

Allocating Costs of Ancillary Services: Contingency Reserves and Regulation

June 2003

**Eric Hirst
Brendan Kirby**

ALLOCATING COSTS OF ANCILLARY SERVICES: CONTINGENCY RESERVES AND REGULATION

Eric Hirst
Consulting in Electric-Industry Restructuring
Bellingham, Washington 98229

and

Brendan Kirby
Power Systems Research Program
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831

June 2003

Prepared by
OAK RIDGE NATIONAL LABORATORY
P.O. Box 2008
Oak Ridge, Tennessee 37831-6285
managed by
UT-Battelle, LLC
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vi
ABBREVIATIONS, ACRONYMS, AND INITIALISMS	vii
SUMMARY	ix
1. INTRODUCTION	1
2. CONTINGENCY RESERVES	3
PURPOSE OF RESERVES	3
DATA RESOURCES	5
REQUIRED VERSUS PURCHASED RESERVES	5
RESERVE PRICES	6
NUMBER AND SIZE OF OUTAGES	6
MAXIMUM OUTPUT OF LARGEST GENERATOR ONLINE EACH HOUR	7
HOURLY OUTPUTS OF ALL LARGE GENERATORS	8
SHARE OF LOAD	10
EFFECTS OF METHODS ON INDIVIDUAL GENERATORS	11
RECOMMENDATIONS	13
3. REGULATION	15
PURPOSE OF REGULATION	15
DATA RESOURCES	17
REGULATION PURCHASES AND PRICES	17
SYSTEM-LEVEL METRIC	17
ALLOCATION OF COSTS TO INDIVIDUAL CUSTOMERS	19
RESULTS	21
RECOMMENDATIONS	23
4. CONCLUSIONS	25
REFERENCES	27

FIGURES

S-1	Contingency-reserve costs for four generators and the major intertie if charged solely on the basis of the number and size of forced outages or the size of the large units online each hour	ix
S-2	Relationship between the share of regulation costs and energy consumption for several customers	x
1	One-minute loads of two large industrial loads, normalized to the same energy use for the hour shown	2
2	Daily pattern of contingency-reserve requirements and purchases	6
3	The time for recovery as a function of outage size for the 27 outages between July 2001 and June 2002	6
4	Allocation of the 27 outages and the 9100 MW of outages among the 13 units with outages	7
5	The percentage of the hours of MW-hr reserve required during which eight generators plus a major intertie were the largest contingencies	8
6	Assignment of responsibility for reserves on the basis of unit output or share of total load served	9
7	The share of annual costs for contingency reserves based on the largest-single-contingency criterion	10
8	The share of annual costs for contingency reserves based on the 5% plus 7% rule	11
9	Components of a hypothetical load on a weekday morning	15
10	Average monthly (top) and hourly (bottom) regulation prices and quantities purchased, July 2001 through June 2002	18
11	Minute-by-minute values of system load and the load-following and regulation components for one hour	19
12	Geometric allocation of individual loads A and B to regulation Total	20

13	Relationship between the share of regulation costs and energy consumption for 15 customers	21
14	The annual bills for regulation for 15 transmission customers based on the method developed here and based on energy use	23

TABLES

1	Allocation of the costs of contingency reserves to generators	12
2	Ratio of share of regulation cost to share of energy consumption for 15 electricity customers	22

ABBREVIATIONS, ACRONYMS, AND INITIALISMS

ACE	area control error
AGC	automatic generation control
CPS	Control Performance Standard
DCS	Disturbance Control Standard
ERCOT	Electric Reliability Council of Texas
MWh	MW of electricity provided for an hour
MW-h	MW of ancillary service provided for an hour
NERC	North American Electric Reliability Council
WECC	Western Electricity Coordinating Council

SUMMARY

In general, the costs of the real-power ancillary services (spinning reserve, supplemental reserve, and regulation) amount to several percent of the wholesale cost of electricity. These costs are typically recovered from transmission customers on the basis of a very simple billing determinant—MWh of energy. However, hourly energy use has nothing to do with the factors that require contingency reserves and regulation.

Because of the high cost of these services and the possible inequities in the current system used to charge transmission customers for these services, we collected and analyzed data on alternative cost-allocation methods. For contingency reserves, we focused on the number and size of forced outages, the sizes of the large generators online each hour, and system load. For a particular control area, we obtained hourly data on the output of the large generating units plus the flows across the major interties, the number and magnitude of forced outages, the amounts of contingency reserves required and acquired, and the prices of spinning and supplemental reserves. We used these data to develop and test alternative ways to assign contingency-reserve costs to individual generators. Although retail customers ultimately pay all the costs of electricity production, transmission, and distribution, substantial efficiency gains might be realized by, in the first instance, charging generators for reserves. Charging generators for contingency reserves gives generator owners and investors incentive to optimize the tradeoff between higher equipment and maintenance expenses versus fewer forced outages (and therefore lower costs for contingency reserves). Therefore, we recommend a method that assigns these costs on the basis of the number and size of forced outages and the sizes of the large generators online each hour. Figure S-1 shows how different generators would fare if charged for reserves on the basis of these two factors.

For regulation, we focused on the standard deviation of 1-minute fluctuations for individual customers. This method charges individual loads (and, in principle, individual generators) on the basis of their contribution to the overall variability of system load. Figure S-2 shows that individual loads vary enormously in their relative uses of regulation and energy.

The results of these analyses suggest that the electricity industry should

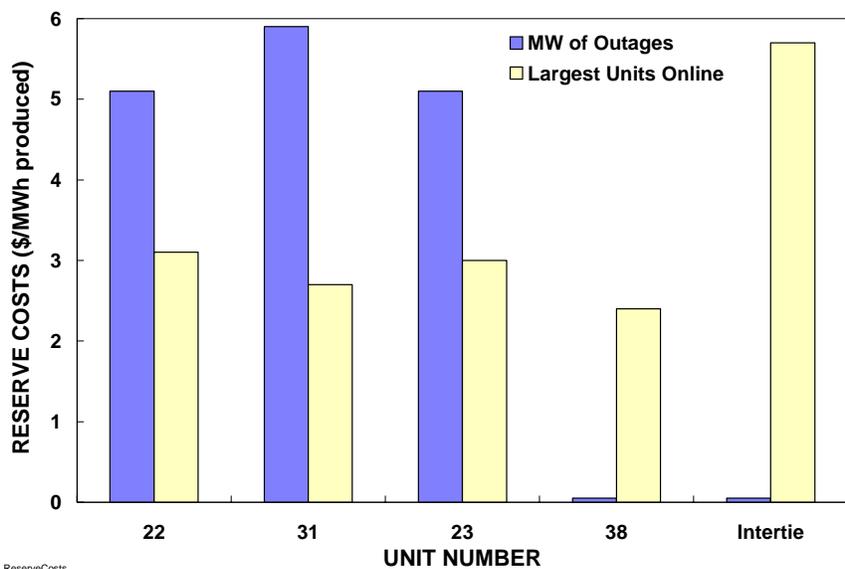
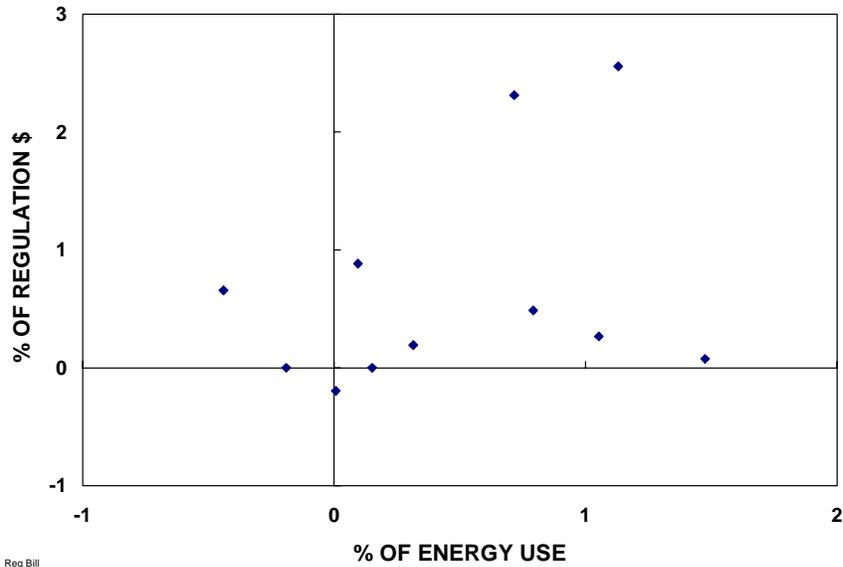


Fig. S-1. Contingency-reserve costs for four generators and the major intertie if charged solely on the basis of the number and size of forced outages or the size of the large units online each hour. The average cost, across all generators, was \$2/MWh of energy production.



Reg Bill

Fig. S-2. Relationship between the share of regulation costs and energy consumption for several customers.

consider the methods developed here to assign costs for contingency reserves and regulation. The current method of charging customers for these services on the basis of their hourly energy consumption bears no relationship to the customer-specific costs that the power system incurs and does nothing to encourage customers to reduce these costs. In other words, the current system is economically inefficient and inequitable.

1. INTRODUCTION

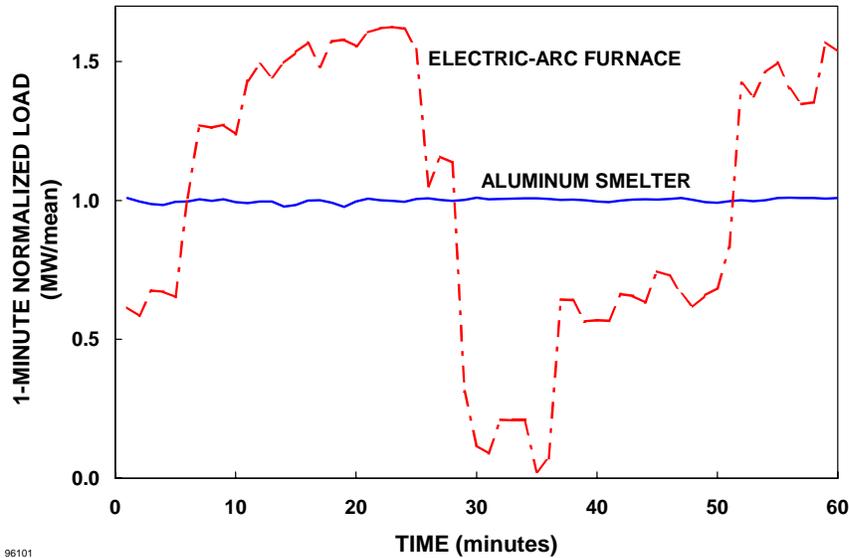
Two of the potential advantages of competitive over regulated markets are the likelihood that prices will accurately reflect costs (and system conditions) and that those who cause costs to be incurred will pay those costs. For these reasons, among others, competitive wholesale electricity markets in North America separate the provision and purchase of the real-power ancillary services from energy. These entities—in New England, New York, the mid-Atlantic region, Texas, and California in the United States and Ontario and Alberta in Canada—created separate markets for energy and each of these services. The ancillary services typically include regulation, 10-minute spinning reserve, 10-minute nonspinning reserve, and sometimes a longer-term (30- or 60-minute) nonspinning reserve.*

Although these entities all have separate markets for various services, transmission customers are generally charged for these services on an energy basis. That is, the costs of ancillary services are collected from customers on the basis of dollars of ancillary service cost per MWh of energy (\$/MWh). For example, the PJM Interconnection (2002) charges customers for regulation and contingency reserves as follows:

- Regulation: “hourly \$/MWh rates calculated as total cost of Regulation in applicable East or West regulation market divided by total real-time load in that market.”
- 10-minute reserves: “daily \$/MWh rates calculated as total cost of Ten-minute Reserves in applicable East or West energy market divided by total real-time load plus exports in that market.”

Although billing customers for these services on the basis of hourly energy use is simple, convenient, and easy to understand, it may not bear any relationship to the costs different customers impose on the system for these services. Consider regulation to illustrate this point. As discussed in Chapter 3, regulation is the ancillary service that tracks the moment-to-moment fluctuations in system load. Figure 1 shows the 1-minute loads of two large industrial consumers, an electric-arc-furnace steel mill and an aluminum smelter. Although they consume the same amount of energy during the hour shown, their regulation requirements are substantially different. The standard deviation of the 1-minute variations of the steel mill is 50 times greater than that of the aluminum smelter. However, under the tariffs currently in use throughout North America, both customers would pay the same amount for regulation, leading to a substantial transfer of money (subsidy) from the aluminum smelter to the steel mill. Why, one might ask, should the aluminum smelter pay for the steel mill’s regulation? Also, the current tariffs provide no incentive for the steel mill to modify its technologies or operations to reduce its regulation requirement even when it might be cheaper to do so than provide the regulation.

*Other ancillary services, including system control, voltage control, and system blackstart, are much less expensive than the real-power ancillary services and are also much less amenable to competitive markets. See Hirst and Kirby (1998) and Interconnected Operations Services Working Group (1997) for additional discussion of ancillary services in general.



96101
Fig. 1. One-minute loads of two large industrial loads, normalized to the same energy use for the hour shown.

service and to charge them on the basis of their unit size and forced-outage frequency.

Recognizing the economic efficiency and equity benefits of assessing ancillary-service charges on the basis of market-participant-specific costs, we investigated the requirements for two key ancillary services, contingency reserves (consisting of 10-minute spinning reserve and 10-minute nonspinning reserve, discussed in Chapter 2) and regulation (Chapter 3). The purpose of these analyses is to identify and analyze equitable ways to collect from the appropriate transmission customers the costs of these real-power services. The results reported here are “fuzzed” slightly to protect the identity of the control areas from which we obtained data, as well as those of the large generators and wholesale customers within its service area. This ~10-GW control area is located within the Western Interconnection.

Similar questions arise for the contingency-reserve services. Again, customers generally pay for these services on a \$/MWh-of-energy basis. But it is the generators, not the loads, that cause these costs to be incurred. The contingency reserves are required to maintain the necessary generation:load balance for loss of supply caused by the sudden trip of a large generating unit or transmission line. Given this reality, it might make sense to charge generators for this

2. CONTINGENCY RESERVES

PURPOSE OF RESERVES

Contingency reserves are the front lines in the defense of bulk-power systems against loss of supply caused by major generation or transmission outages. They are provided by generating units that can increase their output (or, in some cases, by interruptible loads that can decrease their consumption) rapidly. The system operator uses these reserves to restore the generation-load balance after a major disturbance occurs.*

Contingency reserves include two components, spinning reserve and supplemental reserve. Resources providing spinning reserve must be online, synchronized to the grid, loaded to less than the unit's maximum output, responsive to interconnection frequency variations, and capable of responding to an outage *immediately*, with a full response achieved within 10 minutes.# Supplemental reserve (sometimes called nonspinning reserve) is identical to spinning reserve except that it need not begin to respond immediately. That is, resources providing supplemental reserve need not be online, synchronized to the grid, or producing energy at the time they are deployed

The Western Electricity Coordinating Council (WECC) (2002) specifies the minimum amounts of contingency reserve each control area in the West must carry:

Contingency reserve. An amount of spinning and nonspinning reserve, sufficient to meet the [North American Electric Reliability Council, NERC] Disturbance Control Standard [DCS] as defined in 1.E.2(a). This Contingency Reserve shall be at least the greater of:

- (1) The loss of generating capacity due to forced outages of generation or transmission equipment that would result from the most severe single contingency (at least half of which must be spinning reserve); or
- (2) The sum of five percent of the load responsibility served by hydro generation and seven percent of the load responsibility served by thermal generation (at least half of which must be spinning reserve). For generation-based reserves, only the amount of unloaded generating capacity that can be loaded within ten minutes of notification can be considered as reserve.

*Transmission reserves are also required to accommodate the shift in line flow that results from the sudden change in the generation pattern that occurs when contingency reserves replace a failed generator. They are also required to accommodate the sudden shift in line flow that occurs when a transmission line fails. This analysis does not address transmission reserves.

#Spinning reserves respond automatically to changes in system frequency. They also respond to system-operator commands.

Plus (c) **Additional reserve for interruptible imports.** An amount of reserve, which can be made effective within ten minutes following notification, equal to interruptible imports.

Plus (d) **Additional reserve for on-demand obligations.** An amount of reserve, which can be made effective within ten minutes following notification, equal to on-demand obligations to other entities or control areas.

In essence, WECC sets the minimum reserve requirement as the greater of the largest contingency (item 1) or 5% of the load served by hydroelectric resources plus 7% of the load served by all other resources (item 2), with additions related to nonfirm energy imports and firm exports of reserves.* WECC requires that at least 50% of contingency reserves be spinning, primarily to respond to frequency deviations.

These reserves are required to meet the DCS. WECC (2002) describes the DCS as follows: “Each control area or reserve sharing group shall include events that cause its Area Control Error (ACE) to change by at least 35% of the maximum loss of generation that would result from a single contingency. ... Following the start of a disturbance, the ACE must return either to zero or to its pre-disturbance level within the time specified in the Disturbance Control Standard currently in effect in NERC Policy 1 [15 minutes].”

The rest of this chapter discusses several ways the costs of contingency reserves could be assigned to those generators[#] responsible for the need for reserves. The empirical basis for these alternative allocation schemes is data on generator capability and output for the 1-year period from July 2001 through June 2002.[§] None of the methods we discuss considers transmission congestion (i.e., deliverability of contingency reserves).

Although retail customers ultimately pay for all the costs associated with electricity production, transmission, and distribution, substantial efficiency gains might be realized by, in the case of production, charging generators for reserves. Specifically, if a suitable cost-allocation scheme can be developed and applied, it will provide the owners of existing generating units with the incentive to spend the right amount of money on equipment maintenance. That is, owners will optimize the tradeoff between higher maintenance expenses and fewer forced outages (and therefore lower costs for contingency reserves). Similarly, potential investors in new generation will make the appropriate tradeoff between more expensive equipment and fewer forced outages.

*We were unable to find any documentation on the basis for the two-part contingency-reserve requirement, which goes back about 30 years (Davies 2002).

[#]In principle, the costs of contingency reserves should be assigned to both generators and transmission elements. However, this control area experienced no transmission outages large enough to require the use of contingency reserves during the analysis period.

[§]We did not obtain data on the output of customer-owned (behind-the-meter) generation. The sudden loss of such a generator, even though intended exclusively to serve the needs of a particular retail customer, might require the use of contingency reserves if it were large enough and the customer did not design its system so that an equal amount of load was automatically curtailed when its generator failed.

DATA RESOURCES

We received several sets of data for this analysis:

- The capability of the generating units with rated capacity of 40 MW or more.
- The hourly outputs of the large generating units, plus the flows across the key interties, from July 2001 through June 2002.
- The number and magnitude of forced outages that required the system operator to deploy contingency reserves, from July 2001 through June 2002. These data showed, for each outage, the MW loss, the MW deployed, the recovery time, and whether the outage was classified as a DCS event.
- The amount of contingency reserves required each hour and the WECC factor that determined this requirement (i.e., largest single contingency or 5% plus 7% of load).
- Hourly values of the amounts (MW) acquired and the prices (\$/MW-hr)* paid for spinning and supplemental reserves.#

REQUIRED VERSUS PURCHASED RESERVES

During the 1-year analysis period, the amount of contingency reserves required averaged 445 MW; during 90% of the hours, the amount was between 380 and 500 MW. For almost 98% of the hours, the reserve requirement was determined by the 5% plus 7% load rule rather than the size of the largest contingency.

On average, the system operator purchased 459 MW of contingency reserves during this period, 14 MW more than required by the WECC rules. The system operator generally purchases more reserves than required during the early morning (4 to 7 A.M.) and afternoon (3 to 5 P.M.), as shown in Fig. 2. The range in hourly purchases was much greater than the range in requirements; during 90% of the hours, the amount bought was between 400 and 520 MW.

For 18% of the hours, the amount of reserves purchased was less than the amount required, with an average deficit during those hours of 19 MW. These deficits could occur when a reserve provider is unable to meet its commitment (e.g., its unit fails to start or trips offline). Or these deficits could be a consequence of changes in system conditions between the time the reserves were purchased and real time.

*Energy prices are in \$/MWh, while ancillary-service prices are in \$/MW-hr, where MW-hr refers to a MW of ancillary service provided for an hour.

#We imputed prices for contingency reserves (as well as for regulation, discussed in Chapter 3) based on prices in those North American systems with active wholesale electricity markets.

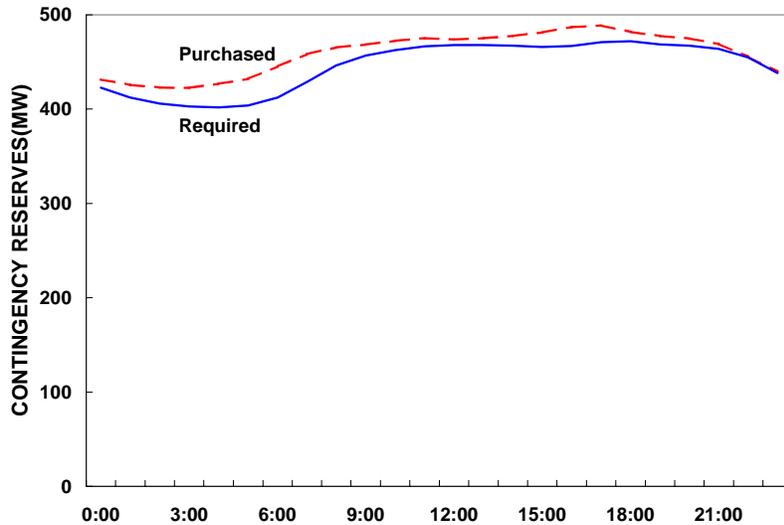


Fig. 2. Daily pattern of contingency-reserve requirements and purchases.

RESERVE PRICES

During the 1-year analysis period, the price of spinning reserve ranged from zero to \$990/MW-hr, with an average of almost \$30/MW-hr. The price of supplemental reserve ranged from zero to \$980/MW-hr, with an average of \$17/MW-hr. Overall, the price of 10-minute reserves averaged \$19.70 during the year. During this year, the system operator spent \$63 million on spinning reserve plus \$49 million on nonspinning reserve for a total of \$113 million.

NUMBER AND SIZE OF OUTAGES

For the year we examined, the system experienced 27 outages that required the use of contingency reserves. Of these 27 outages, 18 were large enough to trigger the DCS criteria. These outages ranged in size from 223 to 409 MW; the average was 337 MW. In every case, the recovery period was less than the NERC maximum of 15 minutes. As shown in Fig. 3, the recovery time was uncorrelated with the size of the outage.

Thirteen generating units (no transmission elements) accounted for these outages. Two units each experienced four outages during the year, one unit had three outages, six units had two outages, and four units had only one outage (Fig. 4).

Figure 4 also shows how the total capacity that suffered forced outages was distributed among these 13 units. Clearly, if the cost of contingency reserves was assigned fully on the basis of the frequency of forced outages, units 22, 31, and 23 would each be responsible for more than 10% of the annual cost.

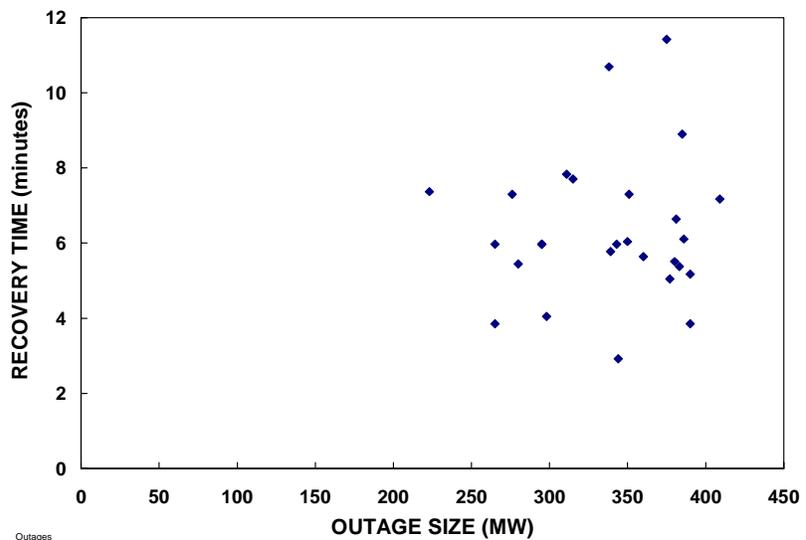


Fig. 3. The time for recovery as a function of outage size for the 27 outages between July 2001 and June 2002.

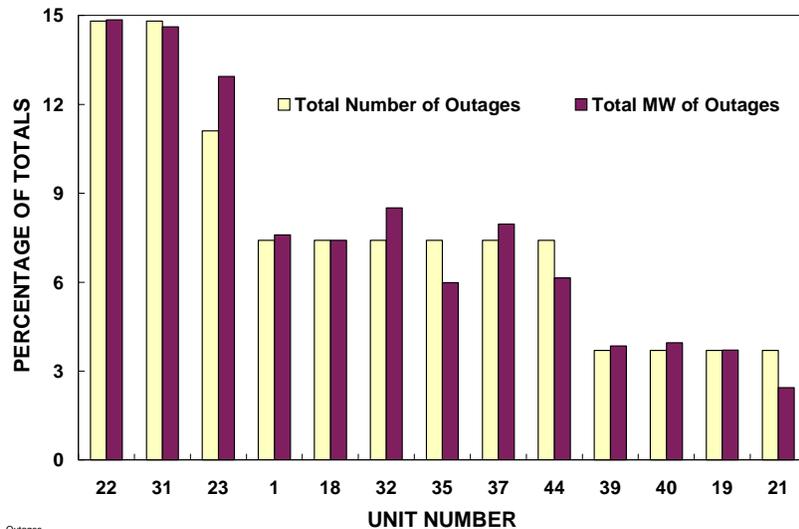


Fig. 4. Allocation of the 27 outages and the 9100 MW of outages among the 13 units with outages.

Assigning the cost of contingency reserves to those generation (and transmission) elements that suddenly trip offline and require the use of reserves seems fair. After all, if the system operator knew that no large elements would suddenly fail, it would not need any reserves at all.

On the other hand, implementing such a system could be difficult for several reasons. First, large outages are, fortunately, rare events; therefore, traditional statistical methods cannot be used to

determine which entities should be charged for reserves. The assignment of costs would likely be erratic, with some generating units paying large charges one year and nothing the next.

Second, the reserve requirements and acquisitions are set prospectively, while the costs are allocated retrospectively. The system operator procures sufficient reserves beforehand to withstand the largest credible contingency and then allocates those costs after the fact to the generators that actually trip offline. This cost-allocation system fails completely if no generators trip during an assessment period: there are costs to allocate but no entity to pay these costs.

It is similarly unfair if a large but reliable generator forces the system operator to continuously procure a large amount of reserves that regularly get allocated to a small but unreliable generator. The small generator would argue that it should be required to pay for only a small amount of reserves (enough to cover its outages). Also, the reserve charges for a particular year could not be determined until the assessment period was over; therefore, generators would not know for a few months how much they owed for reserves.

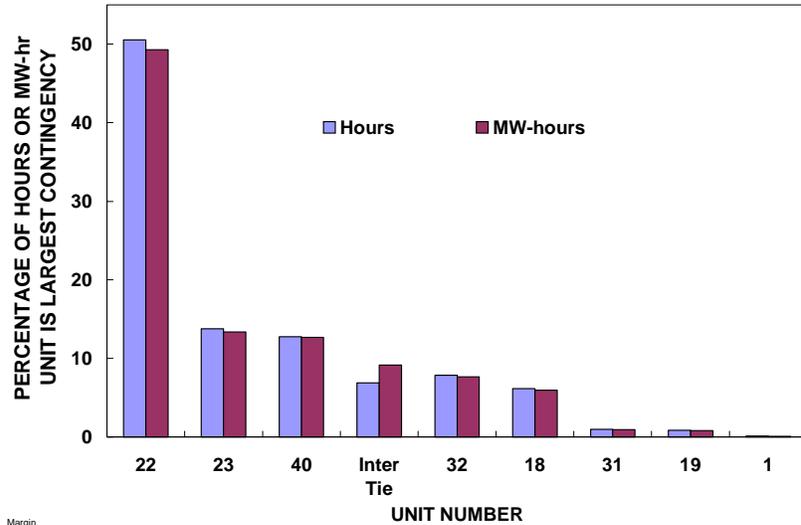
MAXIMUM OUTPUT OF LARGEST GENERATOR ONLINE EACH HOUR

One of the two WECC methods outlined above sets the minimum amount of reserves required equal to the largest contingency online that hour. (This approach is also used in several other NERC regions.) Thus another way to charge generators for reserves is to record the size (in MW) of the largest unit online each hour. Each unit would then be charged for contingency reserves on the basis of the number of hours a year it is the largest unit and how much it was producing each hour when it was the largest single contingency.

This method, as well as the ones discussed after it, define the largest size on the basis of a unit's hourly output, not its nameplate rating. For example, a 500-MW unit producing 425 MW during a

particular hour would be considered a 425-MW unit for purposes of calculating the amount of contingency reserves required. If actual and scheduled hourly outputs often differ substantially, this approach could be difficult to implement.

Five units (including the major interties with neighboring control areas) account for about 90% of either the hours during the year or the MW-hr of the largest contingency (Fig. 5). Comparing Figs. 4 and 5 shows modest overlap between the two methods. Units 22, 23, and 32 are important contributors to reserve requirements in both methods.



Margin **Fig. 5. The percentage of the hours of MW-hr reserve required during which eight generators plus a major intertie were the largest contingencies.**

The primary advantage of this method is its simplicity in both implementation and understanding. It is easy to measure, each hour, the output of the largest generator then online (or the maximum flow through a transmission element).

The primary disadvantage of this method is that it ignores all the large units that are smaller than the largest single contingency. That is, the method incorrectly assigns all the reserve costs to the largest online unit and only to that unit. The following section discusses a more reasonable approach.

HOURLY OUTPUTS OF ALL LARGE GENERATORS

An alternative to the previous approach considered recognizes that the largest unit online requires *incremental* reserves equal to the difference between its hourly output and that of the second largest unit (Hirst and Kirby 1997). Similarly, the first and second units contribute equally to the reserve requirements based on the difference between the output of the second unit and that of the third largest unit, and so on.

Figure 6 illustrates graphically how this method might work. In this example, all the generators, regardless of how small their output that hour, are assigned a share of the total reserve requirement.* All seven units share equally in the requirement for the first 200 MW of reserves (29 MW of reserves each). Four units (A, B, C, and D) are responsible for the next 100 MW (300 MW–200 MW). Their contributions are now 54 MW each (29 MW for their 1/7 share of 200 MW plus 25 MW for their 1/4 share of 100 MW). Two generators (A and B) are responsible for the next 100 MW, bringing

*In practice, the smaller units, those whose outages would not require the use of contingency reserves, probably should not be included in these calculations. The outage data we received suggest a lower limit of about 150 MW.

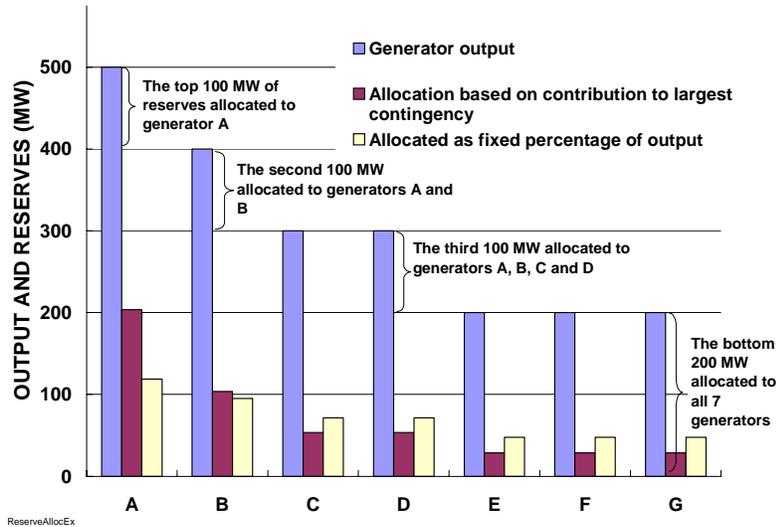


Fig. 6. Assignment of responsibility for reserves on the basis of unit output or share of total load served.

appropriate responsibility for reserves on the larger units that set the contingency-reserve requirements.

Figure 6 also shows how reserves would be assigned to each unit on the basis of the total load, as with the WECC 5% plus 7% rule. In this case, all units, regardless of size, are assigned reserve responsibility. Exhibit 1 discusses the role that small units might play in paying for contingency reserves.

Exhibit 1. Minimum Generator Size

In practice, contingency reserves are not used every time a generator fails. The sudden failure of a small unit can often be dealt with by using the resources available to the real-time energy market and other reserves assigned to load following and regulation. WECC implicitly recognizes this when it sets the minimum requirements for reporting DCS events at 35% of the largest single contingency. The outage data suggest a lower limit of about 150 MW.

It is not appropriate to simply reduce the contingency reserve requirement by, say, 150 MW, however. When a 500-MW contingency occurs, the control area needs to respond with 500 MW of reserves to reestablish the pre-contingency ACE and satisfy the DCS. A 450-MW response is not adequate. So, does a 151-MW generator require 151 MW of reserves and a 149-MW generator require 0 MW of reserves? Probably not. But a 300-MW generator probably does require 300 MW of reserves, and a 100-MW generator probably does not require any reserves.

At least in the West, small units cannot be ignored completely. WECC's 5% plus 7% rule includes small generators. This system typically obtains about 1500 MW of energy from small generators. If these units are ignored, their share of reserves (about 105 MW) must be supplied by the larger generators, raising the larger units' requirements above the 5% or 7% required by WECC. In this case, the 1500 MW of small generation would raise the reserve requirements for the remaining 4900 MW of large nonhydro generators from 7% to 9%.

their total contributions to 104 MW each (29 + 25 + 50). The largest generator (A) is solely responsible for the last 100 MW of reserves.

In this example, the largest unit (A, with output of 500 MW) is assigned 204 MW (40%) of the total reserve requirement, rather than the full 100% the previous method would assign. The second largest unit (B, with output of 400 MW) is assigned 104 MW of reserves, rather than the zero the previous method would assign. This approach places what we believe is the

The remainder of this section applies this method to the larger generators (roughly speaking, greater than 150 MW) operating each hour during the analysis year. The hourly data we received included information on the output levels of 24 generating units plus the flows over the key interties with neighboring control areas. Altogether, these 26 “units” accounted for almost 80% of the generation serving load. Eight of these units each accounted for 5% or more of total generation, while seven units each accounted for less than 1%.

If reserve requirements are calculated on the basis of the largest single contingency each hour, 6 of the 24 units account for 50% of the reserve costs (Fig. 7). Units 18, 19, 22, 23, and 32 are especially important. Some, but not all, of these units appeared near the tops of the lists developed earlier in this chapter.

In part, the units that require reserves are the ones that produce the most energy. Indeed, the share of annual energy production explains 92% of the variation in contribution to the cost of contingency reserves.

Only the intertie stands out as an exception to this rule. Its reserve requirements are much greater than would be expected from its energy delivery alone. The maximum hourly flow across this intertie was more than 800 MW, more than double the next largest contingency. Also, the intertie has a relatively low “capacity factor” of about 15%, compared with 80% for these units in aggregate.

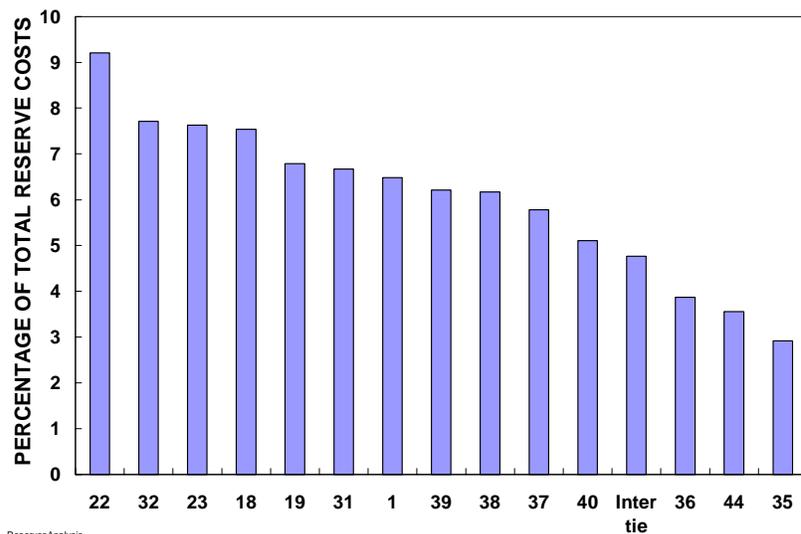


Fig. 7. The share of annual costs for contingency reserves based on the largest-single-contingency criterion.

The key strength of this method is that it accurately captures the contribution of each large generator to the amount of reserves required each hour. This statement, however, is correct only if the amount of required reserves is based on the largest single contingency. For this system, this was true for only 2% of the hours during our analysis year.

SHARE OF LOAD

The second WECC method for determining the total amount of contingency reserves required is based on the load served by each generator. Hydroelectric units are assigned reserves equal to 5% of their load, while nonhydro units are assigned 7%. Presumably, hydroelectric resources require fewer reserves because they are less likely to suffer a forced outage and they have faster frequency response than thermal units. In this system, the hourly reserve requirements are determined by this 5% plus 7% rule 98% of the time.

Unlike the other methods discussed here, this one assigns some reserve responsibility to small generators. For example, the small generators excluded from the analyses discussed previously (and therefore assigned zero reserve costs) are here assigned 22% of the total reserve costs.

Figure 8 shows the contribution to annual contingency costs for 15 units, based on this method of allocating reserves. Ten units account for half the total cost. The most important units (18, 19, 22, 23, 31, 32, and 38) are generally the same units that were most important on the basis of the largest-contingency rule. The similarity of results is not surprising given the high correlation between unit size and energy production.

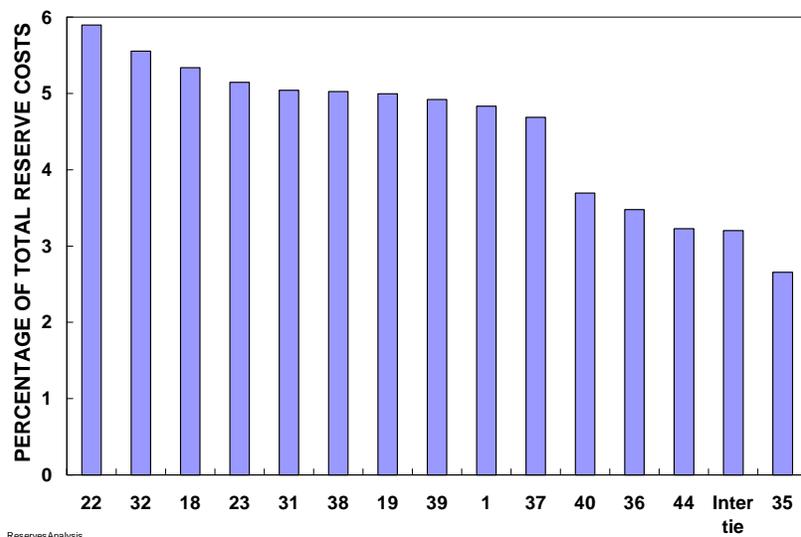


Fig. 8. The share of annual costs for contingency reserves based on the 5% plus 7% rule.

The key strength of this method is its consistency with the WECC rule on determining the amount of reserves to carry each hour. The key disadvantage of this approach, in our view, is the lack of relationship between system load and the frequency and magnitude of outages. Yes, all else equal, a larger system will have more generators online than a smaller system. The more generators that are operating, the more likely it is that one will suddenly fail. On the other hand, it is very unlikely that the number and size of large outages increase linearly with an increase in system load. Thus this approach may require the control-area operator to acquire more contingency reserves than are actually needed to maintain reliability.

EFFECTS OF METHODS ON INDIVIDUAL GENERATORS

The previous sections presented alternative methods for allocating the costs of contingency reserves to individual generators based on the number and magnitude of forced outages they experienced, their size, their contribution to the largest single contingency, and their hourly output. This section demonstrates the implications of these methods. Specifically, we show how costs are shifted among generators as a function of the allocation method used. We consider three cost-allocation schemes:

- Number of megawatts lost during a forced outage. This method considers both the number of forced outages a year and the amount of online capacity lost during each outage.
- Each unit's contribution to the largest single contingency, based on the hourly output of each unit. All these units exceed 140 MW in size.
- Each unit's hourly energy output, 5% for hydro units and 7% for all other generators.

Table 1 shows the implications of these three methods. The left side of the table shows the share of annual contingency-reserve costs for which each generator is responsible under each allocation method. The right side of the table shows the reserve costs to each generator, normalized by energy production (i.e., the costs are expressed in dollars for contingency reserves per MWh of energy production).

Table 1. Allocation of the costs of contingency reserves to generators

Unit #	Share of contingency-reserve costs (%)			Contingency-reserve costs (\$/MWh)		
	MW of outages ^a	Largest contingency	5%/7% of load	MW of outages	Largest contingency	5%/7% of load
22 ^b	14.9	9.2	5.8	5.1	3.1	2.0
31 ^b	14.6	6.7	5.0	5.9	2.7	2.0
23 ^b	12.9	7.6	5.1	5.1	3.0	2.0
1	7.6	6.5	4.9	3.2	2.7	2.0
18 ^b	7.4	7.5	5.3	2.7	2.8	1.9
32 ^b	8.5	7.7	5.5	3.1	2.8	2.0
35	6.0	2.9	2.7	4.0	2.0	1.8
37 ^b	8.0	5.8	4.7	3.4	2.5	2.0
44	6.2	3.6	3.2	3.7	2.2	2.0
39 ^b	3.8	6.2	4.9	1.5	2.4	1.9
40 ^b	4.0	5.1	3.6	1.9	2.5	1.8
19 ^b	3.7	6.8	5.0	1.4	2.6	1.9
38 ^b	0	6.2	5.0	0	2.4	2.0
36	0	3.9	3.5	0	2.1	1.9
Intertie	0	4.8	2.2	0	5.7	2.7
Small units	0	0	22.3	0	0	2.0

^aUnit 21, for which we did not receive hourly data, tripped offline once during this year and accounted for 2.4% of the MW of outages during the year.

^bThe rated capacities of these units are larger than 400 MW; the other units (1, 35, 44, and 36) are between 300 and 400 MW in capacity.

The differences in cost allocation are greatest for units 22, 31, 23, and the intertie, more than \$3/MWh. For example, because the intertie experienced no outages during the analysis year, its cost assignment is zero under the first method. However, if costs are assigned on the basis of contribution to the largest single contingency, the intertie must pay the equivalent of \$5.7/MWh for contingency reserves. The average cost of contingency reserves across all generators was \$2.0/MWh.

The reverse was true for units 22, 31, and 23. These units experienced three or four outages during the analysis year, leading to very high charges for contingency reserves under the first method. These three units would also pay a lot for contingency reserves using the other two methods

because these units are roughly 400 MW in size (therefore being among the largest) and are low-cost producers with high capacity factors (80% or more).

The small units (those with capacity ratings of ~150 MW or less) pay nothing for contingency reserves using the first two methods. However, if contingency-reserve costs are allocated on the basis of the share of load served by each unit, these small units pay for almost one-fourth of the total cost, equivalent to \$2.0/MWh.

RECOMMENDATIONS

We suggest a two-part method for assigning the costs of contingency reserves to generators. The first part would be a function of the number and size of forced outages that required the use of contingency reserves. The second part would be a function of the largest single contingency each hour. Mathematically, the cost assignment is

$$\begin{aligned} \text{Contingency-Reserve Cost}_i (\$/\text{month}) &= a \times \sum_{t=1}^{t=12} \sum_{o=1}^{o=O} \text{MW}_{\text{outages-}i} \\ &+ (1-a) \times \sum_{h=1}^{h=N} \text{Largest Contingency}_{i, h} \end{aligned}$$

where i refers to a particular generating unit, t refers to month, o refers to outages, O is the number of outages each month, h refers to hour, N is the number of hours during the particular month, and a is a constant between 0 and 1.

The first part of this formula charges generators for reserves on the basis of any forced outages they experienced during the past 12 months. We recommend the use of a rolling 12-month average to smooth the effects of what are largely random events. This component would be calculated each month, with $100 \times a\%$ of the total cost of reserves that month assigned to this factor. Any generator that had experienced a forced outage during the preceding 12 months would pay a portion of this dollar amount based on the unit's MW share of the total outages during this period.

The second part of this formula charges generators on the basis of their contributions to the maximum contingency each hour during the month in question.

An alternative to the first part of this formula would involve an upfront specification of the dollar amount to charge for each outage. The money collected each month from those large generators that suffered a forced outage would then be subtracted from the amount owed from the second part of the formula.

Finally, WECC should analyze the technical basis (engineering and economics) of its existing reserve requirements. In particular, we are unsure why the amount of load to be served should determine on a one-to-one basis the amount of contingency reserves that must be maintained.

3. REGULATION

PURPOSE OF REGULATION

Because electricity is a real-time product, control-area operators must adjust generation to meet load on a minute-to-minute basis. Regulation and load following (which, in competitive spot markets is provided by the intrahour workings of the real-time energy market) are the two ancillary services required to perform this function.

Loads can be decomposed into three elements (Fig. 9). The first element is the average load (base) during the scheduling period, 85 MW for the hour shown in this case. The second element is the trend (ramp) during the hour and from hour to hour (the morning pickup in this case); here that element increases from -5 MW at 7 A.M. to $+9$ MW at 8 A.M. The third element is the rapid fluctuations in load around the underlying trend; here the fluctuations range over ± 2 MW. Combined, the three elements yield a load that ranges from 78 to 96 MW during this hour.

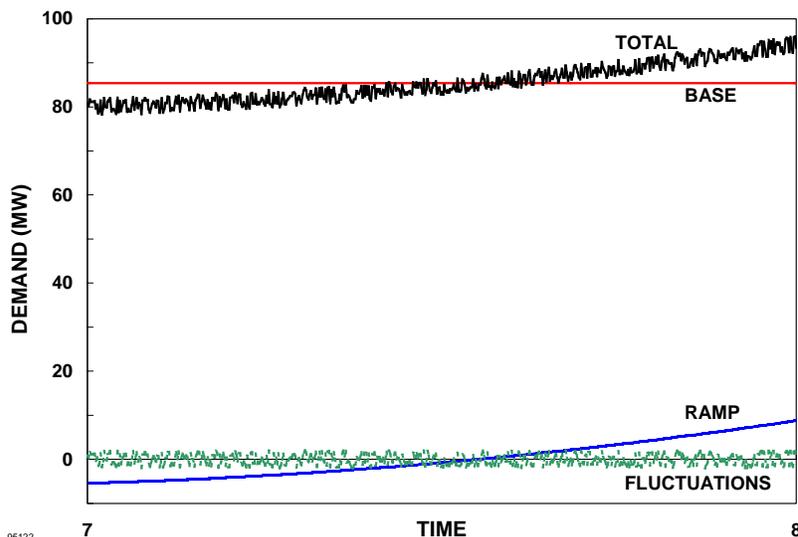


Fig. 9. Components of a hypothetical load on a weekday morning.

The system responses to the second and third components are called load following and regulation. These two services ensure that, under normal operating conditions, a control area is able to balance generation to load. The two services are briefly defined as follows [see also Hirst and Kirby (1998) and Interconnected Operations Services Working Group (1997)]:

- Regulation is the use of online generating units that are equipped with automatic generation control (AGC) and that can change output quickly (MW/minute) to track the moment-to-moment fluctuations in customer loads and to correct for the unintended fluctuations in generation. In so doing, regulation helps to maintain interconnection frequency, manage differences between actual and scheduled power flows between control areas, and match generation to load within the control area. This service can be provided by any appropriately equipped generator that is connected to the grid and electrically close enough to the local control area that physical and economic transmission limitations do not prevent the importation of this power.
- Load following is the use of online generation equipment to track the intra- and inter-hour changes in customer loads.

The key differences between regulation and load following include these:

1. load following occurs over longer time intervals, and regulation occurs *within* the load-following intervals established by the real-time energy market;
2. load following is provided by generators on economic dispatch (i.e., the least-bid-cost mix of generators to meet current demand), while units on AGC are not necessarily the least expensive generators;
3. the load-following patterns of individual customers are highly correlated with each other (e.g., during the morning rampup), while the regulation patterns are random and uncorrelated; and
4. the load-following patterns are generally predictable and similar from day to day.*

In the **PJM Interconnection**, New York, New England, and Ontario, regulation is a 5-minute service, defined as five times the ramp rate in MW/minute. In the Electric Reliability Council of Texas (ERCOT), regulation is a 15-minute service, and in Alberta and California, it is a 10-minute service. Regulation is a zero-energy service in PJM, New York, New England, and Ontario, meaning that, on average, the units providing regulation produce no net energy beyond that which was scheduled. ERCOT and California purchase separate up- and down-regulation services. And Alberta purchases a regulation range, equivalent to up-regulation.

In real time, the system operator dispatches generation (and, perhaps, some load) resources participating in its intrahour energy market to maintain the necessary balance between generation and load. Once every several minutes, the system operator runs an economic-dispatch model to move generators up or down to follow changes in load and unscheduled generator outputs at the lowest possible operating cost. Generators that participate in the system operator's balancing market provide the load-following ancillary service.

To track changes in the minute-to-minute balance between generation and load, the system operator uses its AGC system to dispatch those generators providing the regulation ancillary service. These generators respond to short-term generation:load imbalances that are not addressed by the economic-dispatch process.

The primary purpose of these intrahour resource movements is to maintain ACE within certain limits. ACE is the instantaneous difference between actual and scheduled interchange, taking into account the effects of Interconnection frequency. In plain language, it measures how well the system operator maintains the necessary generation-load balance. NERC (2002) established the Control Performance Standard (CPS) to determine the amount of imbalance that is permissible for reliability purposes. CPS1 measures the relationship between the control area's ACE and Interconnection frequency on a 1-minute average basis. CPS1 values can be either "good" or "bad." When frequency is above its reference value, undergeneration benefits the Interconnection by lowering frequency and leads to a good CPS1 value. Overgeneration at such times, however, would further increase frequency and lead to a bad CPS1 value. CPS1, although recorded every minute, is evaluated and

*In practice, the distinctions between regulation and load following are less clear. In particular, when the units participating in intrahour balancing markets (those providing load following) are limited by their ramp rates (rather than their upper and lower operating limits), the system operator uses the units on regulation to respond to rapid changes in load, even though that response should, in principle, be provided by load-following resources.

reported on an annual basis. NERC sets minimum CPS1 requirements that each control area must exceed each year.

CPS2, a monthly performance standard, sets control-area-specific limits on the maximum average ACE for every 10-minute period. Control areas are permitted to exceed the CPS2 limit no more than 10% of the time. This 90% requirement means that a control area can have no more than 14.4 CPS2 violations per day, on average, during any month.

DATA RESOURCES

We obtained 1-minute data for four 1-week periods between July 2001 and June 2002. For each period, the data included total system load plus loads for 15 individual municipal and industrial customers, flows over the key interties, and outputs from four non-AGC generators. Five of the 15 customers have onsite (behind the meter) generation. We also received data on the amounts of regulation purchased and the price paid, hour by hour, from July 2001 through June 2002.

REGULATION PURCHASES AND PRICES

During the 1-year analysis period, the system operator purchased an average of 163 MW of regulation at an average price of almost \$35/MW-hr. The amounts purchased and the prices paid were between 126 and 217 MW and \$3.7 and \$79.3/MW-hr for 90% of the hours.

During this year, the system operator spent \$50 million on regulation. The amount of regulation purchased varied little from month to month (Fig. 10, top graph). On a daily basis, the amount of regulation bought was highest during the early morning (5 to 7 A.M.) and evening (5 to 9 P.M.) (Fig. 10, bottom graph).

Prices were highest in July and August 2001, roughly triple the prices in January and February 2002. On a daily basis, prices were highest between 10 A.M. and 5 P.M., averaging \$47 during these eight hours.

SYSTEM-LEVEL METRIC

Following the method developed by Kirby and Hirst (2000), we used a 30-minute rolling average to define the boundary between the regulation and load-following services. Specifically, we calculated the rolling average for each 1-minute interval as the mean value of the 14 earlier values of the variable, the current value, and the subsequent 15 values:

$$\text{Load following}_t = \text{Load}_{\text{estimated-}t} = \text{Mean} (L_{t-14} + L_{t-13} + \dots + L_t + L_{t+1} + \dots + L_{t+15}),$$

$$\text{Regulation}_t = \text{Load}_t - \text{Load}_{\text{estimated-}t}.$$

Fig. 11 shows the results, for one hour, of the method we used to disaggregate total load into its regulation and load-following components. The average load during this hour was 6948 MW, with minimum and maximum 1-minute values of 6916 and 6987 MW. The regulation value, by our

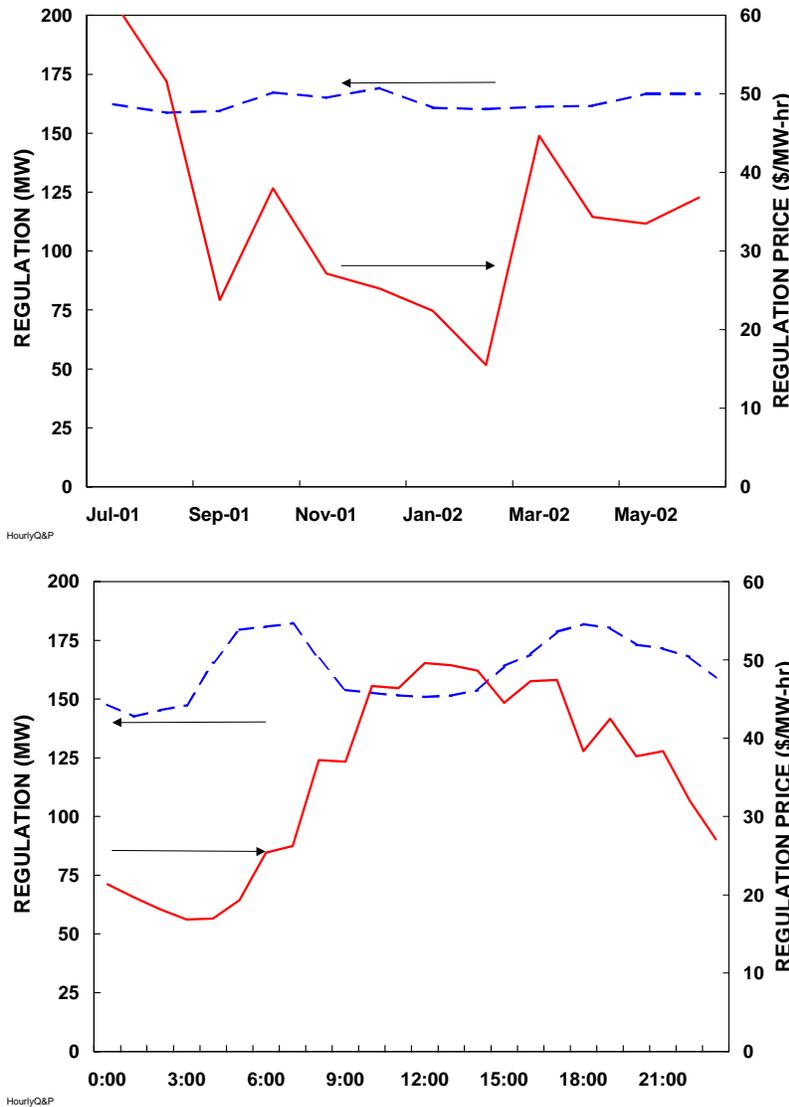


Fig. 10. Average monthly (top) and hourly (bottom) regulation prices and quantities purchased, July 2001 through June 2002.

follow the dispatch signals exactly, and flows over interties that do not exactly match their schedules.

Initially, we sought to define an aggregate regulation requirement based on all the components listed previously. We could not easily implement such an approach because some of the components were negatively correlated with other components. In particular, the flows across a major intertie between control areas were quite volatile and frequently shifted from imports to exports. As a consequence, this intertie (sometimes acting as a load and sometimes acting as a generator) contributed a very large regulation component to the total requirement. We think these large values were caused by two factors. First, the WECC interchange scheduling convention calls for a 20-minute ramp beginning at 10 minutes before the top of the hour and ending at 10 minutes after the hour starts. This 20-minute ramp accommodates load changes over a 60-minute hour, leading to large regulation requirements (to counter the intertie flow) during the interchange-ramping period.

definition, averages zero for the hour, with minimum and maximum values of -34 and $+39$ MW; the standard deviation of these 60 values is 21 MW.

For purposes of this study, we define regulation as the standard deviation of the 60 one-minute values of the regulation component defined previously as applied to the system load. Thus our measure of regulation is hourly and based on a zero-energy service.

Deciding on the total regulation requirement against which to assess the charges for individual customers is not obvious. In both principle and practice, the system operator dispatches generation on AGC to compensate for any moment-to-moment variations in the generation:load balance, reflected in the system ACE. ACE is agnostic about the source of the imbalance, which includes retail loads, non-AGC generators that do not follow their schedules exactly, generators on AGC that do not

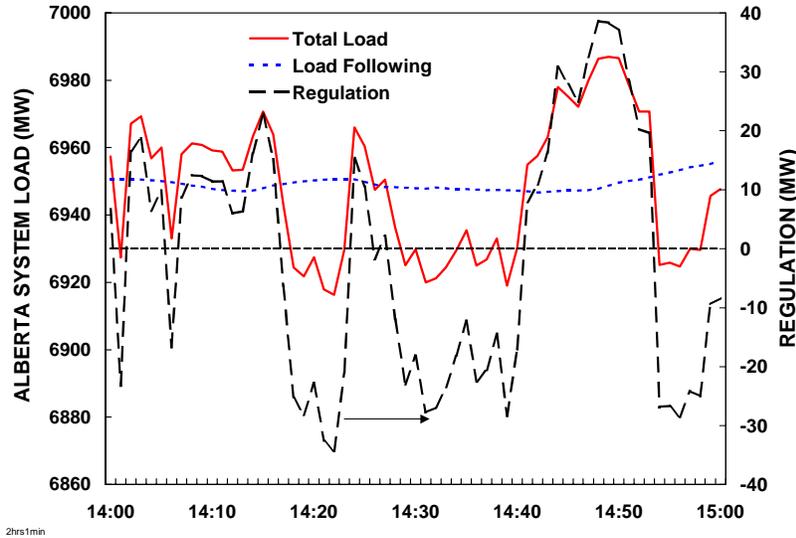


Fig. 11. Minute-by-minute values of system load and the load-following and regulation components for one hour.

As a consequence, the present analysis focuses on allocation of the regulation requirements for the system load among components of that load and does not include either the inerties or the generators.

ALLOCATION OF COSTS TO INDIVIDUAL CUSTOMERS

Because regulation involves the short, minute-to-minute fluctuations in load, the regulation component of each customer's load is largely uncorrelated with those of other customers.* If each customer's load fluctuations (e.g., its standard deviation, σ_i) are completely independent of the remainder of the system, the total regulation requirement (σ_T) would equal

$$\sigma_T = \sqrt{\sum \sigma_i^2},$$

where i refers to an individual customer and T is the system total.

In this idealized case, the share of regulation assigned to each customer would equal

$$\text{Share}_i = (\sigma_i / \sigma_T)^2,$$

and there would be no need to analyze interactions among customer loads in calculating the total regulation burden.

If, on the other hand, the loads are completely correlated with each other [i.e., the correlation coefficient (r) between each pair of loads equals 1], the total regulation requirement is the simple sum of the individual requirements:

*Most of the material in this section is from Kirby and Hirst (2000).

Second, generation changes in response to frequency deviations throughout the entire Western Interconnection are reflected in this inertia.

We had also intended to include unscheduled and uninstructed generator movements in our analysis of regulation requirements and cost allocation. However, the assignment of generators to the regulation service changes from hour to hour, and the data we received on several generators did not show which ones were on AGC and when. As a

$$\sigma_T = \sum \sigma_i .$$

In this idealized case, the share of regulation assigned to each customer would equal

$$\text{Share}_i = \sigma_i / \sigma_T .$$

The question is how to allocate fairly the total regulation requirement between any two loads (and, by extension, among several loads). The allocation method should yield results that are independent of any subaggregations. In other words, the assignment of regulation to load L should not depend on whether L is billed for regulation independently of other loads or as part of a group of loads. In addition, the allocation method should reward (pay) loads that reduce the total regulation burden.

Figure 12 illustrates schematically the method we developed for such allocations. This method works for the two extreme situations discussed, when loads are either completely uncorrelated or perfectly correlated. More important, this method yields reasonable results for the intermediate cases when loads are partially correlated with each other. Consider two loads A and B and the *Total*, with the regulation requirement of each based on the standard deviation of the short-term fluctuations. We use a geometric approach to calculating the contribution of A to the *Total*, based on the projection of A onto the *Total* (shown as X in Fig. 12):

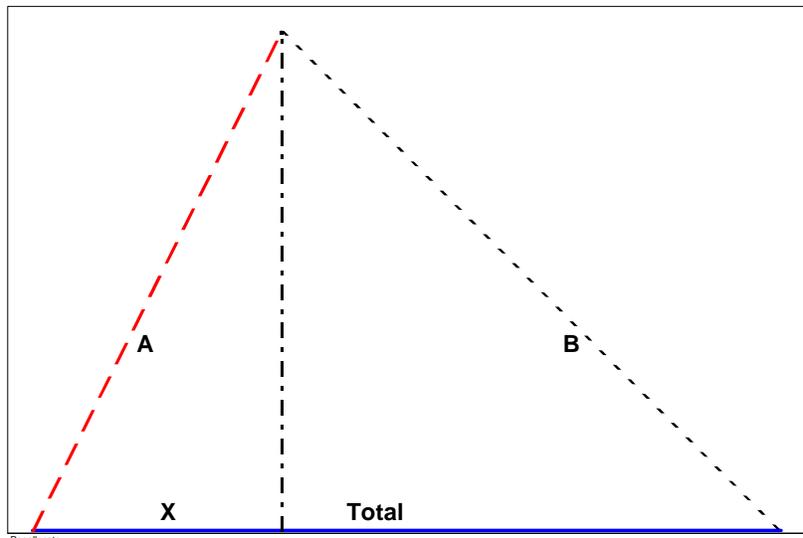


Fig. 12. Geometric allocation of individual loads A and B to regulation $Total$. X is A 's share of the total. B 's share, by subtraction, is $Total - X$.

$$X = (Total^2 + A^2 - B^2) / (2 \times Total) .$$

The contribution of B to the *Total* is then equal to $Total - X$ or

$$Total - X = (Total^2 + B^2 - A^2) / (2 \times Total) .$$

This method can be extended to three or more loads through disaggregation of the total into various components. The only computational requirement is to calculate the standard deviation of each component and of each subtotal (total minus load i). Consider, as an example, a utility that wants to assign regulation charges separately for the residential class, the commercial class, five industrial customers, and the remainder of the industrial class, eight groups in all. The utility would calculate, for each hour, the standard deviation of eight subtotals (total—residential class,

total—commercial class, and so on) as well as the standard deviations of each group of customers and the total, 17 values in all.

This method can accommodate a mix of individually metered loads and subaggregations, such as several large industrial loads that are metered separately and aggregations of thousands of residential and commercial customers. The subaggregations of the nonmetered residential and commercial loads will have the correct share of regulation assigned to them; any cost shifting will occur within the subaggregations and not between the subaggregations and the individually metered loads. This desirable property greatly reduces the need to meter any but the most nonconforming loads.

RESULTS

The 15 loads for which we have customer-specific 1-minute load data account for slightly less than 25% of system load and slightly more than 25% of the area’s regulation requirement. However, as shown in Fig. 13, the individual loads vary tremendously in their “consumption” of energy and regulation. Four of the customers (G, J, Lg, and M) are high regulation consumers but low energy users. Four customers (A, Kg, N, and O) are high energy users but low regulation users. Customers with a “g” have onsite generation.

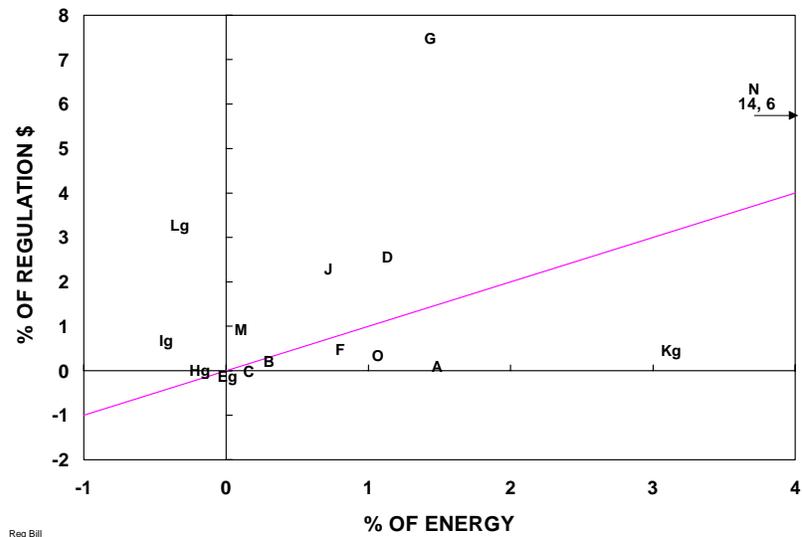


Fig. 13. Relationship between the share of regulation costs and energy consumption for 15 customers. The shares are equal along the solid line.

Table 2 provides additional details on each of these 15 loads for each of the six periods analyzed. The table shows values of the ratio of each customer’s share of the total regulation cost to its share of the total energy consumption. A ratio of 1.0 means that the customer has a regulation-to-load ratio equal to the system average. To illustrate, the roughly 3/4 of the load for which we do not have customer-specific data has a ratio that ranges between 0.95 and 1.04 across the four 1-week periods. Interestingly, these ratios diverge greatly from 1.0, although in the opposite direction, for the two nonweek periods in March and June 2002.

Turning to the individual loads, D, Eg, G, J, Lg, and M are intensive regulation users relative to their energy consumption. On the other hand, A, C, Hg, Kg, N, and O use very little regulation relative to their energy consumption.

How would application of the method developed here affect the regulation costs for individual customers? The total cost for regulation was \$50 million from July 2001 through June 2002. If, as

Table 2. Ratio of share of regulation cost to share of energy consumption for 15 electricity customers^a

	July 2001	October 2001	January 2002	April 2002
Measured load	1.17	1.03	1.06	0.87
Remainder	0.95	0.99	0.98	1.04
A	<i>0.28</i>	<i>-0.01</i>	<i>-0.01</i>	<i>-0.06</i>
B	0.65	0.64	0.85	0.27
C	<i>0.05</i>	<i>0.01</i>	<i>-0.05</i>	<i>0.00</i>
D	2.62	0.51	3.45	2.59
Eg	-41.51	-12.80	-43.15	224.55
F	0.73	<i>-0.05</i>	0.74	1.06
G	4.86	5.14	4.06	7.33
Hg	<i>0</i>	<i>0.00</i>	<i>0.02</i>	<i>-0.04</i>
Ig	-1.54	-1.49	-0.42	-5.4
J	4.41	4.35	2.20	2.23
Kg	<i>0.25</i>	<i>-0.20</i>	<i>0.17</i>	<i>0.21</i>
Lg	-5.28	-15.07	-27.91	5.67
M	7.48	11.80	12.94	5.09
N	<i>0.48</i>	<i>0.32</i>	<i>0.46</i>	<i>0.37</i>
O	<i>0.14</i>	<i>0.25</i>	<i>0.27</i>	<i>0.36</i>

^aThe numbers shown in bold are for loads with a ratio greater than 1.5, and the numbers shown in italics are for loads with a ratio less than 0.5.

occurs in most of North America, customers pay for regulation on the basis of their energy use (i.e., the charge for regulation is in dollars for regulation per MWh of energy), the 15 customers for which we analyzed 1-minute data would have paid a total of \$11.9 million for the regulation service for this year. Using the method developed here, these 15 customers would have paid \$12.2 million, only 2.5% more.

However, the effects on individual customers is quite dramatic (Fig. 13). Customer N, in particular, would pay \$7.2 million under the traditional approach but pay only \$2.9 million using our cost-allocation method. In other words, charging for regulation on the basis of hourly energy use instead of the correlation between customer-load volatility and system-load volatility results in an overcharge for this customer of \$4.3 million a year.

At the other end of the spectrum, customer G would pay \$0.7 million a year for regulation under the traditional approach but pay \$3.8 million under our approach, a shift of \$3.1 million (Fig. 13). If the method developed here accurately reflects the regulation costs imposed by customers on the system, these results suggest that the current tariff requires Customer N to subsidize Customer G, a substantial cost shift.

Indeed, for the 15 customers analyzed here, the total cost shift is almost \$15 million a year, more than the total paid by these customers for regulation in aggregate.

Figure 14 shows a few instances in which the calculated charge for regulation would be negative. Three of the five customers with self-generation (Hg, Ig, and Lg) produce more energy than they consume. Assigning regulation costs on the basis of energy *consumption* yields negative charges for these customers, which makes no sense. In addition, the regulation component from customer Eg is negatively correlated with the overall regulation burden, yielding a negative charge with our method. In this case, customer Eg should receive a credit for regulation because its load volatility reduces the total regulation requirement (i.e., this load is acting like a generator on AGC for regulation purposes).*

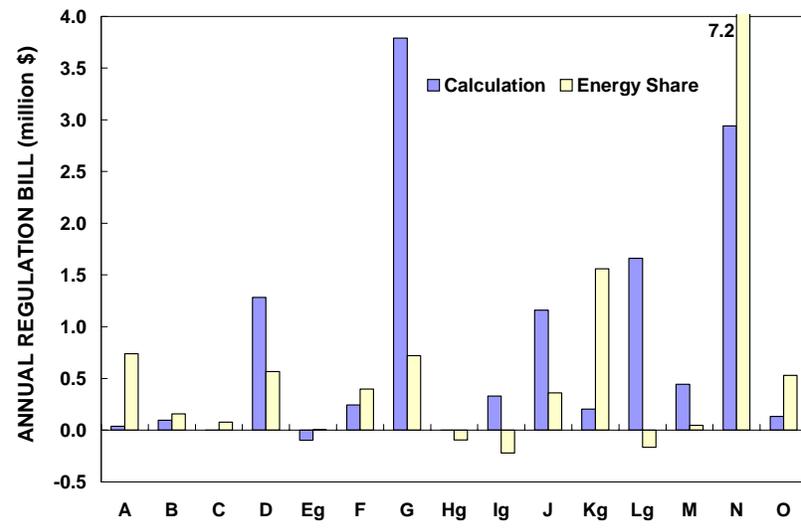


Fig. 14. The annual bills for regulation for 15 transmission customers based on the method developed here and based on energy use.

RECOMMENDATIONS

We suggest using the method developed here to charge customers for the regulation ancillary service. Implementing this method requires data at the 1-minute level on system load plus the loads for which customer-specific requirements are to be calculated. We suggest that, initially at least, this method be applied only for the larger customers, in particular those for whom 1-minute load data are already available. The calculations should be done on an hourly basis because both the price and amount of regulation purchased vary from hour to hour.

*In practice, it is unlikely that the volatility of individual loads will consistently be negatively correlated with the volatility of system load. But when this situation occurs, those loads should receive a credit for the regulation service they provide.

4. CONCLUSIONS

This project collected, organized, and analyzed data on two real-power ancillary services, contingency reserves and regulation. The purpose of these data and analyses was to develop and test alternative ways to allocate the costs of these services to individual customers. In particular, the goal was to find methods that appropriately assign costs to those transmission customers that are responsible for causing those costs to be incurred. (Ultimately, retail customers pay all these costs.)

For contingency reserves, cost causation focuses on generator size and the number and size of forced outages. For regulation, cost causation focuses on the volatility of the loads of individual customers. Specifically, this project analyzed data on the size of large generating units and the number of forced outages for the 1-year period from July 2001 through June 2002. The purpose of this part of the project was to assess the pros and cons, as well as the effects on individual generators, of different ways to allocate the costs of the contingency-reserve services to individual generators. These alternatives included ones that focused on the number and size of large outages, the largest units online each hour, and the share of system load served by each generator each hour.

In a similar fashion, the project analyzed data on the 1-minute fluctuations in system loads and the loads of 15 large industrial and municipal customers. The purpose of this part of the project was to analyze the effects of allocating the total costs of the regulation ancillary service to individual customers on the basis of the volatility of their 1-minute loads and the correlations between customer-specific loads and the total system load. The purpose of this part of the project was to analyze the effects of allocating the total costs of the regulation ancillary service to individual customers on the basis of the volatility of their 1-minute loads and of the correlations between customer-specific loads and the total system load.

For both services, the allocation methods examined here would eliminate the subsidies (cost shifts among transmission customers) implicit in the method currently used to charge for ancillary services. For contingency reserves, one method would require some generators to pay as much as \$5/MWh of energy for reserves, while another method would require other generators to pay nothing; the average price was \$2.0/MWh. For regulation, the method developed here would charge one customer \$17/MWh, while another customer would receive \$0.6/MWh; the average price was \$0.9/MWh. Although some customers would appropriately pay more for these ancillary services and others would appropriately pay less, the average cost for these services would remain unchanged.

The electricity industry should consider seriously the methods developed here to assign costs for the two sets of ancillary services we studied. Specifically, we believe the current method of charging transmission customers for these services on the basis of energy production or consumption has nothing to do with cost causation and is therefore economically inefficient as well as inequitable. Implementing the kinds of cost-allocation methods presented here will improve economic efficiency within bulk-power systems and wholesale electricity markets. In addition, customer charges will be more equitable in that they will reflect the costs individual transmission customers impose on the bulk-power system.

REFERENCES

D. Davies 2002. Personal communication between D. Davies, Western Electricity Coordinating Council, Salt Lake City, UT, and E. Hirst, consulting in Electric-Industry Restructuring, November 5.

E. Hirst and B. Kirby 1997. *Ancillary Service Details: Operating Reserves*, ORNL/CON-452, Oak Ridge National Laboratory, Oak Ridge, TN, November.

E. Hirst and B. Kirby 1998. *Unbundling Generation and Transmission Services for Competitive Electricity Markets: Examining Ancillary Services*, NRRI 98-05, National Regulatory Research Institute, Columbus, OH, January.

Interconnected Operations Services Working Group 1997. *Defining Interconnected Operations Services Under Open Access*, EPRI TR-108097, Electric Power Research Institute, Palo Alto, CA, May.

B. Kirby and E. Hirst 2000. *Customer-Specific Metrics for the Regulation and Load-Following Ancillary Services*, ORNL/CON-474, Oak Ridge National Laboratory, Oak Ridge, TN, January.

NERC (North American Electric Reliability Council) 2002. "Policy 1—Generation Control and Performance," *NERC Operating Manual*, Princeton, NJ, November 21.

PJM Interconnection, LLC 2002. "Billing Determinant Definitions," www.pjm.com, accessed November 14, 2002.

WECC (Western Electricity Coordinating Council) 2002. *Reliability Criteria, Part III—Minimum Operating Reliability Criteria*, Salt Lake City, UT, April.

