A Multiperiod Generalized Network Flow Model of the U.S. Integrated Energy System Part I – *Model Description*

Ana Quelhas, *Member, IEEE*, Esteban Gil, *Member, IEEE,* James D. McCalley, *Fellow, IEEE*, and Sarah M. Ryan

Abstract—This paper is the first of a two-part paper presenting a multiperiod generalized network flow model of the integrated energy system in the United States. Part I describes the modeling approach used to evaluate the economic efficiencies of the system-wide energy flows, from the coal and natural gas suppliers to the electric load centers. Under the proposed problem formulation, fuel supply and electricity demand nodes are connected via a transportation network and the model is solved for the most efficient allocation of quantities and corresponding prices. The methodology includes the physical, economic, and environmental aspects that characterize the different networks. Part II of this paper provides numerical results that demonstrate the application of the model.

***Index Terms*— Generalized network flow model, integrated energy networks, nodal prices, optimization.**

I. Nomenclature

The main symbols used in this paper are described below for quick reference.

A. Parameters

- $b_j(t)$ Supply (if positive) or negative of the demand (if negative) at node *j*, during time *t*.
- $c_{ij}(l,t)$ Per unit cost of the energy flowing from node *i* to node *j*, corresponding to the *l*th linearization segment, during time *t*.
- $e_{ij.max}$ Upper bound on the energy flowing from node *i* to node *j*.
- $e_{ij.min}$ Lower bound on the energy flowing from node *i* to node *j*.
- $\eta_{ij}(l)$ Efficiency parameter associated with the arc connecting node *i* to node *j*, in the *l*th linearization segment.
- $SO2_i(t)$ Emissions rate associated with the fuel consumed by power plant *i*, during time *t*.
- α_i Removal efficiency of to the pollution control equipment installed at power plant *i*. If no pollution equipment exists at power plant *i*, then $\alpha_i = 0$.
- $NSO2$ National SO_2 limit.

B. Variables

- $e_{ij}(l, t)$ Energy flowing from node *i* to node *j*, corresponding to the *l*th linearization segment, during time *t*.

C. Sets

- L_{ij} Set of linearization segments on the energy flowing from node *i* to node *j*.
- M Set of arcs.
- N Set of nodes.
- T Set of time periods.
- G Set of arcs representing electricity generation. This is a subset of M.

II. Introduction

THE movement towards deregulation and competition has lead to the more than ever decentralized and fragmented level of decision making that is happening today. As a result, electric power systems are planned and operated without the conscious awareness of the energy system-wide implications, namely the consideration of the integrated dynamics with the fuel markets and infrastructures. This has been partly due to the difficulty of formulating models capable of analyzing the large-scale, complex, time-dependent, and highly interconnected behavior of the integrated energy networks, while accounting for characteristics unique to each energy system (e.g., coal, natural gas, and electric power). Consequently, each subsystem supports specific procedures and strategies according to their own value system (i.e., economic, technical, political, and environmental context), which may be fragmentary because they are missing the necessary consolidation in global actions or alternative strategies for an efficient overall operation.

Today's industry climate motivates a more integrated study

Manuscript received June X, 2006. This work was partially supported by the National Science Foundation (NSF) under Grant 0527460 and the Fundação para a Ciência e a Tecnologia, Portugal, under Grant SFRH/BD/8079/2002.

A. Quelhas and E. Gil are Ph.D. candidates in Electrical and Computer Engineering at Iowa State University, Ames, IA 50011 USA (e-mail: aquelhas@iastate.edu; tebagil@iastate.edu). J. D. McCalley and S. M. Ryan are professors in the Department of Electrical and Computer Engineering and the Department of Industrial and Manufacturing Systems Engineering, respectively, at Iowa State University, Ames, IA 50011 USA (e-mail: jdm@iastate.edu; smryan@iastate.edu).

of the energy system. First, as the electric power industry becomes more competitive, economic performance of electricity delivery is intensely scrutinized from a national perspective, with electricity delivery price as a key metric. Customers and regulators are questioning electricity markets in which prices are significantly higher than those in other parts of the country, resulting in heavy pressure to identify means to gain economic efficiencies (lower prices) without seriously diminishing the reliability of the system. Second, the percentage of fuel purchased on the spot-market has been increasing with a corresponding decrease in the percentage of fuel purchased under long term contracts. In addition, long term contracts have become shorter in duration, as electric power generators try to pass market risks on to primary energy suppliers (producers and carriers) [1]. This fact increases concern on the part of generation owners that they may be more vulnerable to short or medium term contingencies in fuel supply. Third, there exists increasing awareness of the environmental problems caused by pollution emitted by the electric energy sector, which leads to the intensification of measures to internalize the externalities associated with electric power generation. In particular, the passage of the Clean Air Act Amendments (CAAA) of 1990 [2] forced electric generators to reduce their emissions of sulfur dioxide (SO_2) through the implementation of an innovative tradable permit system. Utilities are endowed with considerable operational flexibility since it is the total quantity of emissions that matters and a utility can achieve its target level through emission controls, fuel switching, conservation programs, or by buying allowances. Depending on the compliance strategies adopted, the impacts of the SO_2 regulations can go beyond the electric power subsystem and affect the energy flows of the fuel networks. Finally, the perception has grown that the national economy relies on a complex, multi-scale, distributed, and increasingly vulnerable and interconnected energy infrastructure [3]. The interconnected and interdependent nature of these infrastructures makes them vulnerable to cascading failures, i.e. the propagation of disruption from one system to the other, with possible catastrophic consequences.

There has been significant work in scheduling fuel deliveries in order to optimize electric energy production [4]. The common denominator of all known fuel scheduling approaches is that they view the fuel system only in terms of delivered prices and associated penalties for possible violations of contracts. In other words, there has been little effort to optimize the electric power system operations with the consideration of the integrated dynamics with the fuel markets and infrastructures, accounting for the fuel production, storage, and transportation costs and capabilities.

A number of energy models have been developed for policy analysis, forecasting, and to support global or local energy planning [5]. An important consideration regarding many of the exiting energy models is that they typically tend to be highly resource intensive, both in terms of expertise requirements to develop the model and support the underlying data and in terms of execution time and other computational resource requirements, reflecting the highly complex algorithmic and programming routines. Although many of these models integrate different energy systems in a modular form, they are not typically designed to illustrate the effects of alternative energy transportation modes. Their methodology usually follows a top-down approach that evaluates a broad equilibrium framework from aggregated economic variables. In contrast, the bottom-up model presented in this paper captures the physical and environmental restriction of the coal, natural gas, and electricity flows in an engineering sense. In addition, due to the typical complexity and high proprietary costs of existing integrated energy models, they are not readily available to the research community. Consequently, many opportunities exist to enrich this field of research and the rather limited technical literature and information available in the public domain.

In this two-part paper, we propose a generalized network flow model of the national integrated energy system that incorporates the production, storage (where applicable), and transportation of coal, natural gas, and electricity in a single mathematical framework, for a medium term analysis. In general, the model can be used to foster a better understanding of the integral role that the coal and natural gas production and transportation industries play with respect to the entire electric energy sector of the U.S. economy. The model represents the major fossil fuel markets for electricity generation (coal and natural gas) [6] and solves for the optimal solution that satisfies electricity demand, deriving flows and prices of energy. Each energy subsystem considers the factors relevant to that particular subsystem, for example, coal transportation costs, or gas transmission capacities. The modeling framework presented integrates the cost-minimizing solution with environmental compliance options to produce the least-cost solution that satisfies electricity demand and restricts emissions to be within specified limits. Despite the relative importance of electricity generation from nuclear energy (roughly 20%), it is exogenously given because of its slow dynamics, which is assumed not to influence the medium term analysis intended. The schedules of electricity generated from renewable energies are also represented as direct inputs into the electric transmission system, due in part by their relative small contribution the generation mix and the lack of emissions restrictions. In addition, most of them cannot be transported as a raw fuel (e.g., wind and sunlight) and therefore represent no energy movement alternative to electric transmission in the way that coal and natural gas do. Water, however, could be endogenously included in the model and formulated with the network flow techniques presented in this paper, as long as data characterizing the hydraulic networks (e.g., reservoir capacities) were available.

Part I of this paper describes the theoretical underpinnings of the modeling approach adopted, the mathematical formulation, and the modeling assumptions. Part II provides numerical results and identifies directions for future work and possible applications of the model.

III. Modeling Approach

A. Network Flow Model

The integrated energy system is readily recognized as a network defined by a collection of nodes and arcs with energy flowing from node to node along paths in the network. Such a structure lends itself nicely to the adoption of network flow programming modeling technique. When a situation can be entirely modeled as a network, very efficient algorithms exist for the solution of the optimization problem, many times more efficient than linear programming in the utilization of computer time and space resources. The network flow problem formulated in this paper falls into the category of generalized minimum cost flow problem and can be solved by applying the generalized network simplex algorithm [7].

The scenario of this generalized minimum cost flow problem is the following. The supply node (source node) has an excess of coal or natural gas, while the nodes with demand require certain amounts of energy. The remaining nodes (transshipment nodes) neither require nor supply the commodity (energy), but serve as a point through which energy passes. The energy flows through arcs that connect the nodes, and there is conservation of energy at the nodes, implying that the total flow entering a node must equal the total flow leaving the node. The arc flows are the decision variables of the network flow programming model. Associated with each arc (i, j) are the following parameters:

i) **Lower bound**, $e_{ij.min}$, (which can be zero) on the flow,
ii) **Upper bound**, $e_{ij.max}$, on the flow (also called capacity),
iii) **Cost**, c_{ij}, per unit of flow,
iv) **Efficiency** parameter, η_{ij}, (also called the gain or the loss factor) which multiplies the flow at the beginning of the arc to obtain the flow at the end of the arc.

The interpretation of the efficiency parameter is the following: when 1 unit of flow is sent on arc (i, j), η_{ij} units of flow arrive at node j. It is a positive rational number that represents losses if $\eta_{ij} < 1$or gains if $\eta_{ij} > 1$. A network in which all arcs have unit gains is called a pure network. If some efficiency parameters have values other than 1 the network is a generalized network. Multipliers substantially increase the flexibility of the network modeling approach beyond that of pure networks. Their ability to modify flows along the arcs makes it possible to represent increases or decreases in flow that actually occur in real world. In the integrated energy model presented, multipliers are used to represent, for instance, natural gas extraction losses, electric transmission losses along power lines, or any other type of efficiency measurement. Furthermore, the application of multipliers is particularly relevant to transform flows along arcs from one unit of measurement to another. Some examples include transformation of short tons of coal to million Btu (MMBtu) or thousand cubic feet (Mcf) of gas to MMBtu.

The goal of the generalized minimum cost flow problem is to satisfy electric energy demands with available fossil fuel supplies at the minimal total cost, without violating the bound constraints, including emissions limits. The costs considered are the fossil fuel production, transportation, and storage costs, the operation and maintenance costs associated with electricity generating units operations, and the electric power transmission costs.

B. Node and Arc Definitions

The network flow model of the integrated energy system comprises the following nodes:

i) **Source node**: The source node is an artificial node that supplies all the energy necessary to satisfy the electric energy demand. Supply can not be specified *a priori*, because it depends on the losses of the entire system, which in turn depend upon the flows.
ii) **Transshipment nodes**: The transshipment nodes represent the primary energy production facilities (coal mines and gas wells), the storage facilities (coal piles and natural gas reservoirs), the energy conversion facilities (power plants), and the North American Electric Reliability Council (NERC) regions and subregions as defined by NERC.

The outgoing arcs of the dummy source node represent the production of coal and natural gas and imports of coal, natural gas, and electricity. In the coal and natural gas subsystems, arcs represent coal transportation routes and major natural gas pipeline corridors. Arcs also represent storage injections and withdrawals, and inventories carried over between two consecutive time periods. In the electric subsystem, arcs from the generators to their respective NERC region or subregion represent electricity generation, and arcs between regional nodes represent bulk electric power trade. Energy losses in the production, storage, and transportation of the primary energy forms, losses in the energy conversion process at the power plants, and losses in tie lines are represented by appropriately chosen multipliers on the arcs. Fuel production costs (extraction and processing charges) are associated with the outgoing arcs from the dummy source node; coal transportation rates and pipeline tariffs are assigned to the respective transportation arcs; storage fees are allocated to the arcs representing storage withdrawals; operation and maintenance costs of power plants are assigned to the arcs connecting the power plant nodes to the corresponding load nodes; and wheeling charges, or transmission costs associated with electric power trade, are allocated to the arcs representing tie lines.

Since the electricity demand is modeled at the level of the NERC regions and subregions, the only transmission lines represented in the model are the tie lines among NERC regions and subregions, whose flows can be considered decision variables since the control areas that operate them have the capability of controlling the imported/exported energy flow with their adjacent control areas. In contrast, the energy flows in the transmission lines within a control area can not be considered decision variables, because they are determined according to the Kirchhoff's laws. As a result, only bulk power (wholesale) transactions are considered.

C. Tie Line Representation

A tie line is an undirected edge, because the energy can flow in both directions. Since the network flow model requires directed arcs, the transformation in Fig. 1, shows an equivalent model with each edge replaced by an oppositely directed pair of arcs.

If the flow in either direction has a lower bound of value 0 and the arc cost is nonnegative, in some optimal solution one of the flows in the directed arcs will be zero, which guarantees a non-overlapping solution.

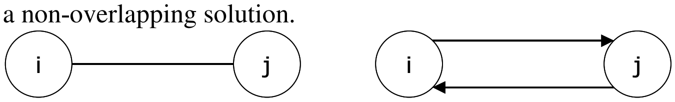


Fig. 1. Representation of transmission lines.

D. Elimination of Nonzero Lower Bounds

A network flow model with directed arcs with nonzero lower bounds can be replaced by an equivalent model with zero lower bounds. The left side of Fig. 2 shows an arc with the parameters lower bound e_{min}, upper bound e_{max}, cost c, and multiplier η. An equivalent representation of the arc with zero lower bound is shown on the right side of Fig. 2. Making this transformation requires an adjustment of the supply at both ends of the arc, i.e. b_i and b_j. This transformation also changes the objective function by a constant equal to $c \times e_{min}$ that can be recorded separately and then ignored when solving the problem.

Finally, it is interesting to note that when arcs have equal upper and lower bounds, i.e. when the flow is fixed, they can be eliminated from the equivalent network because its upper bound on the flow will be zero.

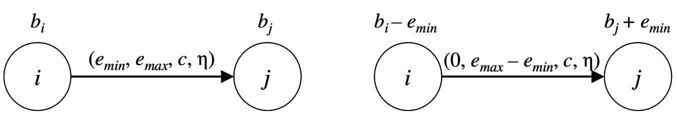


Fig. 2. Removing nonzero lower bounds.

E. Restrictions on Nodes

In a standard network flow model, the only parameters associated with the nodes are the supply or demand. In the integrated energy system, resources such as power plants and storage facilities have restrictions on the flow that can pass through them (e.g., capacities, efficiency rates, and costs), which are parameters associated with arcs in a network flow model.

The transformation into a standard network flow model is done by replacing each of these nodes into a pair of nodes with an arc connecting them. The parameters of this arc dictate the restrictions on the flow that passes through the respective facility. Fig. 3 illustrates this transformation, where the parameters (e_{min}, e_{max}, c, η) refer to the lower bound, upper bound, cost, and efficiency, respectively, of the facility represented by node i.

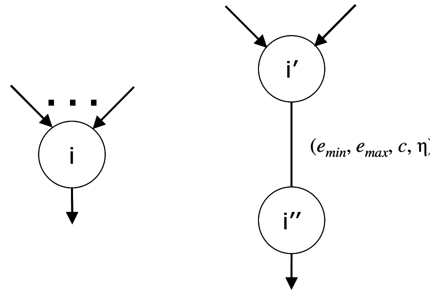


Fig. 3. Representation of restrictions on nodes.

F. Linearization of Costs and Efficiencies

A typical input-output characteristic of a steam turbine generator can be represented by a convex curve [4]. When multiplied by the fuel cost, we obtain the generating unit cost as a convex function of the flow. Total cost functions can then be approximated by piecewise linear functions, which leads to step incremental cost functions. In a network flow representation, each linearization segment is modeled by an arc, with the number of arcs determining the accuracy of the approximation. To illustrate this idea let us consider an arc that carries flow between nodes i and node j. The cost associated to the flow in this arc is a convex function and can be fitted by a piecewise linear cost function. This cost function tells us that the first 20 units of flow have a unit cost of $2.5, the next 10 units of flow have a unit cost of $5, and any additional amount has a unit cost of $10, up to the capacity of 40 units of flow. As shown in Fig. 4, this situation is modeled using a set of arcs, each one for each segment of the piecewise linear cost function. Because the unit costs are increasing, the flow in a given arc will only be positive if all the other arcs with smaller unit costs have reached their capacity limits, which guarantees that the solution is physically possible.

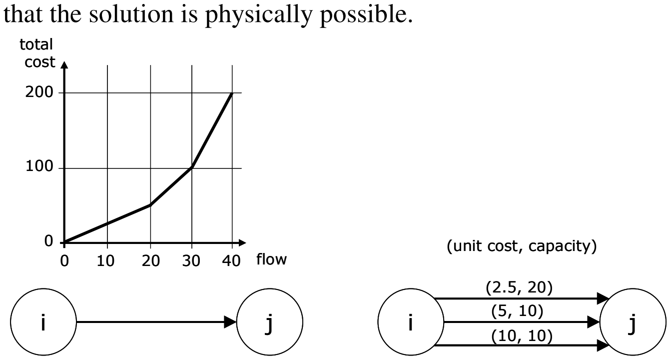


Fig. 4. Representation of convex cost functions.

Nonconvex cost functions, in particular those associated with the input-output characteristics of combined cycle gas turbines, cannot be addressed with network flow programming techniques, and are therefore approximated by linear or piecewise linear convex functions. Although optimization techniques capable of dealing with nonconvexities are available [8], the cost in modeling complexity outweighs the improvement in model fidelity considering the level of aggregation intended, which is mainly dictated by data availability restrictions.

Efficiency parameters may also be modeled using piecewise linear functions of the flow and can be represented by the multiple arc transformation illustrated above for convex cost functions. For example, power losses along the transmission lines are proportional to the square of the flow, and efficiency can therefore be approximated by a piecewise linear function where the slopes decrease with the flow. In this situation, it is guaranteed that the arcs with the higher efficiency parameters (lower losses) will be filled up first, since they require the smallest amount of flow, and thus the smallest cost, for the same energy demanded at the head node.

G. Dynamics of the Model

Static models have no underlying temporal dimension. However, in the case of an integrated energy model, we have to account for the evolution of the system over time, as inventory is carried over from one time period to another.

Multiperiod network flow models may be viewed as a composition of multiple copies of a network, one at each point in time, with arcs that link these static snapshots describing temporal linkages in the system. With this construction, the size of the network is proportional to the number of periods.

If a unique time step is chosen to apply to the entire model, it must be small enough to capture the fastest dynamics of the integrated energy system, which are imposed by the electric energy subsystem. However, this results in unnecessary and counterproductive computations that take place for slower energy subsystems. Alternatively, one can take advantage of the fact that the integrated energy system is composed of different energy subsystems with distinct dynamics, and then define a different time step for each one, thus eliminating the burden of redundant simulation. As a result, different simulation time steps can be used for different energy subsystems [11].

IV. Mathematical Formulation

A. Generalized Network Flow Model

Mathematically, the multiperiod generalized minimum cost flow problem is an optimization model that can be formulated as follows:

$$\text{Minimize } z = \sum_{t\in T} \sum_{(i,j)\in M} \sum_{l\in L_{ij}} c_{ij}(l) e_{ij}(l,t) \tag{1a}$$

subject to:

$$\sum_{\forall k} \sum_{l\in L_{jk}} e_{jk}(l,t) - \sum_{\forall i} \sum_{l\in L_{ij}} \eta_{ij}(l) e_{ij}(l,t) = b_j(t) \text{ , } \forall j \in N, \forall t \in T \tag{1b}$$

$$e_{ij.\min} \le e_{ij}(t) \le e_{ij.\max} \qquad \forall (i,j) \in M \text{ , } \forall t \in T \text{ .} \tag{1c}$$

where z is the objective function.

The objective function in (1a) represents the total costs associated with the energy flows from the fossil fuel production sites to the electricity end users and non-electric natural gas consumers. These total costs are defined as the sum of the fuel production costs, fuel transportation costs, fuel storage costs, electricity generation costs (operation and maintenance costs), and electricity transmission costs. The set of constraints in (1b) represent the conservation of flow constraints (energy balance constraints) for all nodes and for all times. For a particular node, the first term of this constraint is the total outflow of the node (flow emanating from the node) and the second term is the total inflow of the node (flow entering the node). The conservation of flow constraint states that the inflow minus the outflow must equal the supply/demand of the node. The set of constraints defined by (1c) are the flow bound constraints, which state that the flow must satisfy the lower bound and capacity of the respective arcs. The flow bounds represent the flows' operating ranges.

In matrix form, the problem can be represented as follows:

$$\text{Minimize } z = \underline{c}'\underline{e} \tag{2a}$$

subject to:

$$A\underline{e} = \underline{b} \text{ ,} \tag{2b}$$

$$\underline{e}_{\min} \le \underline{e} \le \underline{e}_{\max} \text{ .} \tag{2c}$$

In this formulation, $\boldsymbol{A}$ is an $n \times m$ matrix, where n is the number of nodes and m is the number of arcs. $\boldsymbol{A}$ is called the node-arc incidence matrix. Each column of $\boldsymbol{A}$ is associated with a decision variable, and each row is associated with a node. The column A_{ij} has a +1 in the ith row, a –1 or a $-\eta_{ij}$ in the jth row, and the rest of its entries are zero. An illustrative example of the formulation of the node-arc incidence matrix for a simple integrated energy system is presented in [9].

B. Side Constraint

As mentioned before, the overall objective of this optimization problem is to determine the energy flows that meet the demand for electricity at the minimum operating costs, subject to physical and environmental constraints. The mathematical formulation presented above is suitable to address the physical constraints of the integrated energy system. However, it is not sufficient to guarantee that the SO_2 emissions limit imposed by the CAAA is not exceeded. In addition to the energy balance constraints at all nodes and the flow bound constraints for all arcs, another constraint must be incorporated to impose a national-level limit on emissions. According to the CAAA, the allowances for SO_2 emissions are traded nationwide so the corresponding limit on emissions is actually national rather than regional or unit-level. This national limit is determined by the sum of the allowances allocated to power plants (as defined by the CAAA) and adjusted to capture the exogenously given emissions banking effects. The amount of emissions produced depends on the fuel used, the pollution control devices installed, and the amount of electricity produced. The additional constraint may be represented as follows:

$$\sum_{t\in T} \sum_{(i,j)\in G} SO2_i(t) \cdot (1-\alpha_i) \cdot \sum_{l\in L_{ij}} e_{ij}(l,t) \le NSO2 \text{ ,} \tag{1d}$$

All compliance strategies that can be implemented in an operational time frame – fuel switching (e.g., use low sulfur content coal or natural gas instead of high sulfur content coal), utilization of emissions control devices or abatement technologies (e.g., scrubbers, particulate collectors), revising the dispatch order to utilize capacity types with lower emission rates more intensively, and allowance trading – are now effectively captured by the mathematical model described by equations (1a)-(1d).

The inequality constraint (1d) can be transformed into an equality constraint and incorporated in the matrix equation (2b). This transformation is done by introducing a nonnegative slack variable in the left-hand side of the equation. With the addition of constraint (1d) to equation (2b), some of the columns of the matrix $\boldsymbol{A}$ have now more than two non-zero entries, which makes it no longer a node-arc incidence matrix, but instead a more general constraint coefficient matrix. In linear programming terminology, the constraint (1d) is called a bundle, complicating, or side constraint, which specifies a flow relationship between several of the arcs in the network flow model. The integrated energy system can also be interpreted as a multicommodity flow problem, where energy and emissions are the commodities that flow along the arcs of the network. The complicating constraint ties together these two commodities.

C. Nodal Prices

The Karush-Kuhn-Tucker conditions associated with the constrained linear optimization problem defined above yield the so called Lagrangian multipliers or dual variables. In economic terms, the Lagrangian multipliers are explained as the shadow values related with each active constraint at the optimal solution of the choice variables, and they represent the marginal costs of enforcing the constraints. In a network flow formulation, these shadow prices are also referred to as nodal prices, because each node of the network structure has a Lagrangian multiplier associated to it, as a result of the mass balance constraints defined for the nodes.

For simplicity, and without loss of generality, let us assume that the cost and efficiency parameters associated with each arc are constant functions. This permits the elimination of the parameter l, for notational simplicity. The Lagrangian function for (1a)-(1d) is given by (3), where $\lambda_j(t)$ is the Lagrangian multiplier (or nodal price) associated with the energy balance constraint at node j for time t. $\delta_{ij}(t)$ and $\mu_{ij}(t)$ are the Lagrangian multipliers associated with the lower and upper bound constraints, respectively, on the energy flowing from node i to node j, during time t. Finally, γ is the Lagrangian multiplier associated with the emissions limit constraint.

$$
\begin{aligned}
L = & \sum_{t \in T} \sum_{(i,j) \in M} c_{ij}(t) e_{ij}(t) + \\
& + \sum_{t \in T} \sum_{j \in N} \lambda_j(t) \left[\sum_{\forall k} e_{jk}(t) - \sum_{\forall i} \eta_{ij} e_{ij}(t) - b_j(t) \right] + \\
& + \sum_{t \in T} \sum_{(i,j) \in M} \delta_{ij}(t) \left[e_{ij.\min} - e_{ij}(t) \right] + \sum_{t \in T} \sum_{(i,j) \in M} \mu_{ij}(t) \left[e_{ij}(t) - e_{ij.\max} \right] + \\
& + \gamma \left[\sum_{t \in T} \sum_{(i,j) \in G} SO2_i(t) \cdot (1 - \alpha_i) \cdot e_{ij}(t) - NSO2 \right]
\end{aligned} \quad (3)
$$

For optimality, in a given time period t, the relationship between the nodal prices of two linked nodes i and j, is given by one of the following equations. If $(i,j) \notin G$, that is (i,j) does not represent electricity generation, then:

$$\frac{\partial L}{\partial e_{ij}(t)} = c_{ij}(t) + \lambda_i(t) - \lambda_j(t)\eta_{ij} - \delta_{ij}(t) + \mu_{ij}(t) = 0 \quad (4a)$$

Otherwise, if $(i,j) \in G$, that is (i,j) is an arc representing electricity generation, then:

$$
\begin{aligned}
\frac{\partial L}{\partial e_{ij}(t)} = & c_{ij}(t) + \lambda_i(t) - \lambda_j(t)\eta_{ij} - \delta_{ij}(t) + \mu_{ij}(t) + \\
& + \gamma SO2_i(t)(1 - \alpha_i) = 0
\end{aligned} \quad (4b)
$$

If the inequality constraints are slack, i.e. not binding or not active, the corresponding Lagrangian multipliers are zero. Therefore, from equation (4a) we conclude that if the flow bound constraints are not binding, the cost is zero ($c_{ij}(t) = 0$), and there are no losses ($\eta_{ij} = 1$), then the nodal prices of two linked nodes are the same ($\lambda_i(t) = \lambda_j(t)$). Likewise, from equation (4b) we conclude that the nodal price at a power plant node i is the same as the nodal price at the corresponding electricity demand node j if and only if the flow bound constraints are not binding, the arc cost is zero, there are no transmission losses, and the emissions limit constraint is also not binding. Note that flow bound constraints being binding is equivalent to congestion in the associated arc.

In the context of the electric power industry, the concept of nodal prices has become more and more familiar, as several electricity markets have used the information from nodal prices to improve the efficient usage of the power grid, to perform congestion management, and also to design a pricing structure for the power system [10]. In the power industry terminology, nodal prices are often referred to as locational marginal prices, or LMP. In 2002, the Federal Energy Regulatory Commission (FERC) proposed a standard market design that incorporates a locational marginal pricing mechanism to induce efficient electric power markets. In contrast with a single price mechanism, under a nodal pricing scheme market clearing prices are calculated for a number of locations on the transmission grids called nodes. Prices vary from node to node because of transmission line congestion and losses. At each node, the price represents the locational value of electric energy, including the cost of energy and the cost of delivering it, i.e. losses and congestion. In other words, the nodal price is the cost of serving the next megawatt of load at a given location. Therefore, LMP can be used to determine the value of transmission rights and to provide economic signals for generation and transmission investments.

The concept of nodal prices widely used in the electric power arena is herein expanded to the integrated energy system, by optimizing the energy flows in a generalized network flow model that explicitly represents the electric subsystem together with the various fossil fuel networks in a single mathematical framework [11]. Since all entities involved in the operation of the energy system are fully represented, the nodal prices obtained as a by-product of the optimization procedure provide a means to identify the interdependencies between the fuel subsystems and the electric subsystem. Knowledge and understanding of these interdependencies is expected to induce the most economically efficient use of fuel production, fuel storage, fuel transportation, electricity generation, demand, and transmission resources, through the correct economic signals provided. In addition, because nodal prices monetize congestion costs, they provide clear economic signals that indicate where infrastructure improvements should take place to relieve constraints, thus promoting efficient investment decisions.

V. Modeling Assumptions

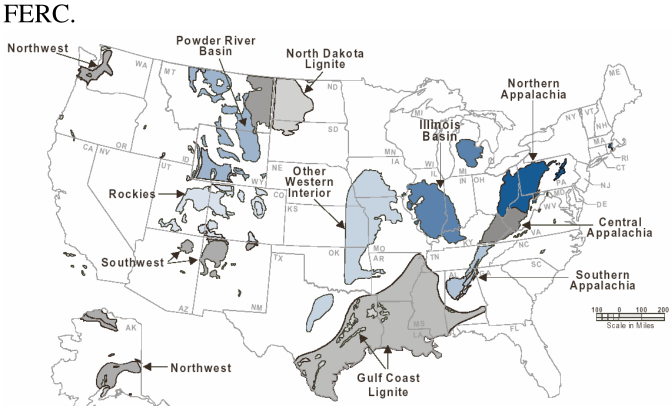
A. Coal Network

The coal network model proposed is defined based on the supply regions depicted in Fig. 5. For each coal supply region a coal production node is defined and characterized by their associated productive capacity, average heat value, average sulfur content, and average minemouth price.

Because coal exports and imports represent a very small percentage of the U.S. coal production and consumption, respectively, international coal trade is not considered. Coal consumption by non-electric consumers is also neglected.

Precise modeling of the over thousands of individual transportation routes used to transport coal from mines to electric power plants would require an enormously detailed and very complex model, using large quantities of data that are not in the public domain. As a result, a simplified approach is adopted, where an arc is established between each coal supply node and all feasible coal-fired power plants. A transportation link is considered not feasible, and therefore not included in the model, when it represents an either economically or physically impractical route. Arcs connecting coal production nodes with coal-fired plants are characterized by a lower bound that represents existing contractual agreements and a transportation cost.

Coal data are gathered from various sources, namely the Energy Information Administration (EIA), the Mine Safety and Heath Administration of the Department of Labor, and FERC.



Source: Energy Information Administration

Fig. 5. Coal supply regions.

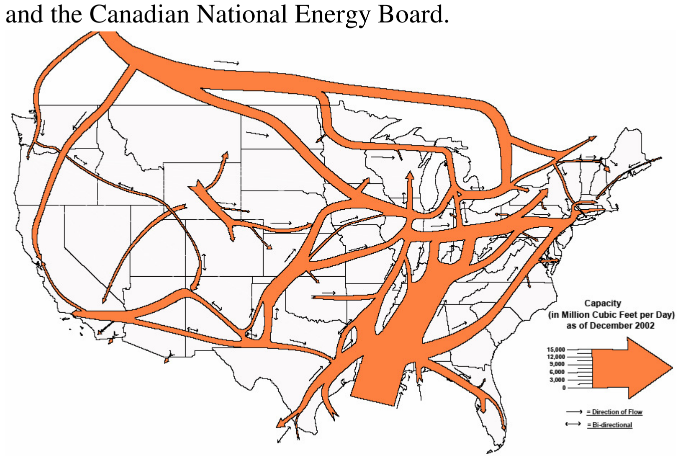
B. Natural Gas Network

Natural gas production nodes are defined taking into account the geographical distribution of natural gas reserves and production and data availability. Natural gas production nodes are characterized by their effective productive capacity, average wellhead price, and an efficiency parameter that accounts for extraction losses.

Fig. 6 shows the major pipeline transportation corridors. Given the complexity of this system and data availability restrictions, representation of the actual physical system is prohibitive. A simplified approach is therefore adopted, where the lower 48 states are divided into transmission regions, each region containing one transshipment node and one storage node. Transshipment nodes represent a junction point for flows coming into and out of the regions. Arcs connecting the transshipment nodes represent interregional flows. Flows are further represented by establishing arcs from the production nodes to the correspondent transshipment node. Similarly, arcs are also established between the transshipment nodes and storage nodes and from the transshipment nodes to the appropriate gas-fired power plant nodes. Imports and exports with Canada and Mexico are also represented. Natural gas consumption by non-electric end-users is represented as an exogenously given demand in the natural gas transshipment nodes.

Natural gas transportation arcs are characterized by a capacity, a loss factor, and a transmission markup. Arcs representing natural gas storage injections are characterized by an injection capacity and arcs representing storage withdrawals are assigned withdrawal capacities and a cost parameter to account for the storage cost of service. Arcs denoting natural gas carried over between two consecutive time periods are characterized by a lower bound, which represents the cushion gas, and an upper bound, which corresponds to the total storage capacity of the region.

Natural gas network modeling assumptions are derived from various sources, namely EIA, FERC, the Minerals Management Service of the U.S. Department of the Interior, and the Canadian National Energy Board.

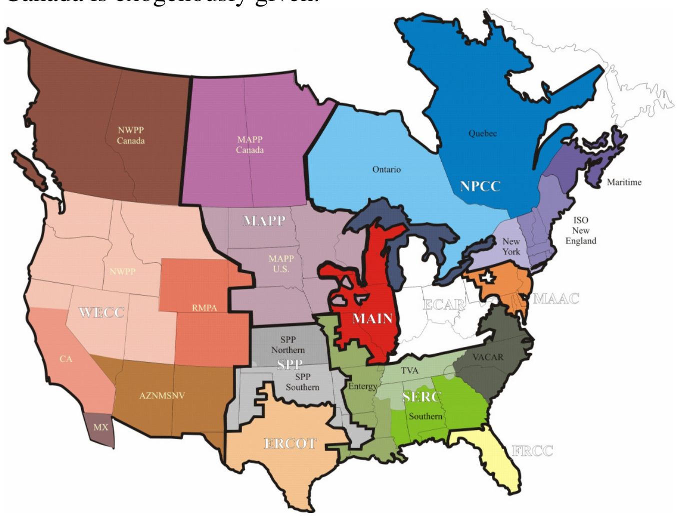


Source: Energy Information Administration

Fig. 6. Major pipeline transportation corridors.

C. Electricity Network

The electric power sector is modeled at a regional level. The regions considered are the NERC regions and subregions in the contiguous U.S., as depicted in Fig. 7. This aggregation level is based on the topology of the electrical grid and operating constraints, such as transmission bottlenecks, and is an adequate simplification of the physical and institutional complexity of the electric power industry. For each region, a transshipment node is defined and assigned a given demand. Arcs between the transshipment nodes are established to represent interregional transmission paths composed of one or more parallel tie lines connecting adjacent control areas in interconnected neighboring regions. These arcs are characterized by interregional total transmission capabilities, transmission costs, and loss factors. International trade with Canada is exogenously given.



Source: North American Electric Reliability Council

Fig. 7. NERC regions and subregions (as of December 2004).

Within each region, generating units with similar characteristics are clustered into equivalent power plants with a combined capacity and weighted average characteristics that are representative of all the units comprising the equivalent plant. Equivalent power plant nodes are differentiated by fuel type and prime mover. Coal-fired power plants are further disaggregated by the type of installed SO_2 pollution control device, i.e. the flue gas desulfurization technology used, if any, and assigned a corresponding removal efficiency rate.

Data characterizing the electric power network are mainly obtained from EIA, FERC, NERC, and the U.S. Environmental Protection Agency.

VI. Conclusions

Although economic and physical performances of individual subsystems are well studied and understood, there has been little effort to study its global characteristics. The study presented in this paper has been motivated by the hypothesis that the current fragmented decision making environment in which coal, natural gas, and electricity firms operate leads to potential inefficiencies. Given the critical role that these infrastructures represent and their great interdependency, it is of vital importance to keep an overall system perspective, both during planning and in all stages of operation. To the extent that traditional tools and simulation models do not allow for a comprehensive analysis capable of handling the complex dynamics of highly integrated energy systems, individual decision makers support specific procedures and strategies according to their own value system (i.e., economical, technical, organizational, political, and environmental context), which may lead to efficiency losses.

In order to address these issues, this paper has presented a multiperiod generalized network flow model of the U.S. integrated energy system. The model focuses on the economic interdependencies of the integrated system, in the sense that it represents multiple energy networks (electric, coal, and natural gas), along with a detailed characterization of their functionalities (supply, demand, storage, and transportation), within a single analytical framework that allows for their simultaneous study. The methodology includes the technological, economic, and environmental aspects of the different energy subsystems considered. The benefits of using a network flow modeling technique rather than a more general linear programming approach are associated with the fact that more efficient solution procedures can be used, which is of importance due to the high dimension that characterizes an integrated energy system.

Simulation results are presented and analyzed in part II of this paper. Part II also suggests exciting areas of further research and possible model applications.

References

[1] Energy Information Administration, "Challenges of Electric Power Industry Restructuring for Fuel Suppliers," DOE/EIA-0623(98), Washington, DC, September 1998.

[2] U.S. Congress, House of Representatives, "Clean Air Act Amendments of 1990 Conference Report," U.S. Government Printing Office, Washington, DC, 1990.

[3] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE Control Systems Magazine*, vol. 21, pp. 11-25, December, 2001.

[4] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation and Control*, 2nd ed. New York: John Wiley & Sons, Inc, 1996.

[5] N. van Beeck, "Classification of energy models," Tilburg University and Eindhoven University of Technology, May 1999.

[6] Energy Information Administration, "Electric Power Annual 2004," DOE/EIA-0348(2004), Washington, DC, November 2005.

[7] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1993.

[8] J. M. Arroyo and A. J. Conejo, "Optimal response of a thermal unit to an electricity spot market," *IEEE Transactions on Power Systems*, Vol. 15, No. 3, pp. 1098-1104, August 2000.

[9] E. M. Gil, A. M. Quelhas, J. D. McCalley, and T. V. Voorhis, "Modeling integrated energy transportation networks for analysis of economic efficiency and network interdependencies," in *Proceedings of the 33rd North American Power Symposium*, October 2003.

[10] F. Schweppe, M. Caramanis, R. Tabors, and R. Bohn, *Spot Pricing of Electricity*. Norwell, MA: Kluwer Academic Publications, 1988.

[11] A. M. Quelhas, E. Gil, and J. D. McCalley, "Nodal prices in an integrated energy system," *International Journal of Critical Infrastructures*, Vol. 2, No. 1, pp. 50-69, 2006.

Ana Quelhas (M'00) graduated from the Faculdade de Engenharia, Universidade do Porto, Portugal, in Electrical and Computer Engineering in 1999. She received an M.S. degree in Electrical Engineering and an M.S. degree in Economics, both from Iowa State University, U.S.A., in 2001 and 2004, respectively. She is currently a Ph.D. candidate in the Electrical and Computer Engineering Department at Iowa State University. Ms. Quelhas has interned at the Market Operation Department of the California ISO and at the Directorate General for Research of the European Commission, Brussels, Belgium.

Esteban Gil (M'99) was born in Santiago, Chile. He obtained his B.Sc. and M.Sc. degrees in Electrical Engineering from Universidad Técnica Federico Santa María (UTFSM), Valparaíso, Chile, in 1997 and 2001 respectively, and a M.Sc. degree in Statistics from Iowa State University in 2006. He is a Fulbright fellow and is currently a PhD. candidate in Electrical and Computer Engineering at Iowa State University. His research interests focus mainly on reliability and optimization in electric and energy systems.

James D. McCalley (SM'97–F'04) received the B.S., M.S., and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, in 1982, 1986, and 1992, respectively. He was employed with the Pacific Gas and Electric Company, San Francisco, CA, from 1985 to 1990 and is currently a Professor in the Electrical and Computer Engineering Department at Iowa State University, where he has been since 1992. Dr. McCalley is a registered Professional Engineer in California.

Sarah M. Ryan received her B.S. from the University of Virginia in 1983, and her M.S.E. and Ph.D. in Industrial and Operations Engineering from the University of Michigan in 1984 and 1988. She is an associate professor of Industrial & Manufacturing Systems Engineering at Iowa State University. Her research interests include stochastic models to study long-term investment decision problems as well as resource allocation problems in manufacturing.