Ancillary Services:
Technical and Commercial Insights

July 2007

Brendan Kirby

Prepared For

WÄRTSILÄ

CONTENTS

Page

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

ACRONYMS

ACE	area control error
AEP	American Electric Power
AGC	automatic generation control
BAAL	balancing authority area control error limit
BPA	Bonneville Power Administration
CAISO	California Independent System Operator
CPM	control performance metric
CPS	control performance standard (1&2)
DCM	disturbance control measure
DCS	disturbance control standard
ERCOT	Electric Reliability Council of Texas
FCM	forward capacity market
FERC	Federal Energy Regulatory Commission
FRR	frequency responsive reserve
GSU	generator step-up transformer
IPP	independent power producer
ISO-NE	Independent System Operator of New England
ISO	Independent System Operator
LGIA	Large Generator Interconnection Agreement
LGIP	Large Generator Interconnection Procedures
LIPA	Long Island Power Authority
LMP	locational marginal price
MAAC	Mid-Atlantic Area Council
MAIN	Mid-America Interconnected Network
MISO	Midwest Independent System Operator
MVA	million volt-amps
MVar	million volt-amps reactive
MW	megawatt
MWh	megawatt-hour of energy
NERC	North American Electric Reliability Council
NPCC	Northeast Power Coordinating Council
NYISO	New York Independent System Operator
OATT	Open Access Transmission Tariff
PJM	PJM Interconnection L.L.C.
PSC	Public Service Commission
PUC	Public Utility Commission
RMR	reliability must run generator
RTO	regional transmission organization
SCE	Southern California Edison
TO	transmission owner
UFLS	under frequency load shedding
UVLS	under voltage load shedding
VAR	volts-amps reactive
WECC	Western Electricity Coordinating Council

EXECUTIVE SUMMARY

This white paper has two primary goals. First, it provides background material including what ancillary services are, why they exist, what is required to provide them, what the cost drives are, and their market prices. Second, it provides modeling results that show how an engine driven generating plant can increase profits by selling ancillary services and energy in dynamic hourly markets. Unlike the paper itself, this executive summary provides conclusions before foundation in order to first highlight the potential value.

Selling ancillary services as well as energy can greatly increase a generator's profitability as shown by the modeling results presented in Table E1. A 100 MW hypothetical gas-fired engine-driven generating plant was modeled for all of 2005 operating in four regions: California, Texas, western New York, and Long Island. The plant was first modeled only selling energy. Next it was modeled optimizing the hourly sale of energy and three ancillary services: regulation, spinning reserve, and non-spinning reserve. Profits increased by 17% to 250%. Results will differ for a real plant but the simple simulation shows that there is significant profit potential. See chapter 5 for details.

Table E1 Selling ancillary services in addition to energy increases profits in each of the regions studied in the 2005 simulation of a 100 MW hypothetical engine driven generating plant.

Annual profits in $ millions	California	Texas	Western NY	Long Island
Energy Only	$1.6	$6.3	$2.9	$16.4
Energy when selling AS	$0	$4.1	$1.1	$14.0
Regulation	$1.7	$2.1	$2.8	$4.4
Spin	$1.5	$1.3	$0	$0.1
Non-Spin	$2.3	$3.9	$0.4	$0.7
Total With Energy & AS	$5.4	$11.3	$4.3	$19.1
Additional Profit	**$3.9**	**$5.1**	**$1.5**	**$2.7**
Increase	**250%**	**81%**	**51%**	**17%**

This increased profit potential comes from the integrated sale of energy and three ancillary services. Seven ancillary services are commercially significant as possible income sources for generators: regulation, load following, spinning reserve, non-spinning reserve, supplemental or replacement reserve, voltage support and black start. All are procured by system operators to support grid reliability. Five are typically traded in hourly markets (regulation, load following, spinning, non-spinning, and replacement reserves).[1] Two are procured through longer term, often negotiated, agreements (voltage support and black start). Two services are used continuously to support normal operations (regulation and load following). Three services continuously stand ready to respond if other power system equipment fails (spinning, non-spinning, and replacement reserves). The services, brief descriptions, and their market properties, including rough price ranges, are shown in Table E2.

[1] Load following is obtained from sub-hourly energy markets.

Table E2 Properties of key ancillary services

<table>
<tr><th rowspan="2">Service</th><th colspan="5">Service Description</th></tr>
<tr><th>Response Speed</th><th>Duration</th><th>Cycle Time</th><th>Market Cycle</th><th>Price Range* (average/max) $/MW-hr</th></tr>
<tr><td colspan="6">Normal Conditions</td></tr>
<tr><td rowspan="2">Regulating Reserve</td><td colspan="5">Online resources, on automatic generation control, that can respond rapidly to system-operator requests for up and down movements; used to track the minute-to-minute fluctuations in system load and to correct for unintended fluctuations in generator output to comply with Control Performance Standards (CPSs) 1 and 2 of the North American Electric Reliability Council (NERC 2006)</td></tr>
<tr><td>~1 min</td><td>Minutes</td><td>Minutes</td><td>Hourly</td><td>35-40
200-400</td></tr>
<tr><td rowspan="2">Load Following or Fast Energy Markets</td><td colspan="5">Similar to regulation but slower. Bridges between the regulation service and the hourly energy markets.</td></tr>
<tr><td>~10 minutes</td><td>10 min to hours</td><td>10 min to hours</td><td>Hourly</td><td>-</td></tr>
<tr><td colspan="6">Contingency Conditions</td></tr>
<tr><td rowspan="2">Spinning Reserve</td><td colspan="5">Online generation, synchronized to the grid, that can increase output immediately in response to a major generator or transmission outage and can reach full output within 10 min to comply with NERC's Disturbance Control Standard (DCS)</td></tr>
<tr><td>Seconds to <10 min</td><td>10 to 120 min</td><td>Hours to Days</td><td>Hourly</td><td>6-17
100-300</td></tr>
<tr><td rowspan="2">Non-Spinning Reserve</td><td colspan="5">Same as spinning reserve, but need not respond immediately; resources can be offline but still must be capable of reaching full output within the required 10 min</td></tr>
<tr><td><10 min</td><td>10 to 120 min</td><td>Hours to Days</td><td>Hourly</td><td>3-6
100-400</td></tr>
<tr><td rowspan="2">Replacement or Supplemental Reserve</td><td colspan="5">Same as supplemental reserve, but with a 30-60 min response time; used to restore spinning and non-spinning reserves to their pre-contingency status</td></tr>
<tr><td><30 min</td><td>2 hours</td><td>Hours to Days</td><td>Hourly</td><td>0.4-2
2-36</td></tr>
<tr><td colspan="6">Other Services</td></tr>
<tr><td rowspan="2">Voltage Control</td><td colspan="5">The injection or absorption of reactive power to maintain transmission-system voltages within required ranges</td></tr>
<tr><td>Seconds</td><td>Seconds</td><td>Continuous</td><td>Year(s)</td><td>$1-$4/kvar-yr</td></tr>
<tr><td rowspan="2">Black Start</td><td colspan="5">Generation, in the correct location, that is able to start itself without support from the grid and which has sufficient real and reactive capability and control to be useful in energizing pieces of the transmission system and starting additional generators.</td></tr>
<tr><td>Minutes</td><td>Hours</td><td>Months to Years</td><td>Year(s)</td><td>-</td></tr>
</table>

* Prices are approximate ranges in $/MW-hr for 2005 and include California, ERCOT, and New York. See Table 3 and Table 4 for more detail.

Essentially all generators can profit by selling ancillary services. Generators with greater flexibility can profit more than less flexible units. Detailed production cost modeling results, shown in Table E3, show that a flexible engine driven generating plant can earn more profits than a less flexible combustion turbine plant. The engines' high part-load efficiency improves its ability to provide regulation. This is in addition to greater profits due to the higher engine efficiency, lower efficiency degradation with increased ambient temperature, and the lower capacity degradation at higher altitudes.

Table E3 The greater flexibility of the engine driven generating plant results in greater ancillary service profits when compared with the combustion turbine driven generating plant.

Annual Profits in $ millions	**12 Engines**	**2 Combustion Turbines**
Capacity at plant site	100 MW	79 MW
Energy Only Profit	$1.5	$0.6
Energy and Ancillary Services		
Energy Profit	-$0.4	-$0.5
Regulation	$1.8	$0.7
Spinning Reserve	$2.3	$1.4
Non-spinning reserve	$2.1	$1.8
*Total Profit**	*$5.8*	*$3.3*
Additional Profit	***$4.3***	***$2.7***

*Includes startup costs for cycling

Ancillary services are capacity services, not energy services. Even the services which trade in hourly markets (regulation and the contingency reserves) are capacity services. Costs are primarily generator opportunity costs based on capacity that must be withheld from the energy market. Ancillary service prices are consequently volatile.

Ancillary service prices are volatile (Chapter 4) and often high. A generator should continually reevaluate its position in the energy and ancillary service markets. Figure E1 shows that the optimized example plant constantly moved between energy and the ancillary services in the modeled year.

While the underlying physical needs of the power system remain constant, reliability and market rules are in flux as restructuring continues. Some rules do not perfectly reflect the power system's physical requirements. When a generating technology supports power system reliability but is not compensated appropriately by current market rules, work to change those rules. Market designers are typically receptive to serious market participants, especially when the market participants are technically correct. Prices paid for regulation, for example, do not currently reflect the quality of regulation provided. Generator owners with units that follow automatic generator control (AGC) signals more closely, and therefore reduce the overall amount of regulation needed, should argue for rule changes that appropriately compensate them.

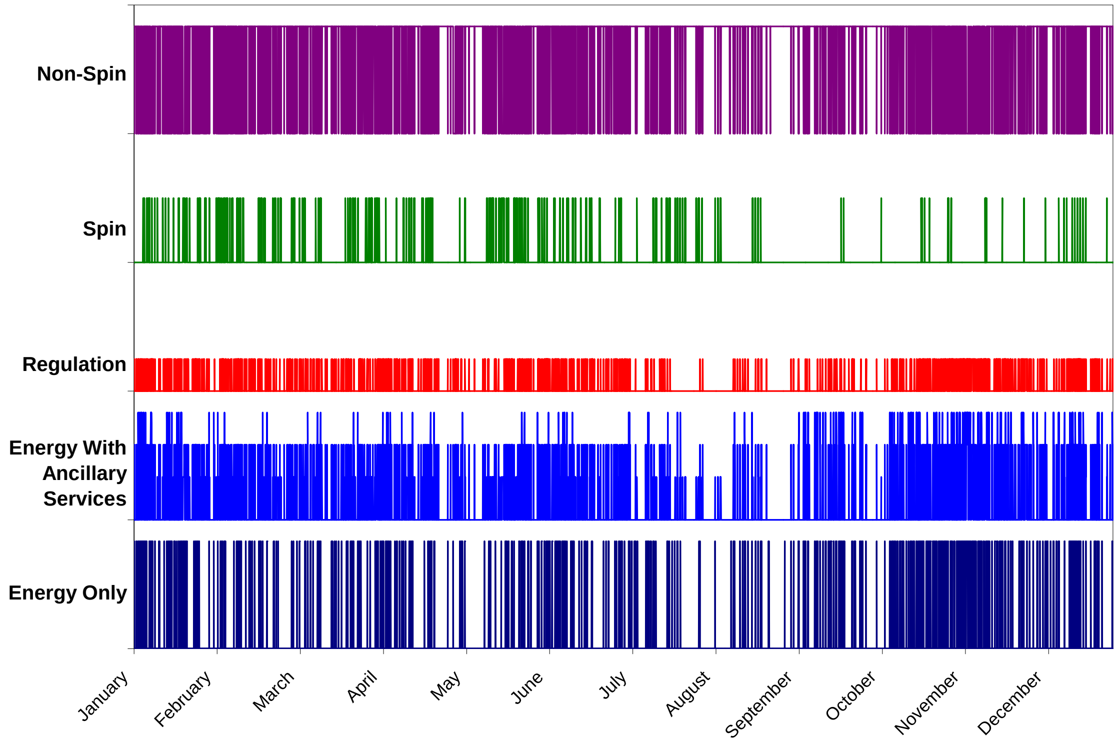

Figure E1 This 2005 California generator constantly moves between volatile energy and ancillary services markets in order to maximize profits.

1. INTRODUCTION

Power systems require ancillary services to maintain reliability and support their primary function of delivering energy to customers. Ancillary services are principally real-power generator control capacity services the system operator uses over various time frames to maintain the required instantaneous and continuous balance between aggregate generation and load. Seven ancillary services are discussed in this report. Two are used continuously under normal conditions: regulation and load following (or fast energy markets). Three only respond intermittently to reliability events but they must continuously stand ready to respond: spinning reserve, non-spinning reserve, and replacement or supplemental reserve. Two additional services are discussed: voltage support and black start. Voltage support is the provision of dynamic reactive power. Voltage support is the only non-real-power service. Black start involves having the capability to start the generator without support from the power grid and subsequently having enough real and reactive power capability and control to be useful in restarting the rest of the power system.

Ancillary services are not new. The functions have been provided by vertically integrated utilities since power systems began to be formed a century ago. With restructuring it has become necessary to more carefully define, measure, and pay for these services. Ancillary services introduce an additional level of complexity and the potential for additional revenue for generation owners. Provision of ancillary services interacts with the provision of energy and capacity. Maximizing revenue requires optimizing the joint production of all available services.

Fundamental power system reliability requirements are quite stable. Ancillary service and energy market rules are not. While this report discusses current market rules its primary focus is on the underlying physical requirements. Understanding the underlying physical power system reliability requirements can help identify when market and reliability rules should be changed to help system operators take advantage of the unique capabilities of maneuverable generators.

This report is organized into an executive summary and seven chapters. Chapter 1 is this introduction. Chapter 2 provides a brief discussion of energy and capacity as a background for understanding ancillary services. Chapter 3 discusses the ancillary services themselves, why the power system needs each, and the physical requirements for supplying them. Chapter 4 discusses ancillary service prices, the cost components for each service, why prices are so volatile, and provides historic hourly price information from several markets. Chapter 5 discusses the interactions among the ancillary services. Example modeling results are presented showing how profits can be maximized by selecting how much of each ancillary service to sell during each hour of a study year. Chapter 6 examines the value of flexibility by comparing profits from a flexible engine driven generating plant with those from a less flexible combustion turbine plant utilizing detailed modeling of one year of operations at a specific example location. Chapter 7 provides conclusions.

2. ENERGY AND CAPACITY: NECESSARY BUT INSUFFICIENT

The electric power system has two unique requirements which must be continuously and exactly satisfied in order to maintain overall system stability and reliability. They are (i) the need to maintain a constant balance between generation and load (there is no storage), and (ii) the need to adjust generation (or load) to manage power flows within the constraints of individual transmission facilities (there is no flow control). [2] These requirements have existed since interconnected power systems started to develop a century ago and vertically integrated utilities have traditionally maintained this continuous balancing act as a normal part of the electricity business. These two principles lead to four important consequences: (i) prices are inherently volatile, (ii) system operations and transmission are communal and must be regulated, (iii) current operations are often restricted by preparations for the next unlikely event, and (iv) response has value.

With restructuring, the multiple functions which the vertically integrated utilities performed as part of their bundled service are being explicitly delineated. The Federal Energy Regulatory Commission (FERC), through Order 888, Order 889, Order 2000, and the continuing reform effort has defined these as “ancillary services” which are “necessary to support the transmission of electric power from seller to purchaser given the obligations of balancing areas and transmitting utilities within those balancing areas to maintain reliable operations of the interconnected transmission system”.

The two basic characteristics and the four consequences underlie the need for and value of capacity and the ancillary services. The basics often get lost in the implemented details of reliability and market rules. Energy is still the basic commodity that is of interest to electricity users. Everything else simply supports the delivery of energy. In its simplest form operating an interconnected power system can be reduced to a few tasks:

- Balance aggregate generation to aggregate load.
  - Under normal conditions.
  - Under contingency conditions.[3]
- Maintain voltages throughout the power system.
  - Under normal conditions.
  - Under contingency conditions.
- Control generation (input locations and amounts) to avoid overloading transmission lines.
- Restart the system after it collapses because you failed to do one of the above.

[2] This is not strictly true. Electricity can be stored and flow can be controlled on a small scale and/or at high cost. The behavior of interconnected AC power systems, however, is dominated by the lack of practical storage and limited flow control. These two characteristics differentiate electricity from communications systems (telephone, cell, radio, and internet), pipe systems (gas, water, oil, etc.), and transportation (air, rail, road, and sea). The result is a very different set of control requirements and ancillary services.

[3] A contingency is the sudden, unexpected loss of a generator or transmission element. Slower events, like load being higher than forecast, are not contingencies.

Generation technologies differentiate based upon their efficiencies, fuels, capital costs, control capability, response speed, and abilities to support the above functions. Ancillary services have been defined to maintain system reliability by explicitly addressing the above requirements.

2.1 ENERGY

The lack of energy storage and the varying needs for electricity result in volatile hourly energy prices. This is true in the vertically integrated environment where the marginal cost of power (system lambda) is optimized in economic dispatch. It is also true in the restructured environment where hourly (and sub-hourly) markets are cleared based on energy bid prices.

The volatility of energy prices is important to the discussion of ancillary services for two reasons. First, opportunity costs in the energy markets drive ancillary service prices. Ancillary service prices are, consequently, typically more volatile than energy prices. Second, varying hourly energy requirements and energy prices complicate consideration of load following (or fast energy markets), which may or may not have high value and may or may not be an ancillary service depending on the mix of generation in the area and the structure of the regional market.

Energy is traded through long-term bilateral contracts and through hourly and sub-hourly (5-15 minutes) markets in many regions. Sub-hourly energy markets allow system operators to do a great deal of balancing of generation and load through energy markets without having to explicitly purchase additional control services. Sub-hourly energy markets may present significant opportunities for flexible generators in the future. We will discuss this in more detail later in the section 3.1.2: Load Following vs. Fast Energy Markets.

2.2 CAPACITY

Capacity is especially important in electric power systems because there is no energy storage. Sufficient generating capacity must always be available to serve immediate load requirements and to compensate for any system failures (contingencies) as shown in Figure 1.[4] Vertically integrated utilities use long-term central planning to obtain needed capacity. In the restructured environment how capacity is obtained differs depending on the market structure.

A major debate continues concerning the advisability of creating explicit capacity markets versus having only energy markets. In California, where there are currently no

[4] The only alternative to having sufficient generating capacity is curtailing load, voluntarily or involuntarily. While robust load response markets could exist they do not now and show no signs of becoming economically significant in the reasonably near future. Involuntary load control (load shedding) is an important reliability tool used as a last-ditch effort to contain cascading blackouts. One third to one half of the entire system load can currently be shed nearly instantaneously with under- frequency, under-voltage, and fast manual load shedding. There are very strong incentives not to use involuntary load shedding, however, so it does not influence the need for, value of, or price of generating capacity.

explicit capacity markets, generators must make sufficient profit in the energy market to cover their capital costs. This means that the energy price a peaking unit receives must be very high if it only runs a few hours a year or it will go out of business. Some argue that this correctly values the cost of energy during the peak hours and that energy prices should not be capped. Unfortunately, capacity reserves, which may never be called upon to run, have no way to be paid. Market designers, especially in the north east (and possibly FERC), worry that insufficient generation will be built to assure reliability if generators must rely only on real-time energy markets. The alternative which PJM, the New York Independent System Operator (NYISO), and the Independent System Operator of New England (ISO-NE) use is to run capacity markets which pay generators to be available and assure that the marginal generators' fixed costs are covered. This reduces the volatility of the energy markets but has the disadvantage of tending to draw low capital cost (and typically high operating cost) generators into the mix. It spreads the cost of meeting the peak across all electricity users in all hours. By changing the generation mix it impacts energy prices at all times. It does help assure that capacity is available when needed, however. Capacity markets themselves tend to be volatile and highly location dependent.

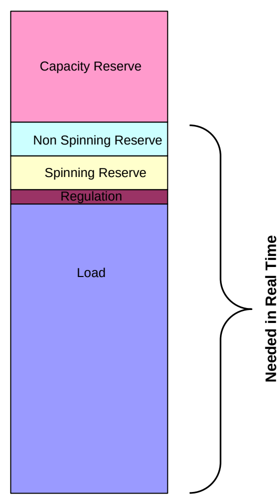

Figure 1 Capacity is required to meet load, provide regulation and contingency reserves, and to cover the unavailability of other generation and load forecast errors (capacity reserves).

The fundamental problem is that capacity markets decide what generation gets built and energy markets decide what generation gets run. With years separating these two decisions it is not possible to optimize the generation mix. Contrast this with the cooptimization of energy and ancillary service markets employed by the same north eastern markets. They have found that running immediately sequential energy and ancillary services markets is unacceptably suboptimal. It is not surprising then that even where there is agreement that capacity markets/payments are necessary to keep generators viable there is little agreement in *how* those markets should work. Market rules are being developed through consensus processes in each of the regions and are in a state of flux. FERC approved ISO-NE's latest Forward Capacity Market (FCM) on June 16, 2006. FCM will conduct annual auctions for capacity obligations three years in the future.

There is a general recognition that location is important for capacity. In New York State, for example, summer capacity markets clear at ~$12/kW-mo for New York City, ~$9/kW-mo for Long Island, and ~$1/kW-mo for the rest of the state. Winter prices in each New York region tend to be around half of summer prices. In contrast, PJM monthly capacity credit market prices dropped from $0.30/kW-mo in January 2005 to just above $0.15/kW-mo in December. All prices are volatile and should be investigated carefully before making investment decisions.

Generators should always seek capacity payments as they are often a significant portion of the generators' total income. Generators receiving capacity payments are typically required to bid into energy and ancillary service markets.

2.3 FREQUENCY CONTROL

System frequency is a fundamental indicator of power system health. It can be observed everywhere on the power system and provides an immediate indication of the balance between generation and load. Frequency drops when load exceeds generation and rises when generation exceeds load. Large frequency deviations result in equipment damage and power system collapse so frequency is tightly controlled in North America, as shown in Figure 2.

Two distinct mechanisms are used to control frequency. Under normal conditions individual generators ignore system frequency; generator governors typically have a ±0.35 Hz dead band. Instead, each balancing authority operator concentrates on balancing generation and load while also watching system frequency. When total system generation and load are in balance system frequency is stable. Rather than measuring generation and load directly system operators concentrate on maintaining the net interchange with their neighbors at the scheduled amount.[5] A frequency bias term is included in the area control error (ACE) equation requiring each balancing area to increase generation when system frequency is low and decrease generation when frequency is high. The bias is established in MW/0.1 Hz and is based on the MW size of the balancing area. This frequency control mechanism is shown in Figure 3.

[5] Accurately measuring net load is difficult because of the large number of dispersed loads. System losses also must be included in net system load. It is much easier to accurately measure the flows on the limited number of tie lines to neighbors. Interchange schedules are commercially arranged.

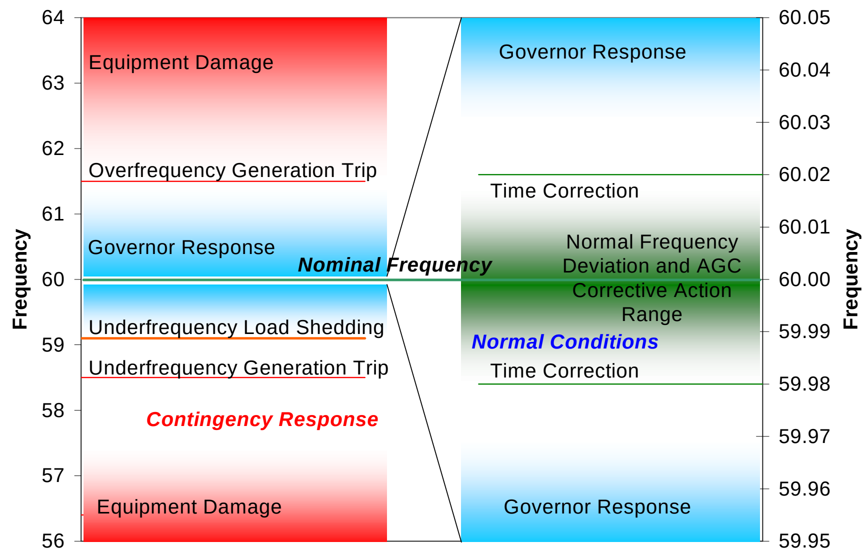

Figure 2 Power system frequency is tightly controlled in North America under normal conditions.

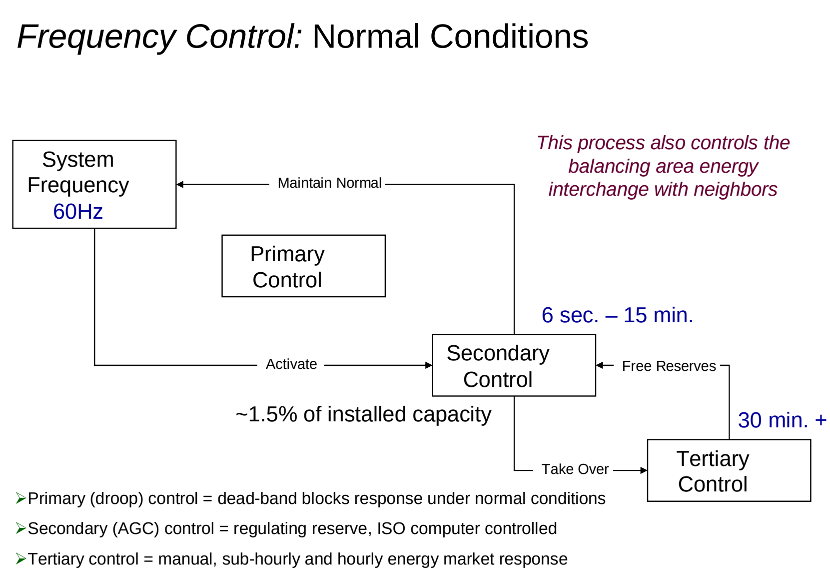

Figure 3 Frequency control is accomplished centrally under normal conditions.

The control mechanism shown in Figure 3 provides excellent frequency control under normal conditions but is too slow to respond to major contingencies. When a large generator suddenly fails system frequency falls rapidly. The generation/load balance must be restored immediately or the power system will collapse. Under contingency conditions, when system frequency moves outside the generator governor dead band, autonomous generator governor action provides immediate response, as shown in Figure 4.

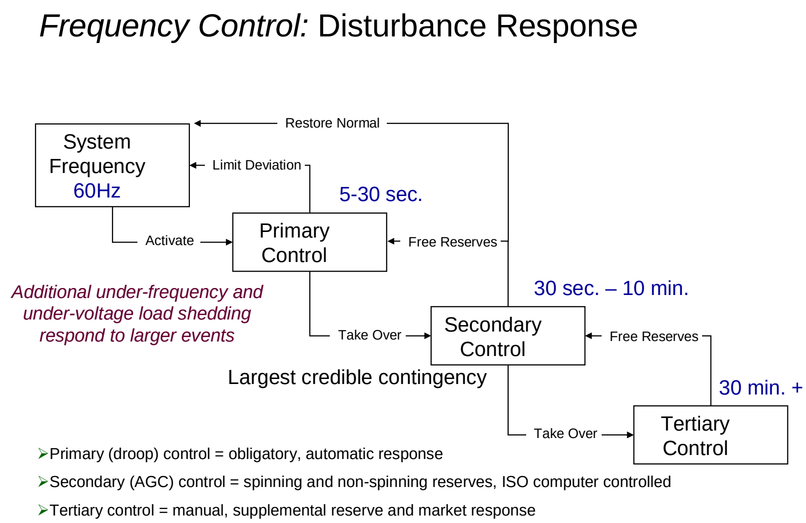

Figure 4 Generator governors provide immediate response to restore system frequency under contingency conditions.

2.4 ANCILLARY SERVICES

Energy and capacity are the basic products customers really use. But energy and capacity alone are not sufficient to reliably operate the power system. A series of ancillary services are required that provide the system operator with the resources needed to maintain the instantaneous and continuous balance between generation and load, to manage transmission line flows, and to implement the control schemes shown in Figure 2 and Figure 3. These services are required under normal conditions and when contingencies happen. Ancillary services also provide the resources needed to restart the power system if the system operator is unable to maintain the generation/load balance and the system collapses. Utilities have been performing these functions for a hundred years as part of their vertically integrated structure. Restructuring and the introduction of competitive generation markets has required that these services be clearly defined and monetized. Markets have been created for several ancillary services in order to minimize the cost of maintaining reliability; they are the subject of the remainder of this report.

3. ANCILLARY SERVICES FOR POWER SYSTEM RELIABILITY

When FERC began the process of restructuring the electric power industry and introducing competition it became necessary to explicitly define the various reliability services that vertically integrated utilities performed in support of delivering energy to end use customers. Most of these services are supplied by generators and generators need to be compensated for providing them. Six of the seven generator supplied ancillary services (all except voltage control) deal with control of real power. As we shall see, the services are distinguished based upon the response time, response duration, and response frequency. Faster, more frequent services get paid more (ancillary service prices will be discussed further in Chapter 4). Response duration, on the other hand, does not translate into higher service price. Response accuracy is not well quantified, but should be.[6] This will be discussed later as well.

The ancillary services are placed into three groups for discussion here: two services which provide continuous response to balance generation and load under normal conditions (regulation and load following), three services which provide reserves that stand ready to respond in the event of a power system contingency (spinning reserve, non-spinning reserve, and supplemental or replacement reserve), and two additional services (voltage control and black start). Table 1 provides brief descriptions of each service and Figure 5 shows how the services are differentiated in response time and duration.

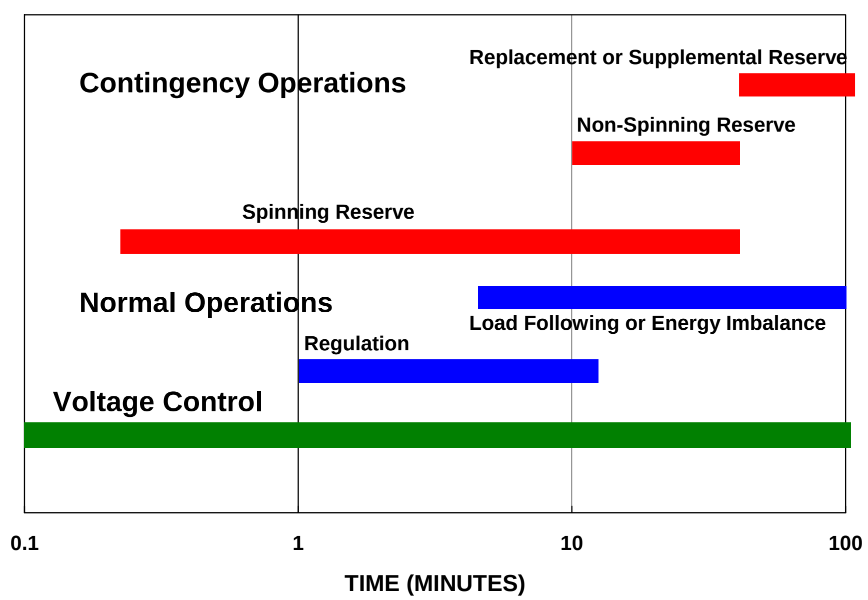

Figure 5 Response time and duration characterize required ancillary service response.

[6] Generators with more accurate response should push for metrics which differentiate, and pay based upon, the quality of response.

Table 1 Definitions of key ancillary services

Service	Service Description		
	Response Speed	*Duration*	*Cycle Time*
Normal Conditions			
Regulating Reserve	Online resources, on automatic generation control, that can respond rapidly to system-operator requests for up and down movements; used to track the minute-to-minute fluctuations in system load and to correct for unintended fluctuations in generator output to comply with Control Performance Standards (CPSs) 1 and 2 of the North American Electric Reliability Council (NERC 2006)		
	~1 min	*Minutes*	*Minutes*
Load Following or Fast Energy Markets	Similar to regulation but slower. Bridges between the regulation service and the hourly energy markets.		
	~10 minutes	*10 min to hours*	*10 min to hours*
Contingency Conditions			
Spinning Reserve	Online generation, synchronized to the grid, that can increase output immediately in response to a major generator or transmission outage and can reach full output within 10 min to comply with NERC's Disturbance Control Standard (DCS)		
	Seconds to <10 min	*10 to 120 min*	*Hours to Days*
Non-Spinning Reserve	Same as spinning reserve, but need not respond immediately; resources can be offline but still must be capable of reaching full output within the required 10 min		
	<10 min	*10 to 120 min*	*Hours to Days*
Replacement or Supplemental Reserve	Same as supplemental reserve, but with a 30-60 min response time; used to restore spinning and non-spinning reserves to their pre-contingency status		
	<30 min	*2 hours*	*Hours to Days*
Other Services			
Voltage Control	The injection or absorption of reactive power to maintain transmission-system voltages within required ranges		
	Seconds	*Seconds*	*Continuous*
Black Start	Generation, in the correct location, that is able to start itself without support from the grid and which has sufficient real and reactive capability and control to be useful in energizing pieces of the transmission system and starting additional generators.		
	Minutes	*Hours*	*Months to Years*

3.1 SERVICES FOR NORMAL CONDITIONS

Regulation and load following or fast energy markets (we will discuss the implications of distinguishing between load following and fast energy markets) are the two services required to continuously balance generation and load under normal conditions. Figure 6 shows a typical daily load pattern with a morning ramp-up, double peak, and evening ramp down. Figure 6 also shows the continuous, random minute-to-minute fluctuation in total system load that is superimposed on the predictable daily load. Regulation is the most expensive ancillary service so we will spend some time discussing it.

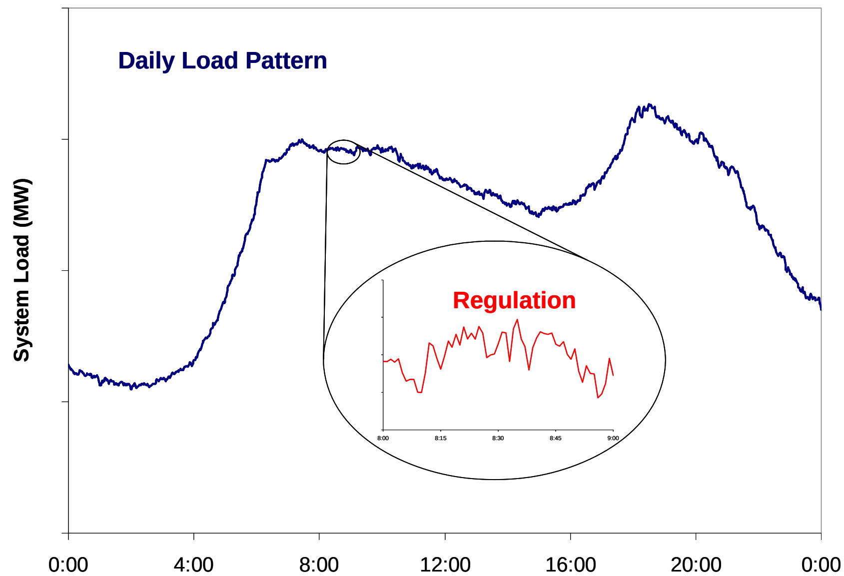

Figure 6 Regulation is a zero-energy service that compensates for minute-to-minute fluctuations in total system load and uncontrolled generation while load following compensates for the slower, more predictable changes in load.

3.1.1 Regulation

Regulation is the use of on-line generation that is equipped with automatic generation control (AGC) and that can change output quickly (MW/min) to track the moment-to-moment fluctuations in customer loads and to correct for the unintended fluctuations in generation. Regulation helps to maintain interconnection frequency, manage differences between actual and scheduled power flows between balancing areas, and match generation to load within the balancing area. Load following is the use of on-line generation, storage, or load equipment to track the intra- and inter-hour changes in customer loads. Regulation and load following characteristics are summarized in Table 2.

Table 2 Comparison of regulation and load following characteristics

	Regulation	**Load following**
Patterns	Random and uncorrelated	Highly correlated
Control	Requires AGC	Can be manual
Maximum swing	Small	10–20 times regulation
Ramp rate (MW/min)	5–10 times load following	Slow
Sign changes per unit time	20–50 times load following	Few

In the PJM region, New York, New England, and Ontario, regulation is a 5-min service, defined as five times the ramp rate in megawatts per minute. In Texas it is a 15-min service, and in Alberta and California it is a 10-min service.

Balancing area operators have not needed to specifically procure load following; it is obtained from the short-term energy market with generators (typically) responding to real-time energy prices. Regulation, however, requires faster response than can be obtained from units responding to market signals alone. Instead, generators offer capacity that can be controlled by the system operator's AGC system to balance the power system. Typical balancing areas require 1-2% of the peak load in regulating capacity.

Balancing areas are not able and not required to perfectly match generation and load. NERC has established the Control Performance Standard (CPS) to determine the amount of imbalance that is permissible for reliability purposes.[7] CPS1 measures the relationship between the balancing area's ACE and the interconnection frequency on a 1-min average basis. CPS1 values can be either "good" or "bad." When frequency is above its reference value, under-generation benefits the interconnection by lowering frequency and leads to a good CPS1 value. Over-generation at such times, however, would further increase frequency and lead to a bad CPS1 value. CPS1, although recorded every minute, is evaluated and reported on an annual basis. NERC sets minimum CPS1 requirements that each balancing area must exceed each year.

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

3.1.1.1 *Up and Down Regulation*

Some regions split regulation into an up and a down service, others treat it as a controllable range. The two methods are equivalent. Figure 7 provides a simple example showing the output from a generator that is averaging 90 MW and regulating between 80 and 100 MW. In PJM this would be 10 MW of regulation on top of 90 MW of energy. In California this would be 10 MW of up regulation and 10 MW of down regulation on top of 90 MW of energy. Because the regulation range in the PJM market is twice that of the range in the California Independent System Operator (CAISO) market the equivalent PJM price is twice the CAISO price.

The example is slightly more complex, but still equivalent, when only one side of regulation is considered. Assume the example California generator was selling 100 MW of energy when it decided to also sell 20 MW of down regulation. The average energy output would now be only 90 MW and the generator would have to obtain 10 MW of

[7] NERC is in the process of developing three new metrics to replace the CPS and DCS metrics; CPM (control performance metric), DCM (disturbance control metric), and BAAL (balancing authority area control error limit). These are not discussed here. The requirements for generators supplying ancillary services will not change significantly if the new metrics are adopted.

energy from the spot market to meet its 100 MW obligation. Similarly, if the California generator was selling 80 MW of energy when it sold 20 MW of up regulation to the CAISO it would be producing an average of 10 additional MW. The excess energy would be sold in the spot market which might be above or below the generators operating cost. In all cases the generator would average 90 MW output and would be controlled by the independent system operator (ISO) within the 80 to 100 MW range. It is not necessary to explicitly calculate the energy market transaction impact when dealing in markets that do not segregate up and down regulation in order to determine the total payment but it is still necessary to consider the opportunity cost, so the net impact is the same.

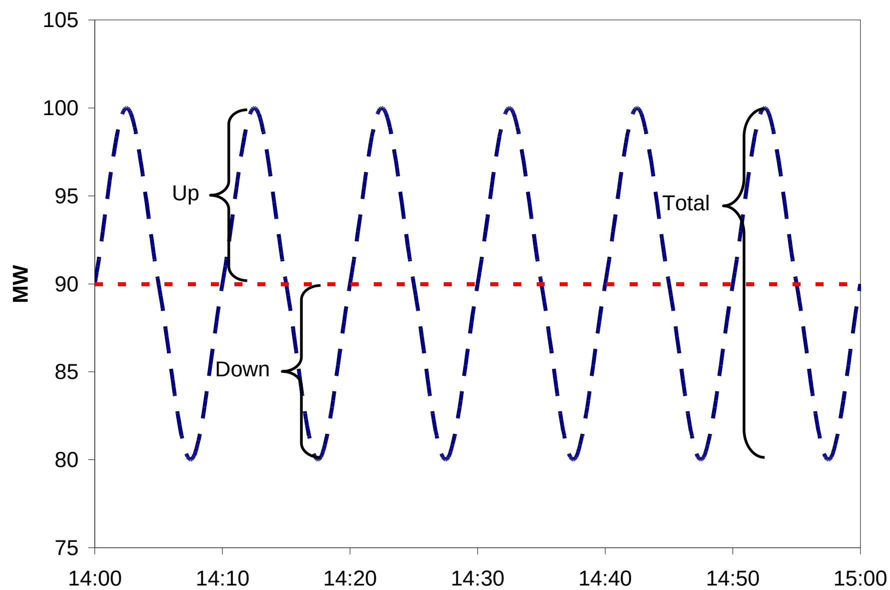

Figure 7 Up/down and total regulation are equivalent.

3.1.1.2 *Response Accuracy*

Interestingly, though there are specific metrics governing how accurately Balancing Areas must match aggregate load with aggregate generation there are essentially no metrics concerning how accurately generators follow system operator AGC commands. Response accuracy *should* differentiate regulation suppliers but it does not do so yet. As shown in Figure 8, conventional generators, especially large thermal generators, do not follow control signals perfectly. The interconnected power system has been designed to accommodate this constraint so reliability is maintained in spite of the deficiency. Balancing Authorities are only required to match aggregate load with aggregate generation within the CPS 1 & 2 statistical limits.

Balancing authorities typically require regulation capacity equal to 1-2% of the peak load to meet the CPS 1 & 2 limits but the exact amount depends on the volatility of the load and the accuracy of the generators providing response; the more accurate the response the less regulation that is required. System operators determine the amount of regulation required empirically. If the CPS scores are not high enough they increase the amount of regulation they are purchasing. If CPS scores are too good they decrease the amount.

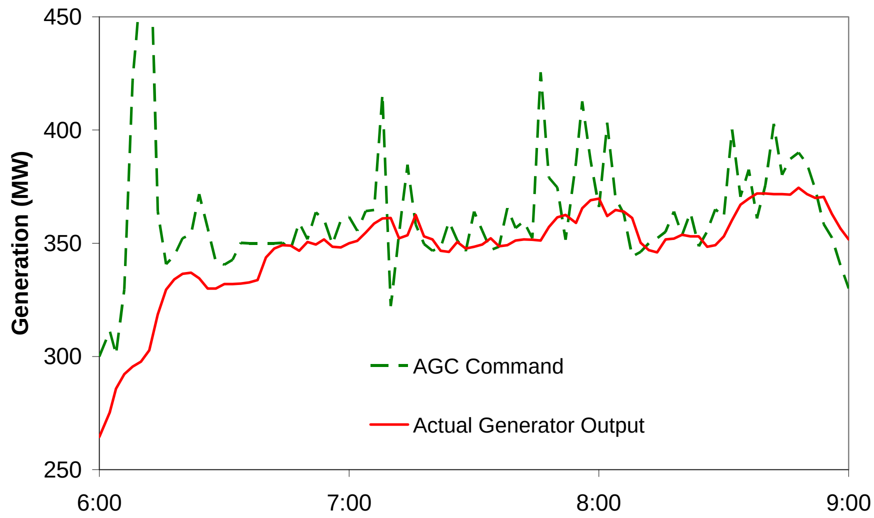

Figure 8 This coal fired power plant follows AGE regulation commands poorly.

Accuracy of generator response is not formally measured because it is felt that there is little a generator owner can do to make the unit respond more accurately. Some regions do establish certification criteria. PJM, for example, uses the response test shown in Figure 9 to certify generators to supply regulation. Generators are tested to assure that they are able to respond. Once certified they can continue to supply (and be paid for) regulation unless the system operator notices that response has degraded unacceptably.

More accurate resources would reduce the amount of required regulation. A 30,000 MW system that requires 450 MW of regulation capacity from its thermal generators might only require 350 MW of more accurate response capability to achieve the same CPS 1 & 2 scores. All things being equal the 350 MW of more accurate response should be paid the same total dollars that the 450 MW of poor response is currently being paid. The $/MW-hr rate for accurate regulation should be higher than the $/MW-hr rate currently being paid for conventional regulation. For this to happen market designers will have to be shown that there is a difference in regulation quality that they can measure and a benefit that they can quantify.

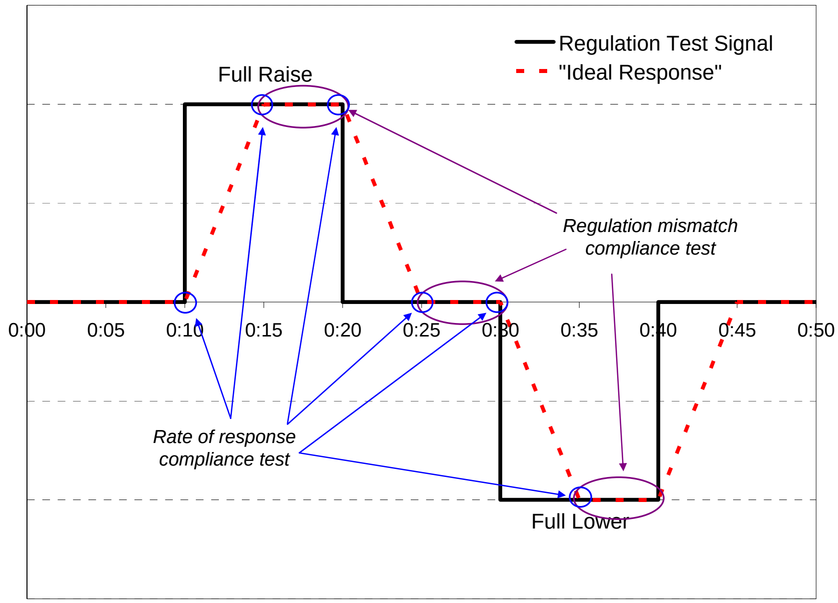

Figure 9 PJM uses a pass/fail test to certify, and pay, generators that supply regulation.

3.1.2 Load Following vs. Fast Energy Markets

When FERC introduced electric industry restructuring in 1996 with Order 888 it did not establish a load following service. Instead, load following has been provided by energy markets. Some load following needs (the morning load pickup and evening load reduction) can be forecast and at least partially addressed by day-ahead markets. Markets clearing at five minute intervals can certainly respond quickly enough to meet the remaining load following requirements. But do fast energy markets appropriately reward generators for their maneuverability? This is an interesting and underappreciated question that has not been addressed by market designers yet.

Note first that the minute-to-minute regulation balancing ancillary service is a capacity service. It is generation capacity held in reserve for use by the system operator to respond to variations in aggregate system load and uncontrolled generation. It is not fundamentally an energy resource. Any net energy which comes out of or goes into a regulation resource is incidental and is paid for separately. Presumably load following would also be a capacity service with (slower) responsive reserves held back to enable the system operator to balance aggregate generation with aggregate load. Any net energy into or out of the load following resource would, presumably, be incidental and settled separately. In contrast, fast energy markets are fundamentally *energy* markets that require an incidental response (ramp) so that the unit is correctly positioned to provide energy for the transaction. This distinction between the basic commodity and the incidental response is at the center of the problem.

Figure 10 presents a typical daily load curve with four classes of generators serving the load. Nearly 20,000 MW can be served from base load generators that can run continuously. The lowest cost generators will be selected as base load units in both the vertically integrated economic dispatch environment and in the market environment. The base load units do not need to have any maneuvering capability in order to successfully meet their energy obligations. Nuclear plants, for example, can meet this need.

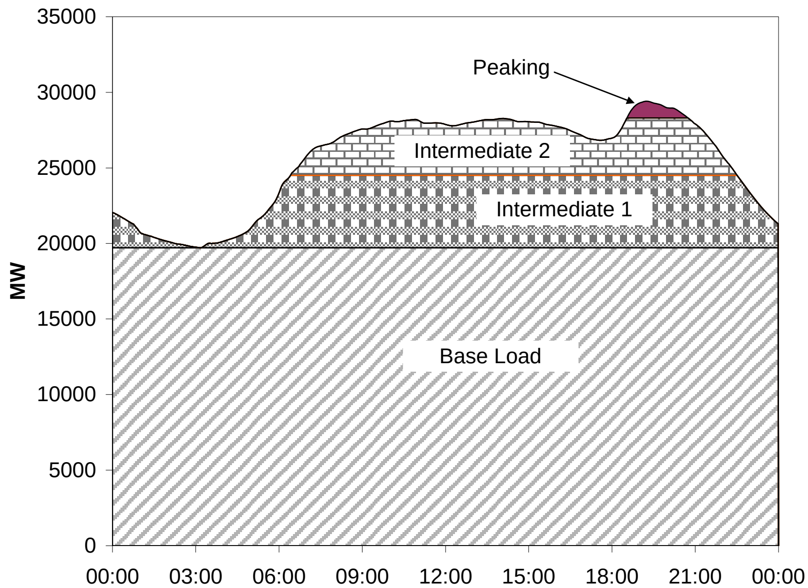

Figure 10. Participation in *energy* markets requires maneuverability for all but base load units.

Something interesting happens when the next generators are selected. Again this applies equally to both the vertically integrated and the market environments. Additional power is not needed all day long. In order to be selected to provide the next block of energy the intermediate units must be able to turn on for the hours when they are needed and off for the hours when they are not needed. It is probably a help to Intermediate Unit 1 that the requirement ramps up and down because the unit may not be able to turn on and off instantaneously. Once on, however, Intermediate Unit 1 has a flat output until it ramps off.

The requirements for Intermediate Unit 2 and the Peaking Unit are more interesting. They must have output flexibility simply to be in the energy market (or available for economic dispatch). The amount of output that will be required in any given hour depends on the overall system load which varies from hour-to-hour, day-to-day, season-to-season, and year-to-year. Regardless of the load following requirements, the last generators in the loading order, the most expensive units, must be flexible in the amount of power they can generate or they will be unable to successfully sell their *energy* output.

Let us belabor this point a bit. These last generators must be flexible concerning their output levels and run times simply to be able to sell energy into a variable market.

Inflexible units such as nuclear plants simply can not serve this part of the load or sell energy into this market regardless of their cost.

The basic question of whether we need load following or fast energy markets can be looked at slightly differently now. Do the generators that are built to meet the intermediate and peaking energy markets (higher operating cost, lower capital cost) inherently have enough response capability to meet the system's load following needs? If so there is little point in creating or paying for a load following service. If the generators do not inherently have sufficient maneuvering capability then a load following service is required or the energy markets will be distorted.

We can hypothesize a system where ramping limits influence energy prices and a load following service seems necessary. Figure 11 shows a system with ample \$10/MWh base load capacity. Unfortunately the base load units can only ramp at 1 MW/minute. When the load ramps from 2550 MW to 2850 MW in 30 minutes at 8:00 the base load units simply can not keep up. Peaking units costing \$90/MWh (the only other generators in this example system) are required to serve load for five hours until the base load units can catch up. In a simple market with no load following service the *energy* price would jump from \$10/MWh to \$90/MWh for those five hours – and it would be paid to all generators by all loads. An alternative would be to let the energy market clear at \$10/MWh, purchase ramping capability (load following) from the fast responding generator, and also compensate the load following unit for the incidental energy it had to supply while following the load.

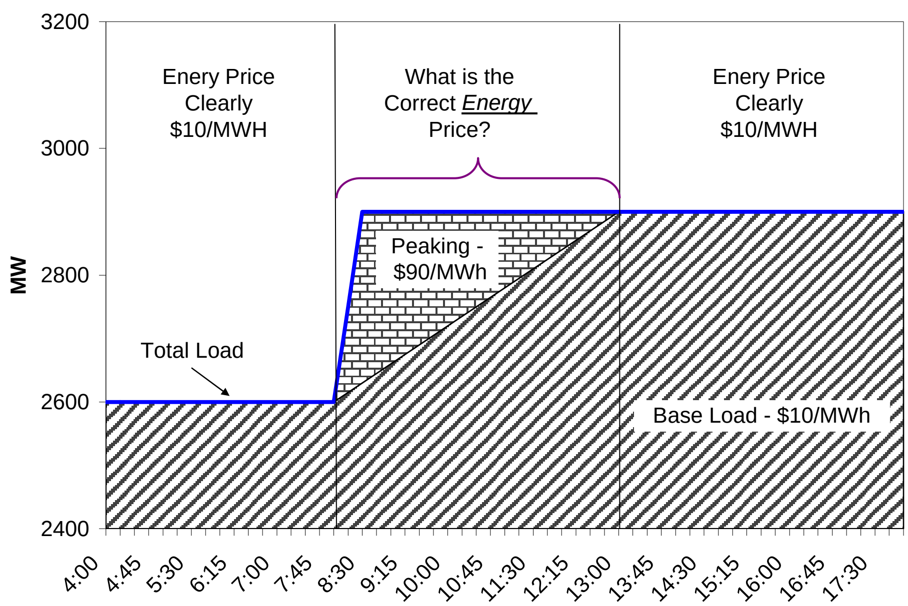

Figure 11. In this simple example load following is required from an expensive peaking generator but energy is only an incidental product.

In Figure 10 and Figure 11 we see two different views of load following based upon the inherent capabilities of the generators that are trying to serve the energy market. If there are ample, reasonably flexible generators on the margin then a specific load following service is probably not justified. If a lack of ramping capability restricts which units can respond then energy markets will be distorted and a load following service would be beneficial.

3.2 SERVICES FOR CONTINGENCY CONDITIONS

Generators and transmission lines can fail at any time. Contingency reserves restore the generation/load balance after the sudden unexpected loss of a major generator or transmission line. Power system frequency drops suddenly when generation trips, as shown in Figure 12. There is no time for markets to react. Frequency-sensitive generator governors responded immediately to stop the frequency drop. Spinning, non-spinning, and supplemental reserves must restore the generation/load balance in order to return frequency to 60 Hz and ACE to zero within 15 minutes in order to meet NERC's disturbance control standard (DCS) requirements. Power systems typically keep enough contingency reserves available to compensate for the worst credible event (contingency). This is typically the loss of the largest generator or the largest importing transmission facility. In Texas, the simultaneous loss of two nuclear plants is credible (as shown by the event recorded in Figure 12), so the Electric Reliability Council of Texas (ERCOT) requires over 2600 MW of contingency reserves.

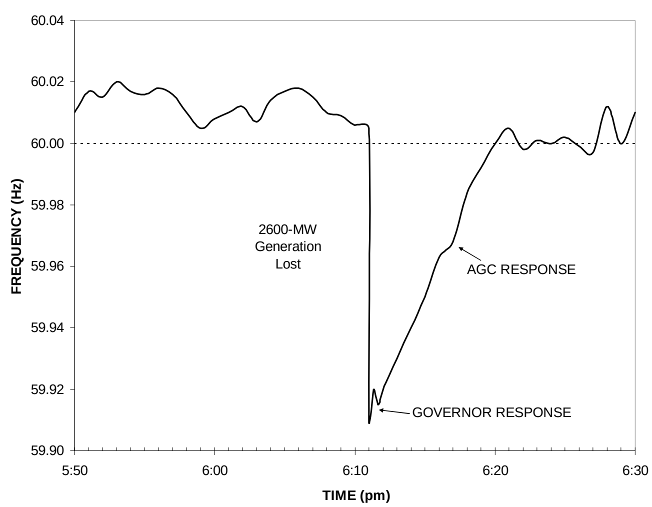

Figure 12 System frequency plummets in response to a major loss of generation and the generation/load balance must be restored quickly.

A series of contingency reserves operate in a coordinated fashion to restore system balance. Figure 13 shows the current contingency reserves we will discuss here (spinning, non-spinning, and replacement or supplemental reserves) as well as frequency responsive reserve. Frequency responsive reserve has historically been a part of spinning reserve and the required additional governor response of all on-line generators. NERC is considering specifically calling out this reserve as a named (and presumably paid for) ancillary service.

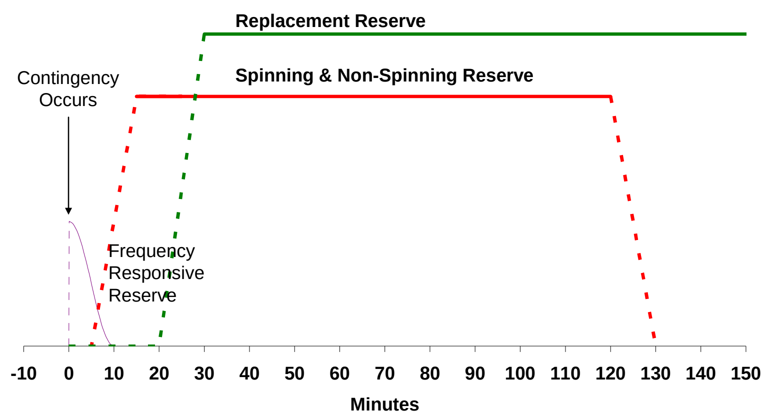

Figure 13 A series of coordinated contingency reserves restore the system generation/load balance immediately following a major contingency.

3.2.1 Spinning Reserve

Spinning reserve is supplied by generation that is on-line, less than fully loaded, begins responding immediately, and is fully responsive within ten minutes.[8] It continuously stands ready to respond to a major loss of generation or transmission. It can be deployed autonomously if system frequency falls or in response to a system operators command. The generator must have a governor to sense and respond to frequency drops and telecommunications (typically AGC) to respond to system operator deployment commands. Spinning reserves must be capable of sustaining the response for, typically, two hours though system operators try to relieve spinning reserves much sooner in order to be ready for the next contingency.

A generator can be limited in ramp rate (MW/min) or in total available generation. Metrics are not well defined; nothing specifies how much of the generation must respond "immediately" or even what "immediately" means.

Generating plants composed of multiple fast-start units (possibly with clutch connected generators) create an interesting possibility for potential greater spinning reserve credit. The overall spinning reserve real and reactive response of a multi engine plant, with some

[8] Responsive load is just beginning to be allowed to supply spinning reserve in some regions.

engines operating and others poised for rapid start, may exceed that of a conventional fossil steam plant. The amount of immediate response might be greater and the time to full output might be less, enhancing power system reliability. If all the generators were connected to their engines through clutches and were spinning the full reactive response of the plant would be immediately available and the stability response of the plant would be enhanced as well.

Regional Reliability Council rules would have to be changed to allow this type of operation. Obtaining this approval will likely require stability studies to fully characterize the engine plant performance and compare it with a conventional resource.

The motivation for the generation owner is to reduce costs and increase revenue by being able to sell additional spinning reserve (which always commands a higher price than non-spinning reserve).

3.2.2 Non-Spinning Reserve

Non-spinning reserve is similar to spinning reserve in that the generator must be fully responsive within ten minutes and it must be capable of sustaining the response for two hours. The generation does not have to be on-line and spinning, it does not have to begin responding immediately, and it does not have to be frequency responsive. The generator must have telecommunications (typically AGC) to respond to system operator deployment commands.

3.2.3 Supplemental or Replacement Reserve

Some regions specify a third contingency reserve called supplemental reserve or replacement reserve. This reserve must be fully deployed in 30 or 60 minutes, depending on the region, and must be capable of sustaining that response for two to four hours. On line generation, off line generation, and responsive load can provide this reserve. The reserve responds to system operator commands to deploy and to restore. The generator must have telecommunications (AGC is not necessary) to respond to system operator deployment commands.

3.2.4 Contingency Reserve Deployment Frequency and Duration

While the power system must have sufficient contingency reserves constantly standing by ready to immediately respond to the sudden failure of the largest generator or transmission facility actual deployment frequency depends on the actual failure rates. Systems also differ in how often they call on reserves or on what types of events justify reserve deployment. NYISO, for example, uses contingency reserves relatively frequently; 239 times in one year. ISO-NE and CAISO use contingency reserves much more sparingly; only 19 and 26 times respectively in 2005. The frequency of reserve deployment depends on the entire generation mix, the number of large units, and their failure rates. It also depends on the mix of responsive generators in the economic pool that are also available for system operator dispatch in the event of a smaller contingency.

In all three cases the response *duration* was typically short (~10 minutes) as shown in Figure 14. But it is critical that the reserves be *capable* of longer response in the event of a truly serious disturbance.

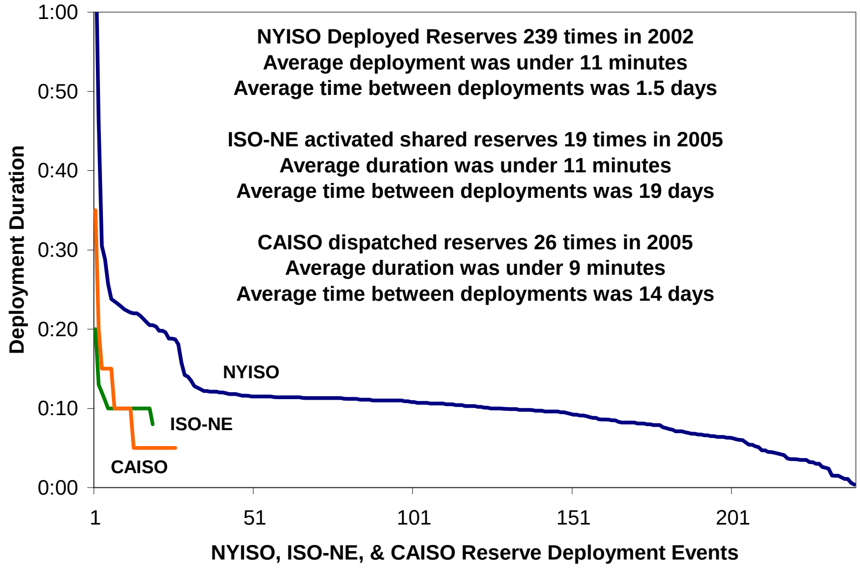

Figure 14 System operators differ in how often they deploy contingency reserves but events are usually relatively short.

3.2.5 Frequency Responsive Reserve

It has been recognized for some time that there is little technical justification for the specifics in the spinning reserve requirements. This is most apparent in the west where the Western Electricity Coordinating Council (WECC) requires maintaining contingency reserves equal to 7% of the load served by thermal generation and 5% of the load served by hydro with at least ½ being spinning. There is no technical basis for the 5%, the 7%, or the ½. The WECC Reserve Issues Task Force is attempting to address this lack by proposing a new Frequency Responsive Reserve (FRR) which does have technical justification. The proposed standard is currently working through the WECC approval process.

FRR would replace, and be very similar to, the current spinning reserve. It is designed to address generator response in the 30 second time frame. Future work will address 10-20 second response. All of WECC would be required to carry only 3200 MW of FRR, an amount equal to the largest credible (category C) double generator contingency that WECC plans to be able to withstand without under frequency load shedding (59.5 Hz). The total requirement would be allocated to balancing authorities based upon a peak load ratio share. Each balancing authority would still be required to carry enough contingency reserves (FRR plus non-spin) to cover the balancing authority's largest single contingency.

The amount of FRR that an individual generator would be credited with carrying depends on how much governor response the unit will provide. Generators are *required* to maintain a 5% governor droop which limits the generator to about 8% of its rated capacity as FRR even if it can ramp faster. This limitation makes little sense. Theoretically a 1000 MW generator could supply 80 MW but a 100 MW generator could only supply 8 MW, regardless of the two generators' maneuvering speed or accuracy. One reason for the limitation is to spread FRR among multiple generators. But that requirement could be better served with an absolute MW limit per generator. A limit that is tied to the generators energy rating makes little sense.

3.3 OTHER SERVICES

Generators supply two additional ancillary services to help maintain power system reliability; voltage control and black start.

3.3.1 Voltage Control

Unlike the other ancillary services listed in Table 1, voltage control is not a real-power service. Instead, it involves the control of reactive power to maintain acceptable voltages throughout the power system under normal and contingency conditions. Reactive power is measured in VARs (volt amps reactive) or MVARs (millions of volt amps reactive). The units are similar to watts and MW except that the voltage and current are out of phase. Power system voltage is sensitive to, and controlled by, the injection and withdrawal of reactive power. Voltages must be maintained within a fairly tight range throughout the power system to protect customer and utility equipment and to prevent voltage collapse. Too high voltage can destroy equipment by breaking down insulation. Too low voltage can make motors stall and equipment overheat. Voltage collapse can occur when a cascading drop in voltage suddenly spreads throughout a region. To protect against these failures and to compensate for the reactive power that loads and the transmission system itself consume the system operator must have reactive power resources available.[9]

Various pieces of equipment on the transmission and distribution system provide relatively inexpensive voltage control and reactive power. Capacitors, inductors, and transformer tap changes are all used as much as possible. But these transmission based solutions are slow to respond. Worse, the reactive support provided by capacitors drops with the square of the voltage so they provide less support when they are needed most.

---

[9] Greater detail can be found in: B. Kirby and E. Hirst 1997, *Ancillary-Service Details: Voltage Control*, ORNL/CON-453, Oak Ridge National Laboratory, Oak Ridge TN, December.

The power system requires a significant amount of dynamic reactive support to control voltage especially during contingencies. Synchronous generators are excellent suppliers of dynamic reactive power. They inherently produce more support during faults and they can be autonomously controlled to maintain local voltage to a coordinated schedule set by the system operator. Reactive support from generators is so important for reliability that the FERC Order 2003 (and supplemental Orders 2003-A, 2003-B, and 2003-C) covering Large Generator Interconnection Procedures (LGIP) and Large Generator Interconnection Agreements (LGIA) requires synchronous generators to be designed and operated such that they can supply reactive power within the range of 0.95 leading to 0.95 lagging while maintaining full rated real power output, "unless the Transmission Provider has established different requirements applicable to all Interconnection Customers." (FERC Order 2003, Article 9.6.1)

A synchronous generator's ability to produce reactive power is related to its ability to produce real power but not linearly, as shown in Figure 15. Reactive power costs and compensation will be discussed in Chapter 4 but it is important to point out here that the electrical generator is typically designed to be larger (MVA) than the prime mover (MW) specifically to provide reactive support capability while operating at full real power. The vertical dotted constant-power line in Figure 15 shows that this machine is capable of operating at a 0.85 power factor when the prime mover is producing full real power.[10] If more reactive power is required from this generator it is necessary to reduce the real power production. The decision concerning how much extra reactive power capability is built into the generator is made at design time and impacts the generators usefulness to support power system reliability for the rest of its life.

The need for dynamic reactive power changes from time to time and from location to location. Changes in transmission system loading change the reactive power consumption of the transmission system itself. As loads' real power requirements change the reactive power requirements change as well. Reactive power requirements are location specific because the inductive impedance of the transmission system is much greater than the resistance – VARs don't travel well.

Synchronous Condensers and Clutches

A generator must be on-line and therefore producing at least minimum power in order to be able to generate reactive power. A synchronous condenser is simply a generator without a prime mover. The synchronous machine operates simply to produce reactive power and control voltage. The machine consumes some real power in order to overcome friction and windage and the electrical losses in the windings (a 60 MVAR synchronous condenser requires about 1.5 MW to operate) but it consumes no fuel. An engine driven

[10] Power Factor is the ratio of real power (MW or watts) to apparent power (MVA or volts times amps). It can be expressed in per unit (0.85) or percent (85%).

generator could be equipped with a clutch allowing it to operate as a synchronous condenser when the system required dynamic voltage support but did not require real power.

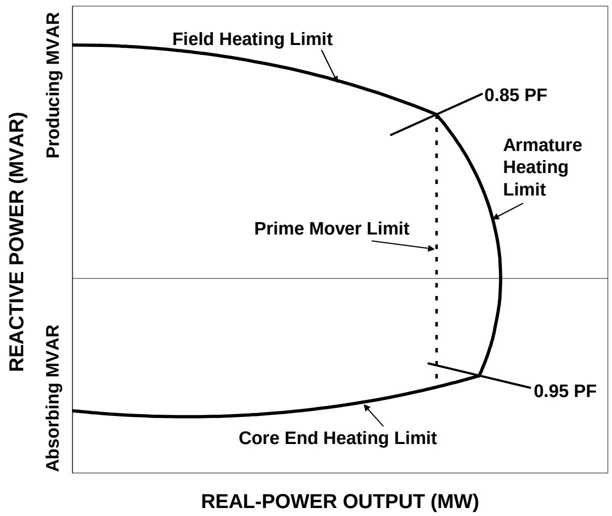

Figure 15 Real and reactive power production capabilities are interrelated in synchronous generators.

3.3.2 Black Start

Black start provides the generation resources necessary to restart the power system in the unfortunate event that a major blackout occurs. Black start generators must be capable of starting themselves quickly without an external electricity source. They must have sufficient real and reactive power capability to be able to energize transmission lines and restart other generators. They must have sufficient ramping and control capability to remain stable as real and reactive loads change. Typically black start generators are at least tens of MW in capacity. They must also have relatively low minimum load capability and a broad operating range. They must be appropriately located in the power system to be useful in restarting other generators and in resynchronizing the interconnection. They must be both able to control frequency and voltage and also be tolerant of off-nominal frequency and voltage. System frequency and voltage can fluctuate dramatically, especially in the early stages of system restoration. They must also have good communications with the system operations control center to facilitate a coordinated restart. Some regions require an on-site fuel supply.

4. ANCILLARY SERVICE PRICES

Hourly markets exist in several regions for up to five ancillary services: regulation, spinning reserve, non-spinning reserve, and replacement reserve. Regulation is always the most expensive service followed by spinning reserve, non-spinning reserve, and replacement reserve. Some markets split regulation into regulation up and regulation down. As discussed previously, this distinction is semantic rather than technical. Prices in split markets can be compared with combined markets simply by adding the up and down prices.

Cost drivers for each ancillary service will be discussed below in a bit more detail but all are driven primarily by opportunity cost. In order to sell into the ancillary service markets, generators must withhold capacity from the energy market. The cost the generator has to charge (or bid) to supply a reserve service is based primarily on the difference between the generator's production cost and the energy sale price for that hour. A generator with a production cost of $50/MWH, for example, would bid $10/MW-hr to sell spinning reserve if the energy price was $60/MWH. At any price higher than $10/MW-hr for spinning reserve the generator makes more profit by forgoing the energy sale and selling spinning reserve. Conversely, at any price below $10/MW-hr for spinning reserve the generator would loose money by staying out of the energy market.

One consequence of this linkage between energy and ancillary service markets is that ancillary service prices are inherently more volatile than energy prices. Contingency reserve prices, for example, are typically zero at night when numerous generators are at minimum load and have capacity available at essentially no cost.

Note that the price unit for reserves is $/MW-hr. This is because the generator is selling one MW of *capacity* (not energy) for one hour. The generator is standing ready to produce but it is not necessarily producing. In fact, if the generator does deliver any energy during the hour the cost of the energy will be settled separately, either at the generator's cost or at the spot energy price. Typically the energy component of the ancillary services is not major. This terminology is not universal but it does make the distinction between the energy and capacity components clear.

4.1 REGULATION COST DRIVERS

The direct costs for generators supplying regulation include a degraded heat rate and increased wear and tear on the unit. The dominant expense, however, is the lost opportunity cost associated with maneuvering the generator in the energy market so that it has capacity available to sell in the regulation market. For example, a 600-MW generator with a full power energy production cost of $15/MWh would have to bid $27/MW-hr of regulation if the energy market were clearing at $30/MWh. This is to compensate the generator for the lost profit in the energy market when it reduces output in order to create maneuvering room to supply regulation and to compensate for the reduced efficiency (increased heat rate) associated with the remaining output's still being sold into the energy market. Figure 16 shows how a generator's cost (and bid price) to supply regulation depends upon the current energy price. Note too that this generator is limited to supplying only about 12 MW of regulation (~2% of its rated capacity). This is because regulation is a quick service and the unit ramp rate, rather than the total available capacity, limits the peak amount of regulation it can provide. For this reason regulation is generally spread across several generators. Opportunity costs similarly dominate contingency reserve prices.

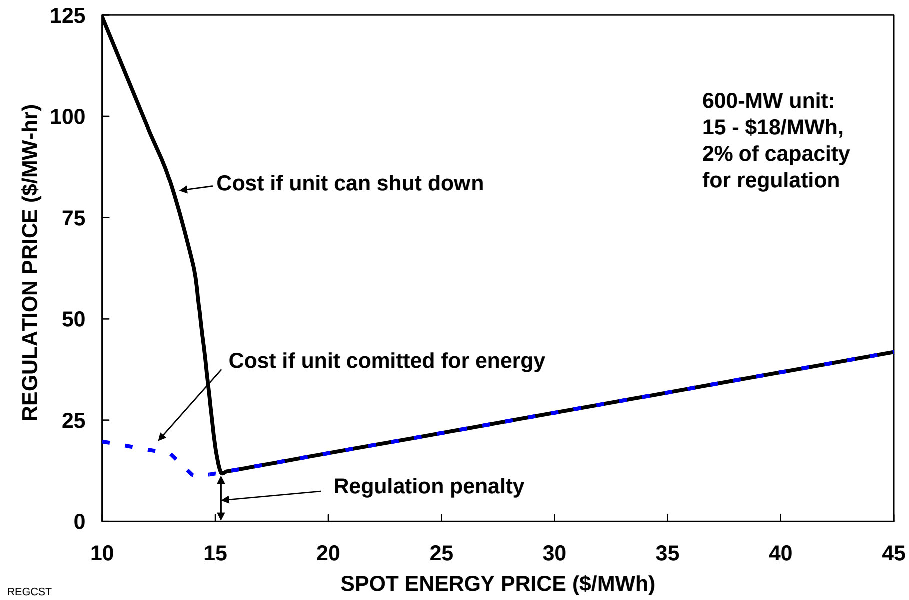

Figure 16 Regulation costs are dominated by generator opportunity costs. Cost at night can be higher than during the day.

There is also an opportunity cost when the energy market price is below the generator's marginal production cost. When energy prices are low (typically at night) and generators are at minimum load, they incur a cost for running above minimum load in order to supply down regulation. For example, a generator with a 150-MW minimum load and an energy production cost of $18/MWh would have to bid $64/MW-hr of regulation if the energy market were clearing at $14/MWh because it would be losing $4 for each of the 162 MWh it must sell into the energy market to get its base operating point high enough to provide room to regulate down.

4.2 CONTINGENCY RESERVE COST DRIVERS

Contingency reserve cost drivers are essentially a subset of the regulation cost drivers. Because contingency reserves deploy infrequently there is no significant degradation in heat rate and no increased wear-and-tear on the unit. Only the opportunity costs are incurred because the unit must withhold capacity from the energy market.

On occasion a generator can incur additional costs to provide contingency reserves. For example, a generator would have to bid a significant price to supply spinning reserve if the generator's energy price was higher than the market energy price and the generator could otherwise shut down. In this case the spinning reserve bid would have to cover all of the losses the generator was incurring in the energy market to operate at minimum load. Similarly, a generator could incur some costs to supply non-spinning and supplemental or replacement reserves if it was necessary to pay plant operators to standby.

4.3 COOPTIMIZATION

Provision of energy and the four reserve services are interrelated. It can be difficult to determine how much capacity to offer into each market and at what price. A generator will naturally want to maximize its profits and sell as much capacity as it can into whichever reserve market is paying the highest price. Alternatively, it will want to forgo the reserve markets and sell as much as possible into the energy market if that is providing higher profits.

The earliest market designs cleared the hourly energy and reserve markets in sequence. The system operator first selected the least expensive set of generators to supply energy. Generators to supply regulation were selected next, followed by spinning reserve, non-spinning reserve, and supplemental reserve. The thinking was that the energy market went first because it was the highest volume and the most important economically. Among the reserves, it was reasoned, it was necessary to procure the most technically demanding services first, followed by services that more generators can provide. This was done to prevent the supplemental reserve market, for example, selecting generation that was needed to supply regulation. This makes sense because a generator that can supply regulation can probably supply spinning reserve, non-spin, and replacement but the reverse is not necessarily true.[11] Unfortunately this can lead to perverse results in some cases. There may be sufficient low cost generation to supply the spinning reserve need, for example, but not enough to meet the need for supplemental reserves once that market is finally cleared. In that case the supplemental reserve market will clear at a higher price than the spinning reserve market. This result is undesirable because it provides an incentive for the technically more agile generators to withhold capacity from the spinning reserve market and offer it into the supplemental reserve market.

A solution was developed whereby the generators simply offer their capabilities and their costs. The system operator then cooptimizes the energy and ancillary service markets, guaranteeing each generator the maximum profit and the system the lowest combined cost. California's initial move in this direction was called the "Rational Buyer" which allowed the system operator to substitute "higher quality" reserves for "lower quality ones" but to pay the higher quality price. California is now going to a full cooptimization. ISO-NE, NYISO, PJM, ERCOT, and CAISO all perform cooptimization of some form. Midwest Independent System Operator (MISO) markets are still developing but it will likely cooptimize as well.

[11] This is actually not true in all cases. Emissions limited generation and responsive load, for example, may be able to supply spinning reserve but be unable to supply replacement reserve or energy due to the response duration. This is a relatively rare complication that is being addressed in some market structures.

4.4 REGULATION AND CONTINGENCY RESERVE MARKET PRICES

Hourly ancillary service market price data is available since September 2000 for California, since October 2001 for New York, and since April 2003 for ERCOT. Monthly averages of hourly prices are shown in Figure 17. Total regulation prices (regulation up plus regulation down) are shown for California and ERCOT to make them comparable to the New York regulation product which is a combined up and down service.

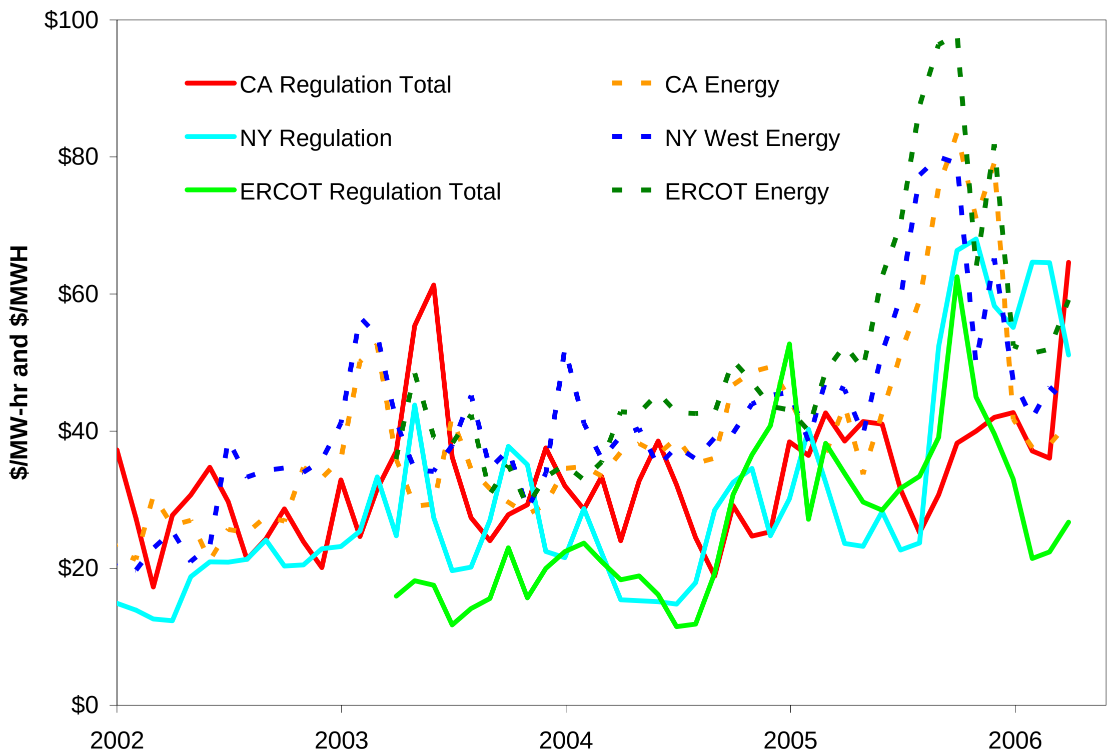

Figure 17 Monthly average regulation prices are typically (but not always) somewhat lower than energy prices.

Figure 17 shows that regulation prices, which include no fuel component, are in the same range as energy prices and are at times higher. Also, both energy and regulation prices are volatile, even on a monthly average basis. Regulation price also tends to track energy price, because of the lost opportunity cost.

Figure 18 compares all of the ancillary services on a monthly average price basis. Again, prices are volatile but over the 3.5 years regulation is always significantly more expensive than spinning reserve which is more expensive than non-spinning reserve and replacement reserve.

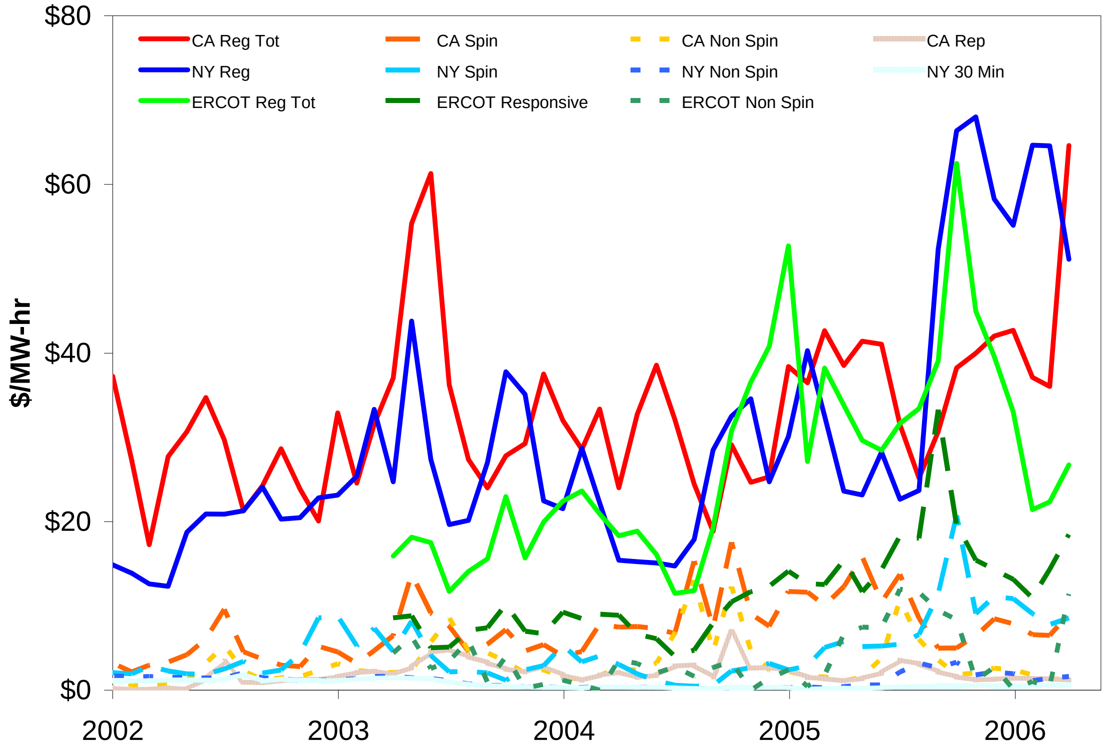

Figure 18 Ancillary service prices in all markets follow a pattern where regulation is most expensive and replacement reserve is least expensive.

Figure 19 expands the view of 2005 making it easier to compare the various services. Figure 20 provides an average daily view from June 2005. Here the typical daily price patterns can be seen. Contingency reserve prices are typically at or near zero overnight when there is significant generating capacity that is backed down. Conventional thermal plants that can not cycle off overnight drive the price of spinning reserve down. Fast start plants keep the price of non-spinning reserve and replacement reserves at zero overnight. The California total regulation price actually rises at night as the regulating units are forced above minimum in order to provide down regulation. Table 3 provides a numerical comparison of the average annual prices for each service in each region.

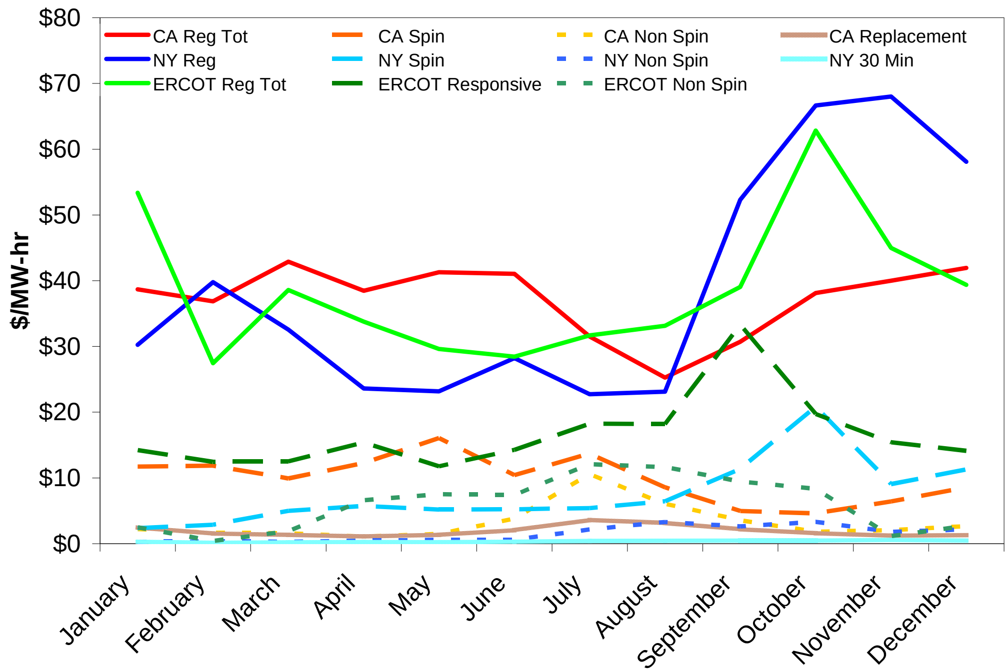

Figure 19 Regulation is always the most expensive ancillary service as shown by these 2005 monthly average ancillary service prices.

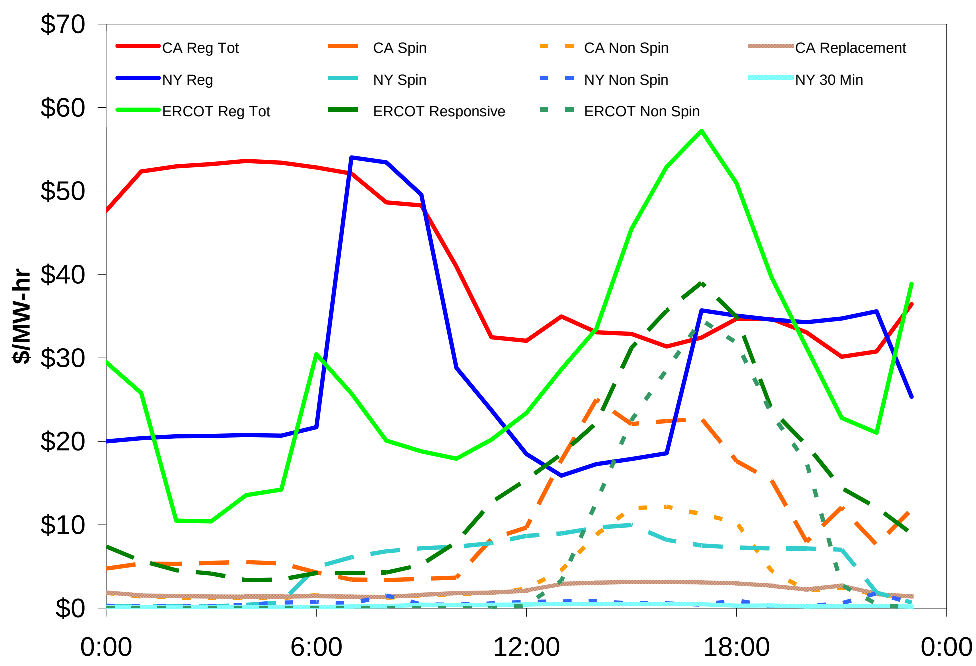

Figure 20 June 2005 average hourly ancillary service prices show a consistent pattern.

Table 3 Annual average and maximum ancillary service prices from four markets for five years.

	2002	2003	2004	2005	2006
	Annual **Average** and Maximum $/MW-hr				
			California		
Regulation	**26.9**	**35.5**	**28.7**	**35.2**	**38.5**
	111	164	166	188	399
Spin	**4.3**	**6.4**	**7.9**	**9.9**	**8.4**
	250	92	125	110	225
Non-Spin	**1.8**	**3.6**	**4.7**	**3.2**	**2.5**
	92	92	129	125	110
Replacement	**0.90**	**2.9**	**2.5**	**1.9**	**1.5**
	80	55	90	36	70
			ERCOT		
Regulation		**16.9**	**22.6**	**38.6**	**25.2**
		177	156	1451	351
Responsive		**7.3**	**8.3**	**16.6**	**14.6**
		150	51	731	351
Non-Spin		**3.2**	**1.9**	**6.1**	**4.2**
		249	400	510	125
			New York East		
Regulation	**18.6**	**28.3**	**22.6**	**39.6**	**55.7**
	99	195	99	250	250
Spin	**3.0**	**4.3**	**2.4**	**7.6**	**8.4**
	150	55	44	64	171
Non Spin	**1.5**	**1.0**	**0.3**	**1.5**	**2.3**
	45	3	3	64	171
30 Minute	**1.2**	**1.0**	**0.3**	**0.4**	**0.6**
	45	3	3	4	31
			New York West		
Regulation	**18.6**	**28.3**	**22.6**	**39.6**	**55.7**
	99	195	99	250	250
Spin	**2.8**	**4.2**	**2.4**	**4.9**	**6.0**
	150	55	44	50	45
Non Spin	**1.4**	**1.0**	**0.3**	**0.6**	**0.9**
	45	3	3	13	38
30 Minute	**1.2**	**1.0**	**0.3**	**0.4**	**0.6**
	45	3	3	4	31

4.5 REACTIVE POWER AND VOLTAGE SUPPORT COMPENSATION

While dynamic reactive power from generation is vital for power system reliability, market based reactive power compensation mechanisms are not yet well established or consistent. Markets are difficult to develop for reactive power because of the locational constraints. Because reactive power can not be moved over great distances there are usually too few generators within a given reactive power area to create a competitive market. There are a range of non-market alternatives within FERC's guidelines to provide compensation to generators for reactive power capability. One approach provides no compensation to any generator and simply requires response within the (typically) ±0.95 power factor range. Another approach compensates all generators based on their revenue requirements for reactive power capability as a capacity payment (the AEP Methodology). A third approach has the system operator develop locational prices for reactive power supply based on typical cost estimates. A fourth approach periodically develops location-based prices for dynamic reactive power supply based on generator and transmission system offers (a near-market solution). There are common features for essentially all compensation systems:

- Generators are required to provide reactive response within the power factor range (typically ±0.95) if the generator is operating.
- The system operator can not order a generator to operate simply to provide reactive support (though the system operator could ask for and pay for response)
- Generators are compensated for lost opportunities if a system operator requires an operating generator to reduce real power production in order to increase reactive power production.
- Cost-based Reliability-Must-Run (RMR) contracts are used in locations where a specific generator is required to run to maintain system reliability.

Generators incur capital and operating costs for supplying reactive power. The initial capital cost of the generating plant is higher because of the required larger generator, larger step-up transformer (GSU), and additional equipment (exciter, voltage regulator, etc.) that are required to generate and control reactive power. Losses in the generator stator, rotor, and step-up transformer are incurred as operating costs. Maintenance is required for the exciter and voltage regulator. The AEP Methodology is one favored by FERC for calculating the revenue requirements associated with reactive power production. It is based on the plant capability, either nameplate or maximum obtainable.

- Reactive cost of the generator/exciter
  - =(Generator + exciter cost) * ($MVAR^2/MVA^2$)
- Reactive cost of the generator step-up transformer
  - =(GSU cost) * ($MVAR^2/MVA^2$)
- Reactive cost of the accessory electric equipment
  - =(Accessory elec. equip. cost) * (Generator|exciter auxiliary load)/(Total plant auxiliary load)
- Reactive cost of the remainder of the plant
  - =(Cost of remainder of plant) * (exciter MW/Generator MW) * (Max MVAR/Nameplate MVAR)
- Total cost of reactive = Sum above components

Some ISOs, RTOs (regional transmission organizations) and TOs (transmission owners) (CAISO, for example) provide no compensation for the supply of reactive power within a designated range.[12] Reactive supply is required of all generators as a condition of interconnecting.[13] Others (MISO for example) do compensate generators [both affiliates of vertically integrated utilities and independent power producers (IPPs)] for providing reactive power within the designated range. The institutional arrangement provides compensation using a cost-based schedule set in advance, usually a payment equal to the generation owner's monthly revenue requirement. In exchange, the generators must be under the control of the control area operator and be operated as dispatched to produce or absorb reactive power. When there is a reduction in real power output due to a request for reactive power production, the RTO will provide an additional payment to compensate the generator for the lost opportunity of delivering real power into the network. Cost-based compensation to generators for providing reactive power supply is regulated by FERC, and all ISOs/RTOs must provide a Schedule 2 tariff for Reactive Supply and Voltage Control as part of their Open Access Transmission Tariff (OATT). Examples of compensation arrangements from a number of ISOs and RTOs are presented in Table 4.

Table 4 Regional Comparison of ISO/RTO Arrangements for Reactive Power Compensation

Region	Method of Compensating Generators for Reactive Power Supply	Provisions for Testing/Confirming Reactive Power Capability of Generators & Other Facilities	Required Power Factor Capability Range for Generators (leading/lagging)	Approximate Annual Payment to Generator
PJM	Payment equal to revenue requirement approved by FERC	Capability test every 5 years	0.95/0.90	$2,430/MVar
NYISO	Capacity	Capability test once a year	0.95/0.90	$3,919/MVar
CAISO	No compensation for operating within power factor range	Tests are not normally run unless ISO detects a problem	0.95/0.90	None
ISO-NE	Capacity	Capability test every 5 years	0.95/0.90	$1050/MVar
SPP	Pass through of revenues collected by control area operators	Control area operators negotiate with generators	Not available	Not available
MISO	Payment equal to revenue requirement approved by FERC	Control area operators negotiate with generators	0.95/0.95	Generator revenues are aggregated by pricing zone
ERCOT	No capacity payment	Capability test every 2 years	0.95/0.95	Paid the avoided cost of DVAR or equivalent equipment

Reactive power compensation schemes are far from mature. They will typically not provide appropriate automatic compensation for additions like clutches and generator operation as a synchronous condenser. Generators with unique and useful capabilities should always contact system operators to negotiate appropriate compensation.

[12] CAISO does provide cost-based compensation to RMR generators for reactive support.

[13] There can be exceptions for generators that are not physically or contractually capable of providing reactive support. Wind generators get a unique "needs" based exception from FERC.

4.6 BLACK START COMPENSATION

Compensation for black start capability is less transparent than compensation for reliability reserves. System planners reevaluate black start plans annually and re-work them every few years. They typically negotiate multi-year contracts with black start capable generators as the plans are updated. As markets for energy and ancillary services mature, market designers are moving to more public processes for obtaining black start. Ultimately open markets may develop in some regions. In the interim, ISOs are starting to publicly call for black start resource proposals and then selecting from the offers received. The selection process is only partially based on the offer price. Location and resource capability are also important.

PJM and NYISO utilize unit-specific cost based compensation methods to procure black start capability. Costs include the generators' black start related capital costs, operating costs, and training costs. [14]

ISO-NE has replaced a unit-specific cost-of-service compensation method with a fixed payment for all black start resources. A stakeholder process resulted in $4.50/kW-yr compensation in 2006 rising to $4.58/kW-yr in 2007 through 2011.

ERCOT procures black start competitively through an annual call for bids. Generators are paid an hourly standby fee that is adjusted for availability.

CAISO currently only compensates black start capability if it is in conjunction with a RMR (reliability must run) generator. Other black start generators are not compensated. CAISO is evaluating alternatives for competitively procuring black start capability.

[14] A. G. Isemonger, 2006, *The Competitive Procurement of Black Start*, CAISO

5. SELLING MULTIPLE SERVICES

The optimal selection of which energy and ancillary service products a generating plant should sell changes from hour to hour as the cost of fuel, and the energy and ancillary services market prices vary. An optimization model was developed to examine how an engine driven generating plant might best respond to changing price signals in four locations: southern California south of path 15, Texas, western New York, and Long Island. These locations were selected because of the availability of hourly electricity and ancillary services price data.[15]

5.1 ANALYSIS METHODOLOGY

A time series model was developed which simulates the operation of a gas fired engine driven generating plant operating within an energy and ancillary services market structure. The model first determines how the plant would respond to the energy markets alone. The plant hourly electricity production cost is compared with the hourly electric market price to determine if electricity sales are profitable each hour. Daily delivered gas prices were used for California, monthly gas prices were used for ERCOT and New York.

Modeling then examined the profit options in each of the ancillary services markets; regulation, spin, and non-spin. The highest profit combination was selected for each hour. Summary results and plots were generated. Interesting statistics include total profit, profit from the sale of each service, plant capacity factor, and run hours. The profitable operating modes are interesting as well. The first three modes are obvious:

- At times it is most profitable to simply sell energy.
- Regulation and spinning reserve can be sold any time a generating unit is on line and partially loaded. While this is an important operating mode to consider in the real world it does not happen in the simplified modeling results because if the production cost is below the sale price for energy the model fully dispatches the plant. If the production cost is above the sale price for energy the model turns the plant off.
- Non-spinning reserve can always be sold when the plant is off line.

Four more operating modes are less obvious but contribute significantly to the modeled plant profits:

- The plant can forgo profitable energy sales in order to sell regulation reserve. This is particularly attractive when the price of energy is only somewhat greater than the production cost. The energy profit per MW of capacity is often lower than the almost pure profit offered from the regulation market.
- Similarly, if the price of regulation is less than twice the price of spinning reserve and if energy profit is marginal, it can be more profitable to sell spinning reserve

[15] California day-ahead, rather than hour-ahead, energy and ancillary service markets were modeled because these resulted in higher total profits for 2005.

than regulation or energy. This is because the plant has twice as much spinning reserve capability as regulation capability (the plant must be able to move up and down when selling regulation). The minimum operating load is lower when selling spinning reserve (45 MW) than when selling regulation (79 MW).

- Conversely, it can be profitable to run the plant at a minimal load even when the energy market does not cover the fuel cost if the regulation market price is high enough.
- As above, it may be more profitable to sell lower priced spinning reserve than higher priced regulation because the minimum plant load required to enable the supply for spinning reserve is lower than the minimum plant load required to sell regulation. The energy market losses are therefore lower when supplying spinning reserve than when supplying regulation.

Lastly, additional profit opportunities would become available if the plant could supply spinning reserve while operating the generators as synchronous condensers.

This modeling is designed to be illustrative rather than definitive. More detailed modeling should be undertaken before serious investment decisions are made. The results are interesting, however, and point to opportunities to exploit.

5.2 ASSUMPTIONS

A 100 MW natural gas fired engine driven generating plant was assumed to be a price taker in the gas, electricity, and ancillary service markets. All natural gas, energy, and ancillary services prices were taken from 2005. Monthly gas prices were taken from Enerfax for ERCOT and New York. Daily delivered gas prices were available for California. Gas prices and are shown in Table 5.

Table 5 Delivered $/mmBTU gas prices for California, Enerfax prices for Texas and New York.

	California SP15: Southern California	ERCOT: Henry Hub	Western NY: Niagara	Long Island: NYC
January	$6.53	$6.88	$7.05	$13.01
February	$6.22	$6.47	$6.77	$7.23
March	$7.21	$7.49	$7.86	$8.25
April	$7.54	$7.91	$8.24	$8.45
May	$6.71	$7.19	$7.47	$7.64
June	$7.04	$7.90	$8.01	$8.45
July	$7.57	$8.31	$8.32	$8.93
August	$8.99	$10.00	$9.95	$10.82
September	$10.92	$13.17	$12.64	$14.48
October	$11.76	$14.17	$14.27	$15.33
November	$8.70	$11.02	$10.26	$11.18
December	$12.16	$13.83	$13.79	$14.94

The plant's minimum load with all engines running is assumed to be 40% or 40 MW. The plant is assumed to be able to ramp with enough speed and accuracy that ramp rate does not limit the supply of regulation or spinning reserve. This results in a plant capable of selling 30 MW of regulation (30 MW of up and 30 MW of down regulation in California and ERCOT – but all results are shown here as combined up and down regulation to simplify comparisons) and 60 MW of spinning reserve. The plant is assumed to be able to start and fully load in under ten minutes such that the plant can sell the full 100 MW of non-spinning reserve.

Startup and shutdown costs are not included in the model. This slightly overstates the plant's flexibility.

5.3 MODELING RESULTS

Table 6 summarizes the modeling results. On a most general level, selling ancillary services increase plant profits by $1.5 to $5.1 million per year or 17% to 250%. Operating hours increased by 41% to 131%. What the modeling shows about the interactions between the gas, energy, and ancillary services markets is probably of the greatest interest.

In California, ERCOT, and western New York the average gas price resulted in an average production cost that was higher than the average electricity price. Only on Long Island was the average all-hours price of electricity slightly higher than the production cost. As a consequence, the energy-only profits on Long Island were significantly higher than in the other areas; $16 million as opposed to $2 to $6 million. The run hours and capacity factors were similarly skewed with the Long Island plant operating two to three times as many hours as the plants in the other regions when selling only energy.

When the plants move into the ancillary service markets they reduce their energy profits by $1.6 to $2.4 million. Total run hours increase because the plants operate some of the time at partial load so that they can sell regulation and spinning reserve. Total capacity factor rises because additional hours of non-economic energy production are also added to enable further sales of regulation and spinning reserve. Sales of regulation provide $1.7 to $4.4 million in additional profit while sales of spinning reserve add $0.05 to $1.5 million. Non-spin sales add $0.4 to $3.9 million.

Clearly sales of ancillary services can add significantly to plant profitability. Clearly too, sales of energy and the ancillary services need to be optimized to maximize plant profits. A strategy which simply sells ancillary services from the excess capacity still available after all profitable energy sales are made misses significant opportunities.

Table 6 Selling ancillary services increases profits by 17% to 250% in this 2005 modeling simulation.

	CA SP15	**ERCOT**	**NY West**	**Long Island**
All Hours Production Cost $/MWH	$82.37	$91.66	$91.89	$102.15
All Hours Energy Price $/MWH	$55.08	$66.30	$62.92	$103.88
Energy Only				
Operating Hours	836	1,271	760	2,648
Capacity Factor	10%	15%	9%	30%
Energy Profit	***$1,554,898***	***$6,269,977***	***$2,861,163***	***$16,365,702***
Energy & Ancillary Services				
Total Operating Hours	1908	1787	1753	4205
Energy Only Operating Hours	234	554	327	1491
Capacity Factor	14%	20%	20%	39%
Energy Profit	*$402*	*$4,139,928*	*$1,091,895*	*$14,011,621*
Regulation				
All Hours Regulation Price $/MW-Hr	$37.18	$38.55	$39.17	$39.17
Regulation Hours	1170	844	1407	2683
Regulation Profit	*$1,652,297*	*$2,074,235*	*$2,794,540*	*$4,355,417*
Spinning Reserve				
All Hours Spin Price $/MW-Hr	$9.92	$16.63	$4.92	$7.61
Spin Hours	504	389	19	31
Spin Profit	*$1,464,737*	*$1,250,498*	*$45,502*	*$82,280*
Non-Spinning Reserve				
All Hours Non-Spin Price $/MW-Hr	$3.24	$6.05	$0.59	$1.51
Non-Spin Profit	*$2,321,954*	*$3,875,014*	*$395,419*	*$652,713*
Total Profit	***$5,439,391***	***$11,339,675***	***$4,327,356***	***$19,102,031***
Additional Profit	***$3,884,493***	***$5,069,698***	***$1,466,193***	***$2,736,329***
Increased Profit	***250%***	***81%***	***51%***	***17%***

Unlike base load nuclear or coal fired plants, gas fired engine driven generators operating response to energy market prices is fairly volatile. Adding ancillary service response increases that volatility. Figure 21 presents a year long view of hourly optimized energy and ancillary service production for plants located in each of the four studied regions. The plants make so many transitions between not selling and selling services that the plots often look like solid bars. This shows the value of the technology's agility. Figure 22 shows a single week of response in May 2005 for the California plant. A plant with significant transition costs or significant minimum run and rest times would be unable to respond fast enough to profit from high hourly prices and avoid low hourly prices. This benefits power system reliability as well since the price signals reflect reliability needs.

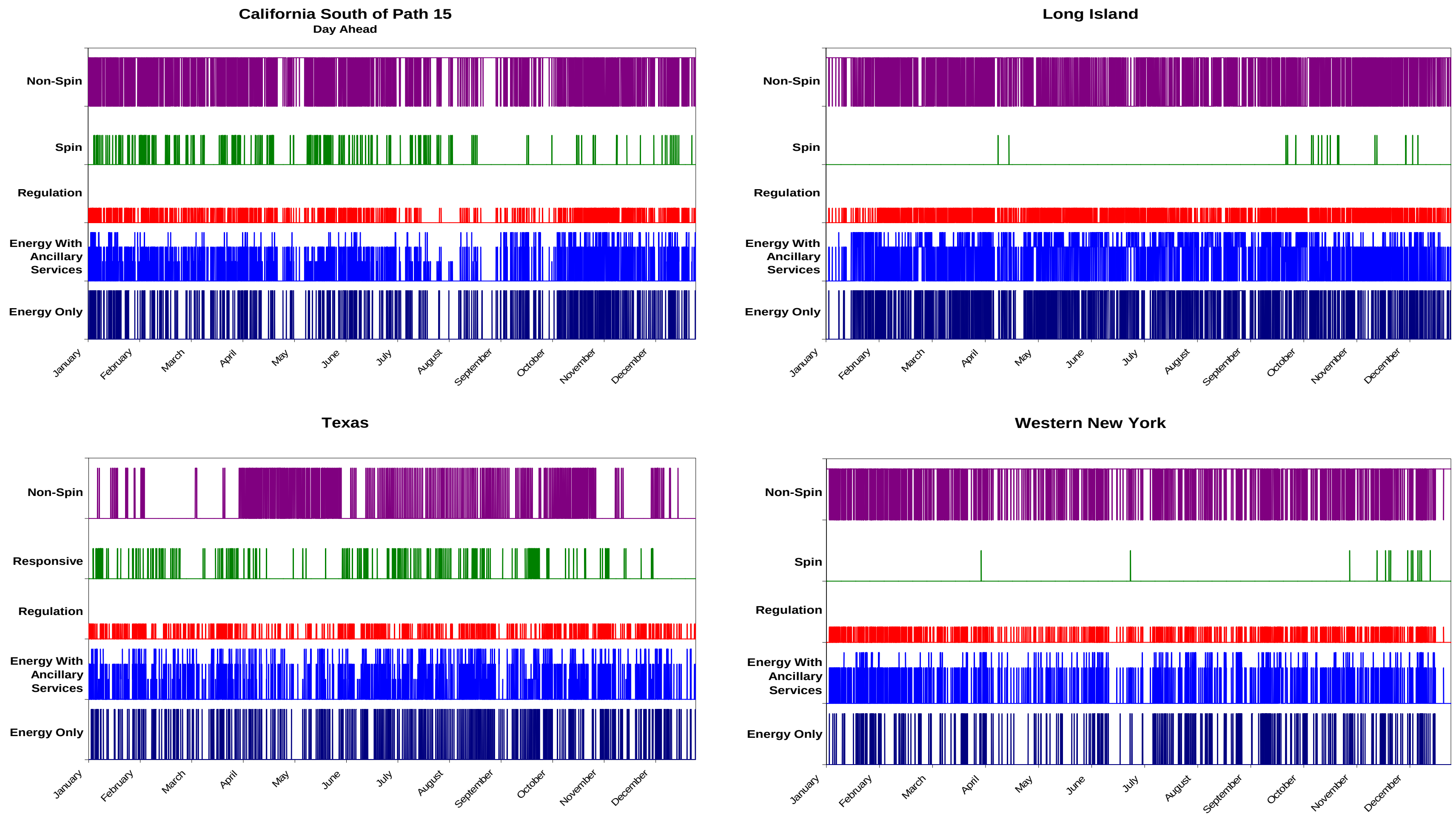

Figure 21 The optimal production of energy and ancillary services varies from hour to hour and is different in different regions as shown in this 2005 simulation.

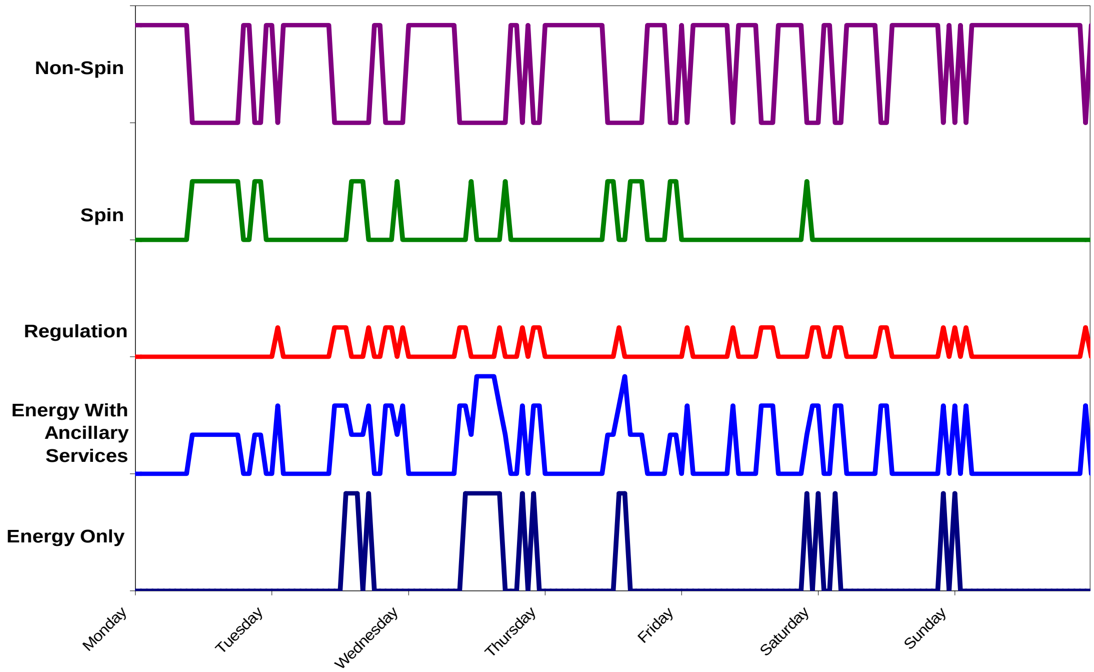

Figure 22 One week of response in May 2005 for the California plant shows a full range of operating modes.

Note that the plant in Figure 22 sold its full output into the energy market for only four hours starting at noon on Wednesday and again for a single hour Thursday afternoon. The rest of the time other arrangements were more profitable.

On Monday at noon, for example, it was most profitable to sell spinning reserve. The production cost was $65.33/MWH and energy was selling for $55.92/MWH. Regulation was selling for $76.68/MW-hr (up + down) and the plant could have sold 30 MW of regulation for $2,300. There would have been a $853 loss for selling 70 MW of energy resulting in a net $1,448 profit. It was better, however, to generate 40 MW and loose $909 in the energy market while selling 60 MW of spin at $65.45/MW-hr for $3,927 resulting in a $3,018 profit for the hour.

On Tuesday at 15:00 the situation was a bit different but the result was the same. The energy price was $69.40/MWH so an energy sale would have been profitable ($190 profit). Regulation was selling for $43.91/MW-hr so that would have been profitable as well (-$68 from energy due to the lower efficiency at lower load but $1,317 more from regulation for a net $1,249 profit). Spinning reserve was still the most profitable even though its price was lower than regulation at $42.33/MW-hr (-$477 from energy due to the even lower plant efficiency but $2,540 from spinning reserve for a net $2,063 profit).

More commonly the higher regulation price resulted in the better option as on Tuesday at 13:00. Energy was selling for $73.91/MWH and selling energy would yield a $640 hourly profit. Regulation was selling for $39.53/MW-hr so backing the plant down and loosing $393 in the energy market was more than made up for by gaining $1,186 in the regulation market for a net profit of $1,433 – more than doubling the hourly plant profit. Spinning reserve was only selling for $5.77/MW-hr and was not attractive to sell that hour, though it was the next hour.

Non-spinning reserve is typically sold whenever the plant is idle and the reserve price is above zero.

Clearly the physical plant will not achieve the same results as the modeled plant. Some operations and maintenance costs likely increase with energy and ancillary service production. Plant efficiency likely declines when providing regulation. Increasing plant maneuvering may increase maintenance costs. Nonetheless, optimizing the provision of ancillary services and energy has the potential to greatly increase profits.

Spinning Reserve From Fast-Start Clutch-Coupled Generators

If fast-start clutch-coupled engine driven generators are allowed to supply spinning reserve while the engine is turned off there is the potential for greater profit. Modeling is complicated because not all of the restrictions are known. Is it necessary for two out of ten engines to be operating in a multi-engine plant? One? Four? Determining this will require both stability studies and negotiations.

One modeling run was made for each plant location to begin to get a feel for the range of increased profit. It was assumed that no engines needed to be running. It was also assumed that spinning the generators without the engines operating required the purchase of 4 MW (for a 100 MW rated plant). The energy purchase requirement means that at times the spinning reserve revenue would not cover the cost of purchased electricity and the plant would decline the sale. It might sell non-spinning reserve if that price were greater than zero. The model did not optimize the sale of spinning reserve from clutch-coupled generators against the sale of any other ancillary service or energy. This results in slightly conservative results but the assumption that no engines need to be running is optimistic. The model shows that profits could be increased by up to $1.7 million in California, $3.9 million in Texas, $1.7 million in western New York, and $1.1 million on Long Island. If this potential is attractive then efforts should be initiated to discuss the concept of supplying spinning reserve from clutch-connected fast-start generators with regional reliability councils. ERCOT might be the best region to start with because potential profits are greatest, because ERCOT has greater autonomy in setting reliability and market rules, and because ERCOT already obtains about half its spinning reserve requirements from responsive load.

6. THE VALUE OF FLEXIBILITY

Chapter 5 demonstrated that selling ancillary services in addition to selling energy can improve the profitability of a power plant under the correct conditions. Providing ancillary services to the power system is not without cost, however. The power plant must have sufficient flexibility to quickly and accurately change output when called upon by the system operator. The plant must be able to provide this flexibility without significantly increasing its production cost or the potential added profits will be lost. Flexibility has value.

To further quantify the value of flexibility a sophisticated production cost model was used to compare the potential response of two different generating technologies in a single example situation; reciprocating engines and combustion turbine. A year-long hourly time-series model was run that optimizes plant participation in energy and ancillary service markets, similar to the model discussed in chapter 5. The model is more detailed than the one discussed in chapter 5 in that it includes plant site altitude, natural gas pressure, gas quality, NOx requirements, water use, ammonia expenditure, and lube oil consumption. In addition to the hourly energy and ancillary service prices the model accounts for site hourly temperature, humidity, and barometric pressure. Startup and shutdown costs are included each time the plant cycles on and off.

6.1 COMPARING TWO POWER PLANTS

The detailed model was used to compare the performance of two power plants to assess the value of flexibility. One plant consists of 12 engine driven generators rated at 101 MW (ISO conditions, 59 F). The other plant consists of two combustion turbines rated at 98 MW. Hourly data (2005 energy, ancillary service, meteorological, etc) was available for a plant site in southern California. Both sea level and 5000 ft. elevation conditions were modeled.

Power plant characteristics at 5000 ft. elevation are compared in Figure 23. Combustion turbine plant capacity is reduced by 17 MW to 79 MW due to the elevation while the engine plant capacity is not significantly impacted. The combustion turbines' capacity and efficiency are also more adversely affected by higher ambient temperatures than are the engines'. These factors impact straight energy production economics as well as ancillary service sales. Part load efficiency is more directly related to provision of ancillary services, especially regulation. Both plants' efficiency is reduced at lower loads but the effect on the combustion turbine plant is greater.

Model results are shown in Table 7. The modeled engine plant operated more hours and generated more profit then the combustion turbine plant when they only sold energy; 815 hours and ~$1.5 million for the engine plant versus ~400 hours and $0.6 to $0.8 million for the combustion turbine plant, depending on elevation. The difference was more pronounced when ancillary service sales were considered. Both plants lost money in the energy market ($424,000 and $428,000 to $513,000 respectively) in order to enable sales of ancillary services. The engine plant increased its total profit to ~$5.7 million by selling ancillary services while the combustion turbine plant increased its total profit to $3.3/$4.4 million. This represents increases of $4.3 and $2.7 to $3.3 million respectively.

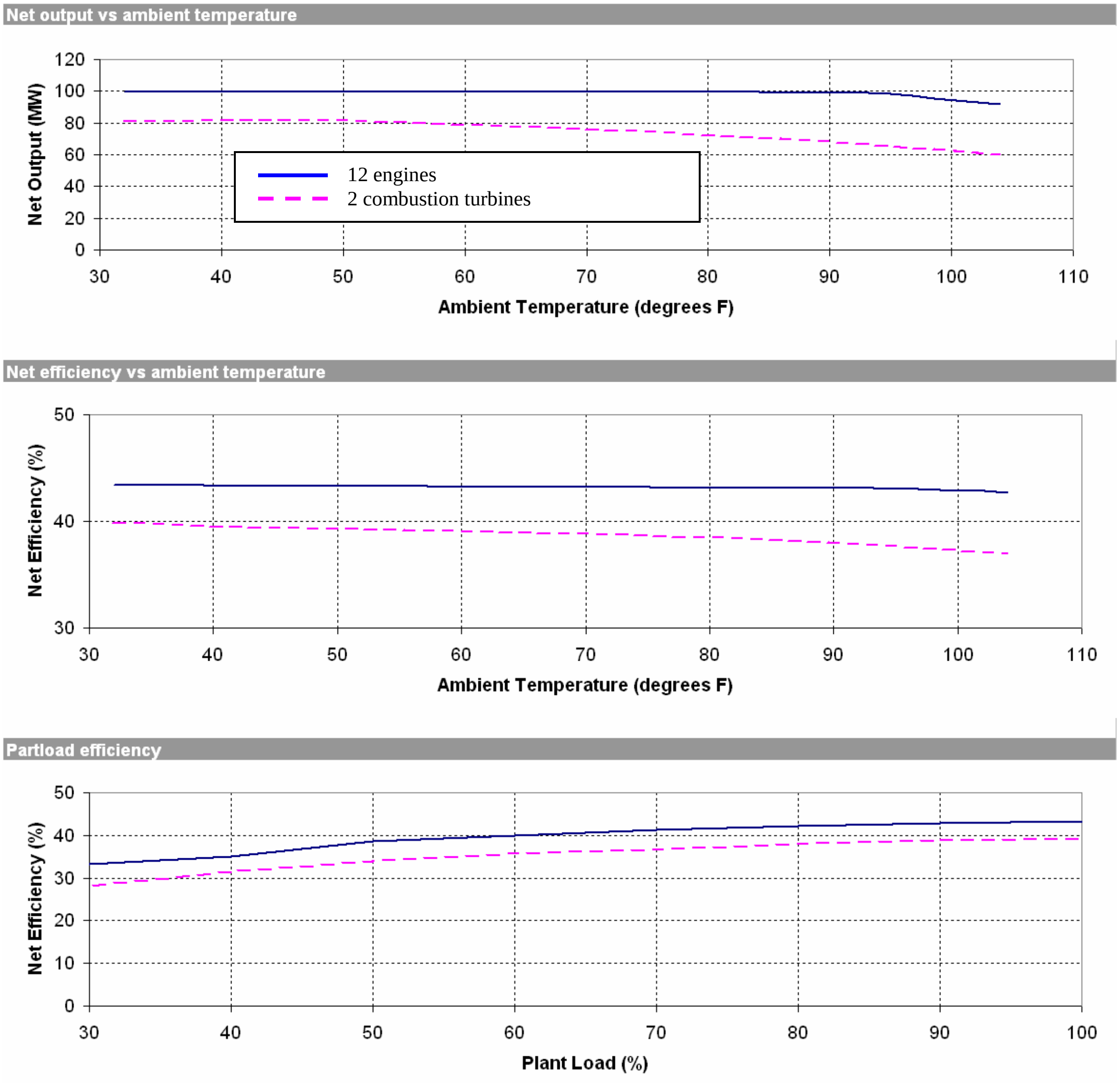

Figure 23 This engine driven generating plant has higher efficiency at all loads and lower efficiency degredation with increased ambient temperature than the combustion turbine driven generating plant.

The value of flexibility is clear when profits from the sales of the various ancillary services are compared. The engine plant received 2.0 to 2.5 times as much profit from the sale of regulation as the combustion turbine plant and 1.4 to 1.7 times as much profit from the sale of spinning reserve. The engine plant also received 1.0 to 1.2 times as much profit from the sale of non-spinning reserve as the combustion turbine plant. The greater flexibility enabled the sale of the faster, higher priced services.

Table 7 The greater flexibility of the engine driven generating plant results in greater ancillary service profits when compared with the combustion turbine driven generating plant.

Annual Profits	**12 Engines – Sea Level or 5000 ft**	**2 Combustion Turbines – Sea Level**	**2 Combustion Turbines – 5000 ft**
Plant Capacity	100 MW	96 MW	79 MW
Energy Only			
Operating hrs	815	408	402
Profit	$1,498,000	$797,000	$639,000
Energy and Ancillary Services			
Operating Hours	2,222	1,244	1,231
Energy Profit	-$424,000	-$513,000	-$428,000
Regulation	$1,752,000	$858,000	$696,000
Spinning Reserve	$2,309,000	$1,663,000	$1,359,000
Non-spinning reserve	$2,141,000	$2,144,000	$1,763,000
Total*	$5,754,000	$4,085,000	$3,324,000
Additional Profit	**$4,256,000**	**$3,288,000**	**$2,685,000**

* Includes startup costs for cycling

7. CONCLUSIONS

Restructuring is resulting in the definition of ancillary services that support power system reliability. Seven ancillary services are potentially of commercial importance to generators. Four of those services are currently traded in hourly markets in a growing number of regions. Modeling shows that generating plants can increase their profits significantly if they optimize their production of energy and ancillary services. Modeling also shows that increased operating flexibility is commercially rewarded in ancillary service markets.

Market and reliability rules are important and can greatly influence short term profitability. Clearly any physical plant must pay strict attention to both and respond in whatever way the rules currently reward. Market and reliability rules are evolving so elicited behavior may change. Underlying fundamental power system characteristics are stable, however. Generator maneuverability is valuable for maintaining reliability. Ancillary service markets are likely to continue to mature and expand, continuing to value response that increases reliability.

Ancillary service metrics are a specific area where improvements are needed. Generator regulation performance should be quantified and paid for. More accurate response will result in the need for less regulation and savings for customers even if the rapid/accurate regulating generators are paid more.