Exploring the Design Space of Balancing Services Markets- A Theoretical Framework

Alireza Abbasy and Rudi A. Hakvoort[1]

***Abstract*—** This paper provides the theoretical foundation which is necessary to properly design balancing services markets, as part of power system ancillary services needed for secure and reliable operation of the electricity grid. The objective is to propose a generic framework that is applicable to design of balancing services markets for different power systems, regardless of electricity markets design differences in different countries/regions. Firstly, in order to analyze the effect of different decisions for each design variable on the performance of the whole market and eventually evaluate the performance of different alternative market designs, a set of performance criteria is proposed. Secondly, this paper identifies the variables on which a decision needs to be made in the design process; design variables identification. Based on the different possible options (states) for each design variable, the entire design space for each balancing service market is identified; design space identification. Finally, the relations between design variables and performance criteria are presented. The proposed framework functions as the basis for evaluating alternative balancing services market designs in future research.

***Keywords*—** Balancing services, electricity markets, market design

I. Introduction

According to the non-storable nature of electricity, the amount of electric power produced must equal the amount of power consumed in a power system at each moment. The balance between electricity supply and demand has a close relationship with the frequency of the whole system. Imbalances can occur due to a wide variety of reasons; load forecasts errors, a generation outage, transmission line outages leading to islanding of the system, etc. Any imbalance between production and consumption will lead to frequency deviation from the synchronous frequency of the system, which in turn can result in serious system stability problems and equipment damage. Therefore, balancing the production and consumption in a power system at every moment plays a crucial role in insuring the security of the whole electric grid. In the new restructured liberalized power systems introduced in the early 90's, the task of operating the system in a reliable and secure way is the responsibility of the Independent System Operator (ISO). Thus, the ISO needs to employ a mechanism to balance the system, known as Balance Management Mechanism. The system operator needs various balancing services from different resources in the power system in order to balance the system in an effective way.

Balancing services markets are employed by system operators in order to procure sufficient balancing services. These balancing service markets are single-buyer markets with the system operator as the only buyer of balancing services. There are two different types of markets for procurement of balancing services:

- *Reserve capacity markets*: These markets aim to procure the required amount of reserve capacity for each type of balancing services to insure the secure operation of the system. Selected reserve capacity providers in the market will receive an option fee for "availability" of their capacity and are obliged to offer the capacity in real-time markets. Therefore, these markets are mechanisms with which the system operator makes sure that enough balancing services will be available in real-time to be employed in order to resolve system imbalances.
- *Balancing energy markets*: These markets are (near) real-time markets in which balancing energy is purchased by system operator to resolve imbalances. Therefore, they are markets for actual delivery of balancing energy and the selected balancing energy providers receive a payment for "utilization" of their offered energy.

Design of balancing services markets significantly differs in different power systems with different designs for electricity markets. These differences originate from fundamental institutional differences, technical differences (e.g. dominant generation portfolio and control technologies), history, etc. While developing a generic framework for design of balancing services markets applicable to different systems, is necessary, most of the existing literature on ancillary services markets design is based on the fundamental design characteristics of specific systems (countries/regions) [1]-[6]. The American definitions and fundamental characteristics are assumed as the basis in [1]-[4], while the focus is on Australian ancillary service market designs in [6]. In addition, a clear definition of ancillary services, that includes balancing services, black start capability, voltage and reactive power control, is needed. Since the trading products in these markets are totally different in nature with different characteristics, the design problem needs to be addressed

[1] Authors are with the faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands (phone: +31 (0)15 27 82040; fax: +31 (0)15 27 83422; e-mail: a.abbasy@tudelft.nl).

978-1-4244-5791-5/09/$25.00 © 2009 IEEE

separately for each of these ancillary services markets. Moreover, the current literature on balancing services markets design mainly focuses on specific detailed aspects of the market design, ignoring the generic decisions that need to be made in designing these markets; ignoring the big picture. The issue of using both capacity and energy bids in procurement of balancing services (the problem of weighting factor) is analyzed in [1], while [2] discusses the simultaneous vs. sequential clearing of different balancing services markets. This aspect is combined with the pricing mechanism and discussed in [3].

In this paper, design of balancing services markets is addressed from a higher level perspective and a theoretical framework is proposed to be employed in design of these markets, irrespective of the design differences in different countries/regions. The analysis is mainly based on the study of the existing designs in Northern Europe (the Netherlands, Germany and The Nordic system) and two examples of the North American designs (PJM Interconnection and California ISO). Numerous examples from these systems are given which help us develop a general framework that can be applied to all these different systems. Section II identifies and introduces the performance criteria. And section III formulates and defines the design variables and the design space for balancing services markets. It also explores the interrelation of various design variables with the introduced performance criteria.

I. PERFORMANCE CRITERIA

In order to evaluate the effect of different design variables on the performance of a balancing service market, and eventually to evaluate different alternative market designs, a set of performance criteria should be identified. By studying and measuring the effect of different decisions in market design on these performance criteria, impact of those decisions on the performance of the balancing service market as a whole can be studied.

Fig. **1** demonstrates the performance criteria related to design of balancing services markets. The diagram starts with the most generic abstract criteria and divides each criterion into several other lower-level and more concrete criteria that can be measured and studied more easily.

A. Operational Security

The objective of balancing markets (market-based balance management) is to balance supply (generation) and demand (load) in the grid by means of a market-based mechanism. Therefore, insuring operational security of the system stands next to a general market-related criterion as the main performance criteria of balancing services markets in general. Operational security performance criterion of balancing services markets is directly related to effectiveness of both reserve capacity markets, in meeting the reserve requirements of the system for each type of balancing services, and balancing energy markets in resolving actual real-time imbalances of the system.

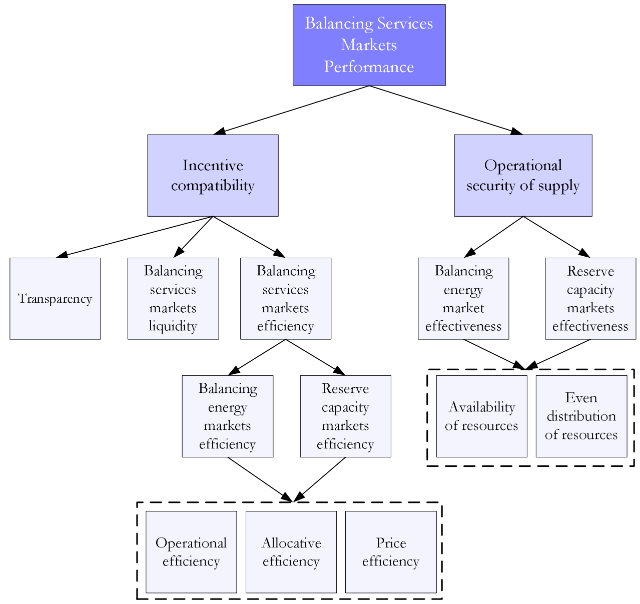

Fig. 1. Performance criteria of balancing services markets design

Effectiveness of these two markets is insured with two main factors; availability and even distribution of resources. In order to achieve effectiveness in reserve capacity and balancing energy markets, enough resources must be available for the ISO to choose. Additionally, because of the considerable interrelation of balancing services procurement (both capacity and energy) with security of the entire system, even geographical distribution of balancing resources plays a critical role in balancing services markets effectiveness. Mainly because of transmission restrictions that may occur in real-time (congestion, transmission line outage, islanding, etc) it is crucial to distribute balancing resources evenly in the system to be able to resolve imbalances in real-time and insure system security.

B. Incentive Compatibility

Incentive compatibility is the generic characteristic of the market design that incentivizes market parties to behave in such a way that best serves the general goal of maximizing societal profit and leads to the globally optimal solution. According to their self-interestedness, agents will pursue any means available to them to maximize their own profit in an open environment. This could lead to undesirable situations and sub-optimal solutions that are more preferable to some agents and not acceptable to others. Therefore, incentivizing agents to behave in such a way that assures the fairness/optimality of the final solution plays a critical role in the design process. Incentive compatibility in balancing services markets design is a high-level performance criterion that includes all the highly interrelated institutional and economical aspects of these markets and it stands as the second main criterion next to operational security. In order to be able to understand and study this high-level abstract criterion, it is divided into more concrete lower-level criteria; transparency, balancing services markets liquidity and balancing services markets efficiency.

Information availability, information symmetry (equal access to information) and clarity of balancing services markets rules will lead to transparency which is a prerequisite of a competitive balancing service market. High level of transparency regarding balancing service market rules, regulation and imbalance prices, volumes of used bids, and imbalance volumes will improve the functioning of the market by enabling market parties to make informed decisions and eventually to encourage new entry and increase competition in these markets.

A market is liquid if there are many buyers and sellers who can access each other easily and have access to information about the market prices. A defining feature of a liquid market is that it can generally absorb the addition or loss of a buyer or seller without a noticeable change in the market price. Since balancing services markets are single buyer markets with system operator as the only buyer, liquidity of these markets relates to the number of balancing services providers and their willingness to offer services in these markets.

"Balancing services markets efficiency" performance criterion refers to the economic efficiency of both reserve capacity and balancing energy markets for each type of balancing services. Considering market efficiency as one of the primary objectives in design of electricity markets in general, and the ambiguity and multi-faceted nature of the concept, this criterion needs to be carefully defined and divided into more tangible and measurable performance criteria.

"Operational efficiency" performance criterion is the aspect of market efficiency that relates to transaction costs of market parties in the market; both reserve capacity and balancing energy markets for each type of balancing services. Therefore, operational efficiency is related to carrying out the market operations with as low a cost as possible. There are various variables in the design of balancing markets that influence this aspect of market efficiency, e.g. method of procurement of reserve capacities (bilateral contracts and auctions), the time horizon of reserve capacity and balancing energy markets (frequency of market clearance), etc.

"Allocative efficiency" is the aspect of market efficiency relating to optimal use of limited available resources. Therefore in case of balancing services markets, this performance criterion aims at meeting system's reserve requirements (through reserve capacity markets) and resolving system's real-time imbalances (through balancing energy markets), with the use of the optimal set of offered available balancing resources. This criterion is influenced by two main factors; system operator's use of the cheapest balancing resources offered in the markets, and incentives of resource owners to offer their capacity as balancing services in balancing services markets.

"Price efficiency" performance criterion relates to cost-reflectivity of prices in both reserve capacity and balancing energy markets. Therefore, price efficiency criterion deals with the issue of market power and competitiveness in balancing services markets. It is also interrelated with balancing services markets liquidity since it is influenced by the number of balancing services providers and trading volumes, while it is not the same as market liquidity.

Table I presents the low-level performance criteria discussed above and the indicators that should be used in order to measure those performance criteria.

TABLE I
PERFORMANCE CRITERIA AND PERFORMANCE INDICATORS OF BALANCING SERVICES MARKETS DESIGN

Performance criterion	Performance indicators
Balancing services markets effectiveness	Frequency deviation, offered and activated balancing services (capacity and energy), Area Control Error
Balancing Services markets liquidity	Number of market parties, volatility of regulation and capacity reservation prices, (cross-border) balancing trade volumes
Operation efficiency	transaction costs
Allocative efficiency	Actual balancing costs/possible minimum balancing costs, balancing costs (capacity and energy), offered/activated balancing services
Price efficiency	Market concentration indices, Pivotal/Residual supplier index, Mark-up index, level and volatility of regulation prices

II. DESIGN VARIABLES AND THE DESIGN SPACE

Identifying all the design variables is the prerequisite for finding possible market designs for balancing services and eventually evaluating those different alternative market designs. For each design variable, there may be different possible options (states) which lead to different possible decisions for each design variable. Therefore, based on the identified design variables and different possible decisions for each variable, the entire design space for each balancing service market can be identified. Additionally, in order to analyze the effect of different decisions for each design variable on the performance of the whole market, the relations between design variables and performance criteria should be understood.

Fig. 2 shows the main design variables for balancing services markets, and also illustrates which performance criteria are influenced by each design variable. The blocks in the central box show the design variables, while the blocks outside the box illustrate the set of performance criteria related to every design variable. The central box also shows the interrelations of different design variables.

A. *Definition of Balancing Services*

The very fundamental variable in design of balancing services markets is the definition and categorization of the services. In order to maintain the balance between generation and consumption in a power system, the system operator needs different types of services with different technical characteristics, e.g. the activation times (response speed), the method of activation (manual/automatic), the minimum deployment time, and their state compared to the system (synchronous/non-synchronous).

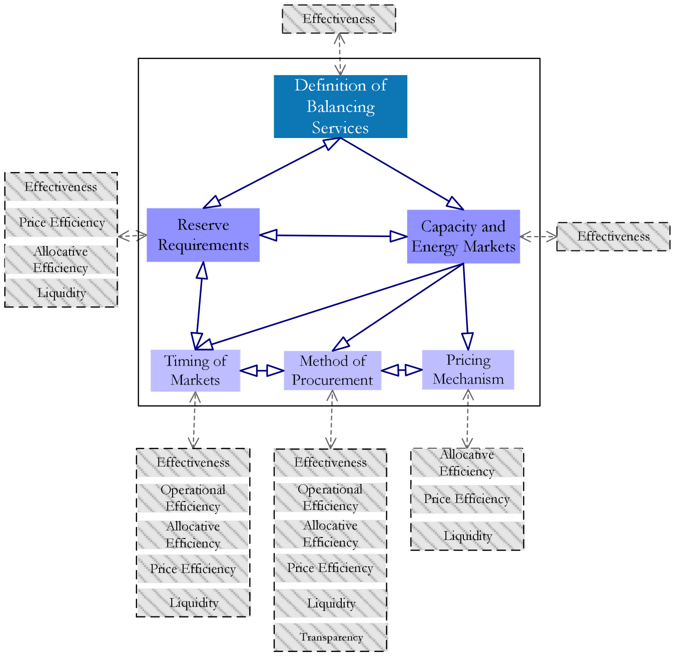

Fig. 2. Design variables and their relation with the performance criteria in design of balancing services markets

Definition of balancing services is fundamentally a technical decision that is made based on technical characteristics of the system such as dominant generation portfolio, technologies used in the control systems, etc. The criterion, based on which the definition of balancing services is done, is the "effective" operation of the entire balancing market as the mechanism employed to keep balance between generation and consumption. Therefore, as Fig. 2 proposes, the performance criterion related to this fundamental design variable is "effectiveness" of the balancing services markets which is directly connected to the operational security of the grid.

B. *Capacity and Energy Markets*

For each type of balancing services, a capacity and/or an energy market may exist in order to, respectively, procure the required amount of reserve capacity to insure system security, and procure balancing energy to resolve imbalances in real-time. The need for establishment of these markets for each type of balancing services highly depends on the characteristics of the corresponding service.

There may be no market at all. For example, provision of the primary control service in the Netherlands and PJM interconnection is the prerequisite for generators to be connected to the grid so the service provision is compulsory and there is no compensation [7]-[10].

There may be only a capacity market and no energy market, e.g. primary control service in Germany is procured via a capacity market and there is no compensation for the actual delivered energy (no utilization payment) [11].

There may be a capacity and an energy market, e.g. regulating power, corresponding to secondary control service in UCTE (an association of system operators in continental Europe) terms [12], in the Netherlands and Norway [10], [13]. In this case, capacity bids selected in the capacity market receive the availability payment and the energy bids selected in the energy market for real-time energy delivery receive a utilization payment.

There may also be one single market that functions both as the capacity and the energy market. For example, in PJM Interconnection, the regulation market (corresponding to secondary control) is actually a capacity market with which the PJM ISO procures the minimum regulation reserve requirement of the system [7], [8]. But there is no other separate energy market for these services and the ISO uses the same bids for activation of regulating power in real-time based on the actual imbalance of the system. So the regulation bids are in \$/MWh, and the service providers take their operating production costs into account in their bids for the regulation market. Therefore once selected, irrespective of whether or not their capacity is actually activated in real time, regulation service providers will be compensated for both availability and utilization of their service.

This design variable (the need for capacity and energy markets for each type of balancing services), as Fig. 2 shows, is hugely influenced by the definition of balancing services which is a higher level design variable. The main performance criterion that is influenced by the decision on this design variable is "balancing services markets effectiveness".

C. *Reserve Requirements*

This design variable is, on one hand, the key to secure operation of the system, and on the other hand, plays a critical role in cost of capacity reservation which is part of the system security costs. The minimum reserve requirements determine the "demand" in capacity markets and therefore, hugely influence the total capacity reservation costs of the system. These requirements are determined based on the system security criteria which are different for different types of balancing services; e.g. in UCTE, available primary control capacity must meet the simultaneous loss of the two biggest generation units, and secondary control capacity is calculated based on the peak anticipated load [12]. There are two factors that affect reserve requirements and eventually capacity reservation costs:

Security criteria: System security criteria determine the amount of minimum reserve requirement to be met by capacity reservation in capacity markets.

Frequency of reserve requirements calculation: Requirements can be calculated for different time horizons; annual, monthly, weekly, daily, on-peak and off-peak hours. Once calculated, the minimum reserve requirements are fixed for the whole time horizon. For example, if the security criterion for secondary control is 1% of the forecasted peak load (PJM example), and minimum reserve requirement is calculated on a monthly basis, it will be different from one month to another, while if the requirements are calculated annually, they will be fixed through the year. Therefore, increasing the frequency of reserve requirements calculation will lead to lower demand in capacity markets and consequently lower capacity reservation costs, while the system security criterion remains the same.

There are three main performance criteria that are affected by the decision on “reserve requirements” design variable: effectiveness, efficiency and liquidity of capacity markets. Its influence on capacity markets effectiveness is obvious because these requirements are aimed to insure secure operation of the system. Since this design variable determines the “demand” in capacity markets, it has a clear influence on the incentive of capacity providers, available resources, and total capacity reservation costs which lead to the “price efficiency”, “allocative efficiency” and “market liquidity” performance criteria (all related to capacity markets).

D. Timing of Markets

Timing of markets for different types of balancing services has two main aspects: timing of the bidding procedure and timing of markets clearance.

1. Timing of the bidding procedure: The time horizon for the bidding procedure of all the capacity and energy markets for all types of balancing services need to be carefully arranged. Time horizon of the bidding procedure for balancing services markets, which is related to the frequency of bidding, highly depends on the characteristics of the service (balancing service type) and the type of the market (capacity vs. energy). After making the decision on timing of balancing services markets (possible from annual to quarter-hourly), the gate opening and closure times of these markets need to be determined. Since, basically generators have the possibility to offer their capacity in wholesale electricity markets (such as day-ahead and intra-day markets) or in capacity or/and energy markets for various types of balancing services, the gate opening and closure times for the bidding procedure in balancing services markets (both capacity and energy) need to be carefully coordinated with gate opening /closure times of other balancing services markets as well as wholesale electricity markets, such as day-ahead and intra-day markets. This aspect significantly influences the interrelations between different markets and plays a critical role in the effect of balancing services markets design on the efficiency and liquidity of other electricity markets.

2. Timing of the markets clearance: Just like the “timing of the bidding procedure”, timing of the markets clearance for balancing services has two main elements; frequency of balancing services markets clearance and coordination of balancing services markets clearance with other markets. Generally speaking, the frequency of bidding (related to the time horizon of the bidding procedure), can be different from the frequency of the clearance of the market. For example, the energy market for secondary control in Germany is monthly and generators submit one price bid and one capacity bid for all the coming month to the market (frequency of bidding=once per month), while the energy market is cleared once for each Program Time Unit, PTU, (15 minutes) based on the real-time imbalance of the system that occurs in each PTU (frequency of clearance=once per quarter an hour). Since energy markets are aimed at resolving system imbalances in real-time using balancing energy bids, the “demand” in these markets is different for different PTUs (based on the system imbalance volume) and therefore these markets need to be cleared for each PTU which leads to the clearance frequency of once per PTU for energy markets. Because the time horizon of capacity markets can theoretically be anything from one year to one PTU, the frequency of capacity markets clearance can have any value greater than once per year and less than once per PTU. The second aspect of “timing of markets clearance” is the coordination of balancing services markets clearances with those of other electricity markets, specifically day-ahead and intra-day markets.

Fig. 3 shows the elements of timing of balancing services markets and their interrelation with day-ahead and intra-day markets. Since energy markets are real-time markets and are cleared at the time when all other markets are closed and cleared, there is no coordination needed between clearance of energy markets and other markets. On the contrary, clearance of the capacity markets may need to be coordinated with clearance of day-ahead or intra-day markets, based on the frequency of market clearance of the capacity markets. In case of a day-ahead capacity market (e.g. secondary control service in Germany, and day-ahead scheduling reserves in PJM), there are two main possibilities in terms of coordination of the capacity market clearance with the day-ahead market; sequential and simultaneous clearance. Another example is coordination of capacity markets clearance with each other (and not day-ahead or intra-day markets). The Californian ISO (CAISO) uses four capacity markets for four different services (regulation service, spinning reserve, non-spinning reserve, and replacement reserves) [14]. CAISO used to clear these markets sequentially (from higher quality services to lower quality ones) and has recently changed it to simultaneous clearance of all four capacity markets because of efficiency and liquidity considerations.

As described above, the decisions that need to be made on this design variable demand meticulous studies of the effect of this variable on the performance of balancing services markets as well as day-ahead and intra-day markets. Since timing of markets (especially coordination with other markets) can influence the offered capacities in balancing services market, this design variable is related to the “balancing services effectiveness” criterion, as shown in Fig. 2. Changing the time horizon of balancing services markets also affects the “operational efficiency” of these markets mainly because the frequency of bidding in different markets influences the transaction costs of the market parties and the market as a whole.

Since timing of balancing services markets (especially coordination with other markets) has a significant influence on the incentives of balancing service providers (both capacity and energy), available resources in different markets, total capacity reservation and balancing energy procurement costs, it also influences “allocative and price efficiency” and “liquidity” of balancing services markets and as well as (possibly) day-ahead and intra-day markets.

E. Method of Procurement

Different methods can be used in procurement of balancing services; compulsory provision, bilateral contracts, auctions (tendering), and a combination of bilateral contracts and auctions. For example, primary control capacity is procured on a compulsory basis in PJM and the Netherlands [9], [10]. Bilateral contracts are used for secondary control capacity procurement in the Netherlands, Sweden and Finland [10], [13]. Auctions are used for secondary control energy in Germany, the Netherlands, Nordic system, etc [10], [11], [13]. In Australia and New Zealand, both bilateral contracts and auctions are used to procure primary control capacity [9].

Because of this design variable’s influence on balancing service providers’ incentives it affects the offered capacities in various balancing service markets, available resources in each market, price bids of service providers, and transaction costs of market parties. So this variable influences balancing services markets “effectiveness”, “efficiency” (operational, allocative and price) and “liquidity” performance criteria. This variable also plays an important role in “transparency” of balancing service markets.

F. Pricing Mechanism

Marginal and pay-as-bid pricing have the advantages and disadvantages and, in case of balancing service markets, the decision on pricing mechanism needs to be made taking into account the characteristics of different services and the type of the market (capacity or energy). Since this design variable has serious impacts on the incentives of service providers, it will influence the bid prices and market prices of different markets and therefore influences “price efficiency”, “allocative efficiency” and “liquidity” of different balancing services markets.

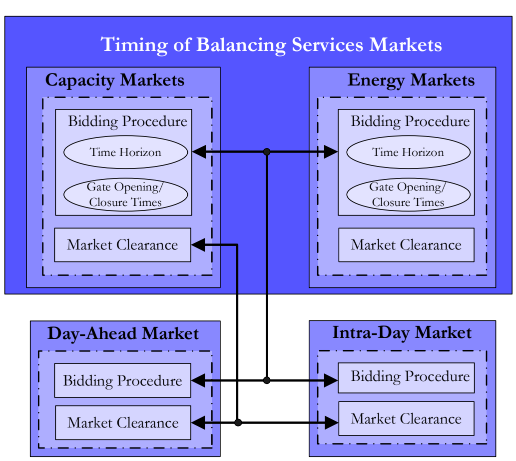

Fig. 3- Timing of balancing services markets and its interrelation with other markets

III. Conclusion

In this paper a generic theoretical framework for design of balancing services markets is proposed. The variables that play a role in design of these markets and different possible options for each variable are identified which leads to the creation of the entire design space for balancing services markets. In addition, the criteria that should be analyzed in evaluating alternative design options are identified (performance criteria) and their relation with design variables is presented. This paper functions as a concrete foundation for a systematic approach towards design of balancing services markets and provides the basis for our next research step which is evaluation of various balancing services market designs by studying the effect of various design variables on the performance criteria using the performance indicators.

References

[1] H. Singh and A. Papalexopoulos, “Competitive procurement of ancillary services by an independent system operator,” *IEEE Trans. Power Systems*, vol. 14, no. 2, pp. 498-504, May 1999.

[2] A. Papalexopoulos and H. Singh, “On the various design for ancillary services markets,” in *Proc. 34th Hawaii International Conf. Systems Science*, Hawaii, 2001.

[3] S. S. Oren, “Design of ancillary service markets,” in *Proc. 34th Hawaii International Conf. Systems Science*, Hawaii, 2001.

[4] K. W. Cheun, “Ancillary service market design and implementation in North America: from theory to practice,” in *Proc. 3rd Int. Conf. Electric Utility Deregulation and Restructuring and Power Technologies, Nanjing,* 2001, pp. 66-73.

[5] M.A.B. Zammit, D.J. Hill and R.J. Kaye, “Designing ancillary services markets for power system security,” *IEEE Trans. Power Systems,* vol. 15, no. 2, pp. 675-680, May 2000.

[6] H. Outherd, “The design of efficient market structures for ancillary services,” in *Proc. 34th Hawaii International Conf. Systems Science*, Hawaii, 2001.

[7] “Scheduling operations,” *PJM Manual 11,* Rev. 37, Nov. 2008.

[8] “Balancing Operations,” *PJM Manual 12*, Rev. 18, July 2008.

[9] Y. Rebours, D. Kirschen, M. Trotignon, and S. Rossignol, “A survey of frequency and voltage control ancillary services—part II: economic features,” *IEEE Trans. power systems*, vol. 22, no. 1, pp. 358-366, Feb. 2007.

[10] TenneT website: http://www.tennet.org/transport_en_systeemdiensten/technische_publicaties/overige_publicaties/index.aspx

[11] S. Riedel, H. Weigt, “German electricity reserve markets,” *Electricity markets working papers*, WP-EM-20, 2007.

[12] UCTE, “Operation handbook policy1: Load frequency control and performance,” final version, July 2004.

[13] “System operation agreement,” Nordel, June 2009.

[14] Department of market monitoring California ISO, “Market issues & performance 2008 annual report,” April 2008.