Renewable and Sustainable Energy Reviews 183 (2023) 113465

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

Reserve and energy scarcity pricing in United States power markets: A comparative review of principles and practices

Mahdi Mehrtash [a,*], Benjamin F. Hobbs [a], Erik Ela [b]

[a] Ralph O'Connor Sustainable Energy Institute, Johns Hopkins University, Baltimore, MD 21218, USA
[b] Power Delivery and Utilization Department, Electric Power Research Institute, Palo Alto, CA 94304, USA

ARTICLE INFO

Keywords:
Operating reserve demand curve
Reserve scarcity
U.S. electricity market
Reserve and energy scarcity pricing
Independent system operator

ABSTRACT

Errors in forecasting load and renewable-based generation in restructured power systems mean that independent system operators (ISOs) must procure sufficient operating reserves to keep the real-time operation of the system reliable and secure. But when procured reserves turn out to be insufficient in real-time due to the lack of resource capacity or ramp capability, operators often set higher prices for reserves and energy to encourage more supply, and to motivate consumers to decrease usage or shift it to other times. This procedure, which is called scarcity or shortage pricing, is a core feature of U.S. electricity markets. It is receiving increased attention from market designers and stakeholders because scarcity will become more important for spot price formation in the future with the increased penetration of zero-marginal cost renewables, and the shrinking role of fuel costs in setting prices.

Scarcity pricing is implemented in various ways by different ISOs. These differences have practical implications for the level of prices and incentives for investment, operations, and demand modification. In this paper, general approaches and specific calculation procedures for reserve and energy scarcity pricing practices and calculations across the seven ISO-based U.S. power markets are reviewed and compared. A consistent terminology is used to facilitate the comparison. Current scarcity pricing practices are grouped into three approaches: (1) imposing an adder after the spot market is run; (2) including stepwise demand curves within market clearing procedures for non-contingency reserve products (e.g., the novel flexiramp product), which tends to yield longer right tails for energy scarcity premium curves; and (3) having stepwise demand curves for traditional contingency reserve products only, which results in shorter right tails in energy scarcity curves. A generic numerical example is presented to highlight the large practical differences among the reserve scarcity pricing approaches and specific implementations. To further investigate factors that contribute the most to demand curves differences among ISOs, a sensitivity analysis is performed. This analysis shows that the largest source of differences among the curves is the scarcity prices assumed in the case of severe scarcity, while the number of steps used and whether flexiramp is considered also yields important differences in scarcity prices.

As renewable penetration increases, it will become increasingly crucial to employ administrative demand curves so that spot prices more effectively motivate supply and demand adjustments exactly when and where they are needed. This study shows that the different assumptions yield very different scarcity premiums for reserves and energy, and are likely to provide divergent incentives for resources to respond to shortages. It is concluded that to promote market efficiency, a reserve shortage demand curve should have at least three features: inclusion of the marginal value of reserve products at each shortage level, consideration of the magnitude and probability of supply contingencies, and avoidance of abrupt price discontinuities that can cause excessively volatile market outcomes.

1. Introduction

Scarcity, in microeconomic terms, occurs when capacity limitations or high supply costs mean that the quantity that can be provided to consumers is less than what consumers would be willing to buy at a low price. It can be defined as occurring when price increases are used to

* Corresponding author.
E-mail address: mahdi.mehrtash@jhu.edu (M. Mehrtash).

https://doi.org/10.1016/j.rser.2023.113465
Received 15 November 2022; Received in revised form 14 May 2023; Accepted 12 June 2023
Available online 23 June 2023

ration scarce supply among consumers, with those prices reflecting either a rising marginal willingness to pay when the quantity demanded is reduced and/or the high cost of extraordinary measures to increase supply.

In the now-classic locational marginal pricing (LMP) framework developed by Schweppe et al. [1], scarcity occurs whenever supply capacity limits lead to price being set by price-responsive load rather than the marginal cost of supply. Although elastic demand can set price under either low or high price conditions, scarcity pricing in power markets is more popularly understood as the use of administratively set high electricity prices to motivate consumers to limit their consumption and to elicit additional supply (e.g., by taking advantage of emergency capacity ratings).

In electricity markets, most consumers do not receive information on the marginal cost of supply in day-ahead (DA) or real-time (RT) markets, and instead pay regulated retail rates. Such rates may have an ex ante dynamic ("time-of-day") component, but only very roughly approximate actual time-varying marginal supply costs. This situation is changing with the growth of demand response programs, stimulated by Federal Energy Regulatory Commission (FERC) Order 719, among other policies [2], but the fundamental situation persists that most demand will not respond to real-time cost variations. Other obstacles to dynamic demand response are features, such as bid and price caps, in many independent system operator (ISO)/regional transmission organization (RTO) markets that limit price variations in DA and RT markets, such that wholesale market prices cannot rise to levels that reflect the actual marginal worth of power when the load is curtailed. This was, for instance, identified as contributing to inefficient management of the August 2020 heat wave event in California [3].

Because of the lack of demand-side participation in price formation in electricity spot markets, market designs have instead implemented administrative shortage prices and constraint violation penalties in market scheduling software, whose influence on prices makes up for customers' inability to affect spot prices by expressing their marginal willingness to pay [4–7].

Multiple approaches have been proposed to implement and reform administrative scarcity mechanisms in ISO-based spot markets. Aligning with recently identified research priorities and opportunities for the U.S. wholesale electricity markets [8], the goal of this paper is to explain and compare methods for quantifying scarcity in electricity markets and adjusting energy and operating reserve prices to reflect that scarcity. The practical implementation of those methods in U.S. ISO markets and offer recommendations for improvement are emphasized.

In particular, key concepts and elements are reviewed for accomplishing three tasks in those markets: detection and calculation of the amount of scarcity; assignment of a marginal economic value for reducing scarcity; and finally, reflection of that value in market prices for reserves and ultimately spot energy. The following is offered as a possibly useful classification of approaches to these tasks and at the same time introduces some standard terminology that will facilitate the comparison of approaches used in different markets.

- Scarcity detection: This task, which can be done hours to days ahead, involves assessment of the availability of resource capacity and comparison of that capacity against needs for energy and operating reserves. This task may involve calculations of the short-term probability that capacity will be insufficient to meet all those needs, which is sometimes termed "loss of load probability" (LOLP).
- Scarcity valuation: This task involves the determination of the marginal value of procuring additional units of reserves or energy, which may be reflected in constraint violation penalties or administrative demand curves for reserves. Determination of those values may often include an ex ante estimate of the "value of lost load" (VOLL, in $/MWh) but other methods are also used. In general, market software includes penalty prices for violation of power balance or reserve requirement constraints (also in $/MWh); when those marginal penalties are increased as the amount of violation grows, the result is a demand curve for the commodity in question. The $/MWh values embodied in penalties and demand curves may be purely arbitrary (e.g., a preset fraction of the offer or price cap in the market). Alternatively, they may be based on a calculation of the probability of having to take certain operator actions and their resulting impacts on consumers, or perhaps factoring in the potential for penalties imposed by administrators for poor performance (i.e., the North American Electric Reliability Corporation (NERC) is empowered by the U.S. Energy Policy Act of 2005 to set such penalties).
- Integration in market scheduling and clearing: To reflect scarcity detection and valuation in market software, there are various approaches, as discussed in the rest of this paper. This can be accomplished through ex post adders that are added to market energy prices, or by endogenous price setting in the software through penalties for constraint violations or demand curves for relaxing those constraints.

The main objective of this study is to perform a comparative review of reserve scarcity pricing practices and calculations across the U.S. power markets. Scarcity pricing is growing in importance in spot market price formation as zero marginal cost renewables increase their output, and fuel prices recede in importance as a determinant of prices. This comparative review shows the similarities and differences among reserve scarcity pricing practices and reveals the features of a correct pricing mechanism. To facilitate the comparison, a consistent terminology is used to describe each reserve product, and three groups are introduced for reserve scarcity pricing approaches: (1) imposing an adder after the spot market is run; (2) including stepwise demand curves within market clearing procedures for non-contingency reserve products (e.g., the novel flexiramp product), which tends to yield longer right tails for energy scarcity premium curves; and (3) having stepwise demand curves for traditional contingency reserve products only, which results in shorter right tails in energy scarcity curves. To highlight the practical differences among the general approaches and specific implementations, a generic numerical example is presented based upon some simplified assumptions. Those results show that the different assumptions yield very different scarcity premiums for reserves and energy, and are likely to provide divergent incentives for resources to respond to shortages.

The paper is organized as follows. A summary of present reserve shortage pricing practice, including the definition of the used terminology and an overview of the basic methods used to parameterize shortage penalties, is presented in Section 2. Section 3 describes the current reserve shortage pricing practices for each of the seven U.S. ISOs, one-by-one. Section 4 includes a generic numerical example to illustrate the practical implications of differences among ISOs in the elements and calculation approaches in their scarcity pricing systems. Conclusions and recommendations are presented in Section 5.

2. Summary of current reserve shortage pricing practices

In this section, a brief overview of electricity markets across the U.S. and their pricing mechanisms when reserve shortage conditions occur is provided. First, a consistent terminology is introduced to make it easier to compare methods used by different ISOs. Second, the general approaches used by U.S. ISOs for implementing shortage pricing are grouped into three categories. Third, three basic approaches to calculating shortage penalty parameters used into those approaches are identified. Finally, a table is presented summarizing the shortage pricing methods and penalty values applied by U.S. ISOs.

2.1. Reserves terminology

Since each market operator has its own terminology, to avoid confusion, a single set of terms is used in this paper to describe the various types of reserves that are procured in spot markets and potentially subject to scarcity pricing.

- Regulation up and down (abbreviated as reg-up and reg-down) are market reserve products that correspond to capacity that can respond to regulation signals sent every 4 s to move up and down within the smallest market interval (e.g., within the 5-min real-time dispatch interval).
- Spinning reserve (abbreviated as spin) is a 10-min reserve product that is synchronized to the system. Usually, but not always, spin is deployed (produces energy) only in response to enumerated contingencies (especially outages of generating or transmission equipment).
- Non-spinning reserve (abbreviated as non-spin) is another contingent reserve that is not synchronized to the system but can be committed and synchronized within 10 min. The non-spinning reserve is sometimes called supplemental reserve by market operators. Spin and non-spin are referred to collectively as primary reserves.
- 30-min reserve is provided by resources that may or may not be synchronized to the system but can respond within 30 min. The 30-min reserve is sometimes called replacement reserve or longer-term contingency reserve in some markets.
- Flexiramp-up and down (also known as flexible ramp product or ramp capability product) is a new market product that some markets have adopted. It is defined as the short-term (e.g., 5–15 min) ramping capability that can be used to meet both forecast and unpredicted net load ramps from one RT market interval to the next.

2.2. Three general computational frameworks: price adders vs. energy-reserves co-optimization (with flexiramp) vs. co-optimization (no flexiramp)

Here scarcity pricing methods are grouped by the basic framework used to compute scarcity premiums. In the first, Ex Post Price Adder, the value of reserves in shortage condition is calculated as a separate scarcity price adder as a function of committed reserves (as is done in the RT market of the Electric Reliability Council of Texas (ERCOT)). The final energy price is the sum of this adder and the locational marginal price calculated by the RT energy market.

The other two frameworks both involve co-optimization of reserves and energy, in which the scarcity premium for energy emerges from the optimization, considering the stepwise demand curves for reserves. These two frameworks differ in terms of whether flexible ramp products (“flexiramp”) or other non-contingency products are considered in the market mechanism or not. Flexiramp is a relatively new approach to managing unexpected net load changes, in which capacity is reserved in one market interval in order to accommodate possible up- and down-net load ramps that could occur between that and the next interval. Unlike more traditional “contingent” reserves, flexiramp capacity can be deployed to meet energy needs in that subsequent interval through economic dispatch like any other supply, whereas other reserves are usually deployed by operator instructions only if certain contingencies or operating conditions are met.

In the second framework, Co-optimized Stepwise Demand Curves with Flexiramp, multi-step reserve demand functions—called operating reserve demand curves (ORDCs)—are used for one or each of several reserve products. These functions include two or more steps with price quantity pairs where the demand value ($/MWh) for the service decreases as the quantity (MWh) of the reserve increases. The systems differ in terms of which reserve products have demand curves, the number of steps, and the levels of the $/MWh values associated with those steps. In addition, flexiramp, as a non-contingency product, is procured and is co-optimized with energy and normal reserves (Midcontinent ISO (MISO), Southwest Power Pool (SPP), and California ISO (CAISO)).

The third framework, Co-optimized Stepwise Demand Curves without Flexiramp, is similar to the second approach, in that stepwise reserve demand functions are used for one or each of several reserve products. However, flexiramp or other non-contingency products are not procured in this market mechanism (e.g., New York-ISO (NYISO), PJM, and ISO-New England (ISO-NE)).

It should be noted that all of these systems automatically trigger high energy prices (at or near the price cap) under extreme shortage conditions (i.e., reflecting prices that are much higher on the left side of the demand curves). This can happen due to several circumstances, for instance when load is curtailed by the operator when reserves are reduced below certain critical levels, or when emergency demand response programs are dispatched by the software. Thus, when scarcity is first detected, reserve and energy prices may be increased by only a small amount (i.e., right tail of the demand curves), but the increases will grow in magnitude as scarcity increases until the extreme conditions trigger the highest prices. The level and rate of increase of scarcity premiums depend on the design of the markets and the level of penalty and demand parameters.

2.3. Approaches to parameterizing shortage penalties

The levels of penalty that a market mechanism applies when reserves or energy are short of the targeted amounts vary greatly among U.S. ISOs, in part because of the different philosophies underlying how they are estimated.

Penalty Method 1. When reserves are short, there is an increase in the probability that the system operator will need to take some action to reduce the risk of uncontrolled load shedding. Thus, a logical approach to calculating a penalty for reserve scarcity is to base it on the product of the cost of having to take that action times the probability of having to take such an action. In particular, the per MW penalty for a reserve shortage should be the marginal change in that product resulting from a 1 MW increase in reserve shortfall.

In Penalty Method 1V (for Value of lost load, VOLL), it is assumed that the operator’s action is not to curtail load involuntarily but in a controlled fashion (e.g., rotating blackouts); the resulting cost is VOLL (in $/MWh) times the amount of unserved energy (UE, in MWh). The overall cost of the reserve shortfall is the product of VOLL, unserved energy, and loss-of-load probability (LOLP), and the marginal penalty is based on the increase in expected unserved energy (UE $\times$ LOLP) from a 1 MW increase in reserve shortage. This, for instance, is the basis of the ERCOT ex post scarcity adder.

However, operators have a number of options instead of shedding a load when the system is at risk of having an energy imbalance. In Penalty Method 1C (for Commitment cost), additional resources are committed by the operator when reserves are short, and the penalty reflects the commitment cost times the probability of having to take the commitment action. The Southwest Power Pool adopts this approach for its flexiramp demand curve.

Alternatively, if there is not enough time to commit such resources, a system may automatically draw power from neighboring systems, resulting in unscheduled flows (area control error). Excessive unscheduled flows not only inflict costs on neighboring systems, but they may also risk the imposition of penalties by NERC. Penalty Method 1L (for Leaning) bases the penalty for shortfalls upon a qualitative assessment of the cost of such flows. The California ISO bases penalty values in its demand curve for flexiramp on an assumption that having a load imbalance due to being unable to meet an unexpected ramp will only result in leaning on a neighboring system, and not load curtailment.

In Penalty Method 2R (for Ranking), neither the probability of having to take an action nor the cost of that action are explicitly estimated. Instead, penalties are assigned based on a ranking judgment of the relative marginal consequences of shortages of the different reserves (in some cases as a function of the level of shortage), relative to a regulatory price or offer cap. Thus, for example, if shortages of spin are more serious on the margin than shortages of 30-min reserves, then the former could have a higher $/MWh penalty. This may depend on categories of reserve that have specific NERC requirements to be held (e.g., contingency reserve) vs. those that are held for other reasons. This is also possible for those in this category, that when the price or offer cap is raised, the penalty for the two reserves would be raised proportionally, preserving their relative rank. This is the basis, for instance, for various constraint violation penalties in the California ISO system. Although it is recognized that the value of lost load is likely much higher than the price cap, the desire to mitigate market power and perhaps lingering sensitivities resulting from the California 2000–2001 power crisis have led to the preservation of a relatively low cap and thus low penalties in that market.

A theoretically appealing approach is Penalty Method 3W (for Willingness to pay), in which the cost of operators shedding load is the consumer's willingness-to-pay to avoid curtailment, as revealed through some market processes. Rather than penalties being set by a purely administrative procedure, this approach uses information revealed by consumer decisions in the market—which is how markets for most goods and services other than electricity work. The market may be a short-term spot market with dynamic prices [1], which provides evidence of how high spot prices need to go before consumers are willing to shift loads or go without. Critical peak pricing is a relatively popular variant of the short-term approach [9], while aggregating consumers with smart thermostats who can control energy use in response to price or other signals is a growing trend [10]. Alternatively, the market may be longer term, in which consumers have contractually agreed to be curtailed under certain conditions. Variants of this approach include legacy appliance load control programs and interruptible rates, while priority pricing [11] and capacity subscriptions [12] are more recent approaches that can be designed to let consumers choose from a menu of reliability alternatives.

2.4. Summary of U.S. ISO shortage pricing practices

Table 1summarizes reserve and energy scarcity pricing mechanisms used by the seven U.S. ISOs, including information on the computational and parameterization methods used (from Sections 2.2 and 2.3, respectively). This table is presented as a precise yet concise summary for reference purposes. As can be seen here, and which is more clearly illustrated by the comparative figures in Section 4, the ISOs have adopted a wide variety of penalties and prices [13]. Principles and further details of shortage pricing practices are presented in Section 3.

3. Description of current reserve shortage pricing practices

In this section, the present shortage pricing practices for each U.S. ISO, and, in some cases, future reforms are discussed. A comparison of the practices using the above standard terminology and classification of approaches is also presented.

It is emphasized below how ISOs differ in their formation of scarcity premiums based on shortage penalties and ORDCs, and the methods they use to derive those parameters. There are other important differences among their operating reserve markets, such as whether reserves are procured zonally or system-wide,[1] the formulation of nesting constraints (in which higher quality reserves also contribute to meeting requirements for lower quality reserves), and particular requirements that a resource must satisfy to qualify to provide a given reserve (such as duration requirements). Those details are omitted to focus on basic price formation processes. Details of practices can be found in the business practice manuals of the ISOs, cited below.

The seven markets are grouped into three approaches of pricing practices: price adder-based (Section 3.1); stepwise ORDCs with flexiramp products (Section 3.2); and stepwise ORDCs without flexiramp or other non-contingency products (Section 3.3).

3.1. Price adder-based approach: ERCOT

ERCOT has a unique method for imposing reserve scarcity penalties: an after-the-fact penalty that is added to the energy price.

3.1.1. ERCOT electricity market summary

ERCOT manages about 90% of total load in Texas by economically dispatching more than 710 generation units every 5 min [14]. The electricity market framework of ERCOT includes a co-optimized DA energy and operating reserve market, and a RT energy-only market. The DA is a financial-only market in which no operational obligations exist. In the RT market, ERCOT relies on high energy price signals during shortage conditions. During such conditions, demand curves are used to set the operating reserve price adders, which will be added to the energy price after the security-constrained economic dispatch (SCED) is run [15–17], as described in detail below.

ERCOT's operating reserve products include reg-up, reg-down, spinning reserve,[2] and 30-min reserve.[3] Operating reserves are procured in the DA market and, if necessary, can be adjusted in the supplemental operating reserve market, which is run only when the operator determines that reserves are likely to be short [14]. Since ERCOT's RT market does not co-optimize energy with reserves, there is no automatic premium on the energy price resulting from opportunity costs and reserve scarcity. Instead, ERCOT adds a premium to the energy price after the market solution based on its ORDCs.

This adder is intended to increase energy prices as the reserve margin is depleted. If the available reserve is 3000 MW or less (the so-called minimum reserve requirement $\boldsymbol{R_{req}}$), the VOLL, which is set to $5000 per MWh, is used as the price adder. If the available reserve exceeds 3000 MW, the price adder is equal to the estimated LOLP multiplied by the VOLL. LOLP depends on assumed distributions for demand and renewable energy forecast errors.

3.1.2. Current methodology for reserve shortage pricing in ERCOT's RT market

In ERCOT, available reserves are grouped as either fast (i.e., reserves that can respond at any time during the operating hour) or slow (i.e., reserves that only can respond in the last 30 min of the operating hour). Thus, two LOLP curves and price adders (i.e., a whole hour price adder $\boldsymbol{Y_{60}}$ and a half-hour price adder $\boldsymbol{Y_{30}}$) are calculated for each operating hour. The steps to calculate these price adders are as follows, and are unique to ERCOT:

1. Run the 5-min RT SCED only for energy (subject to fixed operating reserve requirements, determined prior to RT) to determine the LMPs (i.e., λ). Let $\boldsymbol{R_{60}}$ be the available reserve 1 h ahead, and $\boldsymbol{R_{rt}}$ be the actually realized RT reserve margin. A loss-of-load event is assumed to occur if $\boldsymbol{R_{rt}}$ is less than the $\boldsymbol{R_{req}}$ (which is 3000 MW in ERCOT). Thus, the estimate of LOLP that would be made 1 h ahead of time equals the probability that the ultimately realized $\boldsymbol{R_{rt}}$ turns out to be less than $\boldsymbol{R_{req}}$, conditioned on the known value of $\boldsymbol{R_{60}}$. VOLL is $5000/MWh if $\boldsymbol{R_{rt}} < \boldsymbol{R_{req}}$, whose probability is LOLP.
2. To estimate LOLP 1 h ahead of time, it is assumed that the reserve error $\boldsymbol{RE}$, which is the difference between the $\boldsymbol{R_{60}}$ and $\boldsymbol{R_{rt}}$ (i.e., the decline in reserves from hour-ahead to RT), is normally distributed. Using historical data, RE's mean $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$ are

[1] Consideration of deliverability of reserves recognizing a full network model is proposed for some reserves, such as the CAISO flexiramp, but is not yet implemented.

[2] It is called responsive reserves in ERCOT terminology.

[3] It is called non-spinning reserve in ERCOT terminology.

Table 1
Scarcity penalty and demand curve values as well as computational framework and estimation methods used by U.S. market operators (when reported in ISO documentation). Values are in \$/MW/hr or \$/MWh. "N/A" means that a reserve is either not procured in that market or there is no explicit shortage price. See Sections 2.2 and 2.3 for framework and method definitions. (Source: ISO Business Practices Manuals).

ISO	Computational Framework	Penalty for Requirement Violation/Demand Curve Prices (otherwise) [Penalty Estimation Method]					
		Energy (Where Price or Offer cap limit penalties)	Regulation	Primary: Spin	Primary: Non-spin	30-min	Flexiramp
ERCOT	Ex post Price Adder (RT Only)	\$5000 [1V] (Price cap)	N/A	Non-linear curve from 0 to VOLL = \$5000 [1V]	N/A	Non-linear curve from 0 to VOLL [1V]	N/A
MISO	Multi-step ORDC (for primary reserves)	\$3500 [1V] (Price cap)	Max of \$100 and monthly average peak unit proxy price [1C]	Two steps (for spin requirement, which can be met by spin & regulation): \$98 if scarcity >10% of requirement; \$65 if scarcity is <10% of requirement	Four steps up to \$3500 (for system-wide operating reserve requirement, which can be met by regulation, spin, and non-spin) [1V]	\$100 [2R]	Up and down ramp capability: \$5 [2R]
SPP	Multi-step ORDC (for regulation, primary reserves, and flexiramp)	\$2000 (Offer cap)	Multiple steps up to \$600 [1C]	\$200	Three steps up to \$1100 (for operating reserve requirement, which can be met by both spin and non-spin)	N/A	Six steps up to estimated cost of unit start-up [1C]
CAISO	Multi-step ORDC (for regulation and primary reserves)	\$1000 or \$2000 bid and price cap depending on Western US scarcity conditions	Reg Up \$200 and Reg Down \$700 if scarcity >84 MW, \$600 if > 32 and < 84 MW, \$500 if < 32 MW	\$100	\$700 if scarcity >210 MW, \$600 if > 70 and < 210 MW, \$500 if < 70 MW	N/A	Several steps up to \$247 for both up and down products (RT only) [1L]
NYISO (applicable to NY City region)	Multi-step ORDC (for regulation and 30-min reserve)	\$2000 (Offer cap)	Three steps: \$25 if scarcity <25 MW, \$525 if > 25 and < 80 MW, \$775 if > 80 MW	\$775 [1V]	\$750 [1V]	Nine steps up to \$750 [1V]	N/A
PJM	Multi-step ORDC (for primary)	\$2000 (Offer cap)	N/A	Two steps: \$300 if scarcity < requirement but > largest contingency, \$850 if < largest contingency	Two steps: \$300 if scarcity < requirement but >1.5*largest contingency, \$850 if < 1.5*largest contingency (for primary reserve requirement, which can be met by both spin and non-spin)	N/A	N/A
ISO-NE	Single-step Penalty Factor (RT Only)	\$2000 (Offer cap)	Maximum of zero and \$100 plus Energy Component of the RT LMP	\$50	\$1500	System-wide: \$1000 and \$250 (2-steps); Local: \$250	N/A

estimated, which defines its cumulative distribution function ($\boldsymbol{CDF_{RE}}$).

3. To be conservative, $\mathrm{CDF_{RE}}$ is then shifted to the right by a factor $\mathbf{0.5}\times\boldsymbol{\sigma}$. LOLP is then calculated as the conditional probability of $\boldsymbol{R_{rt}} < \boldsymbol{R_{req}}$, given $\boldsymbol{R_{60}}$ and $\mathrm{CDF_{RE}}$. LOLP thus depends on the reserves and the parameters of RE's distribution:

$$LOLP\left(R_{60}\middle|R_{req},\mu+0.5\times\sigma,\sigma\right)=1-CDF_{RE}\left(R_{60}-R_{req}\middle|\mu+0.5\times\sigma,\sigma\right) \quad (1)$$

4. The portion of the ORDC price adder corresponding to the hour-ahead reserves (defined as $\boldsymbol{Y_{60}(R_{sns})}$) that is added to the RT price $\boldsymbol{\lambda}$ is calculated as (2), where $\boldsymbol{R_{sns}}$ is the sum of hour-ahead fast and slow reserves.

$$Y_{60}(R_{sns})=\begin{cases} VOLL-\lambda \quad ; \ R_{sns}\le R_{req} \\ (VOLL-\lambda)\times LOLP\left(R_{sns}\middle|R_{req},\mu+0.5\times\sigma,\sigma\right) \ ; \ R_{sns}>R_{req}\end{cases} \quad (2)$$

If $\boldsymbol{R_{sns}} > \boldsymbol{R_{req}}$, the adder is the product of (a) the probability that reserves are inadequate in RT (given the amount of hour-ahead reserves), and (b) the difference between $\boldsymbol{VOLL}$ and the LMP, so as that probability increases, the sum of the LMP and this adder approaches VOLL. Note that the resulting RT energy price is therefore conservatively set at $\boldsymbol{VOLL}-\boldsymbol{\lambda}$ when $\boldsymbol{R_{sns}}\le\boldsymbol{R_{req}}$, as if there is a 100% probability of inadequate reserves and load curtailment in that situation, even though it might actually be the case that $\boldsymbol{LOLP}(\boldsymbol{R_{sns}}|\boldsymbol{R_{req}},\boldsymbol{\mu}+\mathbf{0.5}\times\boldsymbol{\sigma},\boldsymbol{\sigma}) < 1.0$. This explains why the vertical segment of the ORDC price adder curve shown in Fig. 1 has a vertical segment at $\boldsymbol{R_{sns}} = \boldsymbol{R_{req}}$.

5. In addition, the first half-hour price adder ($\boldsymbol{Y_{30}(R_s)}$), which is based on the distribution of a 30-min (rather than 60-min) RE with adjusted mean $\boldsymbol{\mu_{30}}$ ($=\mathbf{0.5}\times(\boldsymbol{\mu}+\mathbf{0.5}*\boldsymbol{\sigma})$) and standard deviation $\boldsymbol{\sigma_{30}}$ ($=\mathbf{0.707}\times\boldsymbol{\sigma}$), can be calculated as in (3) below, where $\boldsymbol{R_s}$ is the available fast reserves. The adjustments in the mean and standard deviation are necessary because the original $\boldsymbol{RE}$ is calculated hourly, while only events within the first 30 min are relevant to the calculation of $\boldsymbol{Y_{30}}$. Note that the available fast reserves are the only ones considered to be available to respond to any event that happens in the 30 min prior to real-time.

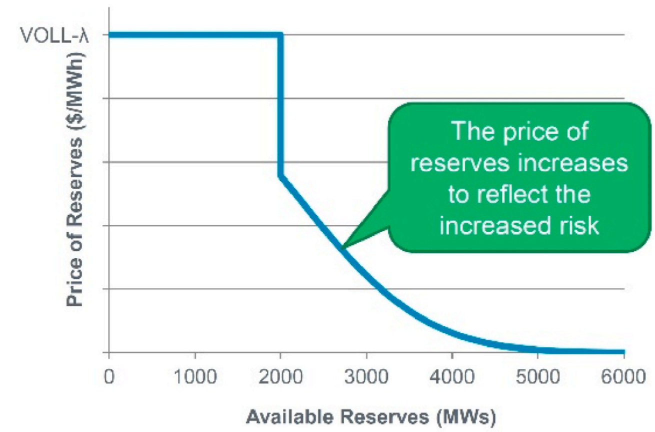

Fig. 1. Shape of the calculated ORDC curve in the ERCOT electricity market (Assumes $R_{req} = 2000$ MW).

$$Y_{30}(R_s) = \begin{cases} VOLL - \lambda; R_s \le R_{req} \\ (VOLL - \lambda) \times LOLP\left(R_s \middle| R_{req}, \mu_{30}, \sigma_{30}\right); R_s > R_{req} \end{cases} \tag{3}$$

6. Finally, the total reserve price adder is then the average of Y_{60} and Y_{30}. The general shape of the final ORDC price adder curve is shown in Fig. 1.

There have been several adjustments in the LOLP portion of the ORDC in recent years. First, based on a regulatory order in January 2019, ERCOT was directed to use a single blended ORDC and to implement a rightward shift equal to 0.25 times the standard deviation in the LOLP calculation (i.e., to be more conservative in the calculation of expected reserves). Second, that shift was doubled to 0.5 times the standard deviation on March 1, 2020. As shown in Fig. 2, these shifts accelerate price increases toward the VOLL whenever there is a shortage in reserves. Third, after the February 2021 cold event in which extreme prices were experienced for many hours, the value of VOLL was decreased from $9000 to $5000.

3.1.3. *ERCOT future plan for shortage pricing*

As of this writing, ERCOT has announced that it will implement co-optimization of energy and operating reserves in the RT market by 2025. That is, the scheduling of resources to provide energy or operating reserve will be optimized in a RT market every 5 min, similar to the other six U.S. ISOs. There will be no ORDC changes before the implementation of RT co-optimization, when the ex post ORDC adders will be retired [14].

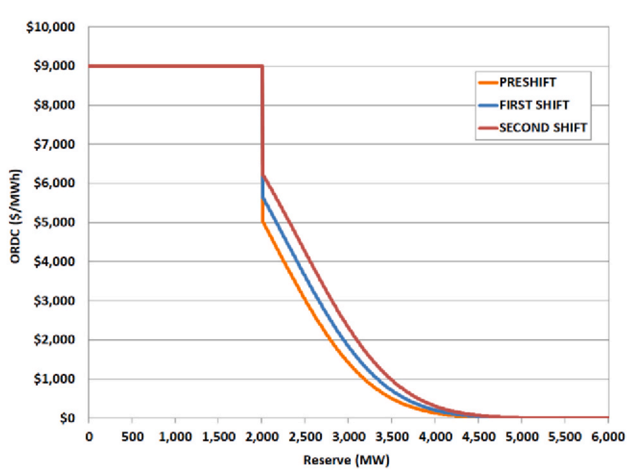

Fig. 2. ERCOT's blended ORDC before and after the 2019 and 2020 shifts ordered by the Texas regulator [14] (the VOLL = $9000 and MRR = $2000 parameters were used before the February 2021 extreme weather event after which they changed to VOLL = $5000 and MRR = $3000).

3.2. *Co-optimized energy and reserve: stepwise ORDCs with flexiramp product*

For the three ISOs that have a flexiramp product, the effect of having penalties or a demand curve for flexiramp is to extend the "tail" to the right for the scarcity premium for energy, relative to those ISOs without such a product (Section 3.3). Those three ISOs determine the penalty for flexiramp shortages in three distinct ways: an arbitrary penalty assuming ramp is the least-important reserve (method 2R, MISO), average start-up cost for short-start generators (method 1C, SPP), and the probability of violating the energy balance due to ramp shortage, times the energy imbalance penalty (method 1L, CAISO).

3.2.1. *MISO*

3.2.1.1. MISO electricity market summary. The Midcontinent Independent System Operator (MISO) manages the bulk power system across fifteen states and the Canadian province of Manitoba [18]. MISO procures six products in both the DA and RT markets, including energy, regulation ("regulating reserve"), spin, non-spin,[4] 30-min reserve, and flexiramp.[5] Three of these (regulation, spin, and non-spin) are subject to nested constraints that allow higher quality reserves to substitute for lower quality reserves; therefore, shortages of lower quality products can induce scarcity prices for the higher quality products. For conciseness, those three products are collectively referred to as "immediate reserves."[6]

The hierarchy of MISO's immediate reserve products is shown in Fig. 3. The 30-min reserve, which is not included in those nested constraints, meets market-wide, sub-regional, and local flexibility requirements.

As shown in the above reserve hierarchy of MISO, immediate reserves are divided into regulation and contingency reserve. Regulation is comprised of generation-based or other market resources qualified for regulation. Contingency reserve consists of spin and non-spin reserves, which both could be based on either generation or other market resources qualified for these products. In MISO, the market-wide requirement for immediate reserves is the sum of the market-wide requirements of regulation and contingency reserves [19,20], and excludes 30-min reserves and flexiramp.

3.2.1.2. Current practices for reserve shortage pricing in MISO. MISO's market-wide ORDC for total immediate reserves in both RT and DA markets is based on the following rules [19,21,22].

- If the immediate reserve cleared in the market exceeds the market-wide requirement, the ORDC price equals zero.
- If the amount cleared is within [89%, 100%) of the market-wide requirement,[7] the ORDC price is set to 1100 $/MWh, which is the sum of the soft cap on energy offers (i.e., 1000 $/MWh for MISO) and price cap of contingency reserve offer (i.e., 100 $/MWh).
- If the amount cleared in the market falls within [4%, 89%) of the requirement, the market optimization automatically applies an

[4] Termed supplemental reserve by MISO.
[5] Called ramp capability product in MISO terminology.
[6] MISO officially refers to these three as operating reserves. To clearly distinguish these from 30-min reserves and flexiramp (which in other contexts are considered "operating reserves"), they are instead called "immediate reserves."
[7] Chosen because the single largest contingency in MISO can be currently addressed by 89% of the market-wide requirement [19].

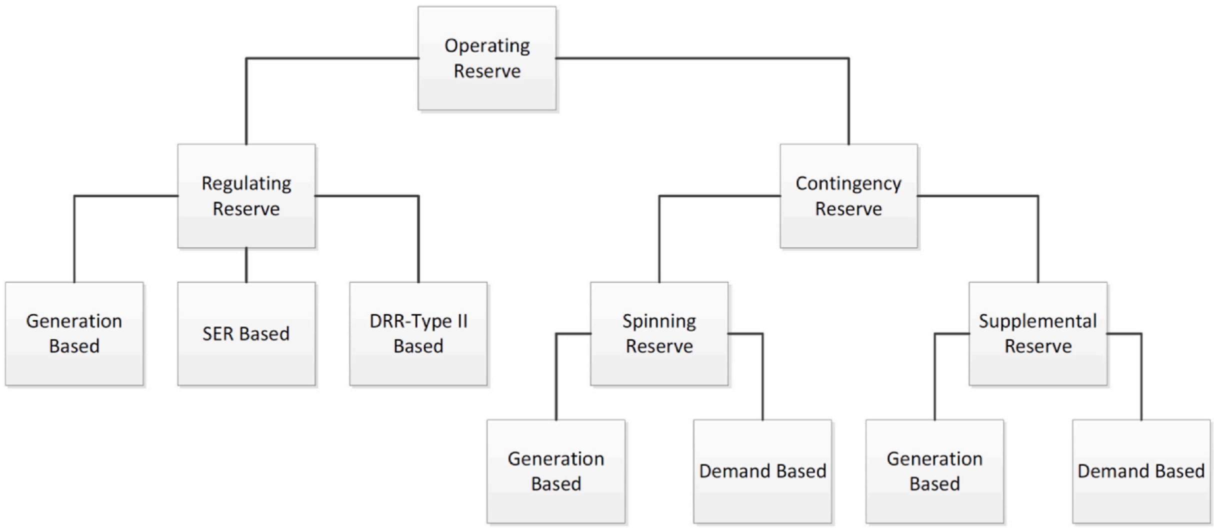

Fig. 3. The hierarchy of MISO's immediate reserve products [19].

ORDC price that is equal to the product of VOLL and the conditional probability that a loss of load occurs due to a single resource contingency, which are inputs to the market software. To calculate this conditional probability, resource contingencies are assumed to have equal probabilities and only contingencies of 100 MW or greater are considered. In addition, the maximum ORDC price is equal to VOLL (i.e., 3500 $/MWh) minus the regulation demand curve price, while the minimum ORDC price is equal to 2100 $/MWh, which is the sum of the hard energy offer cap (2000 $/MWh) and the cap on offers for contingency reserve (i.e., 100 $/MWh).

- If the immediate reserve cleared in the market is less than 4% of the market-wide requirement, the ORDC price equals VOLL minus the regulation demand curve price.

Fig. 4 shows an example of the resulting ORDC in red, assuming that the market-wide requirement is 2500 MW. The figure also shows an alternative ORDC recommended by MISO's independent market monitor (in blue). The monitor has argued that the current ORDC underestimates the value of reserves and VOLL and proposes a revised ORDC based on a VOLL of 10,000 $/MWh and with a different slope. As the figure shows, this would represent a several-fold increase in scarcity prices if the reserves are less than about half of the requirement (here, 1250 MW) [20,23].

Although the independent market monitor favors a significant increase in the ORDC cap, the monitor provides three reasons for suggesting that this value be no more than 10,000 $/MWh. First, as shown in Fig. 4, by far most shortages are minor and result in prices no greater than $1100. Second, since MISO's neighbors have lower shortage prices (see Fig. 4 for the energy settlement price of PJM during shortage), having a shortage pricing of more than 10,000 $/MWh could result in inefficient interchange. Third, having a shortage pricing with higher price levels could cause MISO's dispatch software to inefficiently tradeoffs between reserve procurement and managing flows on network constraints [20,23].

3.2.1.3. Flexiramp in MISO. MISO also defines demand curves for

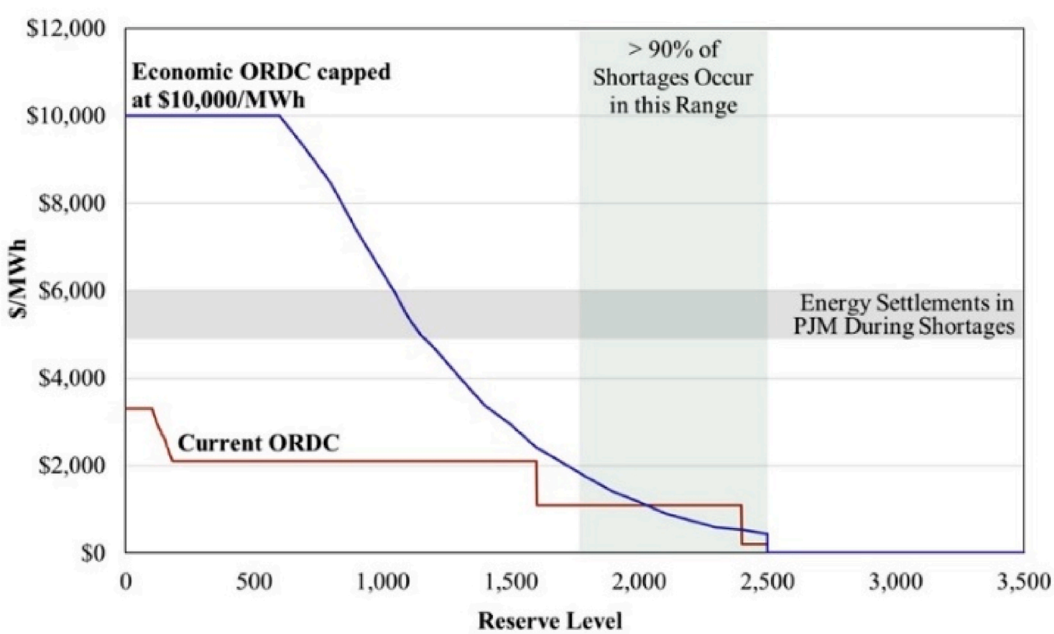

Fig. 4. Comparison of MISO's current ORDC (in red) and the market monitor's suggested economic ORDC (in blue) for MISO [20].

flexiramp products (i.e., the ability of the system to quickly change output), including market-wide flexiramp capability in both the up and down directions. A shortage in flexiramp-up/down, which invokes scarcity pricing, indicates that there is a deficit in the ability of the whole market to adapt to extreme net load ramps.

The MISO-wide flexiramp demand curve is a single-step with a penalty of $5/MWh for not meeting the MISO-wide requirement [19], which is by far the lowest among the three ISOs procuring that product. MISO is re-evaluating flexiramp-up requirement and demand curve based on net input uncertainty and generation outage/derate uncertainties within 30-min and 3-h respectively [23].

3.2.2. SPP

3.2.2.1. SPP electricity market summary. The Southwest Power Pool (SPP) manages the grid and wholesale electricity market of 14 states in the central U.S [24]. SPP co-optimizes energy and operating reserves and settles the DA market on an hourly basis [25] and performs the same function in its RT market every 5 min [26]. SPP reserve products include reg-up/reg-down, spinning reserve, non-spinning[8] reserve, and flexiramp-up/down[9] in both DA and RT markets. Spin and non-spin have a response time of 10 min and are kept for contingency situations [27]. Flexiramp-up/down response times are similarly set to 10 min, and the hourly required amounts will be posted DA by the transmission provider for the SPP balancing authority area.

3.2.2.2. Current practices for reserve shortage pricing in SPP. SPP uses demand curves in both its DA and RT markets to set the reserve price when there is a reserve shortage. Currently, the SPP's reg-up/reg-down demand curves have four steps, which in March 2022 started at $169/MWh[10] with a maximum of $600/MWh.

Meanwhile, the demand curve for operation (contingency) reserve has three steps beginning at $275/MWh for the system-wide target of 1220 MW, and increasing to $1100/MWh if reserves fall below 974 MW (see Fig. 5) [28].

The operating reserve requirement can also be satisfied by reg-up. Thus, since reg-up capacity appears in both the operating reserve and reg-up constraints, the total scarcity premium for reg-up can be as high as $1700/MWh, the sum of the two intercepts in Fig. 5 [27]. Furthermore, SPP has a spin requirement of $200/MWh, which increases the maximum scarcity price up to $1900/MWh.

In 2020, SPP experienced 780 RT intervals (each 5 min in duration) with reg-up shortage, 440 intervals with reg-down shortage, and 120 intervals with operation (contingency) reserve shortage. The number of intervals with reserve shortage condition in SPP's RT market and the monthly average prices for each product are shown in Fig. 6 [27]. Scarcity pricing for flexiramp is discussed next.

3.2.2.3. Flexiramp in SPP. Flexiramp is procured by SPP in both its DA and RT markets. The requirements for flexiramp-up/down are calculated based on forecasted net load changes and historical distributions of net load forecast error. The requirements are posted in DA for each hour of the operating day.

In the situation that there is flexiramp up scarcity in either RT or DA, SPP follows the below rules for defining the demand curve (the same logic and parameters apply to flexiramp down pricing in a scarcity situation) [29].

1. The maximum scarcity price is calculated each month using the average cost per MWh based on the sum of the following energy offer data from all eligible resources for the last three months: cold start-up (in 10 min or less), no-load offer, and energy at minimum cost.
2. The minimum value for flexiramp up price in a shortage situation is $10/MWh (it is $0 for flexiramp down), and the demand curve levels are shown in Table 2.

3.2.3. CAISO

3.2.3.1. CAISO electricity market summary. The California ISO (CAISO), which is responsible for about 80% of California's electric energy system, is among the largest operators in the world. The sequence and timing of CAISO's electricity markets are shown in Fig. 7. Besides a DA market with an hourly time step, in RT has three markets: an hour-ahead scheduling process for imports; a security-constrained unit commitment (SCUC) every 15 min; and an SCED every 5 min to dispatch the cheapest generating units [30].

CAISO's operating reserves include reg-up/reg-down, spinning and non-spinning[11] reserves, and flexiramp. Penalties or demand curves apply if there are shortages of any products, and certain nesting relationships among the products mean that some particular products (and ultimately energy) can have a scarcity premium well in excess of the highest penalty on its own demand curve [31]. All required reserves, except flexiramp, are fully procured in the DA market, co-optimized with energy. Flexiramp, in contrast to spin and non-spin capacity, is procured in one RT interval to be economically dispatched for energy in the subsequent interval, with no contingency-based restrictions. Since flexiramp is not acquired DA, the full amounts are procured in the 15-min market (FMM), rather than just imbalances from the DA position (as SPP does). CAISO procures flexiramp and energy in RT for all other balancing authorities in the west-wide Energy Imbalance Market (EIM), but no other operating reserves. Thus, our discussion of scarcity pricing below applies only to pricing in the CAISO's balancing area, and not areas for other EIM members [32,33].

3.2.3.2. Current methodology for reserve shortage pricing in CAISO. CAISO uses nesting scarcity penalties for both upward reserve products (i.e., reg-up and spin/non-spin) and downward reserve product (i.e., reg-down). Those penalties depend on the maximum allowed energy bid, and are shown in Table 3 [34]. Note that the penalties form multistep demand curves for some products, such as for non-spin which has three levels of penalties and thus a 3-step ORDC for that product. Due to the nesting relationships which make reg-up the highest quality upward reserve, total scarcity premiums of $500, $600, $700, $800, and $1000/MWh can occur for reg-up if any of the upward reserves are short and their penalties contribute to the price. (Section 4's quantitative example illustrates this point.) Because of co-optimization, these premiums translate into similar scarcity premiums for energy over and above the marginal cost of energy supply.

As an example, consider the extreme situation with a non-spin reserve shortage exceeding 210 MW, a reg-down shortage over 84 MW, and the energy offer cap of $2000/MWh. Based on the scarcity reserve demand curve of Table 3, the shadow prices for reg-up, spin, non-spin, and regulation down would then be 2000, 1600, 1400, and 1400 $/MWh, respectively.

It should be noted that under the co-optimization used in the CAISO markets (and most of the other ISOs' markets), reserve scarcity can result in increases in energy prices. For example, if reserves are scarce and the ORDC is setting the price, then when a resource's capacity is scheduled for reserves, diverting it to produce one extra MW of energy

[8] It is called supplemental reserve in SPP terminology.

[9] It is called ramp capability up/down in SPP terminology.

[10] The first step (i.e., base step) for reg-up/reg-down is based on the average estimated cost to commit a resource and is calculated each month using the previous three months historical data.

[11] Spinning and non-spinning reserves are referred as contingency reserves in CAISO terminology.

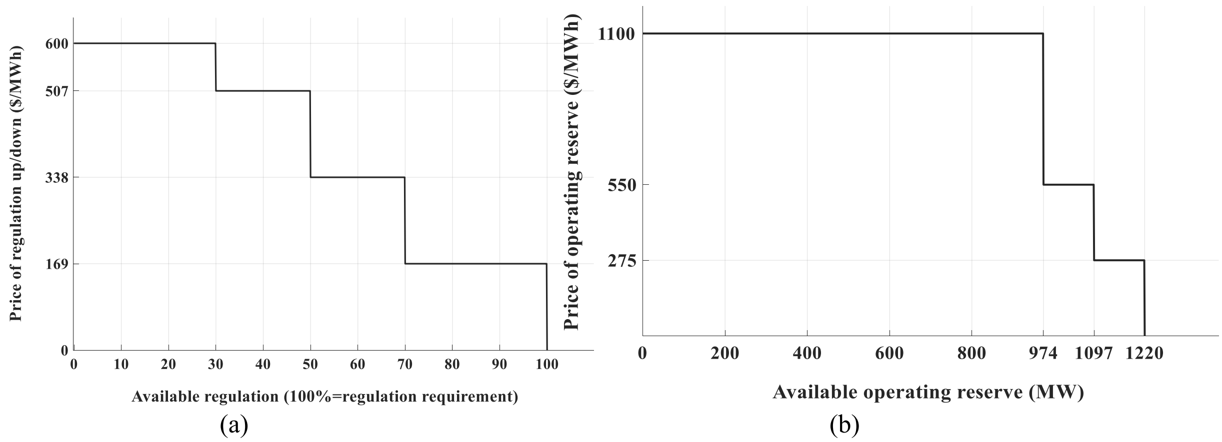

Fig. 5. SPP's reserve scarcity demand curves as of March 2022; (a) regulation and (b) operating reserve (based on the values reported by Ref. [28]).

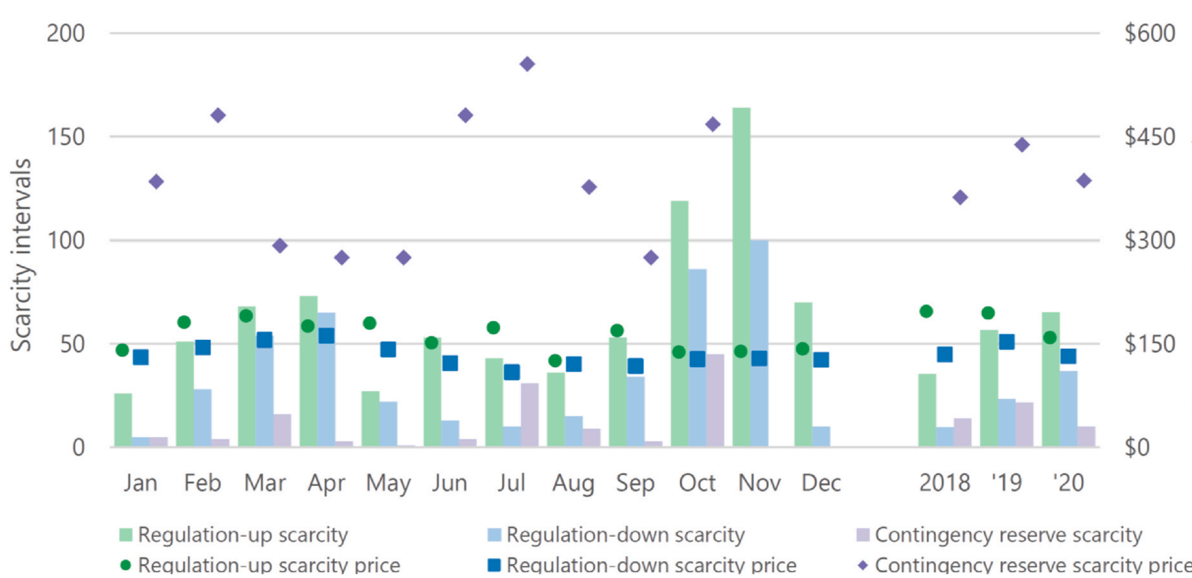

Fig. 6. Shortage intervals and energy clearing prices during those intervals within SPP's RT market in 2020 [27].

Table 2
SPP's demand curve for flexiramp up and down in shortage situation.

Amount of shortage	Scarcity Price[a]
shortage ≤ 5% *of requirement*	1/6 *of maximum scarcity price*
5% *of requirement* < *shortage* ≤ 10% *of requirement*	2/6 *of maximum scarcity price*
10% *of requirement* < *shortage* ≤ 15% *of requirement*	3/6 *of maximum scarcity price*
15% *of requirement* < *shortage* ≤ 25% *of requirement*	4/6 *of maximum scarcity price*
25% *of requirement* < *shortage* ≤ 40% *of requirement*	5/6 *of maximum scarcity price*
40% *of requirement* < *shortage*	6/6 *of maximum scarcity price*

[a] For flexiramp up, the penalty is the maximum of $10/MWh and the value based on start-up, Pmin, and energy offers described in the text.

will incur two costs which summed together equal the energy price: the marginal fuel cost to produce energy plus the value of the capacity for reserves (which is a foregone revenue for the generator, or opportunity cost). Thus, the overall marginal cost of providing energy for that resource has been increased by the reserve shortage price to compensate for the loss of profit in the reserve market which is the shadow price of scarce operating reserve set by the penalties of Table 3.

3.2.3.3. Flexiramp in CAISO. To address shortages of RT ramp capability that may cause price spikes, flexiramp is procured by modeling 5-min ramp needs in both FMM and RT economic dispatch [35].

The flexiramp demand curve is calculated based on the net load ramp uncertainty. Upward and downward uncertainty requirements, which are posted DA, are calculated using a histogram of recent net load forecasting errors for the relevant hour. The 2.5th and 97.5th percentiles from the error histograms, when added to the expected net load ramp from one 5-min interval to the next, define the down- and up-flexiramp requirements, respectively.

The calculation of the flexiramp-up demand curve is described here. Assume that the probability of the forecast error X falling in range i (i.e., $[X_i, X_{i+1}]$) is P_i. Let the forecast net load used in the 5-min market run for the current interval (one interval in the future) and subsequent interval

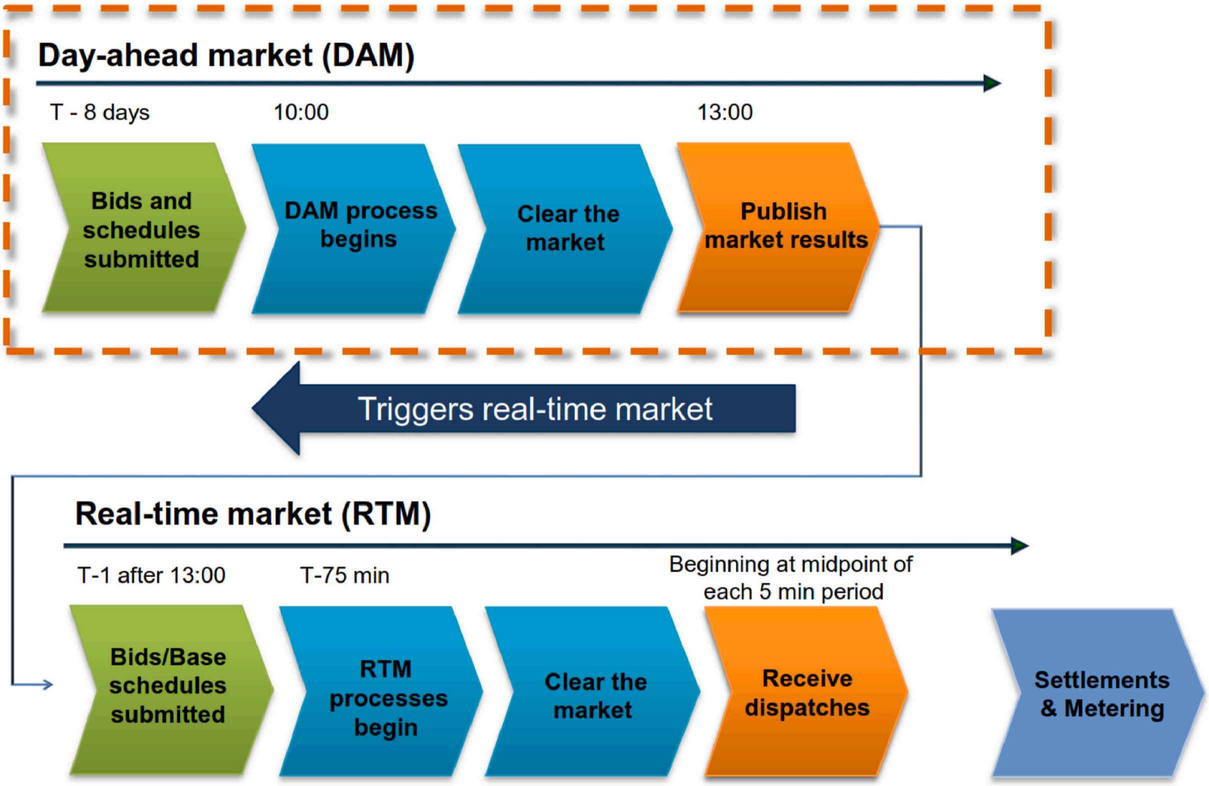

Fig. 7. Timeframe of DA and RT markets in CAISO [31].

Table 3
Scarcity reserve demand curve values in CAISO. (Based on \$1000/MWh offer cap for energy; demand curve values doubled if the \$2000/MWh is the maximum offer, when triggered by west-wide scarcity) [34].

Reserve	Demand Curve Value (\$/MWh)	
	Percent of Max Energy Bid Price	*Max Energy Bid Price = \$1000/MWh*
Reg-Up	20%	\$200
Spinning	10%	\$100
Non-Spinning		
210 MW < Shortage	70%	\$700
70 MW < Shortage ≤ 210 MW	60%	\$600
Shortage ≤ 70 MW	50%	\$500
Upward Sum	100%	\$1000
Reg-Down		
84 MW < Shortage	70%	\$700
32 MW < Shortage ≤ 84 MW	60%	\$600
Shortage ≤ 32 MW	50%	500

(two intervals out) be L_t and L_{t+1}, respectively. The total expected load imbalance penalty (ELIP, in total \$) in the upward direction, which is expected if FRU MW of flexiramp-up is procured, is as follows (assuming a load imbalance penalty \$1000/MWh, which as noted is doubled in times of west-wide scarcity) [35]:

$$\mathrm{ELIP(FRU)} = \sum_{\forall i} 1000 \times P_i \times \max\left\{0, \frac{X_i + X_{i+1}}{2} - (\mathrm{FRU} - (L_{t+1} - L_t))\right\} \tag{4}$$

The total procured flexiramp-up FRU is the total upward movement from L_t that can be accommodated. This upward movement includes the expected upward movement of load ($L_{t+1} - L_t$) plus the 97.5th percentile of the forecast distribution error. The marginal reduction in ELIP per MW of FRU is used to define the demand curve (i.e., operator willingness to pay for FRU), subject to a maximum of \$247/MWh or \$494/MWh (depending on whether the \$1000 or \$2000/MWh energy offer cap applies). That marginal value is simply the derivative of (4), which is \$1000/MWh times the probability that the ramp exceeds FRU. A similar procedure is used to calculate the flexiramp-down demand curve. The flexiramp product demand curve results in a rightward extension of the scarcity premium curve for energy, thus determining the scarcity premium for energy when the total available capacity is slightly short of the amount needed to meet energy and all operating reserves, including the flexiramp requirement.

3.3. *Co-optimized energy and reserve: stepwise ORDCs without flexiramp product*

These three ISO markets lack a flexiramp product. As a result, the long right-side tails on the energy premium curve that the markets of Section 3.2 have been lacking in the below markets. The ISO-NE market differs from the other two in that none of its individual reserve products has a multistep demand curve, although because reserves of different quality are nested in the market model, lower quality reserve shortages will induce multiple steps in the scarcity premiums for higher quality reserves, as illustrated in Section 4.

3.3.1. *NYISO*

3.3.1.1. NYISO electricity market summary. The New York Independent System Operator (NYISO) administers a market that includes 500 generating units and 11,000 miles of transmission lines across the state of New York [36]. In the DA market of NYISO, a SCUC is executed to simultaneously co-optimize energy and reserves [37]. The reserve constraints are nested, such that reg-up and spin together meet the spin constraint; reg-up, spin, and non-spin meet the "10-min reserve" constraint, and all four reserves meet the total demand for reserves that can respond within 30-min [38]. Similar to the DA market, energy and reserves are co-optimized simultaneously in the NYISO's RT market. The

Table 4
ORDC values in NYISO [40].

New York Region	Type	Shortage Relative to Requirement (MW)	Demand Curve Price ($/MW)
NYCA	Regulation	Up to 25 MW	$25
		At least 25 MW up to 80 MW	$525
		80 MW or more	$775
	Spinning Reserve	Any shortage	$775
	10 Minute Reserve	Any shortage	$750
	30 Minute Reserve	Up to 200 MW	$40
		At least 200 MW, up to 325 MW	$100
		At least 325 MW, up to 380 MW	$175
		At least 380 MW, up to 435 MW	$225
		At least 435 MW, up to 490 MW	$300
		At least 490 MW, up to 545 MW	$375
		At least 545 MW, up to 600 MW	$500
		At least 600 MW, up to 655 MW	$625
		655 MW or more	$750
Eastern New York (EAST)	Spinning Reserve	Any shortage	$40
	10 Minute Reserve	Any shortage	$775
	30 Minute Reserve	Any shortage	$40
Southeastern New York (SENY)	Spinning Reserve	Any shortage	$40
	10 Minute Reserve	Any shortage	$40
	30 Minute Reserve	Up to the applicable incremental SENY 30-min reserve quantity (If applicable, a quantity not exceeding 500 MW)	$40
		The applicable incremental SENY 30-min reserve quantity or more (or any shortage if no incremental SENY 30- minute reserve quantity is applicable)	$500
New York City (NYC)	Spinning Reserve	Any shortage	$25
	10 Minute Reserve	Any shortage	$25
	30 Minute Reserve	Any shortage	$25
Long Island (LI)	Spinning Reserve	Any shortage	$25
	10 Minute Reserve	Any shortage	$25
	30 Minute Reserve	Any shortage	$25

RT scheduling system includes a RT commitment subsystem and a RT dispatch subsystem [38,39].

3.3.1.2. Current methodology for reserve shortage pricing in NYISO. Currently, NYISO has stepwise ORDCs for each reserve type differentiated by location (reserve zones). These ORDCs are considered by both the DA and RT markets. Separately, a system-wide regulation demand curve is considered in both DA and RT markets to reflect shortages of that product. The ORDC steps for different reserve products in different NYISO's reserve zones are shown in Table 4. These steps are considered in both DA and RT unit commitment runs and RT dispatch, except when load is shed in RT under the emergency demand response program [40,41].

Fig. 8 summarizes the operating reserve zones and requirements for NYISO. To illustrate the nesting behavior of reserve pricing mechanism in shortage conditions, assume the case that a shortage of 50 MW of 30-min reserve occurs in NYC, SENY, East, and NYCA regions. Since SENY, East, and NYCA regions contain the NYC reserve region (see Fig. 8), a unit in NYC that can provide 30-min reserve will receive a total reserve price of $605/MWh, which is obtained by adding prices for 30-min reserve in NYCA ($40/MWh), East ($40/MWh), SENY ($500/MWh), and NYC ($25/MWh) [37,41]. In addition to the nesting of zonal and system reserve constraints, the nesting of different types of reserves as noted in the previous paragraph can yield much higher scarcity premium for higher quality reserves than shown in the above table. Via co-optimization, these can translate into very high premiums for energy, as noted in the numerical example of Section 4.

3.3.2. PJM

3.3.2.1. PJM electricity market summary. PJM coordinates the electricity market of all or parts of its 13 states and the District of Columbia [42]. In the PJM DA market, the DA hourly schedules and LMPs obtained respectively from the dispatch run and the pricing run of the DA market are binding financial commitments. In the PJM RT market, after forecasting system conditions, a RT SCED program is run to find the least-cost dispatch to balance supply and demand. Then, based on the latest RT SCED solution, a locational price calculator (LPC) program, in which an integer relaxation is applied for designated fast-start units, is run every 5 min to find the RT LMPs and clearing prices for regulation and reserves [43].

In fact, the LPC program is a linear co-optimization model that jointly optimizes energy and operating reserves and determines their prices. Its objective function is the minimization of energy and reserve costs (as bid by the resources) subject to a power balance constraint, reserve requirement constraints, and transmission constraints. Finally, according to the MW deviations of RT dispatch run compared to the DA schedules, and RT prices obtained from the pricing run, the balancing settlement is calculated for each 5-min interval (i.e., RT settlement interval). The relationship of SCED and LPC to separate processes for optimizing regulation/inflexible reserves and advisory intermediate-term scheduling (IT SCED) in PJM's RT decision process is shown in Fig. 9.

PJM markets for operating reserves include regulation, spin, and

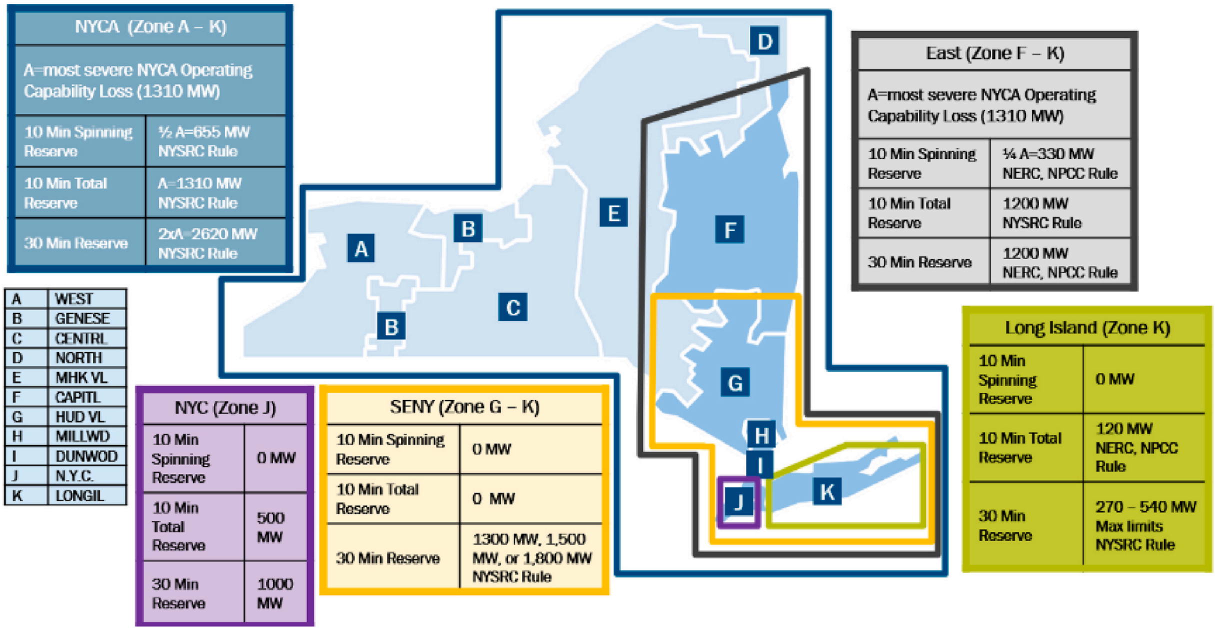

Fig. 8. NYISO's operating reserve zones and requirements [37].

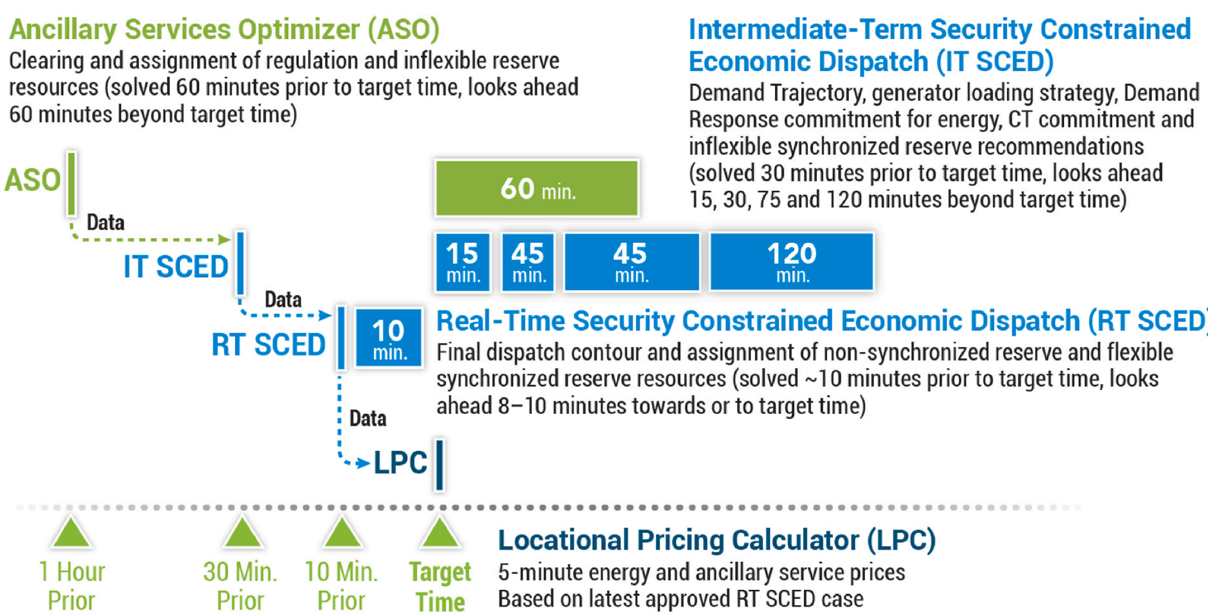

Fig. 9. PJM RT market clearing engines [42].

non-spin[12] markets [44,45]. Co-optimization of energy and operating reserves is first performed by the ancillary service optimizer (ASO) program of PJM in the RT markets 60 min prior to the target interval in the dispatch run (see Fig. 9). The inputs of ASO are the reserve offers as well as energy offers and resource schedules from the market gateway system. Given the inputs, ASO optimizes the RT dispatch and forecasts LMPs to find the commitment and reserve share of flexible and inflexible units to meet operating reserve requirements for 1-h time intervals. Only assignments for regulation and inflexible units for reserves are carried forward. It is worth mentioning that ASO is not responsible for finding the market clearing prices for energy and operating reserves. Then, the RT SCED program (see Fig. 9) co-optimizes energy and operating reserves including the inflexible units committed by ASO and fast-start resources committed by IT SCED. Finally, the LPC program (Fig. 9) calculates the clearing price for operating reserves in 5-min intervals.

3.3.2.2. Current methodology for shortage pricing in PJM. To illustrate how reserve shortages can affect energy premiums, a summary of how LPC works is presented here. First, it determines if either a spin or primary reserve (spin plus non-spin) shortage exists. If there is such a shortage, penalty factors are considered in the RT reserve pricing calculations and, ultimately, energy LMPs due to co-optimization. During a RT reserve shortage condition, all LMP calculations, including scarcity premiums, are based on the LPC.

The LMP during reserve scarcity can go up to the energy offer cap of $2000/MWh plus any penalty factor for reserves that are short. These

[12] PJM calls the last two synchronized and non-synchronized reserves, respectively.

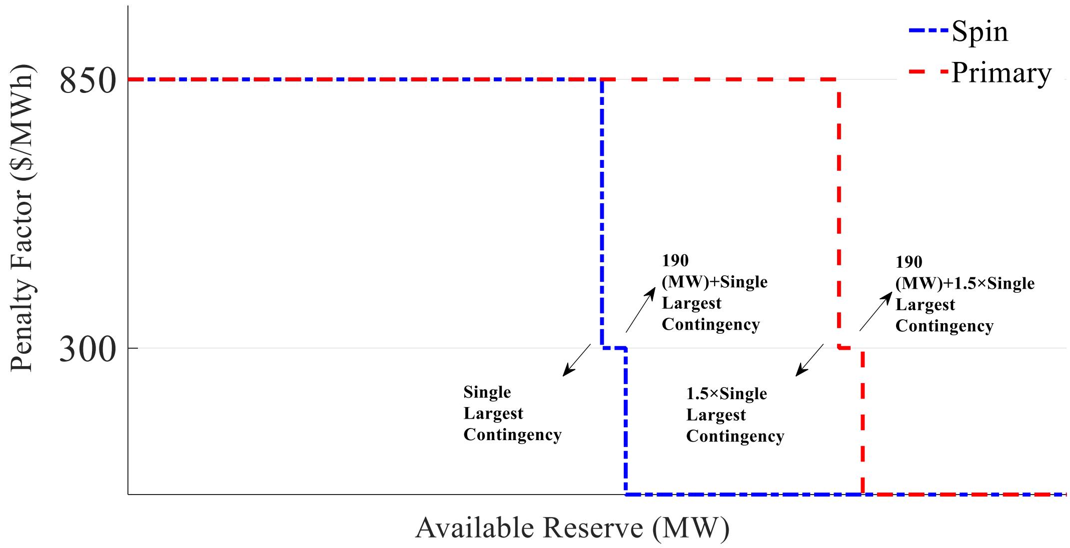

Fig. 10. PJM stepwise penalty factors for spin and primary reserves (based on the values reported by Refs. [43,46]).

Table 5
Fixed penalty factors in ISO-NE market [51].

Requirement	Requirement Sub-Category	Penalty Factors ($/MWh)
System 10-min Spin		50
Total System 10-min Reserve		1500
Minimum Total System 30-min Reserve	Minimum 30-min reserve	1000
Total System 30-min Reserve	Replacement Reserves	250
Local 30-min Reserve		250

values are currently obtained from two-step demand curves for spin and primary reserves [43].

PJM has four different reserve demand curves that are relevant for each product in each reserve zone/sub-zone: RTO primary reserve, Mid-Atlantic Dominion (MAD) primary reserve, RTO spinning reserve, and MAD spinning reserve. The penalty factors on the Y-axis of the reserve demand curves are the same for all four; however, the reserve requirement values on the X-axis are different. Fig. 10 illustrates the general form of the stepwise PJM reserve demand curves. The steps of these reserve demand curves are as follows.

- Step 1: the penalty factor value for the first step is $850/MWh. The MW requirements for spin and primary reserve equal 100% and 150% of the output of the largest individual contingency in the reserve zone/sub-zone, respectively.
- Step 2: the penalty factor value for the second step is $300/MWh. The reserve requirement for the second step is equal to the value of reserve requirement in the first step plus 190 MW, plus any additional reserves that operators decide are needed in case of heavy load conditions [43,46].

Since the reserve requirements in the PJM RT market are based on the single largest contingency, the values on the X-axis of the reserve demand curve, which is used in the market clearing, can change with each market clearing engine run depending on what contingency is relevant at that time.

Due to the imposition of both system and local constraints, the total scarcity penalty for spin can be as much as 4 × $850/MWh or $3400/MWh. However, the total scarcity penalty for spin is capped at 2 × $850/MWh or $1700/MWh. Similarly, due to the imposition of both system and local constraints, the total scarcity penalty for non-spin can be as much as 2 × $850/MWh or $1700/MWh. However, the total scarcity penalty for non-spin is capped at $850/MWh. These premiums propagate over to energy prices due to co-optimization and can raise LMPs by those amounts.

3.3.3. ISO-NE

3.3.3.1. ISO-NE electricity market summary. ISO New England (ISO-NE) is the operator for the region spanning Massachusetts, New Hampshire, Vermont, Rhode Island, Connecticut, and most of Maine [47]. In the DA market, ISO-NE determines the economically optimal way to meet the demand of energy and reserve requirements using SCUC and SCED. The RT market co-optimizes energy together with each of the reserve products. In order of highest to lowest quality, ISO-NE's reserve products are regulation, spin, non-spin, 30-min reserves, and local 30-min reserve [48,49].

The forward reserve market, which occurs twice a year, is for the contracting of non-spin and 30-min reserves requirements within the New England reserve zones. The local forward reserve requirement for each reserve zone is determined with respect to the 95th percentile value obtained from historical requirement data of the last two forward reserve procurement. For the New England control area, the forward reserve requirements are based on the first contingency of 1636 MW and the second contingency of 1263 MW in the winter 2021–2022 forward reserve market [48,49]. Forward operating reserve is procured in the forward reserve market. If a market participant fails to reserve sufficient forward reserve in the RT energy market to meet its forward-reserve obligation, it is assessed a forward-reserve failure-to-reserve penalty for each hour of the operating day during which the failure to reserve occurs. These penalties, however, do not propagate as scarcity premiums to spot market energy prices.

In addition, RT reserves are also procured by ISO-NE during the operating day prior to RT, but providers are paid based on the 5-min RT reserve clearing price for each product in each reserve zone [50,51].

3.3.3.2. Current practices for reserve shortage pricing in ISO-NE. When experiencing a reserve shortage situation, the ISO-NE co-optimization results in adding administratively-set penalty factors to the reserve and possibly energy prices [52]. As in other co-optimized markets, when reserve prices are above zero, it means that they are set by scheduled reserves with the highest opportunity cost unless the penalty factors are lower.[13] The penalty factors for each reserve product are shown in Table 5. The reserve products listed in Table 5 are ordered from highest-quality to lowest-quality. Because of the co-optimization, which makes the energy and reserve requirements interdependent, reserve penalty factors will likely propagate to create a scarcity premium for energy prices.

In common with all co-optimization-based scarcity pricing systems, nesting constraints that allow higher quality reserves to meet requirements for lower quality reserves result in penalty factors being additive. During shortage situations, depletion of lower-quality reserve products affects the prices of energy and higher-quality reserve products. Therefore, the price of all products will move in the same direction together. For instance, assume the case that the marginal offer for energy is \$1000/MWh. In this case, based on the penalties in Table 5, the energy price is \$3550/MWh when spin is short, but only \$2000/MWh when the 30-min reserve is short [51].

3.3.4. *Summary of comparison*

The above comparison of scarcity pricing practices of these seven U. S. ISOs demonstrates that fundamental pricing approaches and specific sources for shortage penalty values significantly differ. While ERCOT adopts an ex post price adder, others use stepwise demand curves for reserve scarcity pricing in their co-optimization procedure. Furthermore, it can be concluded that in current practice, the most widely used way to calculate the marginal reliability value of reserves at any shortage level is to consider the cost of system operators' actions in case of a reserve shortage for reducing the risk of uncontrolled load shedding (see Table 1). These actions can range from committing additional resources (Penalty Method 1C), leaning on importing power from neighboring systems (Penalty Method 1L), or, if there is no alternative, curtailing load involuntarily but in a controlled fashion (Penalty Method 1V). The practical implications of these differences for scarcity prices in the markets are numerically illustrated in the next section.

4. Comparative numerical example

In this section, a generic numerical example is presented to illustrate the practical implications of differences in the elements and calculation approaches for reserve scarcity pricing across the U.S. These scarcity premiums directly translate into energy price premiums (either added after the fact, as in ERCOT, or through co-optimization in the market software), which is portrayed in the figure below. To facilitate the comparison, the assumptions regarding generation technology and reserve needs for each system are as similar as possible. The following subsections present assumptions, graphical results, and a discussion.

4.1. *Assumptions*

A total system load of 40 GW, served by 50 GW of thermal capacity, comprised entirely of generating units each with a capacity of 500 MW costing 50 \$/MWh is assumed. Transmission congestion and zonal shortage values are disregarded, and a minimum contingency reserve of 3 GW is imposed. A short-term forced outage rate (i.e., the chance of having an outage between the RT market run and the actual RT condition) of 2% is assumed. The requirements for reg-up, spin, and non-spin reserves are assumed to be 0.5 GW, 1.5 GW, and 3 GW, respectively. All downward reserves are ignored in this comparison.

For the CAISO, the maximum energy bid allowed is 1000 \$/MWh and 1 GW is the maximum flexiramp-up requirement, which is also the required amount assumed for MISO and SPP. Furthermore, it is assumed that the CAISO's multistep flexiramp demand function can be approximated here as a linear function decreasing from 247 \$/MWh to 0 \$/MWh. For NYISO, the total system requirement (i.e., NYCA) is considered, and the non-spinning reserve requirement is assumed to be only a 10-min requirement. For SPP, the operating reserve requirement is 4 GW, which is the summation of reg-up, spin, and non-spin requirements. For SPP, the assumed maximum flexiramp scarcity price is \$250/MWh. For those ISOs that have a 30-min operating reserve, an additional 1.5 GW on top of the minimum contingency reserve is required. In addition to the 30-min operating reserve requirement, a 100 MW replacement reserve requirement is considered for ISO-NE. Finally, a total operating reserve requirement of 4 GW is considered for MISO, in addition to the flexiramp requirement. The flexiramp shortage price of MISO is \$5/MWh. All cascading reserve products are represented in the curves such that the additive price that impacts the energy market is shown on the chart, with any additional price caps enforced as well. Higher quality reserve products, such as regulation, are assumed not to be short until lower quality reserves, such as non-spin, reach their maximum shortage values.

For ERCOT and MISO, their demand curve calculations involve the estimation of available capacity in RT, based upon the expectation and the standard deviation of that capacity. The following assumptions are considered in those calculations. Given a $\mathrm{p} = 2\%$ outage rate of generating resources, the expected available capacity in RT is equal to $(0.98 \times \mathrm{Cap})$, where Cap is the online capacity when the RT market is run. Having n independent generating units and using the variance of the binomial distribution $(\mathrm{n} \times \mathrm{p} \times (\mathrm{p} - 1))$ means that the standard deviation of the available capacity in RT is equal to $\left(\sqrt{\left(\frac{\mathrm{Cap}}{500}\right) \times 0.02 \times 0.98}\right)$.

4.2. *Base case results and discussion*

Based on these assumptions and the shortage pricing approaches presented in Section 3, the scarcity premiums for energy as shown in Fig. 11 are obtained.

Reserve prices when reserves are between 0 and 3 GW generally depend on the exact penalties or demand curve price levels chosen for the reserve constraint violations. As can be seen in Fig. 11, ERCOT, which uses Penalty Method 1V (see Table 1), has the highest reserve pricing when there is an operating reserve shortage condition (0–3 GW in this example) due to its high VOLL. This high reserve price has a goal of encouraging participation of eligible resources, which would reduce the possible need for controlled load curtailment by the operator. At the other extreme, CAISO has the lowest reserve prices in case of large amounts of scarcity (reserves between 0 and 3 MW) due to its offer cap. A low price reduces the profitability of supplying demand under extreme scarcity conditions and can thereby discourage eligible resources from participating. This is particularly true if other markets can be accessed that can pay more, as happened in California during the August 2020 heat wave. However, it should be noted that although the CAISO scarcity demand curve has the lowest prices when reserves are between 0 and 3 GW, the curve would be approximately doubled in height during western U.S.-wide scarcity conditions, since the CAISO tariff then raises the offer cap from a soft cap of \$1000 to a hard cap of \$2000. This would place it closer to the middle of the other ISOs.

When available reserves exceed 3 GW, a sloped relationship of the premium to the amount of reserves, as in the case of ERCOT, is more realistic than the stepped relationships used by most ISOs. Abrupt discontinuities in reserve prices do not reflect the fact that the probability of scarcity is likely to be a fairly continuous function of reserves and can cause excessively volatile outcomes.

Under our assumptions, ERCOT, ISO-NE, and NYISO have the longest right tail with the highest prices. This long tail encourages eligible

[13] They are called reserve constraint penalty factors (RCPFs) in ISO-NE terminology.

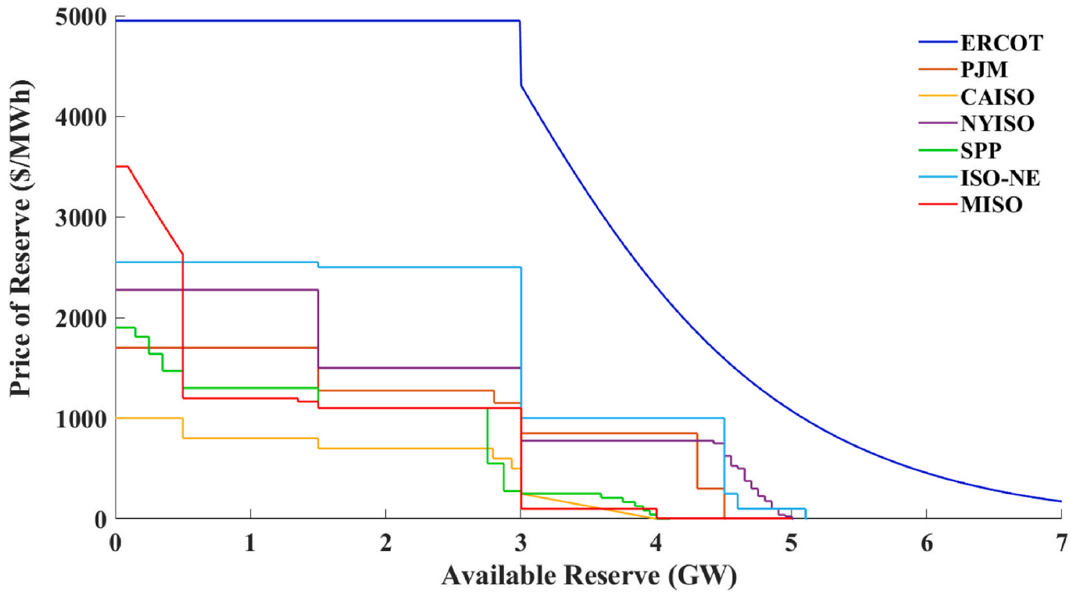

Fig. 11. Results of the comparative numerical example for scarcity premiums across seven U.S. ISOs.

resources to provide reserves and gives operators an extra cushion, reducing the risk of the system moving to more severe reserve scarcity conditions. Such a tail is consistent with the assumption that the consequences of being short of reserves are potentially much more severe than the cost of procuring too much.

4.3. Sensitivity analyses

To further identify the relative importance of various assumptions in determining differences in ORDCs among ISOs, a sensitivity analysis is performed considering three cases.

- Case 1: the penalty for reserve shortages in the first 0.5 GW step is set to be the same as ERCOT's VOLL (i.e., 5000 $/MWh) for all ISOs.
- Case 2: the ORDC step sizes for all ISOs are assumed to be the same. Six steps (i.e., 0.0–0.5, 0.5–1.5, 1.5–3.0, 3.0–4.5, 4.5–5.0, 5.0–10 GW) are considered and the penalty for reserve shortages is assumed to be equal to the largest penalty within each step for each ISO, accounting for nesting constraints.
- Case 3: flexiramp product and its demand curves are ignored in creating the curves.

A similarity index is introduced to quantify the significance of each assumption in making the ORDCs to be more similar to each other. As presented in (5), the index is the sum of the absolute distance of the area

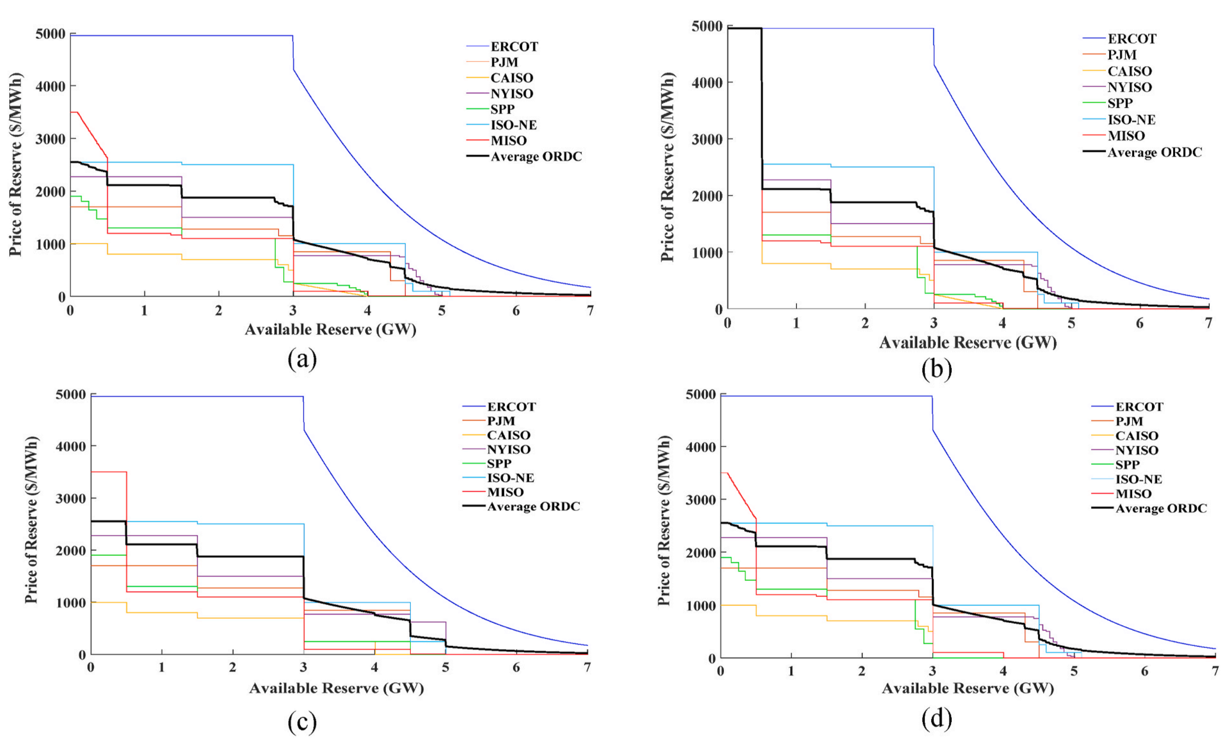

Fig. 12. Sensitivity analysis results for ORDCs across seven U.S. ISOs: (a) original ORDCs and the average ORDC; (b) assuming similar penalty (VOLL = 5000) for first step; (c) assuming similar step sizes; (d) ignoring flexiramp.

below each curve from the area below the average curve, averaged over the 7 ISOs.

$$\text{Similarity Index} = \frac{1}{7}\sum_{\forall iso}\left|A_{iso} - A_{average}\right|; \text{ where } A_{average} = \frac{1}{7}\sum_{\forall iso} A_{iso} \quad (5)$$

Fig. 12 shows the results of these changes in assumptions for all three cases. The original ORDCs are shown in Fig. 12(a), where the average ORDC curve is bolded in black. The results of Case 1, Case 2, and Case 3 are shown in Fig. 12(b), (c), and Fig. 12(d), respectively.

If the penalty values for all ISOs are equal to the values shown by the average ORDC, the similarity index will be zero. Thus, a case with a similarity index close to zero represents the factor causing the greatest divergence in ORDCs. The similarity index (5) for Cases 1, 2, and 3 are 3.862 (M$/h), 4.143 (M$/h), and 4.253 (M$/h), respectively (Fig. 12 (b)–(d), respectively). Therefore, one can conclude that the most important factor differentiating the curves for different ISOs is the assumed penalty values for reserve shortages in the first step (Case 1), followed by the assumed step sizes (Case 2). Less important, but relevant to the shape of the right end of the curves, is the existence and treatment of flexiramp (Case 3).

5. Conclusion and recommendations

In this paper, a comparative review of reserve scarcity pricing practices and calculations across the U.S. power markets is provided. To facilitate the comparison, a consistent terminology is used to describe each of them. The current practices of reserve shortage pricing used by ISOs are classified into three frameworks: (1) the price adder-based approach of ERCOT; (2) the stepwise ORDCs with flexiramp product approach of MISO, SPP, and CAISO; and (3) the stepwise ORDCs without flexiramp or other non-contingency products approach of NYISO, PJM, and ISO-NE. To highlight the practical differences among the general approaches and specific implementations, a generic numerical example based upon some simplified assumptions is presented. Those results show that the different assumptions yield very different scarcity premiums for reserves and energy, and are likely to provide divergent incentives for resources to respond to shortages.

It is concluded that to promote market efficiency, an ORDC should at least have three features. First, it should reflect the marginal value of reserve products at each shortage level, which means that penalties for reserve shortages should be higher when shortages are more severe. It is recognized that in general when load shedding is a possibility, scarcity prices should be higher than is the case for most U.S. markets; on the other hand, if operator actions can prevent curtailment, as is usually the case with minor reserve shortages, scarcity prices should be modest. Ideally, the economic value of avoiding load shedding used in such calculations should rely on consumer preferences as revealed by their decisions as to what reliability they are willing to pay for; although this is not practical now, increased demand participation in markets makes this a possibility in the not-too-distant future.

Second, an ORDC should consider the magnitude and probability of supply contingencies, such as generator forced outages or unexpected reductions in renewable output, that can occur between when the market software is run and actual operations. The calculations should be based on reasonable, if approximate, estimates of probabilities of adverse consequences and the severity of those consequences rather than arbitrary parameters. The calculations should also reflect the latest available information on existing and anticipated weather and operating conditions.

Third, an ORDC should avoid abrupt discontinuities that can cause excessively volatile outcomes, since the probabilities and consequences of inadequate reserves, in reality, are relatively smooth functions of levels of reserves. Fig. 11 shows, in contrast to this recommendation, that scarcity prices often undergo large step changes as reserves shrink. This can cause forecasting price variations difficult, especially in real-time that do not reflect actual changes in fundamental system conditions. Such volatility can encourage large swings in resource output, such as storage, that can be difficult for resource owners and system operators to manage, and inflate non-fuel operations and maintenance costs unnecessarily.

However, following these basic principles in constructing ORDCs to reflect scarcity is just a start. The details matter in implementation, and so the following specific guidance is also offered to market designers.

- Appropriately propagate shortage penalties among all operating reserves and energy, based on energy and reserve co-optimization that correctly considers how different types of reserve substitute for each other and optimizes the allocation of capacity among those commodities. As a result, the software's shadow prices for reserve requirements and energy balances will recognize opportunity costs of devoting resources when scheduling them in shortage conditions. Texas' system does not explicitly do this, for instance, instead relying on market participants to arbitrage separate reserves and energy markets by how they construct their offers into each, which in general can result in less predictability and logical relationships among prices for different commodities.
- Encourage linkages among DA and RT market processes through arbitrage. Although stochastic programming-based scheduling models can explicitly co-optimize over those markets simultaneously [53], computational challenges and conceptual issues in market price formation [54] mean that this will likely not be a practical solution anytime soon for ISO markets. Instead, trading and arbitrage between DA and RT should be encouraged by allowing virtual bidding, harmonizing market rules between the markets, and encouraging resources to offer flexibly in all markets. As a result, offering behavior will translate anticipated shortages in RT to shortage premiums in DA, even if, as in Texas, participation in the DA market is voluntary and there are no ORDC-based shortage mechanisms in that market.
- Consider more granular zones or even network-constrained nodal-based deployment constraints so that operators can be assured that reserves are deployable if needed. It is well-known that in systems with significant congestion, market software will tend to procure system- or zonal-level reserves in generation pockets with low LMPs because that is where the opportunity costs of reserves are lowest. But it is precisely those locations whose transmission congestion will often bottle up reserves so they can't be used to meet needs that arise elsewhere in the system [55]. One approach is to procure reserves at a nodal level, and then include a small set of deployment contingency scenarios in the market software to test the feasibility of dispatching and delivering reserves under the full set of network constraints [56]. However, the ongoing debate over whether to procure day-ahead imbalance reserves nodally or zonally in the proposed west-wide Extended Day-Ahead Market shows that stakeholders hold strong reservations concerning the computational feasibility and effects upon reserve costs of doing so [57].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

This research was supported by the Grid Modernization Initiative of the U.S. Department of Energy (DOE) as part of its Grid Modernization

Laboratory Consortium, a strategic partnership between DOE and the national laboratories to bring together leading experts, technologies, and resources to collaborate on the goal of modernizing the nation's grid. Funding was provided by the Office of Electricity, Office of Energy Efficiency and Renewable Energy, Wind Energy Technologies Office, Water Power Technologies Office, and the Office of Nuclear Energy. The views expressed do not necessarily represent the views of the DOE or the U.S. Government. Suggestions by Todd Levin from Argonne National Laboratory and Miguel Heleno and Alexandre Moreira da Silva from Lawrence Berkeley National Laboratory are appreciated. Furthermore, comments provided by G. Bautista Alderete from CAISO, S. Moorty from ERCOT, T. Zheng from ISO-NE, Y. Chen from MISO, P. Jain from NYISO, A. Giacomoni from PJM, and R. Schoppe from SPP are highly appreciated. The authors bear any responsibility for any errors, omissions, or opinions.

References

[1] Schweppe FC, Caramanis MC, Tabors RD, Bohn RE. Spot pricing of electricity. Springer Science & Business Media; 2013.
[2] FERC. Assessment of demand response and advanced metering. 2021. December 2021. [Online]. Available: https://www.ferc.gov/media/2021-assessment-demand-response-and-advanced-metering.
[3] CAISO. Root Cause Analysis, Mid-August 2020 Extreme Heat Wave. http://www.caiso.com/Documents/Final-Root-Cause-Analysis-Mid-August-2020-Extreme-Heat-Wave.pdf; 2021.
[4] Hogan WW. Electricity scarcity pricing through operating reserves. Economics of Energy & Environmental Policy 2013;2(2):65–86.
[5] Lavin L, Murphy S, Sergi B, Apt J. Dynamic operating reserve procurement improves scarcity pricing in PJM. Energy Pol 2020;147:111857.
[6] Papavasiliou A, Smeers Y, de Maere d'Aertrycke G. Market design considerations for scarcity pricing: a stochastic equilibrium framework. Energy J 2021;42(5).
[7] Bushnell J, Harvey SM, Hobbs BF. Opinion on Market Enhancements for Summer 2021 Readiness. http://www.caiso.com/Documents/MSCOpiniononMarketEnhancementsfor2021SummerReadiness-Mar8_2021.pdf; 2021.
[8] Sun Y, et al. Research priorities and opportunities in United States wholesale electricity markets. Golden, CO (United States): National Renewable Energy Lab. (NREL); 2021.
[9] Faruqui A, Sergici S. Household response to dynamic pricing of electricity: a survey of 15 experiments. J Regul Econ 2010;38(2):193–225.
[10] Zome." http://zomepower.com (accessed June 10, 2022).
[11] Chao H. Price-responsive demand management for a smart grid world. Electr J 2010;23(1):7–20.
[12] Bjarghov S, Farahmand H, Doorman G. Capacity subscription grid tariff efficiency and the impact of uncertainty on the subscribed level. Energy Pol 2022;165: 112972.
[13] EPRI. Wholesale Electricity Market Design in North America–Reference Guide: 2021 Review. https://www.epri.com/research/programs/027560/results/3002021813; 2022.
[14] 2020 state of the market report for the ERCOT electricity markets. Potomac Economics 2021. https://www.potomaceconomics.com/wp-content/uploads/2021/06/2020-ERCOT-State-of-the-Market-Report.pdf.
[15] Liu C, Guo J, Hui H, Fu Y. "Modeling of operating reserve demand curves and system-level price adders in real-time energy-only market,". IEEE Trans Power Syst 2018;33(5):4797–808.
[16] ERCOT. Methodology for Implementing Operating Reserve Demand Curve (ORDC) to Calculate Real-Time Reserve Price Adder. https://www.ercot.com/files/docs/2021/10/31/Methodology_for_Implementing_ORDC_to_Calculate_Real-Time_Reserve_Price_Adder_110121.zip; 2021.
[17] 2019 state of the market report for the ERCOT electricity markets. Potomac Economics 2020 [Online]. Available: https://www.potomaceconomics.com/wp-content/uploads/2020/06/2019-State-of-the-Market-Report.pdf.
[18] MISO. Corporate Fact Sheet 2022 [Online]. Available: https://www.misoenergy.org/about/media-center/corporate-fact-sheet/.
[19] MISO. Business practices manual energy and operating reserve markets. https://cdn.misoenergy.org//BPM%20002%20-%20Energy%20and%20Operating%20Reserve%20Markets49546; October 2021.
[20] PotomacEconomics. 2020 State of the market report for the MISO electricity markets. https://www.potomaceconomics.com/wp-content/uploads/2021/05/2020-MISO-SOM_Report_Body_Compiled_Final_rev-6-1-21.pdf; 2021.
[21] MISO. Scarcity pricing evaluation. https://cdn.misoenergy.org/20210513%20MSC%20Item%20XX%20Scarcity%20Pricing%20Evaluation%20Paper550162.pdf; May 2021.
[22] MISO. Emergency and scarcity pricing update. https://cdn.misoenergy.org/20211201%20MSC%20Item%2004%20Continued%20Reforms%20to%20Improve%20Scarcity%20Pricing%20and%20Price%20Formaitons%20(IR071)606802.pdf; 2021.
[23] MISO. Continued Reforms to Improve Scarcity Pricing and Price Formations. https://cdn.misoenergy.org/20220421%20MSC%20Item%2009%20Continued%20Reforms%20to%20Improve%20Scarcity%20Pricing%20and%20Price%20Formations%20(MSC-2019-1)624101.pdf; 2022.
[24] SPP. SPP 101 an introduction o Southwest Power Pool. https://www.spp.org/documents/31587/spp101%20-%20an%20introduction%20to%20spp%20-%20all%20slides%20print.pdf; 2021.
[25] SPP. Day-ahead and real-time markets user interface business validations. https://www.spp.org/documents/66798/integrated%20marketplace%20market%20user%20interface%20business%20validations_v18.1%2020220314.pdf; 2022.
[26] SPP. 2020 Annual report highlights. https://www.spp.org/documents/65258/2020%20annual%20state%20of%20the%20market%20presentation.pdf; 2021.
[27] SPP. STATE OF THE MARKET 2020. https://www.spp.org/documents/65161/2020%20annual%20state%20of%20the%20market%20report.pdf; 2021.
[28] SPP. Scarcity demand curve. https://marketplace.spp.org/groups/scarcity-demand-curve; 2022.
[29] SPP. Market protocols SPP integrated marketplace, revision 87. https://www.spp.org/Documents/13558/SPP%20Integrated%20Marketplace%20Protocols%20v.0.b%2012-17-10_Word%2097-2003.zip; 2022.
[30] Harvey S. Scarcity Pricing Background Discussion - California ISO. http://www.caiso.com/Documents/ScarcityPricingBackgroundDiscussionHarvey-Presentation-Dec11_2020.pdf; 2020.
[31] Hinman C. Day-ahead market overview. http://www.caiso.com/Documents/Presentation-Existing-Day-Ahead-Market-Overview.pdf; 2021.
[32] CAISO. California Independent System Operator Corporation Fifth Replacement Electronic Tariff: Real-Time Market as of Dec 15, 2021. https://www.caiso.com/Documents/Section34-Real-TimeMarket-asof-Dec15-2021.pdf; 2021.
[33] CAISO. California Independent System Operator Corporation Fifth Replacement Electronic Tariff: California ISO Markets and Processes as of February 15, 2021. https://www.caiso.com/Documents/Section27-CaliforniaISOMarkets-and-Processes-asof-Feb15-2021.pdf; 2021.
[34] CAISO. Business practice manual for market operations, vol. 78; 2021. https://bpmcm.caiso.com/Pages/BPMDetails.aspx?BPM=Market%20Operations.
[35] CAISO. Business practice manual for market operations, vol. 79; 2022. https://bpmcm.caiso.com/Pages/BPMDetails.aspx?BPM=Market%20Operations.
[36] Patton DB, LeeVanSchaick P, Chen J, Naga RP. 2020 state of the market report for the New York ISO markets. Potomac Economics 2021. https://www.nyiso.com/documents/20142/2223763/NYISO-2020-SOM-Report-final-5-18-2021.pdf/.
[37] NYISO. Ancillary Services Shortage Pricing_A Report by the New York Independent System Operator. https://www.nyiso.com/documents/20142/9622070/Ancillary%20Services%20Shortage%20Pricing_study%20report.pdf; 2019.
[38] NYISO. Manual 11_Day-Ahead Scheduling Manual. https://www.nyiso.com/documents/20142/2923301/dayahd-schd-mnl.pdf/0024bc71-4dd9-fa80-a816-f9f3e26ea53a; 2021.
[39] NYISO. Manual 12_Transmission and Dispatch Operations Manual. https://www.nyiso.com/documents/20142/2923301/trans_disp.pdf; 2021.
[40] NYISO. Manual 2_Ancillary Services Manual. https://www.nyiso.com/documents/20142/2923301/ancserv.pdf; 2021.
[41] Jain P. Ancillary Services Shortage Pricing and Reserves for Resource Flexibility: Manual Updates. https://www.nyiso.com/documents/20142/24511213/5%20Proposed%20Manual%20Changes%20ASSP%20and%20RFRF%20-%20BIC%2009142021_Final.pdf; 2021.
[42] PJM. 2020 Annu Rep. https://www.pjm.com/about-pjm/who-we-are/annual-report; 2021.
[43] PJM. PJM Manual 11: Energy & Ancillary Services Market Operations, Revision: 117. https://www.pjm.com/~/media/documents/manuals/m11.ashx; 2021.
[44] PJM. "Ancillary Services." https://www.pjm.com/markets-and-operations/ancillary-services.aspx (accessed December 2021).
[45] PJM. PJM Reserve Market. https://www.pjm.com/-/media/training/nerc-certifications/markets-exam-materials/generation-itp/reserve-market.ashx; 2017.
[46] PJM. Operating Reserve Demand Curve. https://pjm.com/-/media/committees-groups/committees/mrc/2021/20210623/20210623-item-02a-1-ordc-tcpf-issue-charge-presentation.ashx; June 2021.
[47] ISO-NE. Key Grid and Market Stats. https://www.iso-ne.com/about/key-stats; 2022.
[48] ISO-NE. Assumptions and Other Information for the Summer 2022 Forward Reserve Auction. https://www.iso-ne.com/static-assets/documents/2021/07/forward_reserve_auction_assumptions_winter_2021_2022.pdf; March 2022.
[49] ISO-NE. ISO New England Manual for Market Operations, Revision 60. https://www.iso-ne.com/static-assets/documents/2020/10/m_11_revision_60_effective_10_1_2020.pdf; 2020.
[50] ISO-NE. Reserve Market. https://www.iso-ne.com/markets-operations/settlements/understand-bill/item-descriptions/reserve-market; 2022.
[51] ISO-NE. 2020 Annual Markets Report. https://www.iso-ne.com/static-assets/documents/2021/06/2020-annual-markets-report.pdf; 2021.
[52] Chang J, Aydin MG, Broehm R, Yang Y, Sweet R. Shortage pricing in North American wholesale electricity markets. https://www.aeso.ca/assets/Uploads/4.3-Brattle-Paper-Shortage-Pricing.pdf; 2018.
[53] Papavasiliou A, Oren SS. Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network. Oper Res 2013;61(3):578–92.
[54] Morales JM, Zugno M, Pineda S, Pinson P. Electricity market clearing with improved scheduling of stochastic production. Eur J Oper Res 2014;235(3): 765–74.
[55] CAISO. CAISO Energy Markets Price Performance Report. http://www.caiso.com/documents/finalreport-priceperformanceanalysis.pdf; 2019.
[56] CAISO. Flexible Ramping Product Refinements. http://www.caiso.com/InitiativeDocuments/FinalProposal-FlexibleRampingProductRefinements.pdf; 2020.
[57] CAISO. Day-Ahead Market Enhancements. https://stakeholdercenter.caiso.com/StakeholderInitiatives/Day-ahead-market-enhancements; 2023.