Valuation of Reserve Services

Pablo A. Ruiz — paruiz@ieee.org
Peter W. Sauer — psauer@uiuc.edu
Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign

Abstract

Spinning reserve is idle capacity connected to the system with the purpose of balancing power demand and supply in real-time and ensuring reliable system operations in the case of equipment outages. The reserve has an economic value since it reduces the load shedding costs. This value is impacted by the reliability and dynamic characteristics of system components, the variability of the load from its forecast, the system operation policies, and economic aspects such as the risk preferences of the demand. In this paper, we take into account all these aspects to estimate the reserve value. The results require little computational effort, and are useful in the development of reserve demand functions and optimal reserve requirements.

1. Introduction

Operational reliability is provided in part by committing units with total capacity in excess of the forecasted load. This excess capacity, called the *operating reserve* of the system, has the purpose of providing for equipment outages and unexpected load variations. The first objective is attained with the *contingency reserve* while the second is attained with the *regulating reserve*. In some jurisdictions there is a distinction between regulating-up and regulating-down reserve. Regulating-up reserve covers loads increments, and regulating-down reserve covers load decrements. Contingency reserve can be classified into *spinning* and *non-spinning* reserve. The spinning reserve is connected to the system while the non-spinning reserve is provided by fast-starting units which are not connected to the system or by interruptible loads. In this paper, the focus is on spinning reserve.

In the vertically integrated utility structure, the regulatory bodies set minimum required levels of service reliability and the utilities plan and operate their systems to meet the requirements. Reliability requirements may be expressed in terms of reliability metrics, such as the *loss of load probability*, or in terms of minimum quantities of reserve, e.g., the reserve has to be larger than a certain percentage of the forecasted peak load [1]. In the former case, the reliability requirements can be translated into reserve requirements [2]. Under this framework, extensive research has been done to provide a least-cost operation of the system subject to reserve constraints [1].

Competitive electricity markets have been introduced in several jurisdictions to increase the economic efficiency of the industry. Demand-side participation, either directly or by proxy, plays a significant role in the efficient operation of markets and systems. Thus, several independent system operators (ISOs) have implemented *operating reserve demand functions* in their markets [3], [4]. The implementations have the objective of explicitly accounting for the reserve value to the demand in the market clearing process. However, reserve valuation has not received direct attention in the literature. The demand functions implemented are often reflections of reliability requirements which may not fully represent the reserve value.

Spinning reserve is needed to keep the system frequency within specified limits. Large frequency deviations cause vibrations in devices such as steam and gas turbines. These vibrations greatly reduce the device life. Thus, the generating units have under- and overfrequency protection relays that shut the unit down if the frequency deviation is larger than a pre-defined threshhold. Reaching such large deviations may lead to blackouts. Hence, systems operators work to maintain the frequency close to its nominal values. If the frequency decreases below a threshhold, load is shed; if frequency increases above a threshold, generation is promptly reduced.

The economic value of reserve is derived from its ability to decrease the costs of load shedding, either by a reduction in the likelihood, duration and/or magnitude of load shedding. The factors that impact load shedding are the reliability of the system components, the variability of the load from its forecast, the dynamic characteristics of loads and generators, and the system operation policies, such as load shedding policies. These factors have to be taken into account to properly value the reserve.

Several papers, whose main focus is unit commitment, economic dispatch or optimal power flow, consider the value of contingency reserve. Usually, dynamic and operation aspects that may impact the reserve value, such as underfrequency load shedding schemes, are neglected or ignored ([5], [6], [7], [8], [9]). Some of these aspects are considered in [10], [11], at the expense of a substantial increase in the computation effort. Although [5]-[11] do not compute the value of reserve explicitly, their formulations can be used to obtain it by computing the difference between the costs without and with reserve. In contrast to contingency reserve, little work has been documented on the value of regulating reserve and on the determination of regulating reserve requirements. Typically, these requirements are "based upon empirical experience and engineering judgement" [12].

Reserve can be seen as a form of insurance from the demand's point of view. Firms buy insurance to decrease their exposure to *risk*, where risk can be defined as "potential variation in outcomes" [13]. The objective of the insurance is to exchange a risky cash flow by one which is less risky. The insurance buyer pays a *premium* for the reduction in risk exposure. The premium a firm is willing to pay depends on the firm's risk preferences [13]. As the firm's risk aversion increases, the value of the insurance to the firm increases, and so the firm is willing to pay higher premiums. An analogy can be made between insurance and reserve. The risky cash flow is analogous to the (actual) cost of load shed if the system is operated without reserve. The less risky cash flow is analogous to the cost of load shed with reserve. The insurance premium would be the amount paid for the reserve. In the same manner as the premium a firm is willing to pay depends on the firm's risk preferences, the amount willing to be paid for reserve depends on the aggregate demand's risk preferences and the change in risk exposure.

In this paper, we study the economic value of contingency and regulating-up reserve, defined as the reduction in the outage costs provided by the reserve. All the aspects mentioned in the previous paragraphs are explicitly considered using simple models. The estimated value of contingency reserve is obtained in closed form, whereas the regulating reserve value is obtained as definite integrals, which are solved numerically. The computational effort required is limited, and the results are useful for the construction of reserve demand functions in competitive markets and the determination of optimal reserve requirements. Illustrative numerical examples are provided.

The remainder of the paper has five sections. The next Section presents the general models used. Section 3 gives expresions for the value of reserve and their use in the construction of demand functions and the determination of reserve requirements. Sections 4 and 5 discuss the value of contingency reserve and regulating-up reserve, respectively. Section 6 concludes with the summary and the directions for future work.

2. Mathematical models

The time frame $[0, T]$ considered is between several minutes up to a few hours, where T is the *lead time*.

2.1. Power System Aspects

Transmission constraints are usually of a thermal nature, which allow lines to be overloaded for a few minutes after a supply shortage to ease the recovery of the system. Moreover, the units are dispatched to serve the forecasted load taking into account all network constraints. Thus, we neglect the network aspects. The system considered can be thought of as a single-area system with ample transmission capability to accommodate the desired transactions.

The aggregated load is modeled as the sum of the forecast $l(t)$ and an uncertain component $L(t)$. The deviation $L(t)$ may be positive or negative, but the total load $l(t) + L(t)$ is always positive (usually, $l(t)$ is two orders of magnitude larger than $L(t)$). This model is illustrated in Fig. 1, which gives $l + L$, l and L as a function of time for entire New York ISO (NYISO) system on March 1^{st}, 2007. The *energy demand* up-to time t due to L is

$$D(t) = \int_0^t L(\tau) d\tau. \tag{1}$$

The particular models used for L are introduced in the subsequent sections. For simplicity, the load is modeled as frequency-independent.

There are $\overline{n}$ spinning generators in the system. Each generator is subject to failures, and can be fully functional (on, state 1) or off line (off, state 0). Since the mean repair times are usually much larger than T, the possibility of repair is neglected. Generator i's mechanical power at time t is denoted by $p_i(t) \geq 0$, and the total mechanical power is denoted by

$$P(t) := \sum_i p_i(t). \tag{2}$$

Each generator capacity is limited:

$$p_i(t) \leq \overline{p}, \tag{3}$$

with $\overline{n}\,\overline{p} \geq l$. Moreover, the generators are subject to ramp-up constraints,

$$\dot{p}_i(t) \leq \overline{\dot{p}}_i. \tag{4}$$

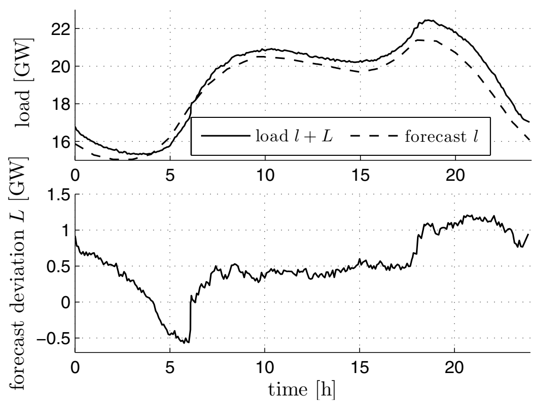

Figure 1. Total load $l + L$, load forecast l and load forecast deviation L for the New York ISO system on March 1st, 2007.

We ignore the ramp-down constraints as the focus is on generation outages and load increases, which cause capacity deficiencies, not surpluses.

We use a lumped, single-machine model for the power system dynamics. The system is operated so that the average system frequency $f(t)$, referred to as *frequency* in the remainder of the paper, is never below a specified threshold $\underline{f}$. Whenever the frequency is at or below $\underline{f}$, the system operator sheds load in an amount $S(t)$ equal to the power shortage,

$$S(t) = \begin{cases} l(t) + L(t) - P(t) & \text{if } f(t) \leq \underline{f} \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The unserved energy in the period $[0, t]$ is denoted by $U(t)$,

$$U(t) = \int_0^t S(\tau) d\tau. \quad (6)$$

Let f_o be the nominal frequency (in the U.S., f_o would be 60Hz.). Using Newton's second law of motion and noting that $f \approx f_o$, the frequency dynamics are given by

$$\frac{2HP_{base}}{f_o} \dot{f}(t) = \big(P(t) - l(t) - L(t) + S(t)\big). \quad (7)$$

where H is the system's coefficient of inertia, in seconds [14], and P_{base} is the base power.

The model of the generators operation is as follows. Whenever there is a departure from the desired frequency, each generator changes its output at a rate equal to its ramp limit so as to return to the desired frequency as quickly as possible. This simple model is used because the focus is on large disturbances, e.g., generation outages, which very likely constrain the generation response.

The rotating masses of the set of available generators store kinetic energy that is used to serve the load for very short periods of time in cases of power shortage. Due to the load shedding scheme, the minimum frequency is limited by $\underline{f}$, and so not all this stored energy can be used to serve the load. The stored energy $E(f, N)$ that can be used to serve the load is computed as follows. Let t_1 and $\underline{t}$ be such that $f(t_1) = f_1$ and $f(\underline{t}) = \underline{f}$. Assume, without loss of generality, that there is no outage in the period $[t_1, \underline{t}]$, so that H is constant on $[t_1, \underline{t}]$. Then, $E(f_1, n)$ is the energy that was supplied by the available generators' rotating masses in the period $[t_1, \underline{t}]$. This is equal to the negative of the integral of the power mismatch on $[t_1, \underline{t}]$,

$$E(f_1, n) = \int_{\underline{t}}^{t_1} \big(P(t) - l(t) - L(t) + S(t)\big) dt. \quad (8)$$

Using (7) to express the mismatch as a function of f and $\dot{f}$, and expressing E explicitly as a function of time, we obtain

$$E(t) = 2H\left(f(t) - \underline{f}\right) \frac{P_{base}}{f_o} \text{ MWs.} \quad (9)$$

Note that E has a discontinuity at each point where a generator changes its state, while f is a continuous function of t. $E(t)$ satisfies

$$0 \leq E(t) \leq 2H(f_o - \underline{f}) \frac{P_{base}}{f_o} =: \overline{E}. \quad (10)$$

The generation response and the load shedding scheme after a generation outage are illustrated with an example in Fig. 2. At $t = 1$s, one of the two spinning generators is outaged and so the energy E immediately decreases to $2H(f_o - \underline{f})\frac{P_{base}}{f_o}$. The remaining generator increases its mechanical power at a rate $\overline{p}$ to compensate for the shortage. However, the response is relatively slow: it takes until $t = 5$s for the shortage to be zero, until $t = 6$s for $P = \overline{p}$ to be reached, and until $t = 7$s to restore the normal operating frequency. The frequency decreases until it reaches $\underline{f}$, remains at $\underline{f}$ until load equals generation, and then starts to increase until f_o is reached.

2.2. Economic Aspects

The end users of electricity pay $\$w$ for every MWh consumed and lose $\$v$ for every MWh of shed load, with $v >> w$. The costs in the period $[0, t]$ are denoted by $C(t)$. The rate of change of the cost at time t, in \$/h, is

$$C_t(t) := \frac{dC(t)}{dt} = wP(t) + vS(t). \quad (11)$$

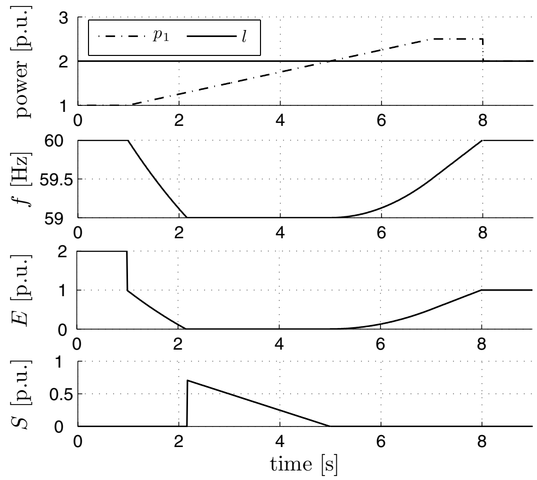

Figure 2. Generator 1's mechanical power p_1, frequency f, energy E and load shed S as a function of time, when generator 2 is outaged at $t = 1$. Parameters: $\overline{n} = 2$, $l = 2$, $L \equiv 0$, $\overline{p} = 2.5$, $\overline{\dot{p}} = 0.25$, $H = 2$ and $\underline{f} = 59$Hz. The control is a threshold policy with $f_o = 60$Hz.

The cost $C(t)$ is given by

$$C(t) = \int_0^t C_\tau(\tau) d\tau, \tag{12}$$

The *uncertain* cost to end users for the period of interest is $C(T)$.

Systems are designed and operated so that the probability of load shed is very small. Thus, usually

$$\int_0^T P(\tau) d\tau \approx \int_0^T L(\tau) d\tau = D(T), \tag{13}$$

which is exogenous (cannot be controlled by the ISO). Moreover, $v >> w$. Using (6), (11)-(13) we obtain that

$$C(T) \approx wD(T) + vU(T). \tag{14}$$

The first term in (14) is the cost of service, and the second term is the cost of not having service, or cost of load shed. The reserve impacts the second term in (14) but not the first. Both terms are uncertain, and

$$E\left[C(T)\right] = wE\left[D(T)\right] + vE\left[U(T)\right], \tag{15}$$
$$\text{Var}\left[C(T)\right] = w^2 \text{Var}\left[D(T)\right] + v^2 \text{Var}\left[U(T)\right]. \tag{16}$$

We now establish a basis for the comparison of different uncertain costs. An organization is indifferent between the uncertain costs $C(T)$ and a deterministic *equivalent cost* $\mathcal{C}$. This deterministic cost depends on the probability distribution of $C(T)$ and the organization's risk preferences, and a convenient model for this dependence is [15]

$$\mathcal{C} = E[C(T)] + k\sqrt{\text{Var}[C(T)]}. \tag{17}$$

The parameter k depends on the organization's risk preferences: for a risk-neutral entity, $k = 0$; for a risk-averse entity, $k > 0$; and for a risk-loving entity, $k < 0$.

3. Reserve Value, Demand Curves and Reserve Requirements

The reserve can be expressed as a function r of the number of generators $\overline{n}$, their capacity $\overline{p}$, and their ability to change their outputs $\overline{\dot{p}}$. Let the equivalent cost $\mathcal{C}$ be a function of r. We define the *decremental value* $\mathcal{V}^-_{\Delta r}(R)$ and the *incremental value* $\mathcal{V}^+_{\Delta r}(R)$ in \$ of a discrete change Δr in the reserve from its initial value R as

$$\mathcal{V}^-_{\Delta r}(R) \;:=\; \mathcal{C}(R - \Delta r) - \mathcal{C}(R), \tag{18}$$
$$\mathcal{V}^+_{\Delta r}(R) \;:=\; \mathcal{C}(R) - \mathcal{C}(R + \Delta r). \tag{19}$$

$\mathcal{V}^-_{\Delta r}(R)$ gives the value of having increased the reserve to R from $R - \Delta r$, i.e., is the contribution of the last reserve increment. $\mathcal{V}^+_{\Delta r}(R)$ gives the value of an additional Δr of reserve, i.e., is the value of the next increment. The unit values of reserve, in \$/MW, are $\mathcal{V}^-_{\Delta r}(R)/\Delta r$ and $\mathcal{V}^+_{\Delta r}(R)/\Delta r$. If r can be varied continuously, the marginal value in \$/MW of a change in r at $r = R$ is

$$\mathcal{V}_r(R) \;:=\; -\frac{d\mathcal{C}(R)}{dr}. \tag{20}$$

Equations (18)-(20) take into account the economic value of the reduction in both expected costs and risk brought by the increment in reserves.

The value of reserve ultimately depends on the change in $\overline{n}$, $\overline{p}$, and $\overline{\dot{p}}$ for the change in reserve. Therefore, to compute $\mathcal{V}^-_{\Delta r}(R)$, $\mathcal{V}^+_{\Delta r}(R)$ and $\mathcal{V}_r(R)$ the *direction* of change in $\overline{n}$, $\overline{p}$, and $\overline{\dot{p}}$ has to be given. Moreover, if the change in r entails a change in the operation policies of the system, this change has to be specified.

Reserve demand curves, assuming that the aggregate demand does not exercise market power, give the value of the reserve to the demand. Thus, if the change in r is continuous, then $\mathcal{V}_r(R)$ can be used as the demand curve. If the change in r is discrete, $\mathcal{V}^-_{\Delta r}(R)/\Delta r$ or $\mathcal{V}^+_{\Delta r}(R)/\Delta r$ can be used instead. For the demand curve to be representative of the aggregate demand, the different parameters involved in the calculations, such as the direction of change in $\overline{n}$, $\overline{p}$, and $\overline{\dot{p}}$, or the procedures to compute them, have to be agreed upon by the ISO

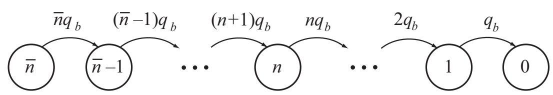

Figure 3. Transition rate diagram for the Markov process model of the generation state. The number in the circle indicates the number of available generators.

and the LSEs. The reserve value may be different for each organization. The reserve demand curve could be thought of as a weighted average of the value assigned to reserves by the different organizations.

The reserve value can be also used as an aid in the determination of optimal reserve requirements. The reserve value (18)-(20) and generation cost data are used to balance costs and benefits and determine an optimal reserve level.

4. Contingency Reserve

Contingency reserve is used to generate in cases of loss of supply. Thus, the load is assumed to be deterministic, i.e., $L \equiv 0$. For simplicity in the explanation, we assume that all generators are identical. The transitions between the functional and the breakdown states are modeled as continuous-time Markov processes, and the breakdown transition rate is denoted by q_b. We assume that the state of each generator is independent of the state of the other generators. The number of available generators at time t is denoted by $N(t)$, with $N(0) = \overline{n}$. The transition rate diagram for the Markov process characterizing the total number of available generators is depicted in Fig. 3.

The state transitions (outages) can be divided into three types according to the effect they cause in terms of load shedding:

- *type 1*: there is no load shed,
- *type 2*: load is shed temporarily, to arrest the frequency drop,
- *type 3*: load is shed for the remainder of the period, due to the lack of available generation.

The different types of transitions are illustrated in Figure 4. In this example, $P(0) = l$ and the outage of the first generator at $t = 10s$ does not lead to load shedding, i.e., it is of type 1. The outage of a second generator, at $t = 40s$, leads to temporary load shedding because the units cannot respond fast enough, i.e., it is of type 2. The outage of a third generator, at $t = 70s$, leads to permanent load shedding because of a lack of reserve, i.e., it is a type 3 transition. The frequency does not return to 60Hz after the third outage because load shedding is exactly equal to the shortage. To allow the return to 60Hz, load needs to be shed in excess of the power shortage.

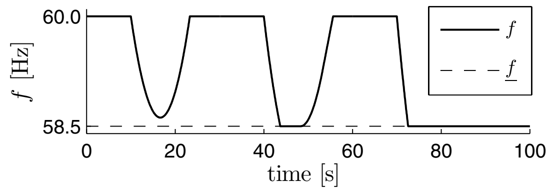

Figure 4. Frequency as a function of time. The frequency decreases abruptly each time a generator is outaged.

After a type 1 transition, all the load is served. After a type 3 transition, the load shed is approximately constant and equal to $S = l - N\overline{p}$. Consider a type 2 transition. When a generator becomes unavailable, P jumps to $Nl/(N+1)$, and the energy E immediately decreases to

$$\overline{e}(N) = 2(NH_g + H_l)(f_o - \underline{f})\frac{P_{base}}{f_o}. \quad (21)$$

where H_g is the coefficient of inertia of each generator, and H_l is the coefficient of inertia of the load, so that $H(t) = N(t)H_g + H_l$. The frequency decreases until it reaches $\underline{f}$, remains at $\underline{f}$ until load equals generation, and then starts to increase until f_o is reached. The generation increases at a rate $N\dot{\overline{p}}$ until either $f = f_o$ or $P = N\overline{p}$ is reached. If f_o is reached first, the total power immediately decreases to l, while if $N\overline{p}$ is reached first, the stored energy increases at a constant rate $N\overline{p} - l$ until $f = f_o$ is reached. We can show that the unserved energy due to the type 2 transition is

$$\mathcal{U}_n = \overline{e}(n) - \frac{D^2}{2n(n+1)^2\dot{\overline{p}}}. \quad (22)$$

Note that $\mathcal{U}_n$ is an increasing function of n. As n increases, $\mathcal{U}_n$ becomes positive, which means that the transition is type 1 (there is no need to shed load).

Let the probability of having i available generators at time t be denoted by

$$\pi_i(t) := P\left[N(t) = i\right]. \quad (23)$$

We can show that

$$\pi_i(t) = \sum_{k=j}^{\overline{n}} a_{jk}^{\overline{n}} e^{-q_b k t}, \quad i = 0, 1, \ldots, \overline{n}, \quad (24)$$

where

$$a_{jk}^{\overline{n}} = \frac{(-1)^{j+k}\overline{n}!}{j!(i-k)!(k-j)!}. \quad (25)$$

We now compute $E[U]$ and Var $[U]$, or, equivalently, $E[U]$ and $E\left[U^2\right]$, since

$$\text{Var}\,[U] = E\left[U^2\right] - E\left[U\right]^2. \quad (26)$$

Conditioning on the number of available generators at time T,

$$E[U] = \sum_{i=0}^{\overline{n}} E[U|N(T)=i]\, P[N(T)=i], \quad (27)$$

$$E\left[U^2\right] = \sum_{i=0}^{\overline{n}} E\left[U^2|N(T)=i\right] P[N(T)=i]. \quad (28)$$

We now make use of the different types of transitions. For the type 1 transitions, clearly

$$E[U|N(T)=i] = 0, \quad (29)$$

$$E\left[U^2|N(T)=i\right] = 0. \quad (30)$$

Consider a type 2 transition. If $N(T) = i$, there were $\overline{n}-i$ transitions during the period $[0,T]$. The time spent in any state is usually several orders of magnitude larger than the time needed for the available generators to restore the frequency to f_o after an outage. Thus, we can assume that the system frequency when an outage occurs is f_o, so that

$$E[U|N(T)=i] = \sum_{j=i}^{\overline{n}} \mathcal{U}_j, \quad (31)$$

$$E\left[U^2|N(T)=i\right] = \left(\sum_{j=i}^{\overline{n}} \mathcal{U}_j\right)^2. \quad (32)$$

where $\mathcal{U}_j$ is obtained from (22).

For type 3 transitions we need some more analytical developments. Let the time at which there is a transition to state n be denoted by t_n. This time is the realization of a random variable T_n. Clearly, $0 = t_{\overline{n}} < t_{\overline{n}-1} < \cdots < t_n < \cdots < t_1 < t_0$. The unserved energy given the transition times and $N(T) = i$ is

$$U_i(t_{\overline{n}}, \ldots, t_i) = \sum_{j=i+1}^{\overline{n}} (t_{j-1} - t_j) S_{j-1} \quad (33)$$

where

$$S_j = \begin{cases} 0 & \text{if } j \text{ is type 1} \\ \frac{\mathcal{U}_j}{t_j - t_{j+1}} & \text{if } j \text{ is type 2} \\ l - j\overline{p} & \text{otherwise.} \end{cases} \quad (34)$$

The joint probability density function (p.d.f.) of $T_{\overline{n}}, T_{\overline{n}-1}, \cdots, T_0$ is in (35), and the conditional density of $T_{\overline{n}}, \ldots, T_i$ given that $N(T) = i$ is in (36). The conditional expectations needed for (27) and (28) are computed in (37) and (38). These can be easily expressed in closed form for each final state i, and so the computational burden of their computation is small.

Table 1. Contengency reserve decremental value for the test system, in \$/MW/h.

	11th unit	12th unit
reserve	100 MW	200 MW
risk-neutral, $k = 0$	\$55.9/MW/h	\$0.4/MW/h
risk-averse, $k = 1$	\$635.0/MW/h	\$31.0/MW/h

4.1. Numerical Example

We compute the value of reserves for a test system with a load $l = 1000$ MW. Each generator has an installed capacity of $\overline{p} = 100$MW and a ramp constraint of $\bar{\dot{p}} = 10$MW/min. Thus, the minimum number of generators required to serve the load is 10. The nominal frequency is $f_o = 60$Hz, and the minimum frequency is $\underline{f} = 58.5$Hz. The breakdown rate is $q_b = 10$1/year, the value of lost load is $v = \$10,000$/MWh, and the time horizon is $T = 1$h. The reserve is $r = \max\{N\overline{p} - l, 0\}$. We use the costs (18) to obtain the decremental value of the 11th and 12th units, and so $\Delta r = 100$MW. Having 11 units spinning satisfies an "n-1" security criterion while 12 spinning units satisfies an "n-2" security criterion. The per unit reserve value $\mathcal{V}^-_{\Delta r}/\overline{p}$ are computed for risk neutral and risk averse demand with $k = 1$. Since $T << 1/q_b$, the likelihood of having multiple outages in the period of interest is very small. Thus, for each n we compute $\mathcal{C}$ considering up-to 3 outages (the contribution of the third outage to $\mathcal{C}$ is two orders of magnitude smaller than the contribution of the first outage). The results are shown in Table 1. We note that the value of reserve for the risk averse demand is between one and two orders of magnitude larger than the value of reserve for the risk neutral one. Also note that the reserve value decreases as the reserve increases, as expected.

5. Regulating Reserve

Regulating reserve is used to provide for the continuous and uncertain load variations. Thus, in this Section the load is assumed to be uncertain and continuously changing, while the generators are modeled as perfectly reliable. Generators providing regulating reserve can change their output relatively quickly for small changes in output. Since the aggregate load does not usually have very large unpredicted variations, and since the regulating reserve is provided by many generators in parallel, we neglect the effects of ramp constraints. Since ramp constraints are neglected, the frequency only departs from the desired frequency when all the

$$f_{T_{\overline{n}}\cdots T_0}\left(t_{\overline{n}},\ldots,t_0\right)=\begin{cases}\overline{n}!q_b^{\overline{n}}e^{-q_b\sum_{i=0}^{\overline{n}}t_i} & \text{if } 0=t_{\overline{n}}<t_{\overline{n}-1}<\cdots<t_n<\cdots<t_1<t_0\\ 0 & \text{otherwise.}\end{cases} \tag{35}$$

$$\tilde{f}_{T_{\overline{n}}\cdots T_i}\left(t_{\overline{n}},\ldots,t_i\right)=\begin{cases}\frac{\overline{n}!q_b^{\overline{n}-i}}{i!\pi_i(T)}e^{-q_b\sum_{k=i}^{\overline{n}}t_k}e^{-iq_bT} & \text{if } 0=t_{\overline{n}}<t_{\overline{n}-1}<\cdots<t_i\leq T\\ 0 & \text{otherwise.}\end{cases} \tag{36}$$

$$E\left[U|N(T)=i\right]=\int_0^T\int_0^{t_i}\cdots\int_0^{t_{\overline{n}-2}}U_i(t_{\overline{n}},\ldots,t_i)\tilde{f}_{T_{\overline{n}}\cdots T_i}(t_{\overline{n}},\ldots,t_i)dt_{\overline{n}-1}\cdots dt_{i+1}dt_i. \tag{37}$$

$$E\left[U^2|N(T)=i\right]=\int_0^T\int_0^{t_i}\cdots\int_0^{t_{\overline{n}-2}}U_i^2(t_{\overline{n}},\ldots,t_i)\tilde{f}_{T_{\overline{n}}\cdots T_i}(t_{\overline{n}},\ldots,t_i)dt_{\overline{n}-1}\cdots dt_{i+1}dt_i. \tag{38}$$

capacity providing reserve is generating and the load increases. These assumptions allows us to work with the aggregate regulating reserve, rather than modeling separate generators.

The generators are scheduled to serve the (deterministic) load $l(t)$ and to provide regulation services so as to account for $L(t)$. In this Section, we focus on the uncertain component $L(t)$ of the load, and the generation $P_r(t)$ that provides regulation services. As the generators are assumed to be perfectly reliable, $P(t)=l(t)+P_r(t)$. The regulating reserve is denoted by R, so that

$$P_r(t)\leq R. \tag{39}$$

We model $L=\big(L(t),t\in[0,T]\big)$ as a Brownian motion [16] with parameter σ^2. Brownian motion models the observed irregular motion of inert small particles, such as polen and dust, in still water. The movement is attributed to a very large number of independent, random collisions with water molecules [16]. In power systems, the deviation L from the load forecast l in the short-term can be thought of as resulting from a huge number of independent, "small" events, such as users deciding to turn their devices on or off. By short-term we mean periods of the order of an hour, where temperature and other important parameters can be considered constant.[1]

The computation of $\mathcal{C}$ requires the computation of $E\left[C(T)\right]$ and Var $\left[C(T)\right]$, which in turn requires the computation of $E\left[D(T)\right]$, $E\left[U(T)\right]$, Var $\left[D(T)\right]$ and Var $\left[U(T)\right]$. The mean and variance of $D(T)$ are

$$E\left[D(T)\right]=0, \tag{40}$$

$$\text{Var}\left[D(T)\right]=\frac{\sigma^2T^3}{3}. \tag{41}$$

[1]In the medium- to long-term, this model may not be very appropiate because the random time-variation of variables such as temperature makes the decision of an end consumer to turn devices on or off to be correlated to that of another consumer.

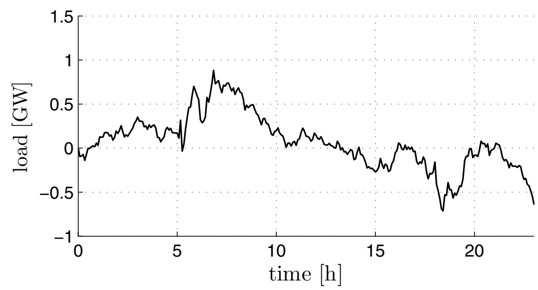

Figure 5. Sample path of a Browian motion with σ chosen as the maximum likelihood estimate given the available data from NYISO.

The mean and variance of $U(T)$ are not straightforward to obtain. However, we can obtain upper bounds for $E[U(T)]$ and Var$[U(T)]$ if we ignore the possibility of energy storage. Taking the limit as $H\rightarrow 0$ in (7) we obtain that the limiting load shedding policy is

$$S(t)=\begin{cases}L(t)-R & \text{if } L(t)\geq R\\ 0 & \text{otherwise.}\end{cases} \tag{42}$$

Then, we can express the expected load shed as

$$E[U(T)]=\int_0^T\int_{\overline{P}}^{\infty}\frac{\ell-\overline{P}}{\sqrt{2\pi\sigma^2t}}\,e^{\frac{-\ell^2}{2\sigma^2t}}d\ell dt \tag{43}$$

For $t<\tau$, the autocorrelation function $R_S(\cdot,\cdot)$ of S is

$$R_S(t,\tau)=\int_R^{\infty}\int_R^{\infty}\frac{(\ell-R)(\lambda-R)e^{\frac{-\ell^2(\tau-t)-(\lambda-\ell)^2t}{2\sigma^2t(\tau-t)}}}{2\pi\sigma^2\sqrt{t(\tau-t)}}d\lambda d\ell \tag{44}$$

In terms of R_S, $E[U(T)^2]$ is

$$E[U(T)^2]=2\int_0^T\int_0^{\tau}R_S(t,\tau)dtd\tau \tag{45}$$

Var$[U(T)]$ is computed using (26), (43) and (45).

We take the derivatives of $E[U(T)]$ and $E[U(T)^2]$ with respect to the reserve $r=R$ to compute the marginal reserve value (20). The derivative of the expected unserved energy with respect to the reserve yields the known result

$$-\frac{dE[U(T)]}{dr}(R) = \int_0^T \int_R^\infty \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{\frac{-\ell^2}{2\sigma^2 t}} d\ell dt = LOLP \cdot T \quad (46)$$

where $LOLP$ is the *loss of load probability*, i.e., the probability that there is load shed in the period $[0, T]$. These integrals can be computed using numerical methods.

We now address the question: how close to the bounds (43) and (45) are the actual $E[U(T)]$ and $E[U(T)^2]$? Or, how valid is to let $H \to 0$ for the purpose of computing the unserved energy? The usual values of the inertia coefficient H are between 2 l^{peak} and 10 l^{peak} s, and a typical value would be $H = 6\ l^{peak}$ s, i.e., the typical total energy stored in the rotating masses is capable of serving the peak load for 6 seconds. With load shedding, since the maximum frequency change is reduced about 100 times, the energy in the rotating masses that can be used to serve the load is also reduced about 100 times to obtain $\overline{E}$. Thus, for a period of one hour, the total energy demand is $60,000$ times greater than the energy stored $\overline{E}$. Another important quantity to take into account is the standard deviation of the energy due to the load L in the period $[0, T]$. For $T = 1$h, the standard deviation of $D(1)$ is of the order of 62 l^{peak}MWs. Load shedding typically starts after the frequency falls about 0.7 Hz for systems with $f_o = 60$ Hz. With these typical values, the maximum stored energy that can be used is $\overline{E} = 0.14\ l^{peak}$MWs. Thus the standard deviation of $D(1)$ is two orders of magnitude larger than $\overline{E}$. Since storage is so little compared to both the expected total energy demand and the standard deviation of the energy demand, we conclude that the bounds (43) and (45) are quite tight.

5.1. Numerical Example

We illustrate the reserve value computed using (20), (43), and (45) in Table 2. The parameter are $k = 0$, $\sigma = 100\text{MW}/\sqrt{\text{h}}$, $v = 10000$\$/MWh and $l = 10000$MW. As for contingency reserve, the value of reserve for the risk-averse demand is between one and two orders of magnitude larger than the value of reserve for the risk-neutral one. Note that the regulation value for zero reserve is $v/2$. The reason is that L may be positive of negative with equal probability, and at zero reserve this means load shedding with a probability of $1/2$.

Table 2. Regulating reserve marginal value for the test system, in $/MW/h.

reserve	0 MW	200 MW
percentage of load	0%	2%
risk-neutral, $k = 0$	$5,000/MW/h	$58/MW/h
risk-averse, $k = 1$	$71,622/MW/h	$3548/MW/h

6. Summary and Future Work

In this paper, we have discussed the valuation of spinning contingency reserve and regulating-up reserve. We fully take into account the reliability and dynamic characteristics of system components, the uncertainty in the demand, the system operation policies, the underfrequency load shedding schemes, the economic impacts of outage costs and the risk preferences of the demand. The reserve value is obtained either in closed form (contingency reserve) or in the form of definite integrals (regulating reserve), which are solved numerically with limited computational effort. The results are useful for the development of reserve demand functions in competitive markets and the determination of optimal reserve requirements.

The work in this paper can be extended in several directions, the most natural being the consideration of non-identical generators and the valuation of regulating-down reserve. Another extension is the modeling of the frequency-dependency of the load to capture its economic value as demand-side reserve. The valuation of non-spinning contingency reserve, with the consideration of failure-to-start probabilities, constitutes another direction for future research. Yet another important direction for future work is the extension of the work to interconnected systems, incorporating network aspects. Finally, the extension of the techniques developed in this paper to reactive reserve studies will provide a deeper understanding of the value of such services.

References

[1] B. Wollenberg and A. Wood, *Power Generation, Operation and Control*, 2nd ed. New York, NY: John Wiley, 1996.

[2] L. T. Anstine, R. E. Burke, J. E. Casey, R. Holgate, R. S. John, and H. G. Stewart, "Application of probability methods to the determination of spinning reserve requirements for the Pennsylvania-New Jersey-Maryland Interconnection," *IEEE Trans. Power App. Syst.*, vol. PAS-82, no. 10, pp. 726–735, Oct. 1963.

[3] *Ancillary Services Manual*, Manual 2, v. 3.5, New York Independent System Operator, Schenectady, NY, May 2006. [Online]. Available: http://www.nyiso.com

[4] *Ancillary Services Market Business Rule*, Final Version, Midwest Independent System Operator, Carmel, IN, Jan. 2007. [Online]. Available: http://www.midwestiso.org/publish

[5] A. Merlin and P. Sandrin, “A new method for unit commitment at Electricite de France,” *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 5, pp. 1218–1225, May 1983.

[6] C.-L. Tseng, S. S. Oren, A. J. Svoboda, and R. B. Johnson, “Price-based adaptive reserve requirements in power system scheduling,” *Electrical Power and Energy Systems*, vol. 21, no. 2, pp. 137–145, Feb. 1999.

[7] J. Chen, J. S. Thorp, R. J. Thomas, and T. D. Mount, “Locational pricing and scheduling for an integrated energy-reserve market,” in *Proc. 36th Hawaii Int Conf Syst Sciences*, 2003.

[8] F. D. Galiana, F. Bouffard, J. M. Arroyo, and J. F. Restrepo, “Scheduling and pricing of coupled energy and primary, secondary, and tertiary reserves,” *Proc. of the IEEE*, vol. 93, no. 11, pp. 1970–1983, Nov. 2005.

[9] M. A. Ortega-Vázquez, D. S. Kirschen, and D. Pudjanto, “Optimising the scheduling of spinning reserve considering the cost of interruptions,” *IEE Proc. - Gen., Trans. Dist.*, vol. 153, no. 5, pp. 570–575, Sept. 2006.

[10] J. W. O’Sullivan and M. J. O’Malley, “A new methodology for the provision of reserve in an isolated power system,” *IEEE Trans. Power Syst.*, vol. 14, no. 2, pp. 519–524, May 1999.

[11] D. S. Kirschen, K. R. W. Bell, D. P. Nedic, D. Jayaweera, and R. N. Allan, “Computing the value of security,” *IEE Proc. - Gen., Trans. Dist.*, vol. 150, no. 6, pp. 673–678, Nov. 2003.

[12] *Transmission and Dispatching Operations Manual*, Manual 12, New York Independent System Operator, Schenectady, NY, Sept. 1999. [Online]. Available: http://www.nyiso.com

[13] C. A. Williams, M. L. Smith, and P. C. Young, *Risk Management and Insurance*, 7th ed. New York, NY: McGraw-Hill, 1995.

[14] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*. Upper Saddle River, NJ: Prentice Hall, 1998.

[15] B. Allaz, “Oligopoly, uncertainty and strategic forward transactions,” *Int. Jour. Industrial Organization*, vol. 10, no. 2, pp. 297–308, June 1992.

[16] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed. New York, NY: Springer, 1991.