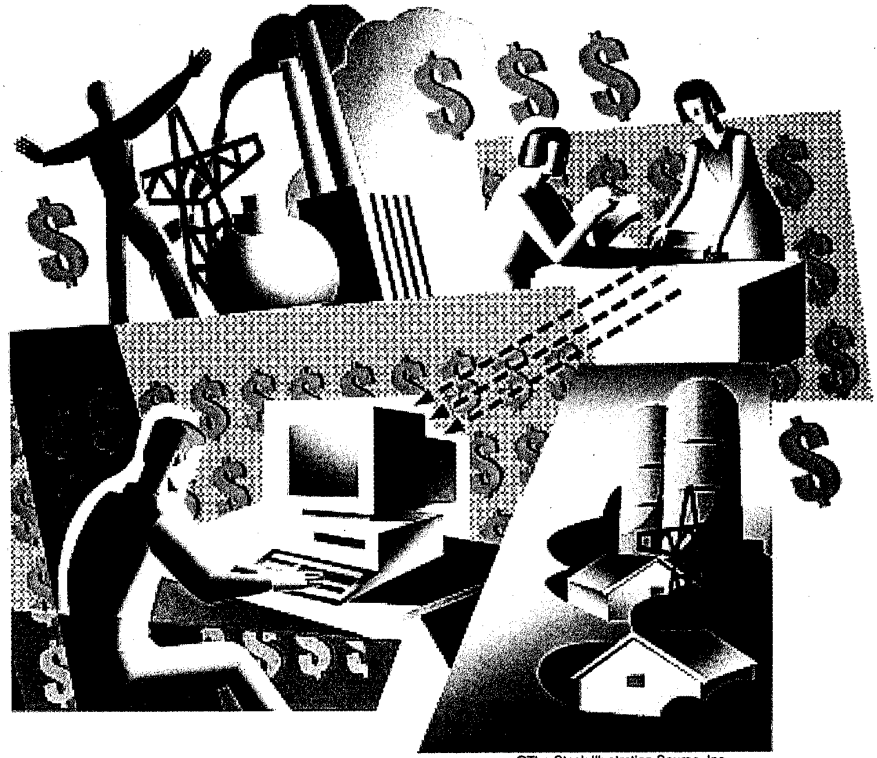
FEATURE



©The Stock Illustration Source, Inc.

Electric Power Systems Operation by Decision and Control The Case Revisited

By Marija Ilic
and Petter Skantze

Recently we have witnessed sweeping changes in the organization of the electric power industry. This has had a profound effect on the industry's system operations and planning rules and has created a need to revisit its modeling, decision-making, and control principles.

Historically, the decision and control of large-scale electric power systems has developed as the system interconnection has grown, not always following systematic control design principles capable

Ilic (ilic@mit.edu) and Skantze are with Massachusetts Institute of Technology, 1 Amherst St., Cambridge, MA 02139, U.S.A.

1053-5888/00/$10.00©2000IEEE

of ensuring a prespecified performance for the specified range of variations of system inputs and system topology. The reasons for this are many, ranging from system complexity through lack of incentives to employing the most effective technology. Competition generally brings opportunities for technology advances. With this in mind, we particularly stress decision and control criteria and their potential value in the evolving industry.

A brief review in the next section of present operating principles indicates that major decision making is an open-loop activity based on the anticipated system demand and equipment conditions. This is enhanced by the ingenious, simple, automatic generation control (AGC) method for balancing each control area's generation and demand for the scheduled exchange with the neighboring control areas, resulting in negligible stationary frequency deviations. In addition, each large generator is equipped with decentralized proportional-integral- derivative (PID) controllers (governor and excitation system) for local frequency and voltage stabilization to its stationary value. The tuning of these controllers has been system specific and has worked quite well in response to small deviations in load demand. Several serious problems have occurred over large portions of the power system interconnection, however, under certain unexpected large equipment outages, leading to further sequential system disintegration and the infamous blackouts. To prevent this from happening in the future, the current industry practice has become to operate suboptimally under normal conditions (with all equipment functioning as planned) to have sufficient control (primarily generation) when something major happens unexpectedly so as not to affect electricity users. This is the well-known $(n-1)$ reliability criterion. Although not the most efficient, this criterion has been followed faithfully by the current industry. The systematic control of large power systems in response to major faults is effectively nonexistent; instead, methods used are expert-system based and somewhat specific to different parts of the interconnection. They are, by rule, not automated, and the inclusion of a human decision maker is critical.

This article begins by assessing principles of operation by decision and control for today's fully regulated industry. Following this summary, several major ongoing changes are interpreted from a systems point of view. It turns out that major challenges-economic, policy, and technical-are inherently systems problems. To illustrate this claim, new decision-making problems essential to having a successful generation business in a competitive industry are defined in the section "Decision Making in the Newly Evolving Generation Business." This section is very detailed and describes the challenge created by introducing profit-based and risk management issues in the context of under the highly uncertain, often volatile, market price of electricity. Following that, the modified objectives of and decision making by future power delivery companies are described; this problem is complex because of its dependence on the type of ownership and regulatory expectations set upon the future transmission system provider. The discussion shows how these regulatory issues directly impact the hierarchies and complexity of the related dynamic systems problem.

Moreover, the so-called load serving entities (LSEs) or energy service providers (ESPs) will be critical to the technologies potentially useful to specific customers. The idea of differentiated power quality served at different prices so that the users see direct benefits can only be implemented by carefully defining the performance objectives of the LSEs and ESPs and then designing everything else around the need to meet the specifications. This concept is qualitatively different from the top-down thinking in the regulated industry characterized by uniform power quality service and average cost.

Power System Operation by Decision and Control in the Regulated Industry

To briefly review the decision and control approaches in the regulated industry, we consider two types of electric power system architectures: 1) isolated systems comprising a single utility (control area) and 2) an interconnection consisting of several horizontally structured subsystems (utilities, control areas) electrically interconnected via tie-lines. These two designs are conceptually different because a single utility case is characterized by two-level decision and control designs, one being a component level and the other the entire system, whereas in the case of an interconnection consisting of several control areas, the general structure has three levels: component, control area (subsystem), and the interconnection. We start with the simpler case.

Single Control Area Case

The main objective of the integrated utility is to plan enough generation and to design a sufficiently strong delivery system to meet its total load demand at the lowest average cost possible. This general problem of controlling system inputs could be posed as a single dynamic decision-making problem [1], [2]. To do this, consider an electric power system with n nodes whose net generation/demand is controllable and the remaining n_d nodes whose power injections are uncertain load demands. A coordinated operations and planning problem is a combined problem of short-term scheduling of power generated P_i and investment in new generation I_i^C and transmission I_i^T to balance load deviations ranging from hourly through seasonal and long term, and to do this at the lowest total cost, while observing operating and control capacity constraints. A possible mathematical formulation is as follows [2]:

$$\min_{I_i^T,\, I_i^C,\, P_i} E\left\{\sum_i \int_{t0}^{T} \left(e^{-\rho t}(c_i(t, P_i(t))) + C_i^C\left(K_i^C(t), I_i^C(t), t\right)\right) dt + \sum_i \int_{t0}^{T} C_i^T\left(K_i^T(t), I_i^T(t), t\right) dt\right\} \tag{1}$$

subject to equations defining rates of investment

$$\frac{dK_l^T(t)}{dt} = I_l^T(t), \qquad K_l^T(t_0) = K_{l,t_0}^T \tag{2}$$

$$\frac{dK_i^G(t)}{dt} = I_i^G(t), \qquad K_i^G(t_0) = K_{i,t_0}^G, \tag{3}$$

technical limits on transmission line flows in each line l

$$F_l\left(P^G(t), P_L(t)\right) \le K_l^T, \tag{4}$$

generation capacity limits

$$P_i(t) \le K_i^G, \tag{5}$$

and the spot electricity price $p(t)$ dynamics driven by the market demand excess

$$\frac{dp(t)}{dt} = c_{\text{spot}}\left(\sum_{i=1}^{n} P_i(t) - \sum_{j=1}^{n_d} P_{Lj}(t)\right), \qquad p(t_0) = p_0. \tag{6}$$

Here $K_l^T(t)$ stands for the amount of installed transmission capacity for line l; $K_i^G(t)$ is the amount of installed generation capacity at node i; $C_l^T(K_l^T, I_l^T, t)$ is the cost of investment in line l; $C_i^G(K_i^G, I_i^G, t)$ is the cost of investment in generation at node i; $P_i(t)$ is the power generation production at node i, at time t; $P^G(t) = [P_1(t) \cdots P_n(t)]$; $c_i(t)$ is the cost of this production, excluding capacity costs; $P_{Lj}(t)$ is the uncertain (uncontrolled) load demand at node j at time t; $P_L(t) = [P_{L1}(t) \cdots P_{Ln_d}(t)]$; $F_l(P^G(t), P_L(t))$ represents the power flow on transmission line l as a function of generation and demand system inputs; $p(t)$ is the spot electricity market price at time t; and ρ is a discount rate. Note that (6) could be interpreted two different ways: either as a mismatch of mechanical power outputs and the load consumed at time t [9] or as the mismatch between the expected demand the next hour and the power generation bids this hour [7]. The equation does not imply a mismatch between the electric power injected and the power taken out of the system, which always balance. We clarify this because it is a potential cause of confusion. Lagrangian coefficients in this optimization problem are $\mu_l(t)$ associated with the inequality constraint of (4), and $\sigma_i(t)$ with the inequality constraint of (5).

The problem defined in (1)-(6) is a very complex optimization problem in which decisions range from very long term (generation and transmission capacity expansion) to shorter term scheduling of the available generation P^G for the anticipated demand P_L. The optimization is subject to control limits on generation (P_i, $i = 1, \ldots, n$) as well as to output variable constraints (line power flows $P_{ij}^{\max}$) and transmission network, load-flow-type constraints requiring that the electrical power injected into each power system bus balances instantaneously with the electrical power flows from the bus to the rest of the system. This basic online scheduling of generation to meet forecasted load demand is generally based on optimization tools available in major utility control centers, such as unit commitment for turning units on and off and security-constrained economic dispatch [3] for changing the power output of the units that are on to follow anticipated load demand variations. This process takes place every 5-15 minutes in many control centers. The system estimators are critical in preparing system data [4].

This problem formulation is not actively used for dynamic decision making in the regulated industry. Instead, it is assumed that the problem can be decomposed into a shorter term scheduling of $P_i(t)$ and a long-term investment problem. The conditions under which this is valid have never been studied. They could be explored by formulating this problem as a singularly perturbed stochastic control problem [2] and using techniques developed in [5] and [6] to establish sufficient conditions for a meaningful separation of the single complex problem into investment and scheduling subproblems.

Short-Term Generation Scheduling

Assuming that network and generation are known over the entire time horizon, a zeroth order control subproblem becomes a decision-making process about which units to turn on and off and how to adjust the power generated. In this case, network topology and design $K_l^T(t)$, as well as the capacity of generation plants $K_i^G(t)$, are given. Assuming, furthermore, that the daily spot market is at its moving equilibrium [each day there is enough generation to meet load demand and power is sold at the optimum clearing price $p(t)$], a short-term optimization problem becomes [2]

$$\min_{P_i[kT_H]} E\left\{\sum_{k=0}^{k=\frac{T}{T_H}} \sum_{i=1}^{i=n} c_i\left(P_i[kT_H], P_L[kT_H]\right)\right\} \tag{7}$$

subject to the constraints

$$\sum_{i=1}^{i=n+n_d} H_{li}\left(P_i[kT_H] - P_L[kT_H]\right) \le K_l^T[nT_S] \tag{8}$$

$$P_i[kT_H] \le K_i^G[nT_S] \tag{9}$$

and

$$p[(k+1)T_H] = p[kT_H] + c_{\text{spot}}\left(\sum_{i=1}^{n} P_i[kT_H] - \sum_{j=1}^{n_d} P_{Lj}[kT_H]\right). \tag{10}$$

This formulation follows from the full problem (1)-(6), assuming decoupling of the long-term investment decision-making problem evolving each season sample T_S and the short-term generation scheduling each hour T_H (this rate is representative of the new hourly spot markets). H_{li} above is the l,ith element of the so-called distribution factor matrix H that relates the vector of transmission line flows F to the vector of power injections P into all nodes, so that $F=HP$. If the load is assumed uncertain, even this short-term scheduling problem requires dynamic programming tools [3]. Assuming the load is known for the next hour (or estimated [3], [4]), however, this stochastic control problem becomes a deterministic optimization problem subject to physical constraints. At the optimum, assuming the load is known, the nodal price $p_i(t)$ is related to the locational price $p_i(t) = dc_i(t)/dP_i(t)$ at different network buses i (spatial aspects) as [8]

$$p_i(t) = p(t) - \sum_{l=1}^{L} H_{li}\mu_l(t). \tag{11}$$

We consider two types of electric power system architectures: 1) isolated systems comprising a single utility and 2) an interconnection consisting of several horizontally structured subsystems electrically interconnected via tie-lines.

The term $\sum_{l=1}^{l=L} H_{li}\mu_l$, where L is the total number of transmission lines, reflects locational differences in optimal electricity prices caused by the active transmission "congestion" [see (8) above]. These formulas provide the basis for the so-called nodal or location-based marginal cost (LBMC) transmission pricing, or spot pricing [8].

Primary Control for Stabilization

The net real-time mismatch between generation produced and the demand consumed at each instant t results in generally small system-wide stationary frequency deviations. Different systems regulate these stationary deviations in frequency in different ways; single control area systems with much flexible generation (such as hydro systems in Northern Europe) are capable of correcting for the cumulative frequency deviations by manually changing the set point values of the generation-turbine-governor units when so-called time error (proportional to the integral of frequency deviations) exceeds a certain acceptable threshold, without necessarily automating the process. In multiarea systems, such as the one in the United States, this regulation is a system-wide scheme, and it is described later in this section.

Finally, the fastest frequency variations around the values resulting after AGC has acted each τ_s seconds are stabilized by the local controllers, the governor and excitation systems in particular. At present, these controllers are entirely decentralized, constant-gain output controllers. It can be shown that the dynamics of any power plant i, no matter how complex, can be modeled as

$$\dot{x}_i(t) = \tilde{f}_i(x_i(t), u_i(t), y_i(t)) \tag{12}$$

where $x_i(t)$, $u_i(t)$, and $y_i(t)$ are local states, primary control, and local output vectors, respectively. A local continuous primary control $u_i(t)$ (governors and excitation systems) is typically designed to stabilize a local error signal

$$e_i(t) = y_i(t) - y_i[k\tau_s]. \tag{13}$$

With a local controller of this form, the closed-loop local dynamics of a typical power plant connected to bus i takes on a general form

$$\dot{x}_i(t) = f_i(x_i(t), y_i[k\tau_s], y_i(t)). \tag{14}$$

Typical controllers of this type are governors and excitation systems on power plants. If a local controller is of the switching type, one obtains instead a closed-loop model defined as as in (12) with a local control law

$$u_i[(k+1)\tau] = u_i[k\tau] - d_i r_i(e_i[k\tau]), \tag{15}$$

where d_i is a control increment at each step k, acting only at discrete times $k\tau$, $k = 1, \ldots$, where $r_i()$ is a relay-type function. Capacitor/inductor switching and on-load tap-changing transformers are typically used for load voltage control [10]. (Generally there is no explicit relation between τ, the timing at which primary controllers are switched, and the rate at which the set-point values of these controllers are changed at each control area level, τ_s.)

In summary, in the single-control area system, one can identify two levels of generation control: an online system-wide generation scheduling level for meeting total anticipated demand and a very fast stabilization level at each individual generator (component). These two levels are implemented by physically changing the set points of the governors and excitation systems at a rate T_H and in real-time stabilizing the fastest deviations, so that the frequency and voltage of each generator are kept close to their set point values [11]. Observe that the separation of the two control levels has been driven by the temporal separation of load demand deviations evolving at significantly different rates, one anticipated for each hour T_H and the other being much faster, real-time load dynamics.

Multiarea Control Case

The same reasoning as when separating a single control area into two levels allows for spatial separation of an interconnected power system (consisting of several control areas (utilities) and electrically interconnected regions) into three control levels. This leads to the basis for hierarchical operation by decision and control within an interconnection comprising horizontally structured subsystems (control areas) in support of operating and planning of the system (see Fig. 1). Clearly, these hierarchies hold only if there is not much interaction between the subsystems. The main step is deciding if the system is in the normal operating mode, which is based on relatively small changes and weak interactions (measured in terms of deviations of the power flow exchanges between the areas around their scheduled values).

The three-level control concept in the regulated U.S. industry has been based on giving autonomy to each subsystem to plan and schedule generation to meet its own (connected) load demand for some assumed tie-line flow exchange with its neighbors. Each subsystem, having its own control center, then employs the short-term scheduling methods described earlier in this article, with the only difference being that each subsystem attempts to meet agreed-upon tie-line flow schedules with its neighbors. These agreements have historically been bilateral, with the parties cooperating in observing the $(n-1)$ reliability criterion for the entire interconnection. The overall (interconnection, tertiary) level is not coordinated online; instead, each subsystem has a preassigned participation in time-error correction resulting from the cumulative system-wide frequency deviations. To avoid excessive time-error, each subsystem (control area) is equipped with its own fully decentralized AGC (secondary-level control). The principles of this scheme are ingenious and are briefly summarized next. What is important for the purposes of this article is that the fully decentralized AGC works perfectly only when its tuning is done very carefully. Finally, each generator is equipped with the primary stabilization designed the same way as in the single control area case described earlier.

Automatic Generation Control

The remaining generation-demand imbalance created by unpredictable, typically smaller and faster, load demand variations as well as the inertia of turbine-generator-governor units not capable of producing the scheduled electrical output instantaneously has been controlled in an automated way by means of AGC. This very simple, powerful concept is based on the fact that in stationary operation, power system frequency is observable at each location and reflects the total system generation-demand imbalance. Until recently, the accepted industry standard has been to rely on AGC, which is effectively a decentralized output control scheme. The ACE of each area I is the output variable of interest defined as

$$ACE^I[k\tau_s] = F^I[k\tau_s] - 10B^I \frac{\omega[k\tau_s]}{2\pi} \tag{16}$$

where ω is the deviation of system frequency from its nominal value (60 Hz), $F^I[k\tau_s]$ is the deviation of net power flow from the area I, and B^I is known as the area bias, which is chosen as close as possible to the so-called natural response of the area β^I.

Assuming the system is at equilibrium, the basic power to frequency static characteristic is

$$\omega^I = \frac{P_I^G}{\beta^I}. \tag{17}$$

The principle of entirely decentralized AGC in which each subsystem (control area) regulates its own ACE^I relies on the fact that if the frequency bias B^I is chosen to be close to the natural response of the area β^I, then each area will effectively balance its own generation-demand and the entire interconnection will be balanced. When the system is presented with a load demand change $P_d[k\tau_s]$, however, it can be shown that the stationary frequency changes driven by this disturbance can be modeled as [9]

$$\omega[(k+1)\tau_s] = -\frac{(B^I + B^K)}{\beta^I + \beta^K}\omega[k\tau_s] - \frac{P_d[k\tau_s]}{10(\beta^I + \beta^K)}. \tag{18}$$

The main observation here is that the system frequency depends on the *sum* of the subsystem biases ($B^I + B^K$). This is the foundation for so-called dynamic scheduling in today's industry, in which power plants from one area can participate in frequency regulation of the other area. This no longer implies, however, that each area balances its own supply and demand.

Note that the regional automatic voltage control (AVC) in several European countries is based on principles similar to AGC, except that a single frequency measurement is replaced by a set of so-called pilot-point load voltages [12]. No

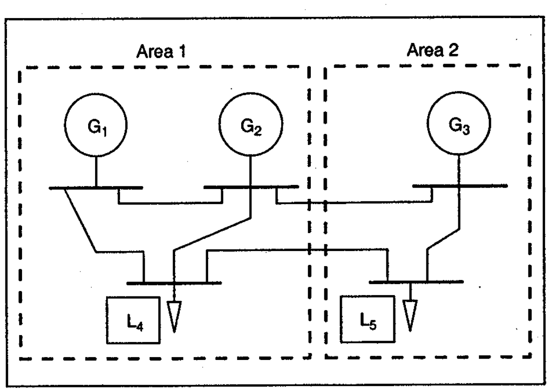


Figure 1. *Generalized structure of decision making and financial flows.*

such automation is in place in the U.S. interconnection. An important observation for comparison with the control problems in the changing industry is that tuning of the frequency bias (and the corresponding control gain in the AVC) is difficult; as the operating conditions vary, the natural response of each subsystem changes. For this reason, it is practically impossible to guarantee a prespecified quality of system frequency.

Moreover, when the $ACE(t)$ is tuned as described above, entirely decentralized AGC regulates deviations of tie-line flows back to their scheduled values and brings system frequency very close to its nominal value under the stationary conditions. Consequently, at present there is no even minimal closed-loop tertiary-level coordination; instead, a cumulative frequency error is corrected for by each area participating in eliminating so-called time error (frequency is used to keep track of time).

Under deregulation, we assume producers act in their own self-interest.

Open Problems in the Regulated Industry

Several important assumptions underlie the operations by decision and control in today's industry. First, a decision is made concerning the "mode" of operation. The famous classification introduced by Dy-Liacco [13] and modified by Carlsen and Fink [14] implies a clear-cut separation between normal and abnormal operations, in a deterministic sense. Several researchers have worked on a probabilistic notion of a system's ability to serve load under equipment outages. The analytic tools for posing the problem this way and, particularly, for solving it using stochastic optimization tools for a typical large-scale power system are nonexistent. Consequently, the industry has adopted a conservative preventive mode of operation in which, for example, generation scheduling is done in a way that ensures sufficient reserve and time to supply load demand in case of *any* single equipment outage, without relying much on secondary control during the outage. This turns out to be extremely costly, and it requires a standby generation reserve on the order of the largest power plant on the system.

This situation is further complicated by potential transient stabilization problems in response to certain large outages. Despite a significant theoretical development in the area of nonlinear control design for power systems [15], none of the primary controllers in real-life operations employ this development. This is unfortunate because nonlinear controllers are potentially capable of stabilizing system dynamics in response to outages that are currently considered critical [16]. Chapter 12 of [15] presents a detailed assessment of open questions concerning the design of primary controllers for stabilizing power systems under stress.

Some of the specific problems described in [15] are:

- Problems related to using a linear system design for controlling nonlinear dynamics.
- Problems caused by inaccurate state variable measurements; consequently, output feedback rather than full state feedback is implemented on the system.
- Problems caused by a lack of coordination of primary controllers. A pragmatic approach to the primary control of power systems has been based on decentralization.

It is critical to assess the performance of a particular design while keeping in mind that these issues are qualitatively different in nature and that potential problems should be related to the correct causes. The reasons for introducing supplemental generator control in the form of power system stabilizers (PSSs), for example, can be attributed to the first two causes listed. The problem of coordination of primary controllers (excitation systems, PSS, and active control devices such as FACTS) primarily relates to the effectiveness of purely decentralized control designs in large-scale electric power systems.

The emphasis of this article is on tertiary- and secondary-level control problems, primarily because many believe that these functions could become market based and would require development of frameworks for guaranteed technical performance in an environment driven by profit/benefit maximization and risk management. Thus it is important to reiterate that the regulated industry uses a preventive mode approach at the scheduling (and investment) level, instead of multistage decision making under uncertainties for optimizing whichever criteria are in place. The result is economic inefficiency, critically caused by observing the $(n-1)$ reliability criterion unconditionally.

Although the $(n-1)$ security criterion is still the industry practice, there has been significant research work done recently toward relaxing it. A relaxed criterion would allow the system to operate under conditions where adequate system controls can eliminate problems caused by contingencies. In the next section, we assess organizational changes underlying the new industry and their impact on the evolving decision-making and control paradigms.

The Changing Industry

Although current operating and planning industry practice has been based on the well-understood spatial and temporal hierarchies (at least in normal operation), the industry is undergoing fundamental structural changes that require new decision and control methods.

The following is an incomplete list of these changes, as viewed from a systems perspective:

- A functional and/or corporate separation of power generation, delivery, and load-serving entities is under way (this is nonuniform, state-dependent within the U.S. interconnection). This organizational change implies, in turn, distributed decisions for supply and demand scheduling, as well as for power delivery.

- A horizontal structure comprising control areas (utilities) as subsystems within the interconnection with non-time-varying boundaries has been losing its basic role under the rules of open-access service: Under retail competition even relatively small customers will have the legal right to purchase electricity from neighboring utilities and not necessarily from the utility to which they are physically connected.
- If system control is based on the existence of control area boundaries, it becomes obvious that the financial and physical processes may not be well related. The financial markets do not actively observe today's utility boundaries.
- The system is generally in a stationary but constantly changing mode driven by the electricity market price dynamics. The nature of imbalances created by these activities is changing, and consequently the hierarchical control based on the principle of each subsystem balancing its own native load is not necessarily the best way to control the system without fixed horizontal structures.
- In the new industry there is an obvious need for decentralized optimization under uncertainties, mainly caused by financial processes. This calls for the development of powerful stochastic optimization methods for market participants.
- Risk-management decision making ranges from very short through long-term time horizons, relevant for planning and investments. Thus multistage decision making becomes a critical tool.
- Relations between the financial and physical processes are often difficult to establish. In particular, system reliability and the $(n-1)$ criterion need new paradigms, characterized by the ability to differentiate quality of service (rate of interruptions or variable levels of voltage variations as seen at the customer's end, for example) at a price.
- Generally, many performance "standards" will be dictated by the load demand itself. This is in sharp contrast to the current top-down standards (control area level and the like).
- Furthermore, although there is no coordination online at the interconnection level for normal operation in today's industry, this may become necessary in the new industry. This is primarily because of active power flow rescheduling between the control areas, which can be interpreted as a stronger connection between the subsystems within a given system [33]. To ensure system reliability under decentralized decision making, some coordination at a system level may be necessary.
- The problem of reliable operation under unplanned outages will require more than voluntary cooperation; typically major technical limits are determined by the equipment in one control area causing operating problems in other areas. The problem of congestion flow control in large power systems [18] and its coordination are basically open research problems.
- The need to maximize profits/benefits of individual businesses and manage risks at the same time is becoming a dominant industry problem. One could value delivery service at various risk levels. Depending on the type of business (risk averse versus opportunistic), the decisions will be qualitatively different.

Fig. 2 presents a generalized structure of decision making and financial (price) flows among the new entities, such as generator operating companies (GenCos), LSEs, and regional transmission organizations (RTOs [19]). As can be seen in this figure, decisions by each business are generally affected by their own profit/benefit maximization and risk management, as well as by the payments made to (or received by) the other businesses that must be internalized. For example, a successful GenCo does its own distributed profit maximization/risk management (as described in the next section), and at the same time must take into account the effect of power delivery, its price p_{priority} at a chosen priority of service, specified in terms of amount of power delivery [20] and quality (amount of power delivery interruption P_{CT}, if the service is nonfirm).

An LSE, on the other hand, maximizes its benefit in a distributed way, while at the same time paying a certain priority access price for the amount of power access P_{access} into the RTO.

An RTO (or any other power delivery entity) also has its financial decision making, in addition to the technical objectives. It is shown in Fig. 2 that its typical cash flows are payments received for serving GenCos at certain priority P_{PAP}, minus the factored-in payment of the capital cost p_{capcost} and minus the payments to the GenCos for providing control services such as AGC and/or AVC and minus the price of operating and maintaining transmission equipment $p_{O\&M}$.

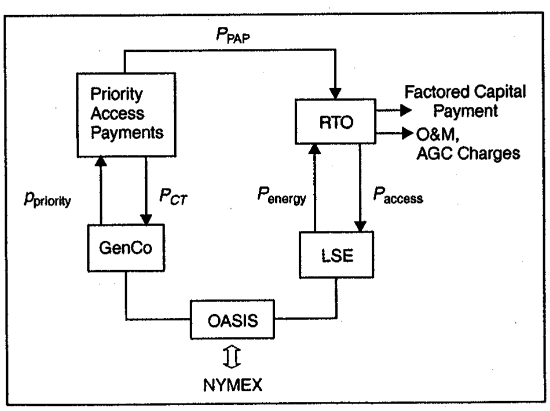


Figure 2. *Real-time transmission pricing.*

In addition to these cash flows, entities such as GenCos and LSEs are also likely to have access to public online information about the status of the power delivery system, such as the presently implemented Open Access Same-time Information System (OASIS) [21]. The most adequate type of information is briefly discussed in the section on power delivery; as described in this section, this is subject to much research. Business entities would also seek active information on the energy spot markets and forward markets (the New York Mercantile Exchange (NYMEX), for example), described in the following section.

Decision Making in the Newly Evolving Generation Business

The ownership of a generator in a deregulated electricity marketplace translates into the ability to convert one type of commodity (oil, gas, or coal) into another commodity: electricity. The cost of the fuel used, coupled with the efficiency of the generation technology, determines the cost of producing electric power. To model the behavior of power producers, we need to define their individual objective functions. Under deregulation, we assume producers act in their own self-interest. Specifically, they will not take action to preserve system reliability or improve power quality unless they are financially compensated for such a service. Therefore, a major component of the objective function is the profit. The profit earned for a given hour k, as a function of market price (p_k) and generator cost (C^G) as a function of output (P_k^G), is

$$\pi_k = p_k P_k^G - C^G\left(P_k^G\right). \tag{19}$$

In a marketplace, future spot prices are uncertain, and thus the profit over a given time period of length n is a sum of random variables

$$\Pi = \sum_{i=1}^{n} \pi_k. \tag{20}$$

A simple objective function would entail maximizing the expected profit. In reality, however, this may not be applicable. GenCos have limited risk tolerance. They may be willing to accept lower expected profits in return for a decrease in the associated financial risk. We model this preference as a risk premium (r). The objective function J for an independent power producer is then written as

$$J = E\{\Pi\} - r\sigma_\Pi \tag{21}$$

where σ_Π is the standard deviation of the total profit. Next we define the inputs, or control variables, to the optimization process. To do so, we first need to specify the market rules under which electricity is traded. Such markets fall into two categories: 1) physical markets, conducted by pools, power exchanges (PXs), and independent system operators (ISOs), trade commitments to produce and consume power; and 2) financial markets, conducted through commodity exchanges, trade financial contracts and derivatives based on the underlying physical markets.

Physical Markets for Electricity

Physical power is traded under many different market structures in the United States, ranging from power pools to power exchanges to independent system operators. Most of these market structures involve a centralized auction mechanism to allocate which generating units should be used to meet the demand. Some areas, such as California, also allow bilateral trades between load and generation. In this section, we will describe the rules governing the California Power Exchange (CalPX). Although somewhat different in structure than its Eastern U.S. counterparts, the CalPX still serves as a good example for understanding the decision process facing producers in the deregulated marketplace. A producer wishing to sell power through the CalPX submits a bid curve to the exchange. The bid curve describes the willingness of the producer to deliver power as a function of market price. For example, a producer may be willing to supply a total of 50 MW if the price is \$20/MWh and may offer to supply a total of 100 MW if the price increases to \$30/MWh. Bid curves are supplied on a day-ahead basis, and a different bid curve may be specified for each of the 24 operating hours. Specifically, a supplier wishing to produce power tomorrow must submit all 24 bid curves by 7:00 a.m. today. The PX gathers all the bids from power producers and similar bids from consumers. The bids are used to compile aggregate supply and demand curves for each hour. The intersection of the supply and demand curves determines the market clearing price (MCP). All supply bids with a price less than the MCP are accepted and the bidders are paid the clearing price. Similarly all demand bids with price higher than the MCP are accepted, and the bidders are charged the clearing price. This ensures that demand and supply commitments match perfectly, as well as that the PX remains revenue neutral.

Financial Markets

Fueled by the physical markets, several financial exchanges have emerged, allowing participants to trade financial contracts and derivatives based on electricity prices. Electricity contracts currently trade on NYMEX and on the Chicago Board of Trade (CBOT). The exchanges trade a number of standard contracts, including forwards, futures, and simple (vanilla) options. In addition, exotic, nonstandard options can be traded over the counter, in bilateral fashion. Derivatives are traded with the electricity spot price as the underlying asset. The spot price, however, varies greatly with the physical location on the network. To get around this problem, the exchange defines a specific node, or bus, on the network on which the price of power is based. Such nodes are known as hubs. The most well-known hubs in the California system are COB (California Oregon Border) and Palo Verde. In the case when contracts specify physical delivery of power, the party that is long power (net supplier) is obligated to purchase transmission rights from the point of injection to the hub. The party that is short power is obligated to purchase transmission rights from the hub to the point of extraction. This type of virtual handoff of power allows a larger number of people to trade in a single hub, thus increasing liquidity in the market.

Market Strategies for an Independent Power Producer

Here we will address the bidding and risk management problem from the perspective of a small- to medium-sized power producer that can be modeled as a price taker. For suppliers with significant market power the decision problem becomes richer and involves game theoretic modeling to determine optimal strategies. See [29] for a discussion on dynamic game-based modeling of electricity markets.

Having described the structure of the markets available to suppliers of electricity, we can now pose the optimization problem facing an independent power producer (IPP). To judge the performance of a strategy, we use the risk-discounted expected profit

$$J = E\{\Pi\} - r\sigma_{\Pi}. \tag{22}$$

In determining the optimal strategy, the model used to describe the generator's variable costs and operating constraints has a profound effect-specifically, the inclusion of startup and shutdown costs as well as minimum run time and downtime for the generator. To illustrate this point, we will first pose the problem with a simple cost function. We later expand the model to include startup and shutdown costs and show how this alters the optimal bidding behavior of the IPP. For an in-depth discussion on the unit commitment problem and the dynamic programming algorithm, see [3] and [24]. In our simple model, the operating cost of the generator for a given hour k is a linear function of the output

$$C^G\left(P_k^G\right) = bP_k^G + c. \tag{23}$$

Next we make the assumption that the producer does not posses market power. Market power can be described as the ability to manipulate price to increase profits. The absence of market power is implicitly included in the model by allowing price p_k to be independent of the production level of the IPP. Our final assumption relates to the tradeoff between risk and expected profit in the objective function of the IPP. Recall that the producer has two markets available to trade in: physical and financial. We will assume that the producer attempts to maximize his profit in the physical market and uses the financial market to manage his risk. Applying this assumption, we can pose the subproblem of designing a profit-maximizing bid strategy for the day-ahead physical market. Recall that the market rules require the supplier to submit bids for all 24 hours at the same time. The profit earned in hour k is given by

$$\pi_k = p_k P_k^G - bP_k^G - c. \tag{24}$$

This results in a profit-maximizing output level as a function of price given by

$$P_k^G = P_{\max}^G \text{ if } p_k > b, \text{ and } 0 \text{ otherwise}, \tag{25}$$

which is equivalent to bidding the unit's marginal cost. Next we recognize that in our current model of the generator, there is no coupling between costs in consecutive time periods. Whether the unit is on or off in hour k has no impact on the cost of running the unit in hour $k+1$. This decouples the problem, and we can solve the 24-hour optimization by solving each hour separately. In this way, the profit-maximizing strategy for an independent power producer with no market power and no startup or shutdown costs is to bid its marginal cost for each hour. It is interesting to note that one does not need to know the characteristics of the price process to formulate the profit-maximizing bid strategy. Submitting the marginal cost as its bid ensures the supplier that the power exchange will automatically dispatch the unit when it is profitable to run and reject its bid when price is below cost. As a result, the total profit is bounded from below by the fixed cost c

$$\pi_k = \max\left\{-c, \left(p_k P_{\max}^G - bP_{\max}^G - c\right)\right\}. \tag{26}$$

Managing Risk Through the Financial Markets

Having defined its profit-maximizing bid strategy, the supplier faces an uncertain future revenue stream. In this section, we show how financial contracts can be used to manage this risk factor. One of the most common derivatives used to hedge volatile positions is the European Call Option. The buyer and seller of the option specify the underlying asset, in this case, the spot price of electricity at a given hub. They also specify the quantity of power (Q), the maturation (delivery) date (T), the strike price X, and the premium c^E (the purchase price of the option). The payoff for the buyer of the option, if exercised, is the difference between the spot price of electricity and the strike price of the option. If the spot price is lower than the strike price, the option will not be exercised and the payoff is zero. Although options can entail physical delivery of the underlying commodity, we will here consider the case where they are settled financially. The cash flow (f) for the buyer of the option can then be written as

$$f = \max\left\{-c^E, \left(p_k Q - XQ - c^E\right)\right\}. \tag{27}$$

This formula takes on the same form as the expression for the profit of an IPP bidding its marginal cost. If we set the quantity of power to equal the maximum output of the generator and the strike to equal the marginal cost, we arrive at a payoff

$$f = \max\left\{-c^E, \left(p_k P^G_{\max} - bP^G_{\max} - c^E\right)\right\}, \tag{28}$$

which is identical to that of the IPP except for the fixed cost component. Now consider what happens when a supplier adopts the marginal cost bidding strategy and simultaneously goes short on (sells) the options contract described above. The net payoff for a given hour is

$$f = \max\left\{c^E - c, c^E - c\right\} = c^E - c, \tag{29}$$

which is deterministic. The options contract is designed to mimic the uncertain portion of the supplier's original position and is therefore known as a replicating portfolio. Taking a short position in the replicating portfolio cancels out the uncertainty and leaves the supplier with a certain payoff. To hedge its position over a given time period (day, week, or month), the IPP sells a series of options for each hour of the period. The quantity and strike price are identical for all options. In reality, the expected profit from the unhedged position may be considerably higher than the certain profit offered by the replicating strategy. The supplier may therefore want to implement a partial hedge to reduce the risk to tolerable levels. He does this by varying the quantity of options traded. Fig. 3 describes the tradeoff between risk, measured by the standard deviation of returns, and expected profit. The hedging ratio, h, is the ratio of Q to $P^G_{\max}$. Π_U is the expected profit from the unhedged position, and Π_H is the certain return from the fully hedged portfolio. We then add the supplier's objective function given by

$$J = E\{\Pi\} - r\sigma_\Pi. \tag{30}$$

The intersection of the objective function and the payoff of the hedged portfolio determines the optimal hedge ratio for the IPP.

Valuation

The ability to construct a portfolio of contracts in the financial market that accurately replicates the cash flow in the physical market is significantly beyond the scope of risk management. The issue is that for every decision process in the physical world there exists a dual problem in the financial world. In our case, the process of building and operating a power plant is equivalent to purchasing and exercising call options. Therefore, a parity in the valuation of physical assets and the pricing of electricity options must exist. Estimating the value of either the option or the physical asset requires knowledge of the stochastic behavior of electricity prices. To price a European call option on the price of electricity in hour k, one needs to know the probability density function (pdf) $f_p(p)$ for the random variable p_k. The value of the option is the integral of the pdf times the payoff of the option at each price level. Recall that the payoff is zero if p_k is below the strike (X), and is equal to $p_k - X$ otherwise. The risk-neutral value of the call option is given by

$$c^E = \int_X^\infty (p_k - X) f_{p_k}(p_k)\, dp_k. \tag{31}$$

Recall, however, that in examining the IPP's position in the physical market, we were interested not only in the expected profit, but also in its variance. Specifically, we considered the variance of the total cash flow over the course of a day. This is equivalent to finding the variance on the payoff from a portfolio of 24 separate call options. Calculating the variance of this portfolio requires knowledge about the correlation of electricity price at different time periods. To resolve this problem, we need to postulate a model for the stochastic evolution of electricity price.

Modeling Electricity Prices

Numerous papers have been written proposing stochastic models for electricity prices, underscoring the crucial importance of such models in the planning and operations of assets, as well as in the pricing of electricity based derivatives (see [3], [23], and [25]-[27]). Candidate models face

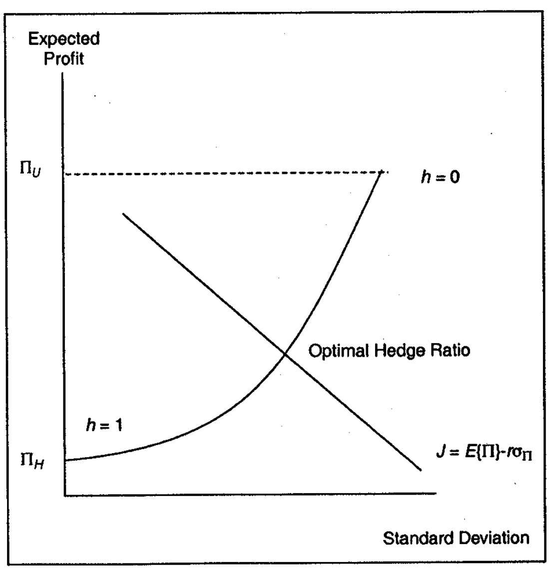


Figure 3. *Determining the optimal hedge ratio for power producers.*

two conflicting performance objectives. They need to be detailed enough to capture the complex behavior of electricity markets, notably the high price volatility and the daily, weekly, and yearly seasonality. At the same time, the model must be simple enough to lend itself as an input to derivative valuation schemes as well as dynamic optimization of asset operations. One of the well-studied processes that seems to meet these requirements is the mean-reverting process, here expressed in continuous time

$$\frac{dp}{p} = \alpha(\ln\mu - \ln p)\,dt + \sigma dz. \tag{32}$$

Here μ denotes the time-varying mean to which price tends to revert at rate α. The uncertainty is driven by a Wiener process Z. The model is attractive because the uncertainty in price at any point in the future is lognormally distributed. This greatly simplifies the integration required in the option valuation process. Although, in practice, the model is somewhat oversimplified (in reality multiple sources of uncertainty are needed to model market realities), it serves as a good example of the types of processes that can be applied in decision and valuation algorithms. From the spot price process we can derive a process for the forward curve. We define the forward price $F(t,T)$ to be the price of a forward contract at time t, in $/MW, for delivery at time T. Under the risk-neutral assumption, we can define the forward price as the expected value of the spot price at maturity

$$F(t,T) = E_t\{p(T)\}. \tag{33}$$

This allows us to derive an expression for the stochastic evolution of the forward price

$$\frac{dF}{F} = e^{-\alpha(T-t)}\sigma dz. \tag{34}$$

The link between the forward price process and the spot price process is crucial in formulating a risk management strategy. The mean reverting property of the spot price predicts that the effect of a spot price change on the forward price will decay exponentially with the maturity of the forward price. The relative effect of a spot market change on the forward price at different maturities is known as the term structure. Accurate knowledge of the term structure allows a trader to use forward contracts to hedge spot market risk or vice versa.

A Realistic Model for the IPP

In our first attempt at describing the decision process of a supplier, we presented an oversimplified model of the cost structure of the generator. The intent was to show the strong link between valuations in the physical and financial markets. An actual supplier, however, faces a more complex decision process. In expanding the cost structure to a more realistic scenario, we show that it is still possible to offset the risk incurred in the physical market with financial contracts. In designing the new replicating strategy, however, we require new types of derivatives and the use of dynamic programming to value them. The form of the model used to describe the cost of the IPP is dependent on the nature of the unit. In this section we consider a fossil plant with significant startup and shutdown costs. Later we contrast this with the decision making for the operator of a hydro power unit. The two examples illustrate the profound effect of generation technology on both the physical scheduling and financial hedging of the asset. The first extension of the cost model is to introduce a quadratic term

$$C^G(P_k^G) = a(P_k^G)^2 + bP_k^G + c. \tag{35}$$

The profit-maximizing output for a given price is now

$$P_k^G = \frac{p_k - b}{2a}. \tag{36}$$

This change in itself presents a problem with our replicating strategy. The generation level is now a linear function of the market price. To mimic this behavior, we would need to purchase a series of options with linearly increasing strike prices. The second modification to the cost structure is the introduction of a startup cost (S) and shutdown cost (T). The effect of introducing these costs is to link the operation of the power plant in consecutive time periods. The cost of operating at a certain output in hour $k+1$ now depends on whether the unit was on or off in hour k. As a result, the temporal decoupling assumption we employed in the previous case is no longer valid. To pose this problem, we introduce the state variable x_k, which is equal to zero if the unit is off and equal to one if the unit is on. We also introduce the decision variable u_k. If u_k is one, the unit is bid at marginal cost for that hour, and if u_k is zero, no bid is submitted. We can now formulate the decision process as a dynamic programming problem (see [3] and [24]). The profit for the producer in stage k is given by

$$\pi_k = u_k\left[p_k P_k^G - C^G(P_k^G) - (1 - x_{k-1})S\right] - (1 - u_k)[c + x_{k-1}T]. \tag{37}$$

The IPP needs to determine the bid strategy $\{x_1,\ldots,x_{24}\}$ that will maximize the total expected profit over the next 24 hours. The input to the optimization is the stochastic price process

$$\frac{dp}{p} = \alpha(\mu - \ln p)\,dt + \sigma dz. \tag{38}$$

Replicating Strategies

Creating a replicating portfolio for the expanded IPP model is a difficult problem. First, the volume traded in each hour is now a function of market price. Second, the actual cost of running the unit in a given hour is not known since it depends on the state of the unit in the previous hour. There is no way to design a perfect replicating strategy using only simple call options. This leaves us with two possibilities. First, we can design a strategy using exotic options (see [22]). This may include swing options, which gives the buyer flexibility in the quantity and timing of the exercise. Such options contracts could be tailored to match the characteristics of the unit. This still leaves open the question, however, of how to value such an option. A second possibility is to use a set of standard options contracts to attempt to approximate the cash flow from the generator. To accomplish this, one must use knowledge of the stochastic behavior of the price to estimate when the unit will be running and at what output levels. In essence, the three problems-optimal bidding of the generator, pricing of the exotic option, and finding a replicating strategy using standard options-are all equivalent. The optimal exercise scheme for the swing option is equivalent to the optimal bid strategy of the generator. All require us to solve the dynamic programming formulation.

Operation and Hedging for Hydro Power

In the previous section we saw how the introduction of startup and shutdown costs into the cost function of a plant coupled the decision process across operating hours. For hydro power with reservoir constraints, this effect is even more profound. The operator of the plant must decide on the optimal time to use the water stored in the reservoir. His decision will be a dynamic programming type formulation, based on the current reservoir level, the expected rate of water flow into the reservoir, as well as projected future spot prices. The time scale over which the optimization is solved will depend on the size of the reservoir. The problem becomes more complex if multiple reservoirs are cascaded, or if additional flow constraints are introduced for environmental or security reasons. An in-depth formulation of the multireservoir scheduling problem is presented in [28]. From a risk management perspective, the problem takes on a different form depending on the time frame of the observer. In the short run the hydro unit gives the operator a high level of flexibility. It can take full advantage of intradaily swings in electricity prices with its high ramping speed. On a longer time scale, however, the unit is critically dependent on precipitation to maintain reservoir levels. An unusually dry season may cripple the economic viability of a hydro unit. This effect is amplified by the fact that rainfall is not a traded commodity (though this may change as weather derivatives gain momentum). Thus the owner of the unit has no means to effectively hedge this long-term risk. A closely related problem is the scheduling and hedging of pump storage units. Similar to the hydro problem, pump storage introduces the additional constraint of being both a load (in pumping mode) and a supplier of power. The pump storage device fits nicely into our hedging scheme since it is effectively the sum of two types of derivatives. By being both a consumer and supplier of power, the unit seeks to exploit the temporal, intraday, difference in the price of power. This can be modeled as an option on the spread between on- and off-peak power prices. Furthermore, the plant has some flexibility (depending on reservoir size) to carry a reserve across multiple days. This is equivalent to a swing option on power. The exact combination of these two derivatives used to model the plant will depend on the physical parameters-reservoir size, pump efficiency and rated capacity.

Hedging with Block Forwards

An effective replicating strategy, even under the simple cost structure, requires the purchase of individual hourly options. Although the physical markets operate on an hourly basis, the financial markets do not. Options and futures are sold based on 16-hour on-peak blocks and 8-hour off-peak blocks. Contracts are further grouped into weekdays versus weekends. A typical contract on COB electricity therefore would be sold on a 5×16 (five weekdays, 16 on-peak hours) basis. This type of averaging is used to create more liquidity in the market. For the IPP, however, it makes the hedging problem much harder. The supplier must now relate the volatility of the hourly spot price to the volatility in a 16-hour block. He must then take positions in the block contract to hedge the risk imposed by the startup and shutdown costs. Such discrepancies between the physical and financial markets put pressure on power marketers to issue custom-made derivatives contracts to power producers. To price these contracts, and later offload the risk by trading in the liquid standardized options and futures markets, the marketers need to develop effective pricing algorithms. One of the key elements in the pricing of these contract is the selection of accurate price models. These models allow the user to interpolate volatilities in exotic contracts by observing the spot markets and the liquid financial markets. The second stage, however, is to solve the dynamic programming problem posed for the physical markets and replicated in the financial contracts.

Decision Making by the Power Delivery Entities

Power delivery under competition is a very difficult problem. Although the basic objective is to build enough transmission to facilitate power delivery from the points of supply to the points of consumption, as in the regulated industry, the objectives of transmission provision functionally (and often corporately) separated from the generation production and consumption processes are subject to huge uncertainties concerning the actual demand for transmission (the overall availability of power plants is market driven and not easily observable by a transmission provider). In addition, because of environmental constraints, it has become increasingly difficult to build new transmission. Instead, the need has increased for "smarter" transmission capable of a *direct line flow control* [15]. This is particularly important since transmission providers are not likely to always have physical control over generation. Moreover, controlling transmission line power flows using both direct flow control and generation inputs is known to cause problems with system-wide control coordination.

The industry organization raises many basic questions concerning electric power delivery and its control. Some of these are listed here and briefly addressed.

Single Control Area Case

A simpler problem concerns power delivery in an electric power system consisting of only a single control area. Depending on the specific energy market structure, decisions concerning power delivery are either made by a separate entity or are bundled with the energy market. So-called mandatory Poolcos are equivalent to the present tight power pools, except that cost functions are replaced by the bid functions described earlier. A more novel case is where the energy market is based on the decision making described under "Decision Making in the Newly Evolving Generation Business," and transmission management is carried out by an independent system operator. Many variations are possible, in which generators and consumers can bid to participate in the so-called transmission congestion management and be compensated competitively [30] or are allowed to trade as long as they observe a technical constraint defined by the ISO [31].

This is a good place to illustrate the dependence of the type of control variables and the definition of the objective function on the industry organization. In a mandatory Poolco, all generation and demand bids are used to optimize what amounts to the performance criterion given in (1) above, except that costs are replaced by the bids.

In a voluntary bidding for participating in transmission congestion management, separate bid functions are used to eliminate transmission limit violations, at the minimal cost. Energy market decisions are made separately.

Finally, in a multilateral market all economic decisions are internalized to the market decision makers and the objective of a transmission provider is to compute active technical constraints only. For a comparison of the three approaches, see [32, chap 2].

Unbundling the Power Industry into a Transmission Service and Its Users

By formulating the dual problem to the problem defined in (7)-(10), it becomes clear that the values $p(t)$ and $\mu_l(t)$ are coordinating variables for dispatch of power. We will assume that the energy price $p(t)$ will result naturally from information exchanges between generators and consumers. The entire optimal control problem (7)-(10) then reduces to the choice of transmission prices $\mu_l(t)$ and capacities $K_l^T(t)$. It is critical to recognize that the two decisions are mutually dependent. Thus a transmission pricing scheme cannot ignore the investment policy, and the investment policy should be dependent on the amount of congestion on the grid. The main issue in this formulation of a transmission pricing scheme as a coordinating activity resides in the *information structure* of the problem. The question of the minimal information flow necessary for coordination, under decentralized decision making by transmission users, is essential. Not only is a transmission provider unable to predict with perfect certainty the future values of demand, but the provider does not know the cost structure of generators. The transmission industry should be structured in a way that enables the incorporation of this information in the appropriate time frame. An interesting question is how to achieve near-optimal decision making by the users themselves, instead of having to depend on the transmission provider. Fig. 4 illustrates the basic feedback role of the signal $\mu_l(t)$.

Multiarea Case

The problem of power delivery by a transmission system affected by the financial processes, illustrated in Fig. 2, raises many systems-type questions. To start with, each control area is currently required to post its available transfer capability (ATC) [32, chap. 2, 3]. This is perceived as necessary to avoid intentional barriers to entry to those attempting to sell or buy outside their own control areas. This issue simply requires an estimate of the control area as an aggregate. Moreover, because the ATC computations typically use data about one specific control area, this implies suboptimality relative to the transfer capability computations obtainable at the entire interconnection level, as well as the danger of possible reliability problems caused by lack of coordination among the control areas when attempting to compute how much each area could transfer. These types of issues have prompted U.S. federal regulators [19] to recommend the creation of much larger transmission-serving entities, or so-called regional transmission organizations (RTOs). Defining the natural borders for these RTOs, their size, and other relevant criteria are clearly ques-

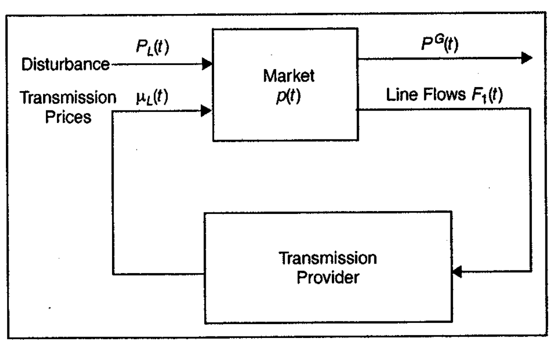


Figure 4. *Multiarea network.*

tions concerning the interplay of tertiary- and secondary-level scheduling processes so that power delivery is not a major obstacle to implementing the energy trades across large geographical areas.

Conceptually, it is possible to have two qualitatively different approaches to ensuring coordinated transmission service:

- Create RTOs freely and have an overall market for scheduling power exchange between them. A tertiary-level control model advocated in [11] several years ago naturally lends itself to implementing this idea in a systematic manner. This would result in a minimal online coordinating mechanism, not currently in place in the regulated industry, based on partial cooperation.
- Aggregate large groups of transmission system users into so-called congestion clusters using their relative impact on flows in the most likely and most critical congested lines. These clusters generally change on a seasonal basis, and could be posted as part of the OASIS mechanism.

This public information mechanism could play a critical role in inducing evolution of active transmission markets, as described in [2]. Moreover, it is conjectured here (without a formal proof) that if the clusters are determined systematically, then the bottom-up processes of internalizing the priority service values by the system users could lead to technical conditions similar to those associated with the process of a transmission company projecting the need for transmission demand and investing in the right places while selling according to the priority-based service [20].

Most appealing is the idea of probabilistic multistage decision making, much in the same spirit as the decision making described for the generation business. This approach could lead to the development of software for valuing financial transmission rights, which could be interpreted in terms of the physical capacity of the transmission lines delivering power. This is a wide-open area and potentially very lucrative as a performance-based transmission business begins to develop [35].

Conclusions

We have suggested in this article that the range of open decision and control problems in the changing electric power industry is broad. Most important, the operating rules determined by each particular industry structure will require different decision-making and operating paradigms. As a rule, the more competition among various industry entities, the more decentralization in the decision making and the greater the need for stochastic optimization with reduced information. Consequently, multistage dynamic decision making becomes a must for successful businesses.

The relations between financial and physical processes become particularly intriguing and relevant because the control schemes in place (automatic generation control, most notably) are based on the premise that each subsystem (control area, utility) serves its own electricity users only; the interconnections with neighboring subsystems are designed primarily for exchanging help under unusually stressful events. In the competitive industry, the existing utility boundaries are not actively observed by the financial trades, which often take place primarily for economic reasons. As illustrated, relating financial processes and actual physical trades of a nonstorable physical quantity such as electric power requires careful decision making even in the seemingly simplest case of selling electricity by a power producer. An enormous challenge that remains is to establish meaningful paradigms on the power delivery side (characterized by the vast network complexity, temporal and spatial hierarchies, and mapping difficulties between the present control regions and the newly evolving boundaries of the electricity markets) in which financial and physical processes are well understood and could form the basis for future delivery companies.

Moreover, although not often thought of as active decision makers, the electricity users will play a determining role in the success of the competitive electric power industry. The top-down specifications for reliability are likely to be replaced by specifications for power quality (e.g., rate of interruption and acceptable range of voltage variations) by various users in response to the price of electricity. The role of demand elasticity may be very effective on hot summer days when one needs a little incentive to use less air conditioning when there is a temporary shortage of electricity. This would eliminate much of the need for unconditional $(n-1)$ reliability criteria and much of the standby reserve. Society is likely to see tremendous cumulative gain by conserving at a price, particularly in the most developed parts of the world.

Most exciting to us is that to push the operating paradigm for future power systems to this new stage will require a very careful rethinking of the problem as a large-scale dynamic system, driven by the typical stochastic variations in load demand, as well as by an active response to the system status reflected in the price of electricity. The temporal and spatial aspects of this are overwhelming on a full-blown detailed system representation. The criteria for what we once thought of as system aggregation must be carefully understood and should be based on distributed objectives of the newly evolving businesses. The modularity question of the newly evolving units, such as GenCos, LSEs, transmission-providing and/or operating companies (see [19]), in the context of their overall technical and financial objectives, will require careful study from a systems point of view. The ultimate objective is to allow for as much decentralized decision making by these entities as possible, yet to provide them with sufficient flexibility and smart technologies so that they jointly approach theoretical limits achievable in a fully coordinated environment under near-complete certainty.

The role of systems thinking and the facilitating technologies, ranging from smart estimation and control gadgets at the individual LSE, GenCo, and power delivery levels to the most effective computer software for near online implementation of electricity pricing as yet another feedback scheme on a complex system, is huge. The challenge to us could represent either a missed opportunity to put all our systems knowledge to work, by means of today's fast-growing information technology tools, or a unique opportunity for a tremendous success.

Acknowledgments

The authors greatly appreciate the financial support under the EPRI/DoD Program on Complex Interactive Networks, as well as the support by the members of the MIT Energy Lab Consortium on New Concepts and Software for the Evolving Electric Power Industry.

References

[1] M. Ilic, J.-P. Leotard, and C. Rochlin, "On the objectives of transmission pricing under open access," in *Proc. Symp. Bulk Power System Dynamics and Control IV: Restructuring*, Santorini, Greece, 1998.
[2] C-N. Yu, J.-P. Leotard, and M. Ilic, "Dynamics of transmission provision in a competitive power industry," *J. Discrete Event Dynamic Systems: Theory and Applications*, pp. 351-388, 1999.
[3] E. Allen and M. Ilic, *Price-Based Commitment Decisions in the Electricity Market* (Advances in Industrial Control). London: Springer, 1998.
[4] A. Monticelli, *State Estimation in Electric Power Systems*. Norwell, MA: Kluwer, 1999.
[5] A. Bensoussan, "Singular perturbation results for a class of stochastic control problems," *IEEE Trans. Automat. Contr.*, pp. 1071-1081, 1981.
[6] H. Khalil and Z. Gajic, "Near-optimum regulators for stochastic linear singularly perturbed systems," *IEEE Trans. Automat. Contr.*, pp. 531-541, 1984.
[7] D. Siljak, *Large Sale Dynamic Systems: Stability and Structure*. Boston, MA: North-Holland, 1991.
[8] F.C. Schweppe, M.C. Caramanis, R.D. Tabors, and R.E. Bohn, *Spot Pricing of Electricity*. Norwell, MA: Kluwer, 1988.
[9] M. Ilic, P. Skantze, C.-N. Yu, L. Fink, and J. Cardell, "Power exchange for frequency control (PXFC)," in *Proc. Bulk Power System Dynamics and Control IV: Restructuring*, Santorini, Greece, 1998.
[10] J. Medanic, M. Ilic, and J. Christensen, "Discrete models of slow voltage dynamics for under load tap changing transformers," *IEEE Trans. Power Appar. Syst.*, 1987.
[11] M. Ilic and S.X. Liu, *Hierarchical Power Systems Control: Its Value in a Changing Industry* (Advances in Industrial Control). London: Springer, 1996.
[12] J.P. Paul, J.Y. Leost, and J.M. Tesseron, "Survey of the secondary voltage control in France," *IEEE Trans. Power Syst.*, 1987.
[13] T.E. Dy-Liacco, "The adaptive reliability control system," *IEEE Trans. Power Appar. Syst.*, pp. 517-531, 1967.
[14] L. Link and K. Carlsen, "Operating under stress and strain," *IEEE Spectrum*, pp. 48-53, Mar. 1978.
[15] M. Ilic and J. Zaborszky, *Dynamics and Control of Large Electric Power Systems*. New York: Wiley, Mar. 2000.
[16] J.W. Chapman, M. Ilic, and C.A. King, "Stabilizing a multimachine power system via decentralized feedback linearizing excitation control," *IEEE Trans. Power Syst.*, 1992.
[17] J. Zaborszky, K. Prasad, and K.W. Whang, "Operation of the large interconnected system by decision and control," *IEEE Trans. Power Appar. Syst.*, pp. 37-45, 1980.
[18] *Standards and Communication Protocols for Open Access Same-time Information System (OASIS)*, FERC, Sept. 29, 1998.
[19] NERC Operating Manual Appendix 9C, "Transmission loading relief procedure," Nov. 1, 1999.
[20] S. Deng and S. Oren, "Priority network access pricing for electric power," in *Proc. Bulk Power System Dynamics and Control IV: Restructuring*, Santorini, Greece, 1998.
[21] *Regional Transmission Organizations*, FERC order 2000, Dec. 20, 1999.
[22] J.C. Hull, *Options, Futures, and Other Derivatives*, 4th ed. Englewood Cliffs, NJ: Prentice Hall, 2000.
[23] D. Pilipovic, *Energy Risk*. New York: McGraw Hill, 1998.
[24] D.P. Bertsekas, *Dynamic Programming-Deterministic and Stochastic Models*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[25] P. Skantze, M. Ilic, and J. Chapman, "Stochastic modeling of electric power prices in a multi-market environment," in *Proc. IEEE PES Winter Meeting*, Singapore, Jan. 2000.
[26] S. Deng, B. Johnson, and A. Sogommonian, "Exotic electricity options and the valuation of electricity generation and transmission assets," in *Proc. Chicago Risk Management Conf.*, May 1998.
[27] B. Johnson and G. Barz, "Selecting stochastic processes for modeling electricity prices," *Energy Modelling and the Management of Uncertainty*. Risk Publications, 1999.
[28] P. Luh, X. Guan, E. Ni, and R. Li "Optimization-based scheduling of cascaded hydro power systems with discrete constraints," in *Proc. North American Power Symp.*, Nov. 1996.
[29] P. Visudhipan and M. Ilic, "Dynamic game-based modeling of electricity markets," in *Proc. IEEE PES Winter Meeting*, New York, Feb. 1999.
[30] W. Hogan, "Contract networks for electric power transmission," *J. Regulatory Econ.*, pp. 211-242, 1992.
[31] F.F. Wu and P. Varaiya, "Coordinated multilateral trades for electric power networks: theory and implementation," Univ. of California, Berkeley, Power Rep. PWR-031, 1995.
[32] M. Ilic, F. Galiana, and H. Fink, Eds. *Electric Power Industry Restructuring: Engineering and Economics*. Norwell, MA: Kluwer, 1998.
[33] M. Ilic and L. Hyman, "Getting it right the first time: value of transmission and high technology," *Electricity J.*, pp. 8-17, Mar. 1996.
[34] R. Wilson, "Interpretation of priority insurance in power exchange networks," *Energy J.*, 1997.
[35] *Transmission Companies: A Blue Print*, S. Awerbuch, L. Hyman, and A. Vessey, Eds. Public Utilities Report, 1999.

Marija Ilic has been at MIT since 1987 as a Senior Research Scientist in the EECS Department. Since September 1999 she has had a 50% appointment at the National Science Foundation as a Program Director for Control, Networks and Computational Intelligence. Prior to joining MIT, she was a tenured faculty at the University of Illinois at Urbana-Champaign. She is a recipient of the First Presidential Young Investigator Award for Power Systems, an IEEE Fellow, and an IEEE Distinguished Lecturer. She has co-authored several books on the subject of large-scale electric power systems: *Dynamics and Control of Large Electric Power Systems*, New York: Wiley, 2000; (as coeditor) *Power Systems Restructuring: Engineering and Economics*, Norwell, MA: Kluwer, 2000; *Price-Based Commitment Decisions in the Electricity Markets*, London: Springer-Verlag, 1999; and *Hierarchical Power Systems Control: Its Value in a Changing Industry*, London: Springer-Verlag, 1996. Her interest is in control and design of large-scale systems.

Petter Skantze was born in Sandviken, Sweden in 1973. He recieved his bachelor of engineering and master of engineering in electrical engineering from MIT in 1998. He is currently pursuing his Ph.D. at MIT's Energy Laboratory. His research focuses on modeling price dynamics in deregulated electricity markets and its impact on operations, risk management, and investment in the industry.