that which is considered in full in the dual problem, the *dual master*.

At this stage it would be difficult to give any sound recommendation as to which subproblem should be taken as master. The rule of thumb, however, is to take the most complex problem as the dual master, where complexity could simply reflect the size of the problem.

The whole discussion is confused by the fact that the primal decomposition algorithm was applied to the dual of the initial problem. However, we will still describe it as the primal decomposition algorithm. This particular implementation of the algorithm is referred to below as DEC-1.2.

In our experiments CENTR was chosen as the (dual) master subproblem, which means that it acted as a pricing device for subproblem END. The roles of the subproblems and the structure of information exchange are illustrated in Figure 2.

The prices p provided by CENTR are used by DEC-1.2 to price the linking variables in subproblem END, and the proposals x generated in this subproblem are then used to update the approximation of subproblem END in CENTR.

This distribution of roles was based firstly on the difference in the complexity of the subproblems (subproblem CENTR is essentially larger than END) and, secondly, on some preliminary computational experience with a small scale
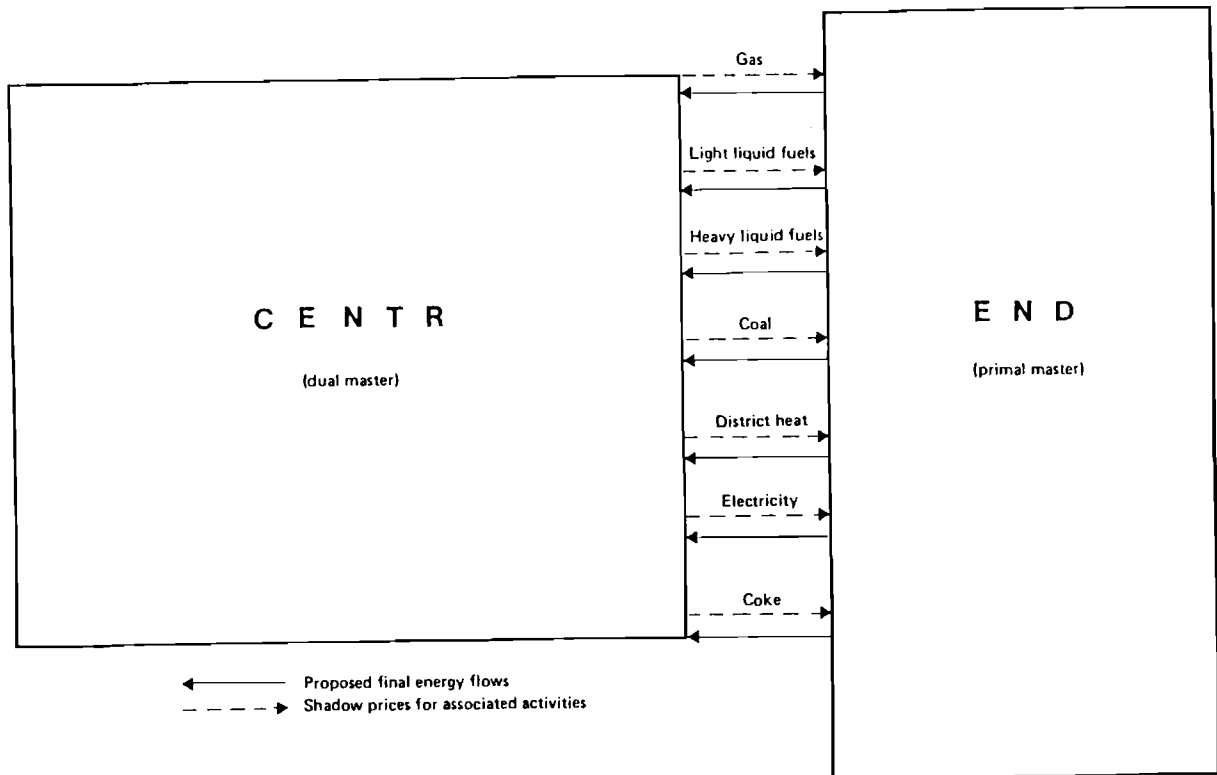
Figure 2.
Interaction between subproblems CENTR and END

version of this problem [5].

The same distribution of roles was preserved for the primal-dual algorithm with the difference that the prices p provided by CENTR were used by END not only to price linking variables but also to update the approximation of subproblem CENTR in END. The implementation of the primal-dual algorithm used in these experiments is referred to below as DEC-2.3.

Both algorithms were first tested on a limited run (12 and 19 iterations , respectively) and the results are shown in Figure 3.

This graph shows the convergence of the upper and lower bounds for the primal-dual algorithm (continuous and dotted lines, respectively), and of the upper bound for the primal algorithm (dashed line). The figure illustrates the relative accuracy (on a logarithmic scale) of each bound, which is calculated as
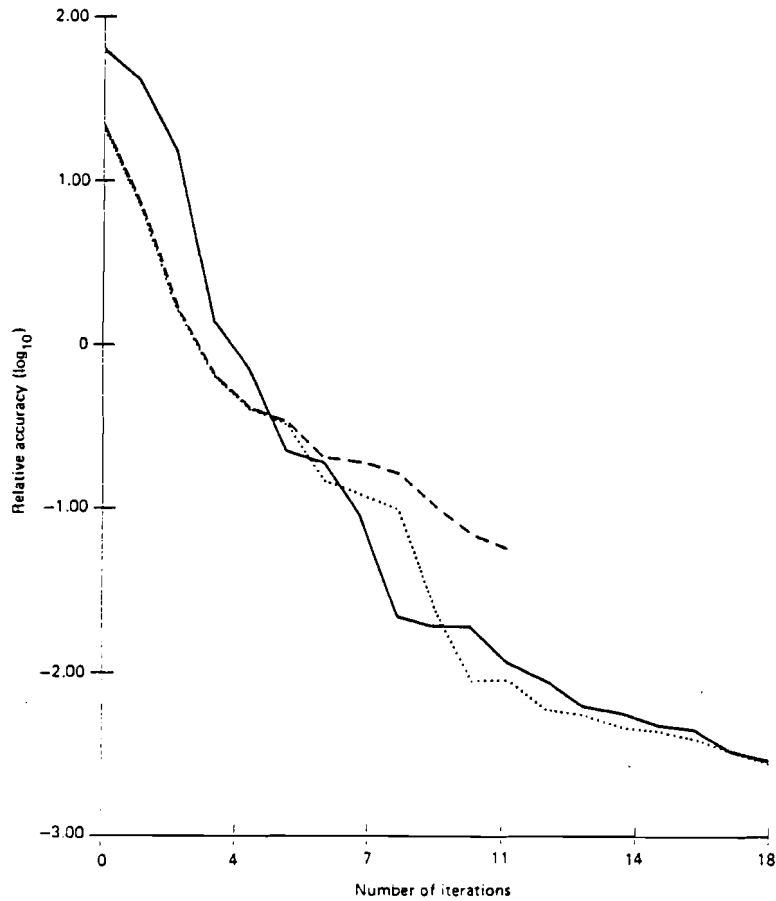
Figure 3
Convergence of DEC-1.2 and DEC-2.3

$$y_i = \frac{|f_i - f^*|}{f^*}$$

where f_i denotes the value obtained for the objective function on the i-th cycle and f^* denotes the optimal value obtained from the subsequent computations.

It was clear that algorithm DEC-2.3 converged more rapidly than DEC-1.2 and so further experiments were carried out with the primal-dual algorithm alone.

The accuracy of the solution (with respect to the objective) is about 5 percent after 19 major iterations. To analyze further convergence the run was extended to 32 cycles, bringing the accuracy to 0.1 percent. The results
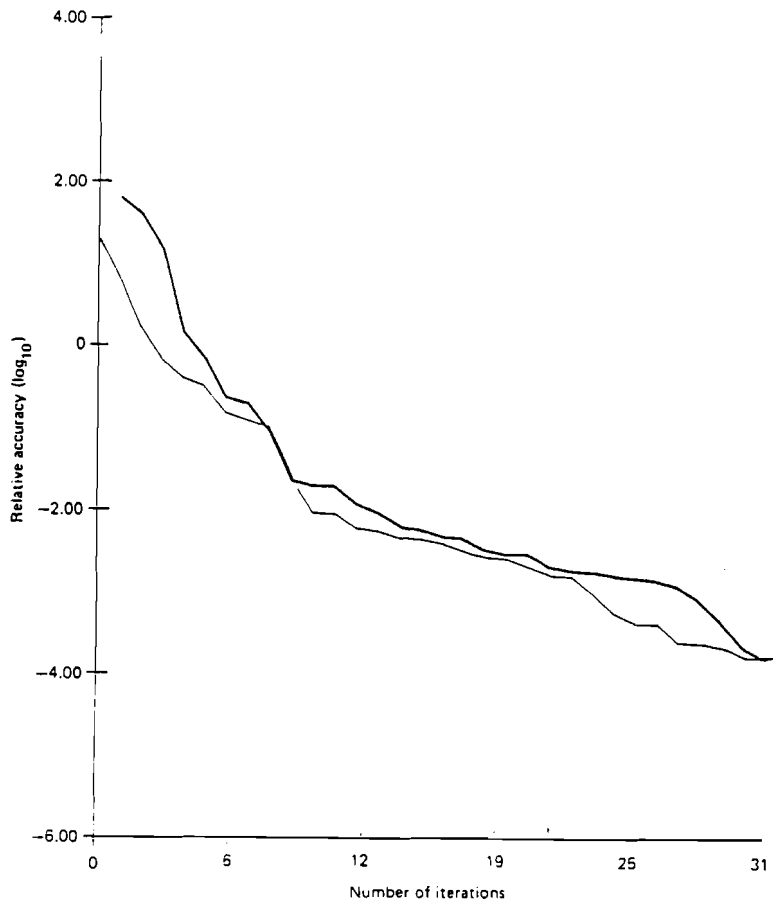
obtained are shown in Figure 4.

Figure 4
Convergence of DEC-2.3

This algorithm displays a geometrical rate of convergence which compares favorably with the long tail of slow convergence in the final iterations typical of the Dantzig-Wolfe algorithm.

The number of local iterations is shown in Figure 5.

Table 5 shows the total number of iterations and computed CPU time for subproblems CENTR and END for both algorithms. The average CPU time per local iteration is 0.891 sec for CENTR and 0.634 sec for END . These values were used to estimate the CPU time for the whole run.

The data for total elapsed time are not meaningful in this case because of
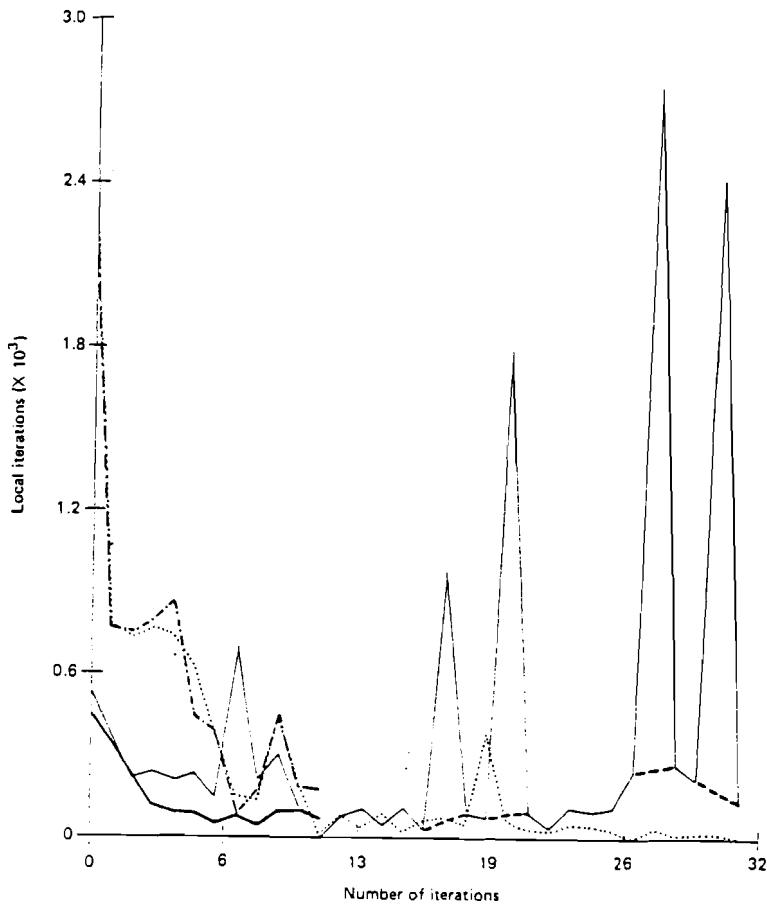
Figure 5
Numbers of local iterations

The dot-dash line and continuous bold line represent the number of local iterations for CENTR and END, respectively, under DEC-1.2. The dotted and continuous lines represent the number of local iterations for CENTR and END, respectively, under DEC-2.3.

Table 5. Total number of iterations and estimated CPU time to solve subproblems CENTR and END, for both algorithms.

Algorithm	Major iterations	Local iterations		User time (est.)	
		CENTR	END	CENTR	END
DEC-1.2	12	7471	1725	6656.7	1091.9
DEC-2.3	19	7760	4283	6914	2715
	32	8496	13147 (est. 6680)*	7570	8335 (est. 4235)*

*Values that would have been obtained if the LP solver had not failed three times.

the large amount of time necessary for read-write operations.

The number of iterations performed by DEC-2.3 for subproblem END was also distorted by repeated malfunctions of the linear program solver, which necessitated starting from scratch on major iterations 21, 28, and 31 (see Table 6).

Table 6. Failed iterations.

Iteration	CENTR	Upper bound	END	Lower bound
21	64	0.941828d + 05	1794	0.937126d + 05
28	39	0.940096d + 05	2765	0.938754d + 05
31	16	0.940026d + 05	2431	0.939661d + 05

These "cold starts" naturally required more local iterations than would otherwise have been the case, but this can be attributed to the fact that the LP solver is not completely reliable rather than reflecting any fault in the algorithm. If the numbers of local iterations corresponding to these failures are replaced by the average number of local iterations obtained in the major iterations on either side, the total number would be reduced considerably (estimated values in Table 5).

The solution of the COMBINED problem required about 10000 iterations and about 24000 sec of CPU time.

5. ANALYSIS OF THE OPTIMAL SOLUTION

The optimal levels of the primal and dual variables obtained using the decomposition algorithm described above are actually the trajectories of the final energy flows and their shadow prices over the specified time horizon. Some of these trajectories are illustrated in Figs. 6 and 7.

The division of the Mexican case modeled via MESSAGE II into two submodels (CENTR and END) chosen here allows us to analyze the sensitivity of the
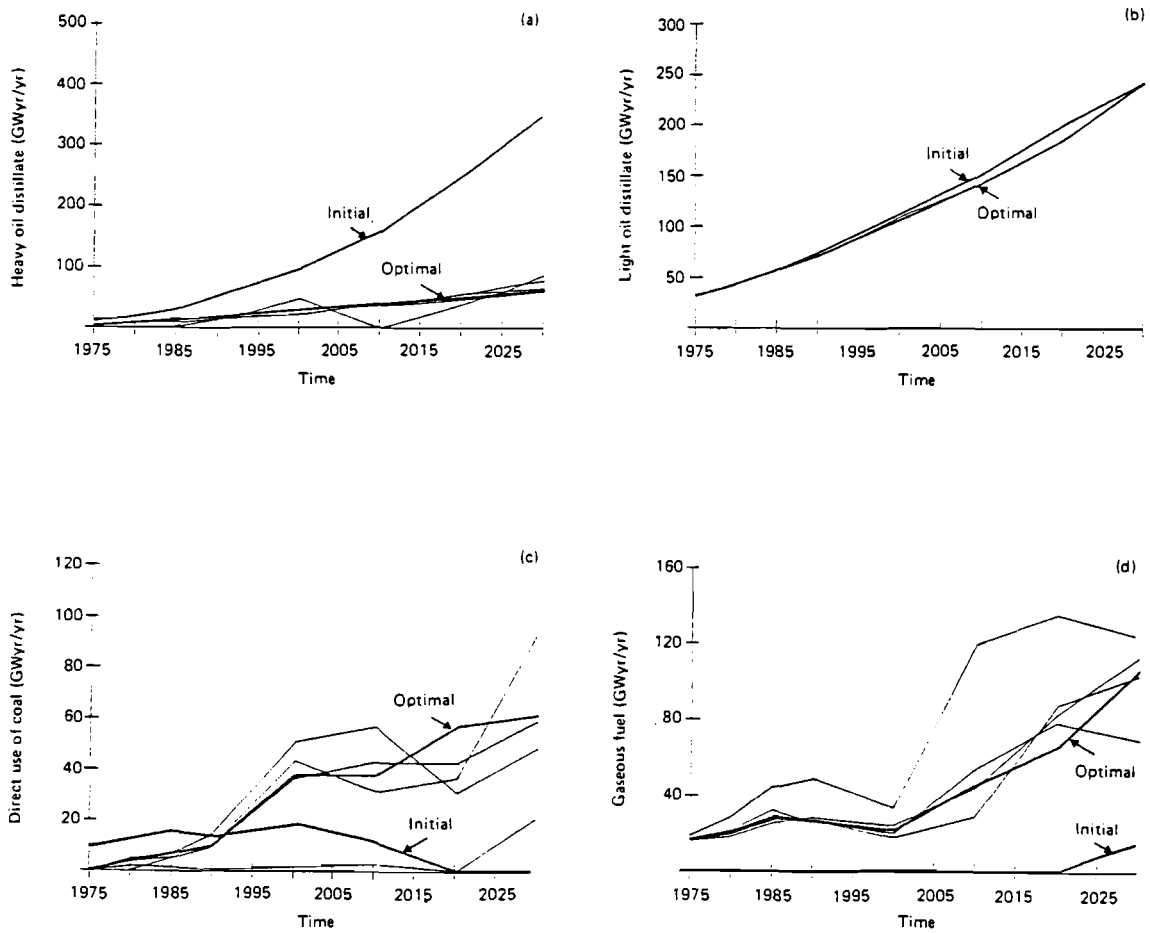
Figure 6
Final energy flows over the planning horizon for selected energy forms

optimal solution of each submodel and the relative contribution of each submo-
del to the optimal solution. This may be achieved by separate consideration of
the two submodels in the environment provided by the joint optimal solution
(optimal levels of activities and shadow prices).

Conceptually MESSAGE II can be considered as a demand driven model, with
the aim of transferring resources via technological chains into a specified
demand vector, and therefore the levels of activities are of primary importance.
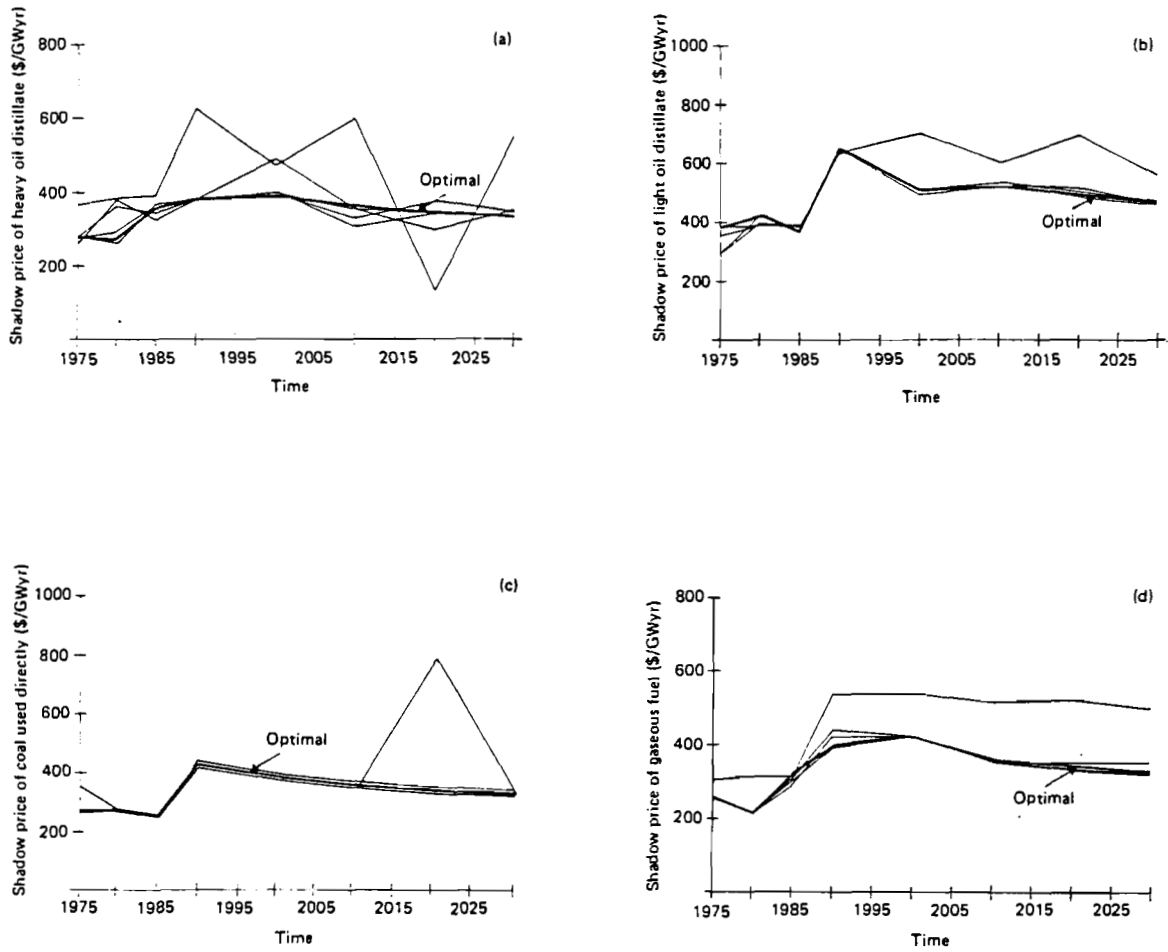
Figure 7
Shadow prices over the planning horizon for selected energy forms

The submodels were therefore run as independent models with the final energy flows (FEFs) fixed at the optimal levels, yielding the shadow prices associated with these constraints. These shadow prices can be called *local shadow prices*. The local shadow prices calculated with fixed optimal FEFs do not necessarily coincide with the optimal prices obtained using the decomposition algorithm. This is a typical feature of linear programming models and the difference between these values provides an estimate of the marginal utility of the FEFs for the submodels. These differences are depicted in Figure 8.
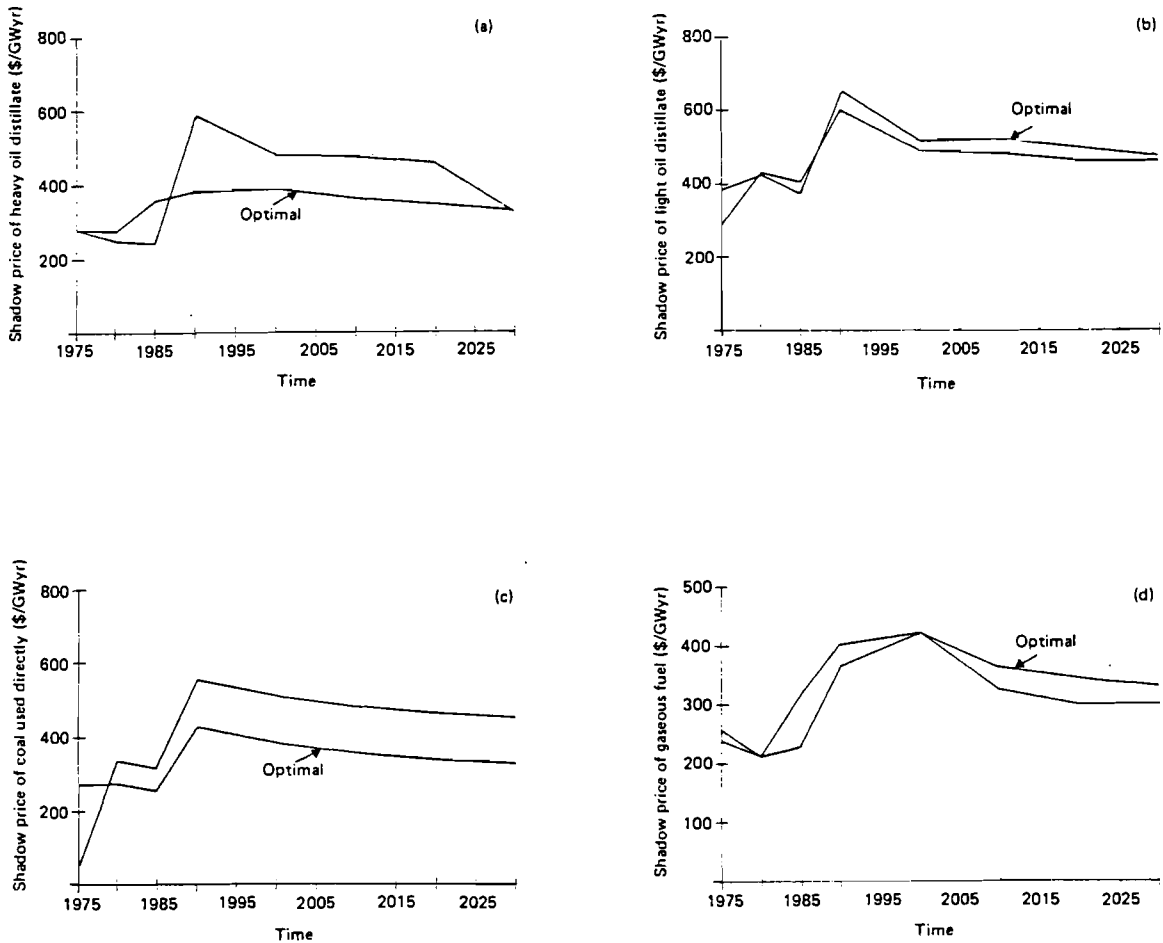
Figure 8
Optimal and local shadow prices for selected energy forms

It can be seen from the figure that these trajectories are very similar, and this fact could be interpreted as follows :

1) The major contribution to the optimal cost is associated with CENTR.

2) END acts as a device transforming demands for useful energy into final energy flows, and has no internal freedom for optimization.

Additional support for the second conclusion is provided by experiments in which END was operating under fixed final energy flows. The subsystem exhibited

slight infeasibility at the level of 10^{-4} which shows that END itself actually has a very small feasible region under these conditions. This is, of course, an undesirable feature and is due mostly to the need to keep the size of the model within practical limits. The decomposition approach allows these bounds to be widened and, as we have shown, is also computationally efficient.

6. CONCLUSIONS

The experiments conducted so far suggest a method of constructing an integrated system of energy models which could provide a detailed representation of the energy supply system itself and its interaction with the major energy-intensive economic subsectors. A thorough investigation of this interaction, in terms of the energy flows represented by the linking variables, could be valuable in determining an internally consistent energy policy for a nation.

The two algorithms studied both converged reasonably fast, with the primal-dual algorithm converging more rapidly in the final stages. However, the accuracy and reliability of the algorithm could be increased by improved implementation.

Acknowledgments

References

1.   Energy Systems Program Group of IIASA (W. Hafele, Program Leader), *Energy in a Finite World* , Ballinger, Massachusetts (1980).

2. L. Fishbone and H. Abilock, "MARKAL. A Linear Programming Model for Energy Systems Analysis," *Energy Research* 5 pp. 353-375 (1981).

3. P. Jadot, "Presentation of the Multinational Linear Energy Model," SE DOC 347.71, Systems-Europe (1977).

4. Brookhaven National Laboratory, "Industrial Process Models, Applications for Market Penetration ," Final Report, Upton, New York (1980).

5. E. Nurminski, "Convergence and Numerical Experiments with a Decomposition Algorithm," WP-82-9, International Institute for Applied Systems Analysis, Laxenburg, Austria (1982).

6. E. Nurminski, "Decomposition Algorithm Based on the Primal-Dual Approximation," WP-82-46 , International Institute for Applied Systems Analysis, Laxenburg, Austria (1982).

7. M. Agnew, L. Schrattenholzer, and A. Voss, "A Model for Energy Supply Systems Alternatives and Their General Environmental Impacts," WP-79-6, International Institute for Applied Systems Analysis, Laxenburg, Austria (1979).

8. S. Messner, "Users Guide to Message II," Technical Report, International Institute for Applied Systems Analysis, Laxenburg, Austria (forthcoming).

9. D.M. Ritchie and K. Thompson, "The UNIX Time-Sharing System," *Bell System Technical Journal* 57(6) pp. 1905-1931 (1978).

10. B.A. Murtagh and M.A. Saunders, "MINOS. A Large-Scale Nonlinear Programming System," Technical Report SOL 77-9, Stanford University (1977).