Economies of Scope and Value of Coordination in the Evolving Electric Power Systems

Marija Ilic, Marcelo Elizondo, Michael Patnik, Zayra Romo, Zhiyong Wu

Abstract- In this paper we address the overwhelming complexity underlying the challenge of technically reliable and economically efficient performance of the future electric power industry. We suggest that the quality of system performance is likely to be primarily determined by the information technology (IT) supported just-in-time (JIT) and just-in-place (JIP) decision making. Because of this, assessments of performance must be based on metrics that account for various economies of scope underlying management of an electric power interconnection whose characterization is very complex over space and time. Examples are given to illustrate these complexities and the magnitude of their effects in some detail. The JIT and JIP electricity service is likely to gradually replace currently ingrained industry paradigm based on the worst-case-design for technical performance and economies-of-scale-based system over-design for cost-savings.

It is proposed that neither strictly technical nor strictly organizational aspects of the evolving industry by themselves exhibit dominant effects on its performance. Instead, it is argued that both technological and organizational innovations require the same breakthroughs concerning: (1) methods for measurably accurate spatial aggregation (decomposition) of the complex network users; (2) methods for temporal characterization and aggregation (decomposition) of network users; and, (3) methods for quality of service (QoS) characterization and aggregation (decomposition) of network users. These methods must be IT-supported and truly dynamic to capture the necessary complexity in order to facilitate JIT and JIP decisions for desired technical and economic performance through IT-based coordination of dynamically aggregated users. The most intriguing is the conjecture that in this environment it becomes possible to define private goods such as delivery service, reactive power/voltage support, reliability service, all of which are currently viewed as public goods; this would, in turn, enable compatible incentives for a variety of services and products and lead to genuine choice.

Manuscript received December 15, 2004. This work was supported in part by the U.S. National Science Foundation EPNES project, Principal Investigator Marija Ilic.

Marija Ilic is a Professor at the ECE and EPP Departments at Carnegie Mellon University, milic@ece.cmu.edu; Marcelo Elizondo is a visiting doctoral student from Argentina, Zayra Romo is a visiting Master student from Mexico, both in the EPP ; Michael Patnik is a Master student in the ECE Department and Zhiyong Wu is a doctoral student in the EPP Department.

I. Introduction

The electric power industry must deal with the fundamental dichotomy brought about by the on-going technological and organizational changes. Namely, both regulatory reforms and new disruptive technologies (distributed generation, controllable demand, FACTS, distributed IT) present the industry with the requirements to implement distributed decision making and control. At the same time, it remains essential to coordinate for system-wide reliability and system-wide efficiency. The inability to resolve these qualitatively different requirements has brought the entire industry to a stand-still.

In this paper we suggest that part of the problem is lack of systematic ways to quantify the industry performance for various types of decentralization. This is true of both technological assessments as well as of organizational. For example, we have no meaningful ways of comparing the performance of (1) an industry architecture supplied by few very large-scale power plants, transmission backbone and passive local distribution networks of today; with the performance of (2) an industry architecture in which major portion of power supply has moved closer to the end users, and is located on local distribution networks. Similarly, we have no clear mechanisms for comparing performance of fully regulated industry organization to the performance of vastly competitive one.

As the industry organizations evolve, reflected in functional and/or corporate unbundling, we need to answer questions concerning : (1) complete product/service design as a function of industry rules; (2) relations between products and industry organization in place; (3) mechanisms to facilitate electricity service unbundling without loosing benefits of coordination.

In this paper we recognize that future solutions are multi-fold with technical and economic outcomes never studied before. We make an attempt to illustrate the criticality of capturing inter-dependencies between the physical operations, industry rules in place and the IT support. Synergies between the physical and economic network previously un-imaginable are plausible by careful design of a supporting IT network.

In this paper we suggest that careful product differentiation and decomposition (aggregation) of users and suppliers are the key determinants of capturing these synergies. One way of quantifying the implications of service unbundling is by defining economies of scope present in today's bundled electricity service, and by analyzing potential loss of economies of scope resulting from the service unbundling.

As this approach is taken, we recognize that there is no well-established body of knowledge documenting economies of scope even for the regulated electric power industry. The existing literature concerning economies of scope in the electric power industry falls short of recognizing unique temporal and spatial characteristics of this industry. In order to bridge this gap, we describe in the first part of this paper several sources of economies of scope potentially relevant in the electric power industry. Electric power delivery, reactive power support and capacity provision are identified as three major components of the unbundled electricity service exhibiting economies of scope.

In the second part of this paper we assess the effects of economies of scope on the industry performance using simple electric power system examples. Various notions of economies of scope are analyzed for several industry structures. We arrive at the general conclusion that significant loss of efficiency could result due to service unbundling. On the other hand, in order for the providers of the unbundled products and services to become sustainable business which provide value in the evolving industry structures, it is necessary to carefully define the unbundled services and products themselves.

To overcome this fundamental dichotomy of having to carefully define differentiated (unbundled) products and services, and to, at the same time, preserve economic efficiency associated with bundling, novel technical, economic and information solutions are needed. We identify the design of compatible incentives for preserving economies of scope in an otherwise unbundled industry as a difficult research and development area. We suggest that one possible way forward is to design IT-supported protocols for acquiring, learning, exchanging and processing information relevant for coordinating decision-making by various industry participants. As the single centralized decision making of the old regulated industry is being replaced by a mix of groups of distributed decision makers, the design of protocols for their interactions for technically acceptable and economically efficient industry performance becomes essential. The basic principles for designing these protocols are discussed.

In conclusion, we have shown using simple electric power networks that components of once bundled electricity services commonly referred to as the "ancillary" services could have significant effects on both technical and economic industry performance. As such, their related products and services must be carefully defined, and managed in order for the industry to make the most out of the available resources and to evolve as new technologies present themselves with the novel values to the consumers.

I. Economies of Scope in Electric Power Systems: Basic Concepts

Operations of electric power systems are highly coordinated. Under ideal assumptions of (1) perfect information; and (2) supply and demand functions of each producer and supplier being independent from what the others are doing, scheduling of power plants for forecast real power demand can be done both in a coordinated and competitive way resulting in the same short-term (static) social welfare [1]. This forms the basis of market designs currently pursued for implementing competitive electricity markets. However, delivering reliable electricity when the demand fluctuates around its forecast, and/or power plants and transmission lines fail requires coordination. This is necessary to ensure that the system, as a whole, operates reliably, but also in order to supply demand as efficiently as possible during these uncertain loads and changes in equipment status. Moreover, unique to electric power systems, in addition to supplying real power, it is essential to support reactive power and voltage throughout the system to compensate for inductive loads and inductive losses in the transmission and distribution grid.

In a vertically-integrated industry the information concerning the status and availability of the system-wide resources is generally assumed known to the system operator of each utility. Today's hierarchically coordinated system operation implies that sharing the available resources results in a fairly reliable and fairly efficient overall performance. (A closer look into current operating and planning practices reveals that much could be improved by more adaptive management of available resources.) Moreover, customers are not differentiated according to how much of specific resources they use. Instead, they are mostly charged according to the class they belong to (commercial, residential, industrial). Large industrial customers are sometimes given an option to install their own reactive power / voltage support, most frequently capacitors banks. This is in exchange for the reactive power charge. Even in this case it is considered to be very difficult to differentiate customers according to their exact impact on the system. More generally, the charges for ancillary services are shared according to some pre-agreed upon rule because it is hard to know exactly who needs what and when these services are needed, and to charge accordingly.

In the changing electric power industry, however one may need more product differentiation and more decentralization, if possible. This can be done by viewing the electric power system as a network of various agents producing and consuming different power at different locations, different times and different quality of service. One could then begin to formalize the notion of economies of scope for the electric power industry. Each given electric power industry structure would have different products and services classification for which one needs to assess presence of economies of scope.

A. Unbundling real power and its delivery

As mentioned above, one of the basic objective in this paper is to use the notion of economy of scope in the electric power systems in order to assess the necessary coordination and communication needs among the generation, transmission and demand. To pursue this goal, we consider in this section the simplest notion related to the fundamental unbundling of transmission services (delivery) from the power production (generation). In this case one could say that if the cost of bundled generation-transmission service to consumer is smaller than the sum of separated costs of providing generation and transmission, then the unbundled service exhibits economies of scope. Expressed in mathematical terms:

$$C(G,T) < C(G,0) + C(0,T)$$

- $C(G,T)$ represents the cost of generation-transmission service bundled for providing electricity to the end user. Generators and transmission work jointly as an intergraded agent with the purpose of meeting the demand at the minimum cost, taking into account all technical requirements and transmission system constraints. In other words, supply and delivery is a single product.
- $C(G,0)$ represents the generation cost of producing real power required by the demand. This cost is borne by certain power generation company. Here we assume, without loss of generality, that there is only one generation company.
- $C(0,T)$ is the transmission cost that certain transmission company is charging for delivering real power to the consumer.

In order to assess economies of scope of this type in the electric power systems, it is necessary to solve an optimization problem. Once this is computed, the following inequality needs to be satisfied for the economies of scope between generation and transmission to exist;

$$\min_{P_{G_i},\theta_i}\left(\sum_i^n C_i(P_{G_i}) + T_T(P_{ij})\right) \le \min_{P_{G_i},\theta_i}\sum_i^n C_i(P_{G_i}) + \min_{P_{ij},\theta_i} T_T(P_{ij})$$

B. Some simple examples

In general, we consider a quadratic polynomial generation cost function for each generator in the system representing its operating cost.

$$C_i(P_{G_i}) = a_i P_{G_i}^2 + c_i P_{G_i} + e_i$$

We consider the transmission cost using the following transmission charge rule:

$$T_T(P_{ij}) = \sum_i \sum_j d_{ij} \cdot P_{ij}^2$$

The transmission cost described by the above equation, represents the utilization of the transmission line, in order words, a flow-based type charge.

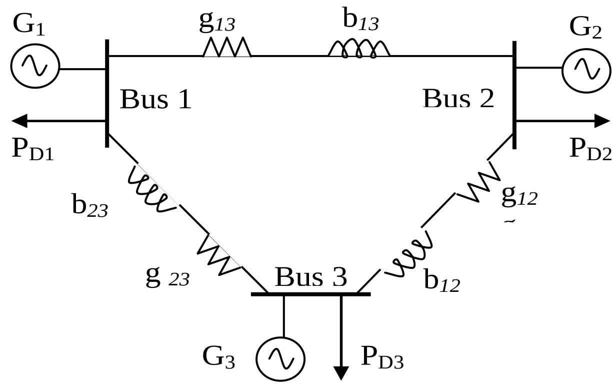

Fig. II.1. A typical small 3-bus electric power system representation

Consider a typical small electric power network shown in Figure II.1 first. For simplicity only a lossless network in the optimization problem is used. In order to study the cost performance of the system, the total system demand is increased; as a result, the unbundled cost presents a slightly higher cost compared with the bundled cost of service.

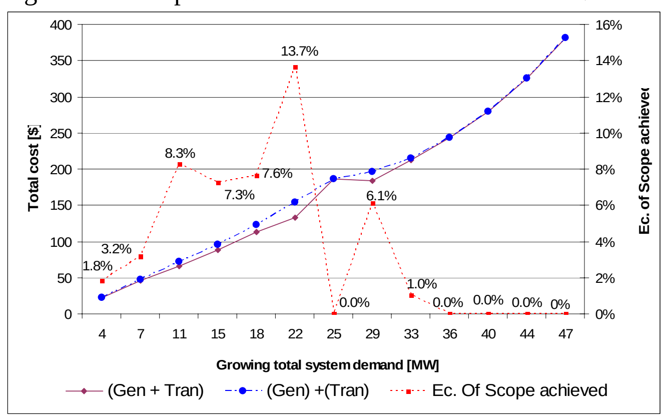

Fig. II.2. Cost comparison between bundled and unbundled services

When the generators G1 and G2 reach their maximum capacity constraints (at 25 MW), the economies of scope are reduced to zero. Then when the transmission line L12 starts reaching its maximum transmission capacity, the economies of scope drop to zero until the system is no longer stable.

A comparison of the nodal price is presented in the Figure II.3-II.4.

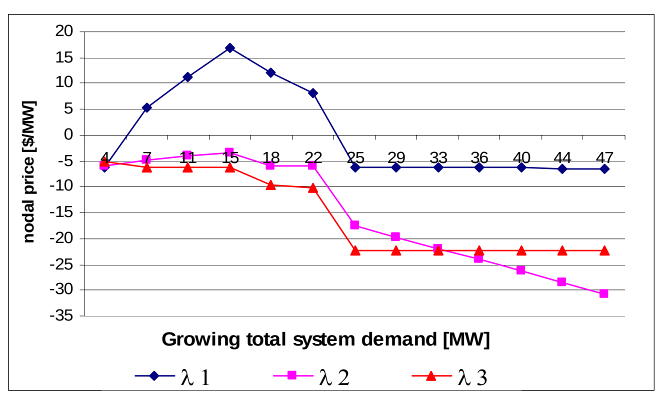

Fig. II.3. Nodal price for bundled service

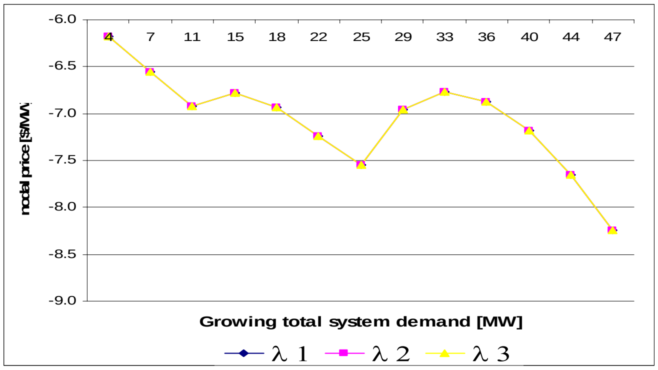

Fig. II.4. Nodal price for unbundled service

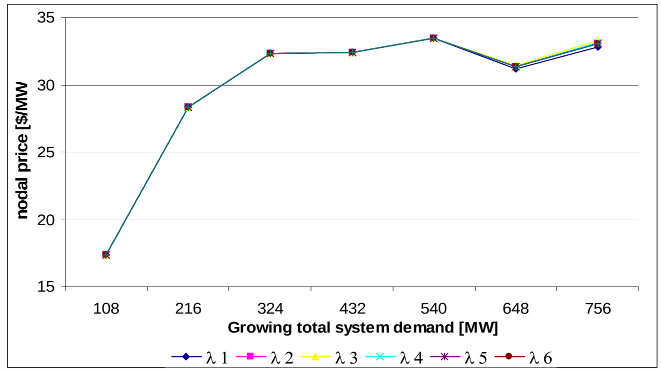

Fig. II.7. Nodal price for unbundled service

Following the same methodology as above, we consider a power system composed of 7 buses, 5 generators and 10 transmission lines. The economies of scope achieved in this case are higher than the three buses example. This case shows a more robust system where more transmission lines are used in order to deliver electricity. Therefore the flow-base charge for transmission becomes to have more meaning in terms of cost.

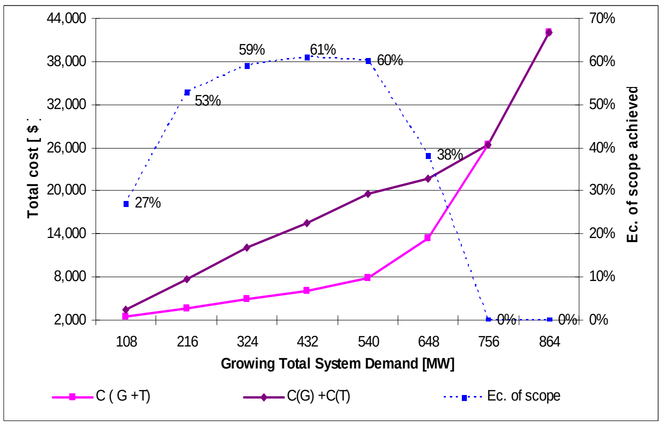

Fig. II.5. Cost comparison between bundled and unbundled services

As a result, the unbundled service presents a higher cost (up to 61%) compared with the bundled service. The savings associated with the economies of scope drop when G2 reaches its maximum capacity constraints (at 180 MW). Then when the transmission line L13 approaches its maximum transmission capacity, the economies of scope tend to zero until the system is no longer stable.

A comparison of the nodal price is presented in the Figure II.6-II.7

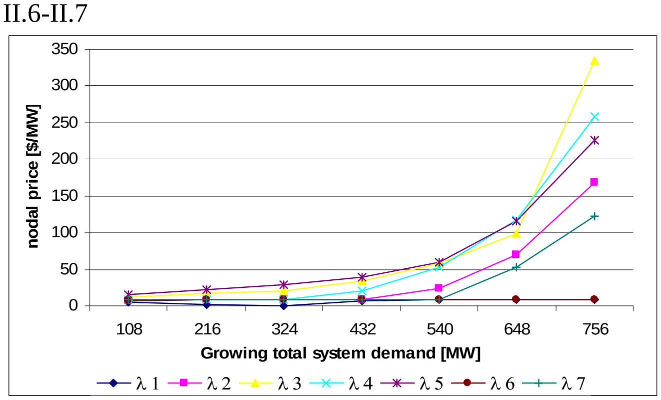

Fig. II.6. Nodal price for bundled service

The results provide numerical elements of economies of scope between generation and transmission, which are translated in cost saving for the industry. It also important to point out that the cost saving associated are hardly depended on the network topology. However, our purpose is not to promote the vertically-integrated of the electric power systems or more over, to return to the former stage. This result expounds the strong value of coordination in the system to ensure reliability and better cost in these utilities which are already unbundled and for those who move into this structure. These results could be used as a measuring method of feasible and efficient unbundling of the services.

These general issues are studied by Kaserman and Mayo (1991), Kwoka (2000), all corroborate that the "vertical economies" could represent cost savings compared within the regulated electric utilities. These works point out that those firms which are nearly fully integrated are more likely to capture the "vertical economies", which is translated in a better price.

Data

$$C_i(P_{G_i}) = a_i P_{G_i}^2 + b_i P_{G_i} + c_i$$

TABLE II.1
3-BUS EXAMPLE GENERATOR CHARACTERISTICS AND LOAD DATA

Node	Demand	Generator	a	b	c
1	0.9	G1	0.002	6.3	0
2	1.8	G2	0.01	5.8	0
3	0.9	G3	0.0014	6.2	0

TABLE II.2
7-BUS EXAMPLE GENERATOR CHARACTERISTICS AND LOAD DATA

1	-	G1	.002	7.62	375.5
2	30	G2	.0014	7.519	403.61
3	9	-	-	-	-
4	30	G4	.0013	7.836	253.24
5	9	-	-	-	-
6	15	G6	.0013	7.537	388.93
7	15	G7	.0019	7.771	194.24

$$T_T(P_{ij}) = \sum_i \sum_j d_{ij} \cdot P_{ij}^2$$

TABLE II.3
D_{IJ} VALUES

	3-bus example	7-bus example
dij	1.2	1

The transmission nodal price considered is the total transmission price divided by the total demand. Therefore, the nodal unbundled cost is:

$$c(P_{Di}) = \lambda_i + T_{Total}(P_{ij}) = \lambda_i + \frac{\sum_i \sum_j d_{ij} \cdot P_{ij}^2}{P_{Dtotal}}$$

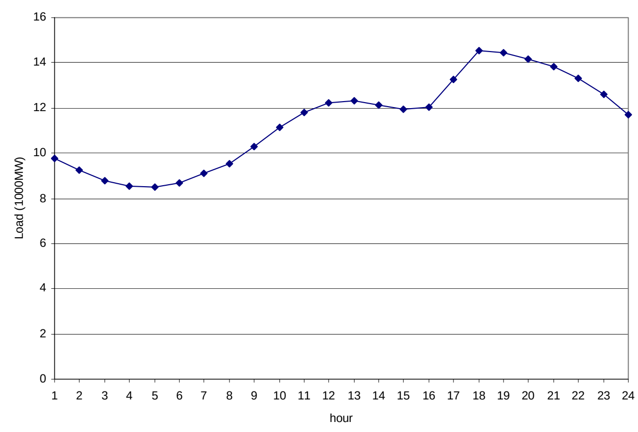

Fig. III.1. 24 hours load profile

II. Effects of Congestion on Generation Dispatch, LMPs, and Generation Profits

In an attempt to better understand the effects of congestion on an electric power network, simple simulations can be run. By varying one parameter at a time, the effect of that parameter on nodal prices can be observed. In these simulations, it is assumed that the demand is inelastic, and that the generators bid only their marginal cost. The sample network is shown in Figure III.1.

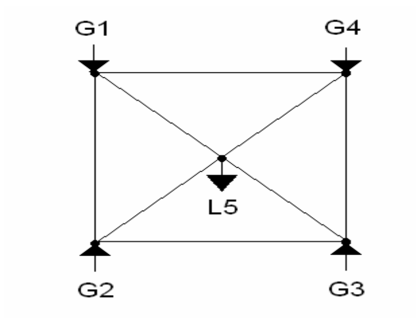

Fig. III.1. A simple 5-bus electric power system

The generator characteristics are shown in Table III.1.

TABLE III.1
5-BUS EXAMPLE GENERATOR CHARACTERISTICS

Generator	P_{min}(MW)	P_{max}(MW)	a	b	c
Gen 1	0	8000	0.001	0.8	0
Gen 2	0	3125	0.002	14.8	0
Gen 3	0	2000	0.003	31.3	0
Gen 4	0	8000	0.004	46.9	0

Bus G1 is arbitrarily assigned as the slack bus. Furthermore, the transmission lines are assumed to be lossless and have a reactance of X=0.001 p. u. The simulated 24 hours load profile is shown in Figure III.1.

The network is first simulated under no congestion, and then a current rating of 4500 MVA is imposed on line 1-5. Also, bidding is limited to the spot market. The generator real power output without congestion considerations and out of merit generation with congestion are shown in Figure III.2-III.3, respectively. It can be seen that, due to congestion, the cheapest generator can not be dispatched, and the power is produced, instead, by the very expensive generator.

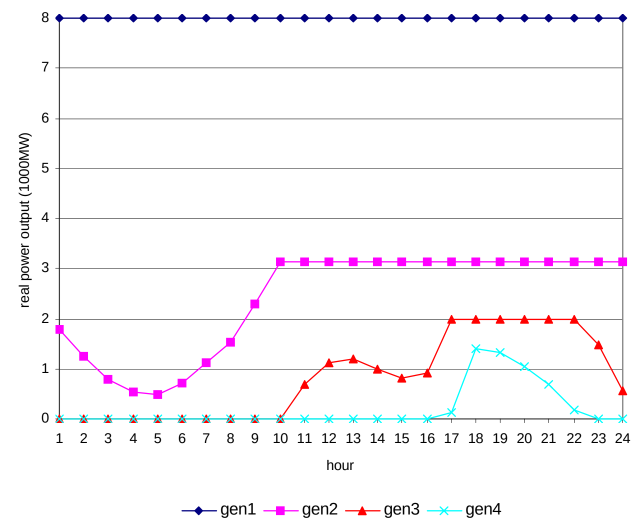

Fig. III.2. Generator real power output w/o line 1-5 capacity limit

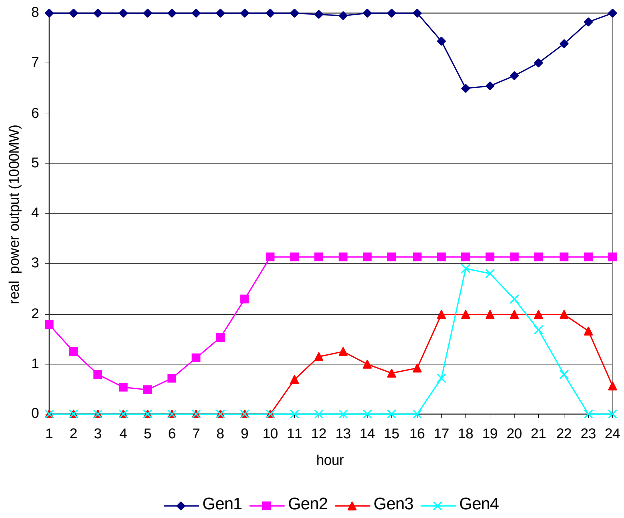

Fig. III.3. Generator real power output w/ line 1-5 capacity limit

The Locational Marginal Prices (LMPs) can be calculated using the method described in [1]. The results are shown in Figure III.4.

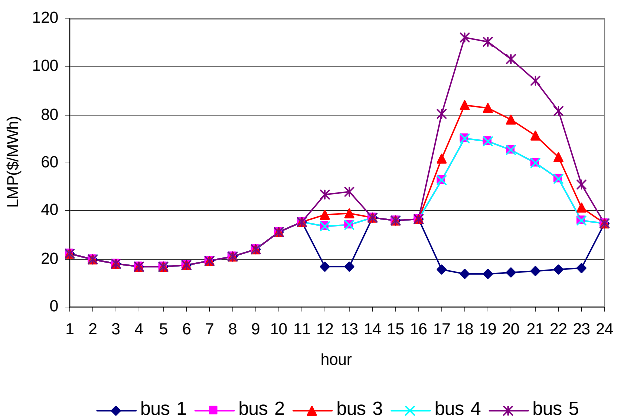

Fig. III.4. LMPs w/ line 1-5 capacity limit of 4500 MVA

A. Spatial Characterization of LMPs: Conjectures and Open Questions

Observing the nodal prices under congestion leads to the conjecture that nodal prices are a function of the node's electrical distance from the congestion. The distance between node i and node j can be defined as the net impedance seen between nodes i and j. This assumption makes intuitive sense, and can be tested by simulating the same network with a different line congested, and by varying the impedance of the lines.

Figure III.5 shows the LMPs when the reactance of line 1-4 is increased from 0.001 p. u. to 0.002 p. u. All other impedances remain at X=0.001.

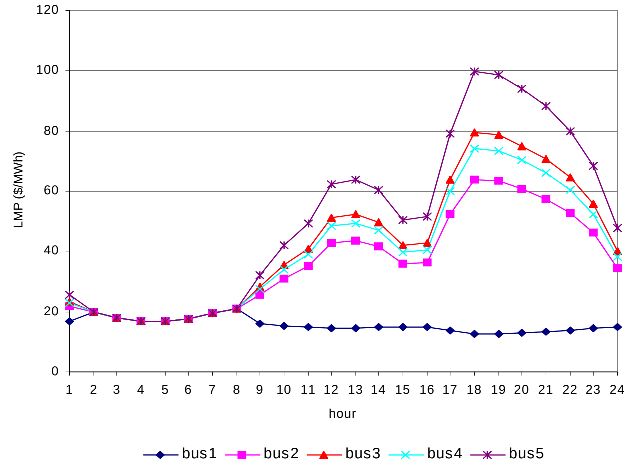

Fig. III.5. LMPs w/ line 1-4 reactance at 0.002 p. u.

These results further support our assumption. The prices at nodes 2 and 4 are no longer equal. This is due to the fact that node 4 now sees a greater impedance between itself and line 1-5 than node 2 does. Notice that the effect of the congestion on nodal prices is still inversely proportional to the node's distance from the congestion.

Next, congestion is created in line 2-5 by imposing a current rating of 3500 MVA on line 2-5. The current rating is removed from line 1-5, and all reactance are once again set to X=0.001. Figure III.6 shows the LMPs with congestion in line 2-5. Once again, these results support our assumption that nodal prices can be expressed as a function of a node's distance from the congestion. Bus 2 is the most affected by the congestion, while prices at bus 1 and bus 3 are affected equally. Bus 4, which is the farthest from the congestion, is the least affected.

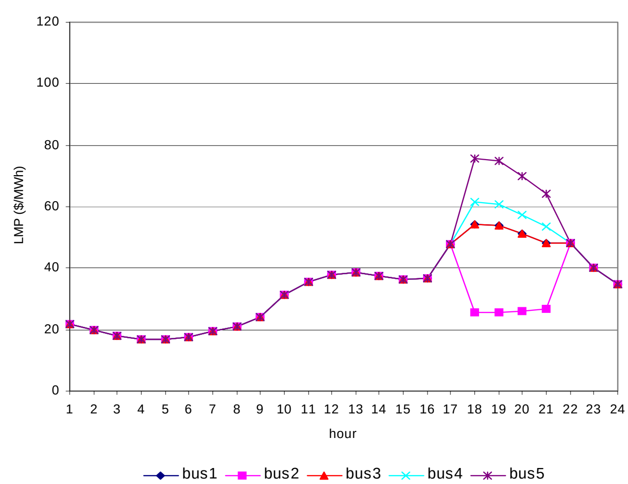

Fig. III.6. LMPs w/ line 2-5 capacity limit of 3500 MVA

III. The Inter-temporal Effects on Industry Performance

In this section, we consider a new market structure referred to as the stratum market [2]. In a stratum market, power is sold in multiple auctions. Each auction will sell contracts of different lengths, i.e. the yearly auction will sell a contract to inject a given quantity over the course of a year. In this paper, we will consider a stratum market in which generators can bid into yearly, monthly, and hourly auctions.

To determine how much power can be sold in the yearly auction, the Independent System Operator (ISO) will forecast the minimum load that the network will incur during the upcoming year, and sell that quantity in the yearly auction. After the yearly auction is settled, there will be a monthly auction. In this auction, the ISO will forecast the minimum load that the network will incur over the next month. The monthly auction will then be held each month, and the quantity sold will be equal to the monthly minimum forecast minus the quantity already sold in the yearly market. The hourly auction will then be held each hour, and the quantity sold will be the actual load for that hour minus the sum of the quantities already sold in the monthly and yearly markets. Figure VI.1 represents a three year load, projected into the stratum market. The data is the actual load profile of ISO New England from 1994-1996.

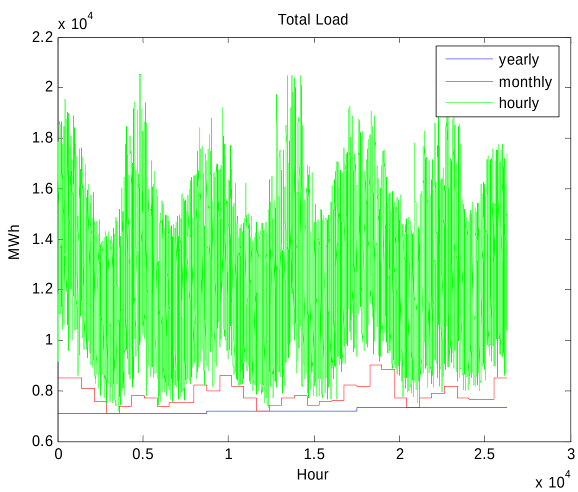

Fig. VI.1. Load profile in a stratum market

The goal of the stratum market is to hedge against the financial risk of recovering capital cost. By allowing generators to bid on long term contracts, they are guaranteed to recover a large portion of their capital cost. Thus, an investor can secure a long term contract and be guaranteed to be scheduled to inject a fixed quantity at a fixed price over the length of that contract.

When trying to simulate the stratum market, it becomes obvious that the generators will have different bidding curves for different markets. These curves will be a function of their marginal cost curve, and a function of the quantities and prices that they have bid in the previous auctions. The revenue that a generator receives per hour from the yearly auction can be defined as

$$RY = QYPY$$

where RY is the hourly revenue from the yearly auction, QY is the quantity that that generator will inject per hour over the course of the year, and PY is the price that the generator will receive per unit injected. Similarly, hourly revenue from the monthly and hourly auctions can be defined as

$$RM = QMPM$$

and

$$RH = QHPH$$

Now, total revenue for each hour can be defined as

$$RT=RY+RM+RH$$

Because of the market structure, generators will be first bidding into the yearly auction, then bidding into the monthly auctions, and lastly bidding into the hourly auctions. Assume that P(Q) is a generator's marginal cost curve. RY can now be defined by the relationship

$$RY=P(QY)QY$$

thus the generator's bidding curve for the yearly market will be

$$PY=P(QY)$$

Next, the generator will bid into the monthly market. The generator will now be producing a quantity of QY+QM. Thus, to satisfy its marginal cost curve, the generator will be forced to bid such that

$$RY+RM=P(QY+QM)\ (QY+QM)$$

substituting for RY and RM we obtain the generator's bidding curve for the monthly market. Note that PY and QY will be known after the conclusion of the yearly auction.

$$PM=1/QM(P(QY+QM)(QY+QM)-P(QY)QY)$$

The same strategy is used for bidding into the hourly market, only now we must consider the quantity being produced in both the yearly and monthly markets. To satisfy its marginal cost curve, the generator will be forced to bid such that

$$RY+RM+RH=P(QY+QM+QH)(\ QY+QM+QH)$$

substituting for RY, RM, and RH we obtain the generators bidding curve for the hourly market. Note, once again, that PY, QY, PM, and QM will be known after the conclusion of the yearly and monthly auction.

$$PH = 1/QH(P(QY+QM+QH)(QY+QM+QH)-P(QY+QM)(QY+QM))$$

One can see that since the generators are assumed to be bidding such that their total revenue matches their marginal cost curve for total power produced, the net quantities injected, and net revenues will be the same for the stratum market as they would be in the spot market with the generators bidding their marginal cost. Thus, we have created a market that is equivalent to the spot market, while allowing for generators to hedge against the risk of their capital cost.

Simulating the Stratum Market

The stratum market can be simulated by using the bidding strategy outlined above. The network used is the same five node network shown in Figure III.1. We will first simulate the market under no congestion. The load used is the load that is shown in Figure VI.1. Once again, it is assumed that the demand is inelastic, and that the generators bid according to the bidding strategies listed above. The monthly and yearly minimum loads have been forecast using load data from the previous two years. The quantities injected by each generator are shown in Figure VI.2-VI.5.

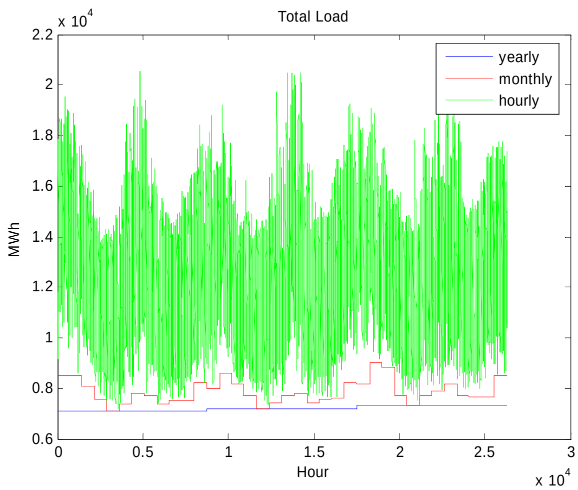

Fig. VI.2. Generator 1 real power output in the stratum market

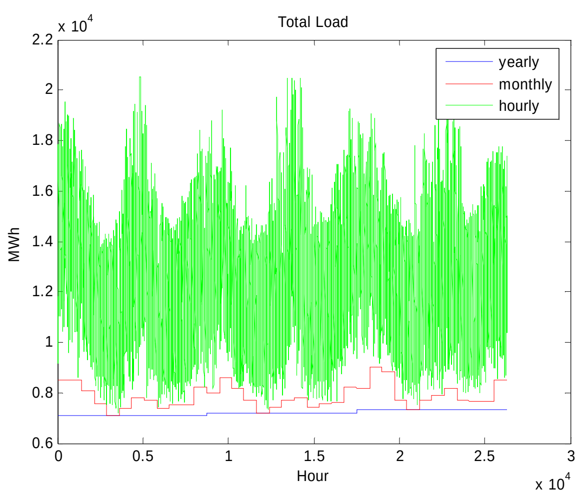

Fig. VI.3. Generator 2 real power output in the stratum market

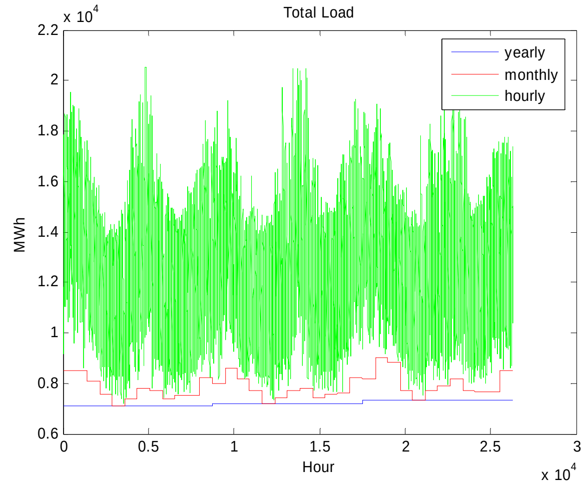

Fig. VI.4. Generator 3 real power output in the stratum market

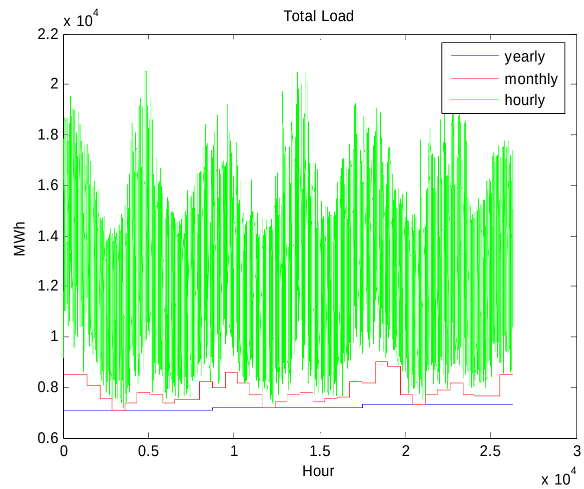

Fig. VI.5. Generator 4 real power output in the stratum market

The prices in each auction, as well as the prices considering only the spot market are shown in Figure VI.6-VI.7.

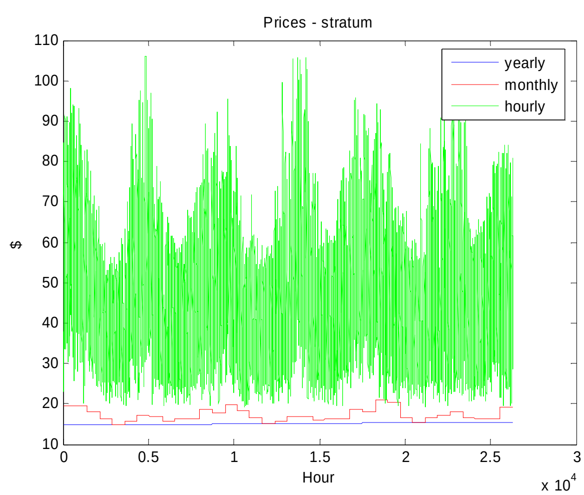

Fig. VI.6. LMPs for the stratum market

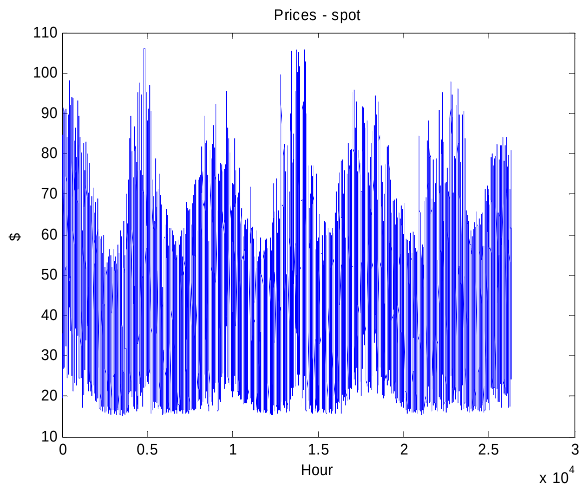

Fig. VI.5. LMPs in the spot market

When examining total revenue, it becomes apparent that the total cost to the load is much lower in the stratum market than in the spot market. This is because there is a discontinuity between the bidding curves of generators 1 and 2, and the bidding curves of generators 3 and 4. This means that the marginal cost curves of generators 3 and 4 will cause them to bid a price that is higher than the price bid by generators 1 and 2 at their maximum capacity. Current market rules allow generators 1 and 2 to be paid this higher price for the power that they are injecting. When power is sold only in the spot market, generators 1 and 2 will be paid this inflated price for the total quantity that they are producing each hour. In the stratum market, this discontinuity is not reached in the yearly or monthly auctions, because generators 3 and 4 are not scheduled. This means that generators 1 and 2 are paid what they bid in the yearly and monthly markets. They are still paid the inflated prices bid by generators 3 and 4 in the hourly market, however the quantities sold at this price afre much lower than the quantities sold in the spot market. Thus, in circumstances where there is a discontinuity in bidding curves, the stratum market will reduce the cost to the load, while still allowing the generators to recover their marginal cost.

IV. Mixed Inter-temporal and Transmission Constraints Effects on Industry Performance: Simulating the Stratum Market Under Congestion

Thus far, we have not considered the consequences of congestion on the stratum market. We will now apply the stratum market auctions to the congestion example described in section III. In this simulation, the quantities forecast in the monthly and yearly market are the same projected minimums used in the uncongested stratum market example. Figure V.1-V.4 shows the power outputs of each generator when line 1-5 capacity is set at 4500 MVA.

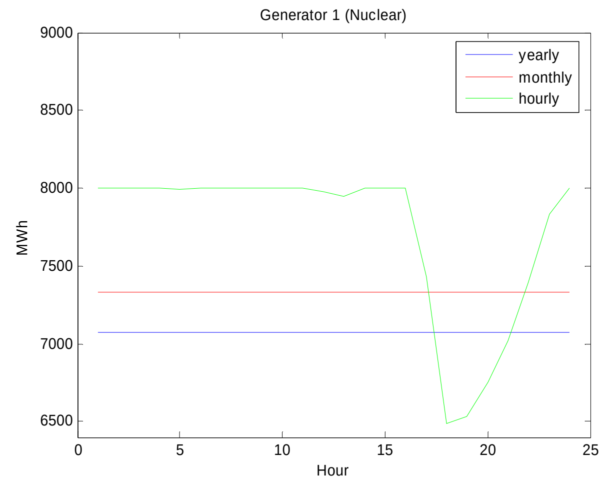

Fig. V.1. Generator 1 real power output in the stratum market

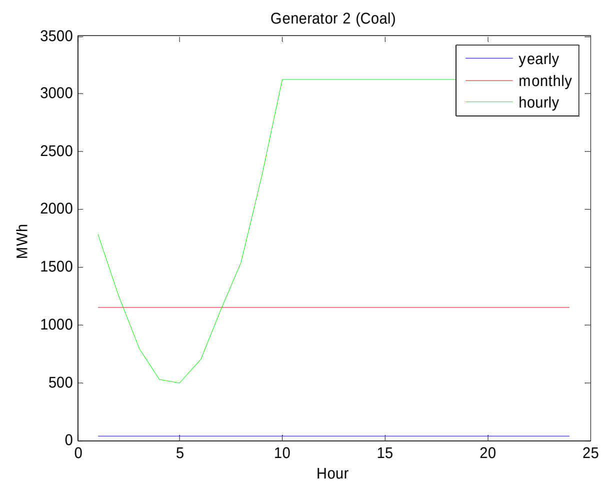

Fig. V.2. Generator 2 real power output in the stratum market

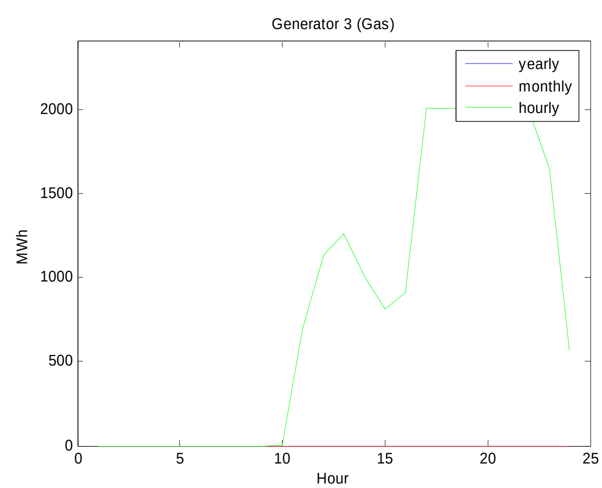

Fig. V.3. Generator 3 real power output in the stratum market

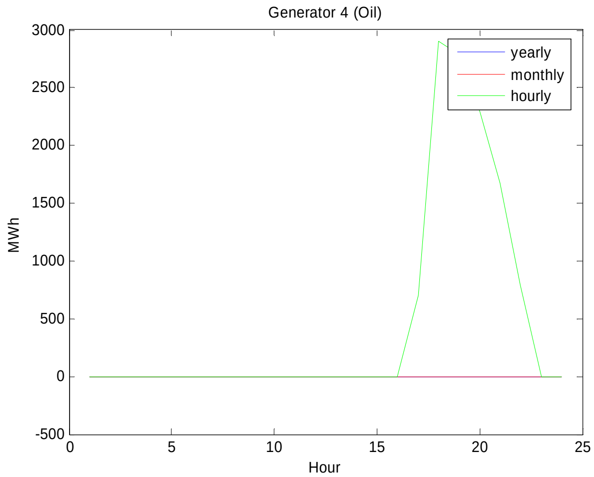

Fig. V.4. Generator 4 real power output in the stratum market

These figures indicate that in multiple markets, such as stratum market, care must be taken of their inter-dependencies. The outputs of generators 1 and 2 show that in order to obey line constraints, and meet the net demand, the generators are forced to inject a quantity that is lower than they have been contracted for. This happens because the longer term auctions do not consider congestion in the hourly market.

The solutions to this problem are many, and need to be researched further. A very simple solution would be to force generators that are unable to meet their obligations to buy power off of generators that are selling power on the hourly market. However, this would create additional risk for the generators bidding into the longer term markets. Since the purpose of the stratum market is to hedge against risk, this may not be the most desirable solution. Another solution would be to have the ISO forecast the maximum load in addition to the minimum load. The ISO could then run a simple DC OPF simulation of the network using the maximum projected load, and determine how the different nodes will be affected by congestion. This would allow the ISO to limit generators' bids in the longer term markets to a quantity that they can consistently produce under congestion. While this is much more appealing than the first proposed solution, further research may lead to more optimal solutions. Another solution is to simply impose the minimum power that must be scheduled in one market based on what had been committed in the other markets. Further work is needed here.

V. The Effects of Reactive Power/Voltage Support on Industry Performance

In this section we consider in some detail the effects of reactive power/voltage support on the electric power industry performance. The same five-bus example described in sections above is adopted. In order to consider these effects, the load at bus 5 is assumed to have a fixed power factor of .8. Bus voltages are allowed to vary between .95 p. u. and 1.05 p. u. The transmission line connecting buses 1 and 5 (line 1-5) has a limited capacity transfer of 4,500 MVA while all other lines have infinite capacities.

A. *Case 1: DC OPF*

In order to assess the effect of non-linear AC power flow on the results, we consider the DC OPF results as the base case. The DC OPF results of LMPs and individual generator real power outputs are shown in Figure III.4 and Figure III.3. The individual generator profits and congestion revenues are shown in Figure VI.1-VI.2.

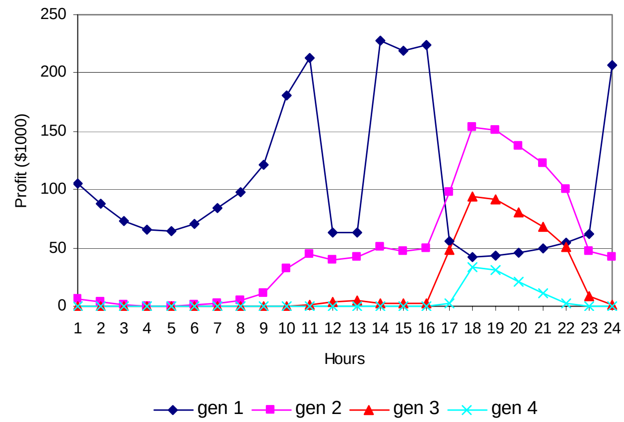

Fig. VI.1. LMPs in DC case

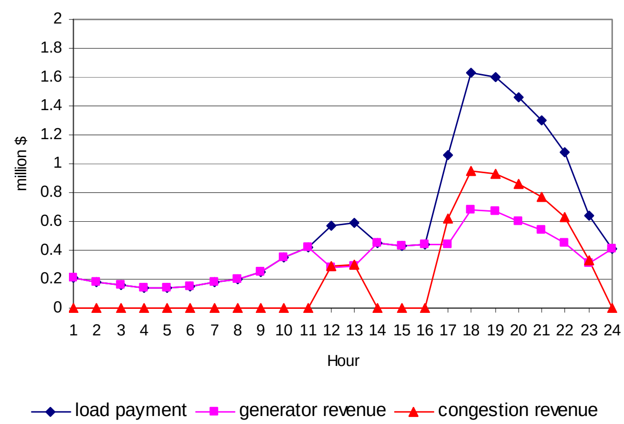

Fig. VI.2. LMPs in DC case

B. *Case 2: AC OPF without reactive power output limits*

Now we consider the results with AC OPF constraints. The objective is also to minimize the total real power generation cost with the decision variables as real power. The equality constraints are real and reactive power load flow equations and the inequality constraints are real and reactive power generation limits, voltage limits and the transmission limits (line 1-5 limit). The optimization problem is formulated as follows:

$$
\begin{array}{ll}
\min\limits_{P_{gi}} \sum c_{gi}(P_{gi}) & \\
\text{such that} & \\
P(V,\theta) - P_{gi} + P_{Li} = 0 & \text{real power balance} \\
Q(V,\theta) - Q_{gi} + Q_{Li} = 0 & \text{reactive power balance} \\
\left| S_{ij} \right| \le S_{ij}^{\max} & \text{power flow limit of lines} \\
V_i^{\min} \le V_i \le V_i^{\max} & \text{bus voltage limits} \\
P_{gi}^{\min} \le P_{gi} \le P_{gi}^{\max} & \text{real power generation limits} \\
Q_{gi}^{\min} \le Q_{gi} \le Q_{gi}^{\max} & \text{reactive power generation limits}
\end{array}
\qquad (5.1)
$$

In this section, we first consider the AC case with no reactive power generation limits. The LMPs difference, the generator real power output difference, generator profits difference and congestion revenue difference between DC case and AC case without reactive power limits as a percentage of DC results are shown in Figure VI.3-VI.6 respectively.

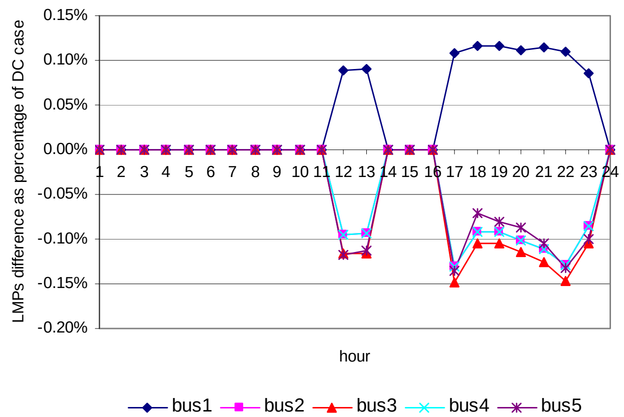

Fig. VI.3. LMPs difference between DC case and case 2

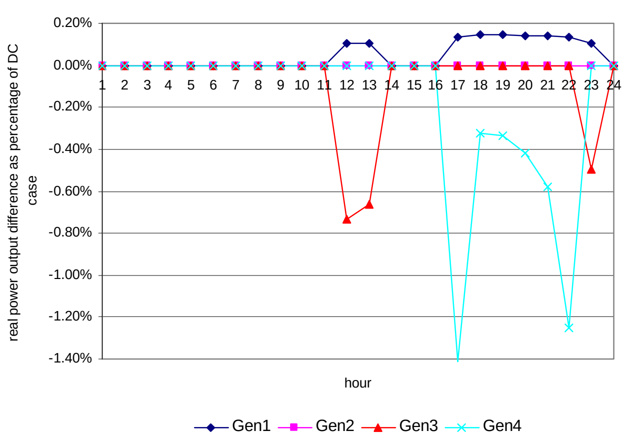

Fig. VI.4. real power generation output difference between DC case and case 2

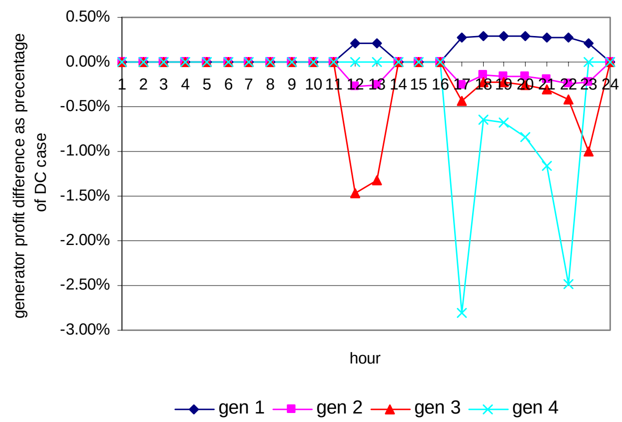

Fig. VI.5. generator profits difference between DC case and case 2

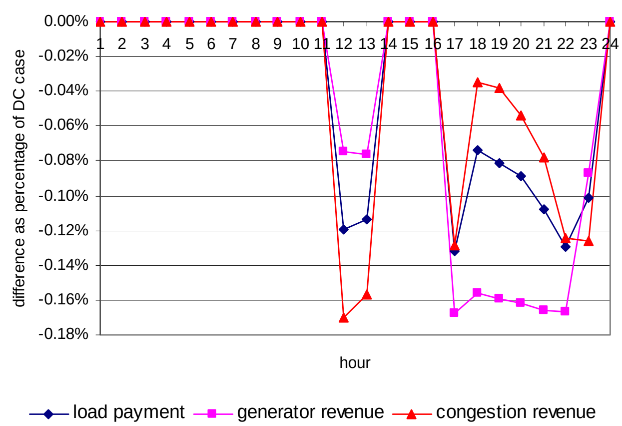

Fig. VI.6. load payments, generator revenue and congestion revenue differences between DC case and case 2

From these plots it is easy to see that the results between the DC case and AC case with no reactive power generation limits are very similar. The difference of the corresponding LMPs is within .15%. The real power difference is within 1.4% and differences of the profits of generators within 3%. The congestion revenue difference is within .18% since the individual generator revenues differences offset each other. The basic conclusion based on this set of simulations is that the differences between the DC OPF and AC OPF without reactive power generation limits are not significant.

C. Case 3: AC OPF with reactive power output constraints

In case 3 the effect of reactive power generation constraints are investigated. The reactive power generation limits, which are determined arbitrarily the same as real power limits, are specified as in Table VI.1.

TABLE VI.1
REACTIVE POWER GENERATION CONSTRAINTS

Gen #	Qmin (MVar)	Qmax (MVar)
Gen 1	-8000	8000
Gen 2	-3125	3125
Gen 3	-2000	2000
Gen 4	-8000	8000

The LMPs results are shown in Figure VI.6.

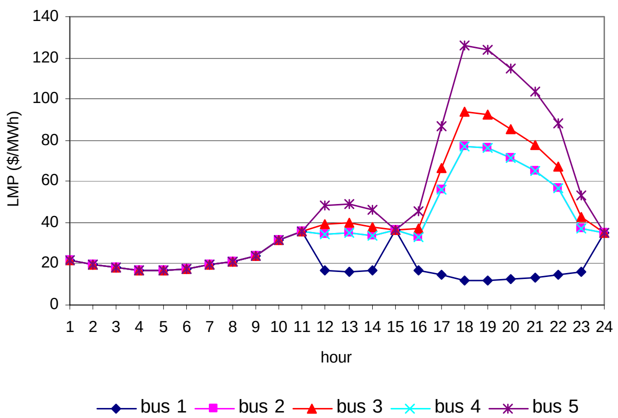

Fig. VI.6. LMPs in case 3

It is interesting to see from Figure VI.6 that even with very loose reactive power generation the LMPs changed substantially. Comparing Figure III.4 and Figure VI.6, LMPs at hour 14 and hour 16 are uniform at all buses in DC case while different in case 3. These LMPs difference is caused by the reactive power generation limits. To look at the difference in perspective, it is helpful to draw the difference as percentage of DC results. The LMPs difference, the generator real power output difference, generator profits difference and congestion revenue difference between DC case and AC case with reactive power generation limits as a percentage of base case results are shown in Figure VI.7-VI.10 respectively.

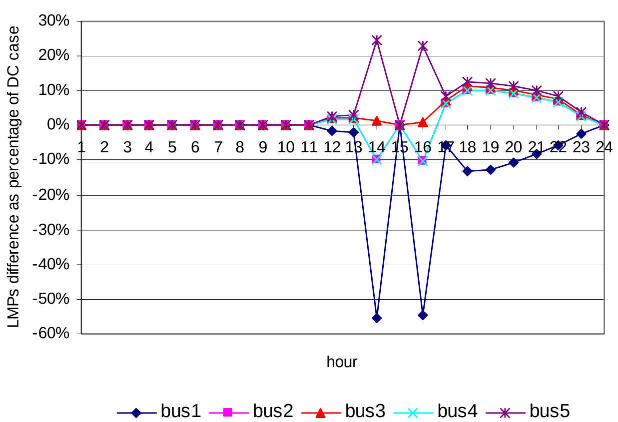

Fig. VI.7. LMPs difference between DC case and case 3

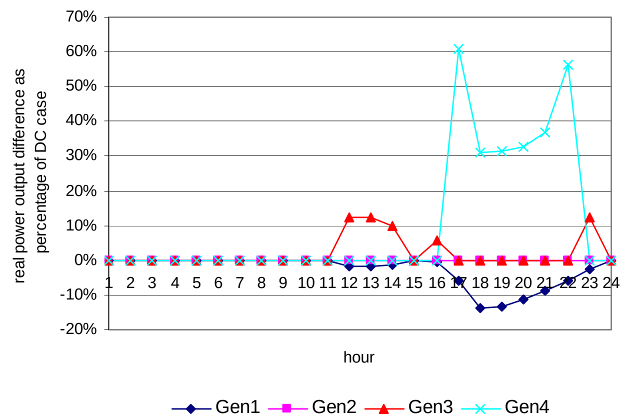

Fig. VI.8. real power generation output difference between DC case and case 3

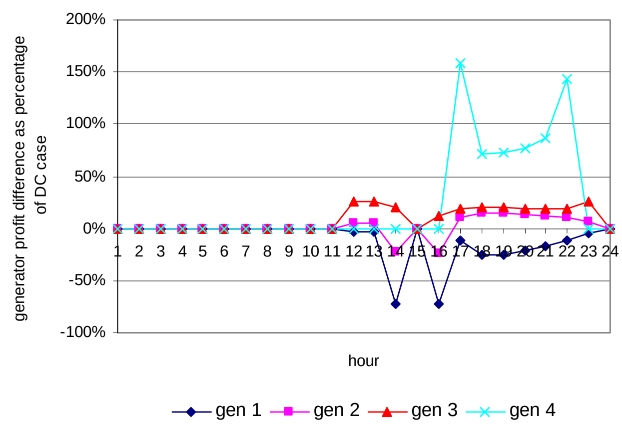

Fig. VI.9. generator profits difference between DC case and case 3

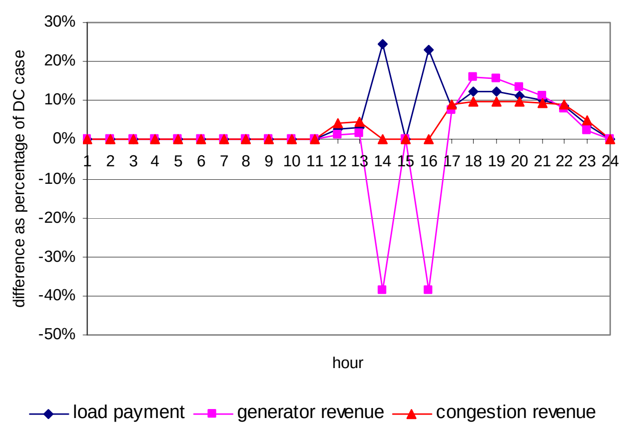

Fig. VI.10. load payments, generator revenue and congestion revenue differences between DC case and case 3

The maximum LMP difference occurred at hour 14 and 16 with a max gain of more than 20% at bus 5 and max drop of more than 50% at bus 1. The real power outputs of base load generator 1 and peak load generator 4 also changed significantly from hour 17 to 23 with a max gain of 60% and max drop 14%. The profits for individual generators changed accordingly with the generator 1 as the biggest loser and generator 4 as the biggest winner. And total congestion revenue increased considerably system-wide during the network congestion hours.

Overall, the reactive power generation limits affect the system LMPs and congestion revenues as well as the individual generators' real power output and profits significantly.

D. Case 4: AC OPF with reactive power output limits and linearly increasing load

The optimization problem (5.1) implies the following constraints may affect the LMPs: load demand, real power generation limits, reactive power limits, voltage stability limits and line capacities limits. In order to study how different kinds of constraints affects the final results as load increases, we revisit the case 3 with a linearly increasing load from 9000MW to 15000MW at bus 5. The power factor remained the same as 0.8.

Again, the DC OPF results are considered as base case. The LMPs and individual generator real power output are shown in Figure VI.11-VI.12.

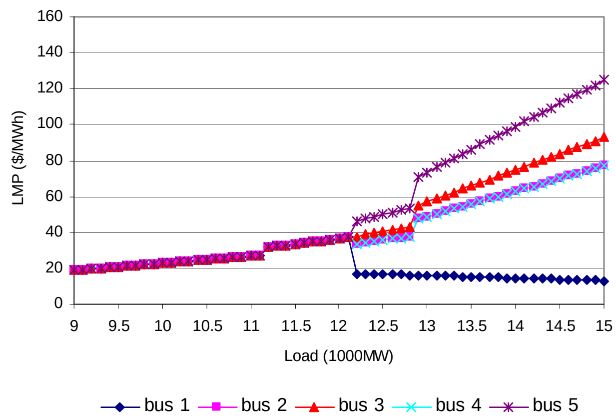

Fig. VI.11. LMPs under DC OPF and linearly increasing load

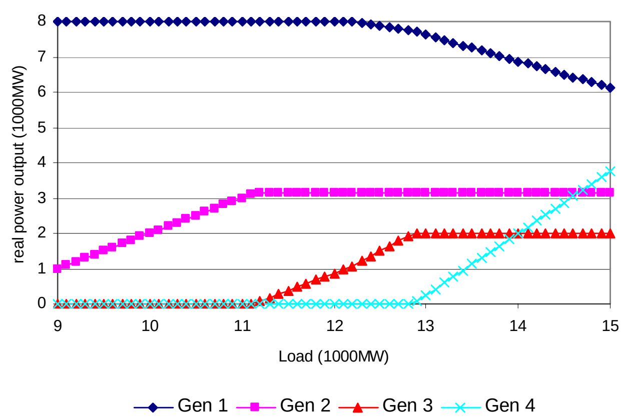

Fig. VI.12. Real power generation under DC OPF and linearly increasing load

The LMPs do not retain a linear relationship with load. There are three discontinuities in Figure 5.16. The first jump occurred at load point 11200 MW when generator 2 reached the maximum real power limit. The second breakpoint occurred at load point 12200 MW when line 1-5 reached the maximum capacity and the third one occurred at load point 12900 MW when generator 3 reached the maximum real power limit. It is also interesting to notice that the cheapest generator 1 began to decrease its real power output at second breakpoint because of the line capacity limit.

The individual generator profits and system congestion revenues results are shown in Figure VI.13-VI.14. The results bear the same non-linearity as LMPs.

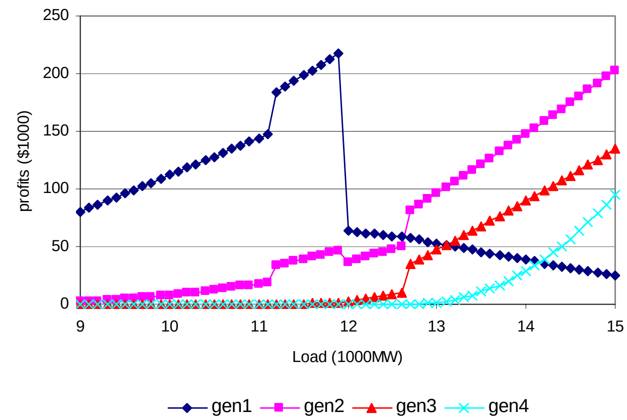

Fig. VI.13. Generator profits under DC OPF and linearly increasing load

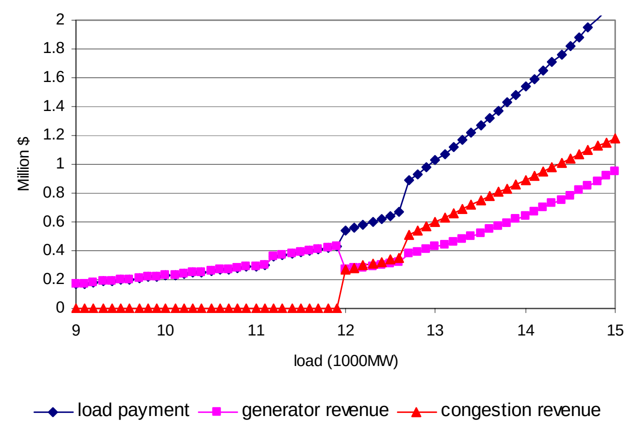

Fig. VI.14. load payments, generator revenue and congestion revenue under DC OPF and linearly increasing load

Now we consider the AC OPF case with reactive power limits. The results of LMPs, individual generator real power output, bus phase angle, individual generator reactive power output and bus voltage are shown in Figure VI.15-VI.19.

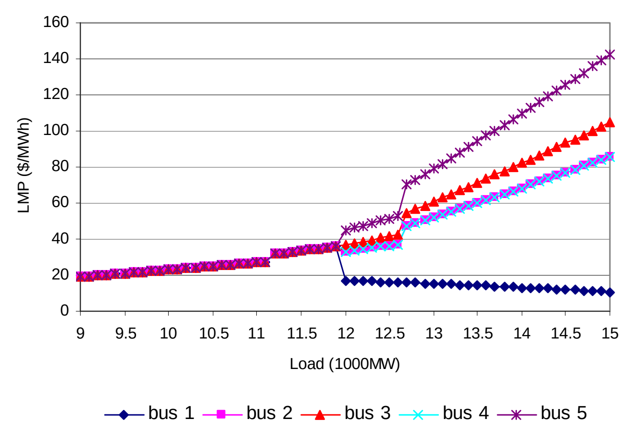

Fig. VI.15. LMPs with AC OPF and linearly increasing load

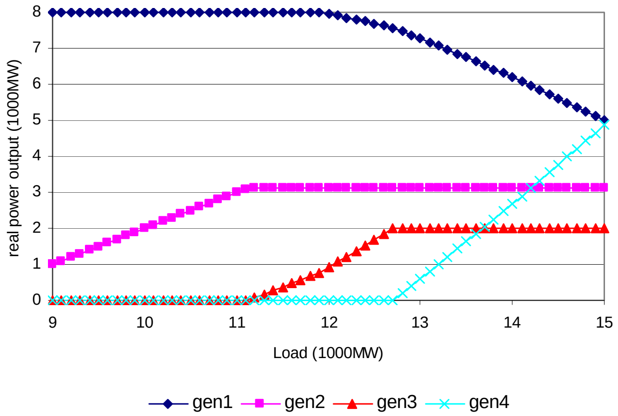

Fig. VI.16. power generation with AC OPF and linearly increasing load

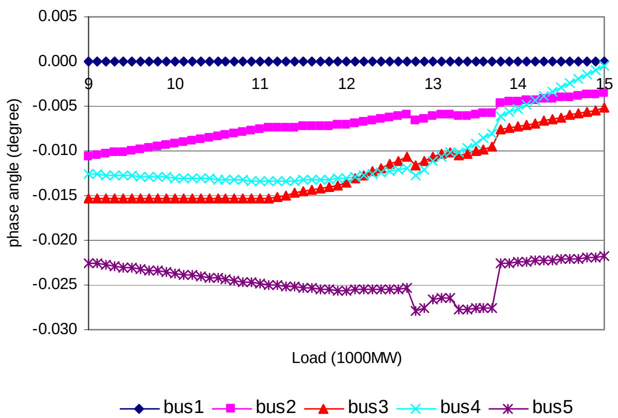

Fig. VI.17. bus phase angle with AC OPF and linearly increasing load

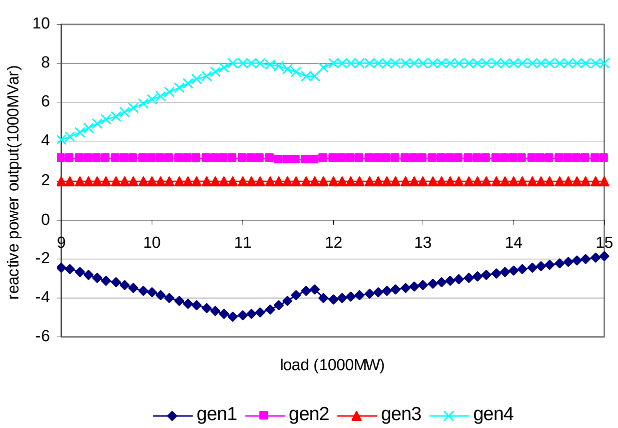

Fig. VI.18. reactive power generation with AC OPF and linearly increasing load

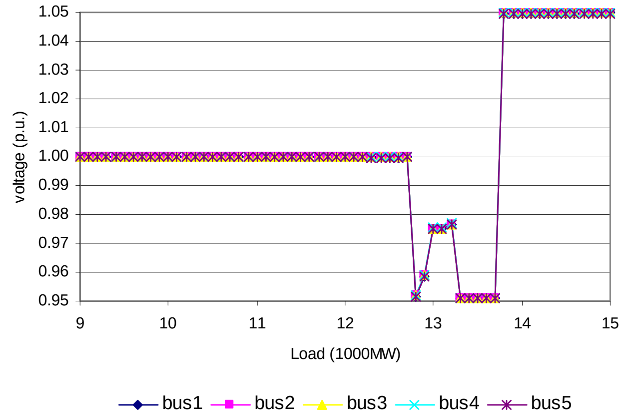

Fig. VI.19. bus voltage with AC OPF and linearly increasing load

There are also three discontinuities in Figure VI.15. The first jump occurred at same load point and caused by the same reason as DC case. The second breakpoint occurred earlier than the DC case, at load point 12000 MW instead of 12200 MW. And it is caused by the joined effect of the line 1-5 capacity constraint and maximum reactive power constraint of generator 1. The third one also occurred earlier than the DC case, at load point 12700 MW instead of 12900 MW, with the same reason as DC case. The reason that LMPs for AC case shifted leftward is that the reactive power flows also account for line capacity occupancy and thus the capacity limit for line 1-5 is reached ahead of time. This explained the LMPs difference for hour 14 and hour 16 between DC and AC cases in the previous section since the load level in hour 14 and 16 are between 12000 and 12200. Also, from Figure VI.18 we can see that the generator 1 reached max reactive power output limit when load is 10900 and 12000. However, this constraint only has effect on LMPs at second load level when the line 1-5 had already been congested. Thus, the reactive power generation constraints do not necessarily affect the LMPs and they have a much bigger impact when the network has already congested.

The congestion revenue difference between DC case and AC case with reactive power limits and linearly increasing load as a percentage of DC results are shown in Figure VI.20.

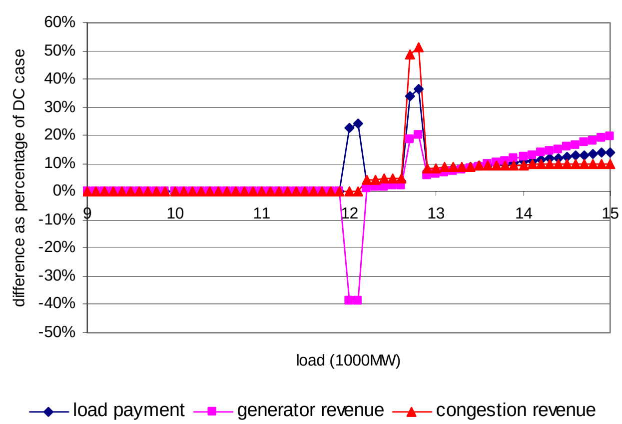

Fig. VI.20. bus load payment, generator revenue and congestion revenue difference between DC case and AC OPF and linearly increasing load case

Generally speaking, real power generation and line capacity constraints have a bigger impact on the LMPs. However the reactive power flow also has an impact on the line capacity and thus contributes to the network congestions. This may cause the LMPs left-shift effect on AC OPF model. We also suspect that the same reactive power generation constraints would also has a more significantly impact on the already congested network than the uncongested network.

Case 5: how reactive power generation limits affect LMPs

In order to examine more specifically how reactive power generation limits may affect the LMPs, we fix the load at 12000 MW level and change reactive power output limits of generator 4 from -10000 MVar to 10000 MVar given all the other generator characteristics the same. The LMPs, congestion revenues and individual generator profits are shown in Figure VI.21 -VI.23.

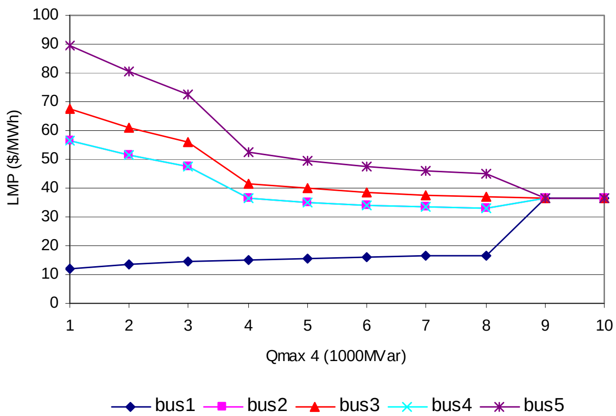

Fig. VI.21. LMPs as reactive power limits of generator 4 changes

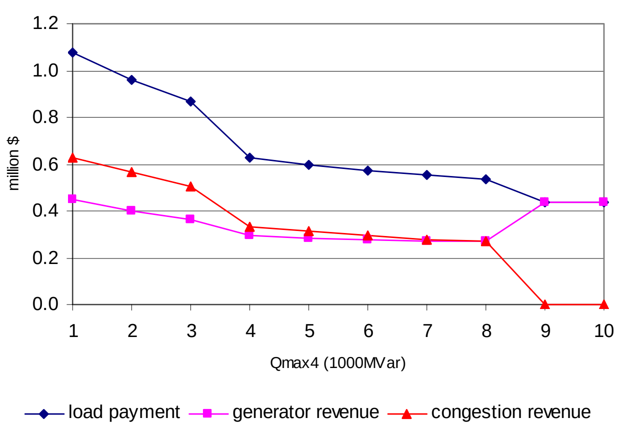

Fig. VI.22. load payment, generator revenue and congestion revenue as reactive power limits of generator 4 changes

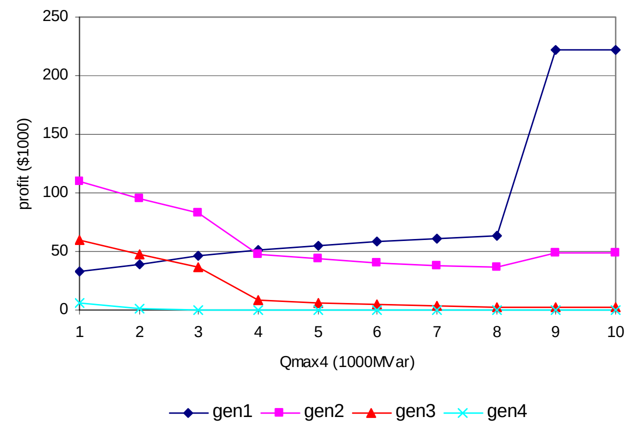

Fig. VI.23. generator profits as reactive power limits of generator 4 changes

The above figures demonstrated that reactive power generation capacity of individual generator alone also has a significant influence on system wide LMPs and congestion revenue as well as every generator's profits using the AC OPF model. As shown in Figure VI.23, the real power output and profit of generator 4 is negligible while its maximum reactive power generation capacity has a huge impact on the system LMPs, congestion revenue and other generator's output and profits.

Case 6: how line impedance affect congestion revenues

Now we relax the reactive power generator constraints for all generators and investigate how line reactance affects the results. The line reactance changes from 0.005 p. u. to 0.01 p. u. with a step of 0.005. The load payment, generator revenue and congestion revenues are calculated as the sum of the results of the same 24 hours load profile used in section III. The results are shown in Figure VI.24.

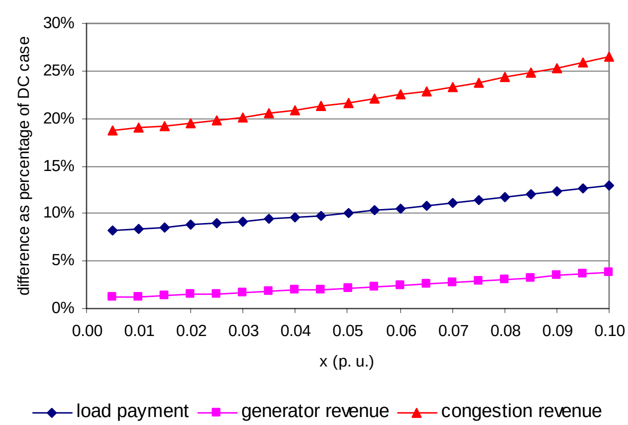

Fig. VI.24. load payment, generator revenue and congestion revenue difference as percentage of DC case as line reactance changes

The line reactance has a substantial impact on the congestion revenue and generator revenue. The relationship is quasi-linear.

VI. EFFECTS OF CONTROL SUPPORT ON THE ECONOMIES OF SCOPE: RESERVES, SYSTEM DYNAMICS, PROTECTION

Finally, we illustrate the most difficult aspects of economies of scope in the electric power networks. These are related to the system dynamics, and are relevant for a variety of reasons currently not addressed in the industry debate.

To introduce the basic notions here, consider a simple 3-bus electric power system shown in Figure VII.1. This system consists on 2 generators, where G1 is a thermal unit, and G2 is a hydraulic unit. Simulations are made using two load models for L2 and L3, constant impedance (z model), and constant power models (p model). Four cases are simulated for the two load models: without any control (cases 1z and 1p), with speed governors only (cases 2z and 2p), with AVR only (cases 3z and 3p), and with both governors and AVR (cases 4z and 4p).

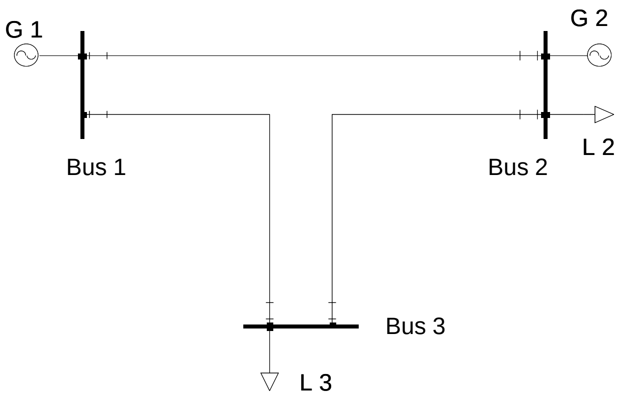

Fig. VII.1. A simple 3-bus power system

During normal conditions the system frequency and bus voltages remain within their limits, $f_{min}<f<f_{max}$, and $V_{min}<V<V_{max}$.

Now, assume that the system active load suddenly increased by 10% (or 10% of total generation is lost) at t=3s. Depending on the type of control devices in place, several fundamentally different outcomes are possible:

If there is no control (cases 1z and 1p), shown in Figures aa to dd, the system becomes unstable and collapses, because neither additional real power nor additional reactive power are injected into the system as a response to the disturbance.

With a governor control only, the system could survive up to about 40 seconds in the case 2z if V_{min}=0.93 and f_{min}=59Hz are accepted by the protection system. While it becomes voltage unstable a few seconds after the disturbance is applied, when a constant power load model is used (case 2p).

With an AVR control only, voltage is kept almost at nominal value until around 20 seconds, but the frequency dramatically decreases in both cases 3z and 3p. This shows that voltage is kept at expenses of transforming kinetic energy into electromagnetic energy, and the frequency decreases as a result of not having additional energy entering the system (Cm=const). Note that in case 3p the system survives longer to voltage instability than it does in case 1p (no control, p load model).

With both governors and automatic voltage regulator control (AVR), the system could survive longer as shown in Figures bb to ee. Here, it can be seen that both frequency an voltage remain near their nominal values ready to be corrected by secondary control actions (cases 4z and 4p).

These different times translate into different transfer limits on transmission lines (non-uniquely, a sign that flow-gates are very difficult to define line by line).

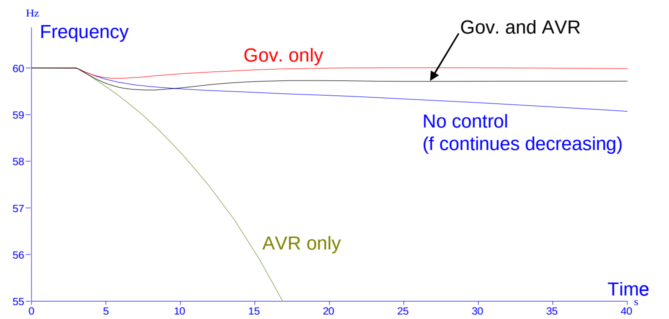

Fig. VII.2. Frequency as a function of time for the 4 cases simulated (1: no control, 2: governor only, 3: AVR only, 4: governor and AVR) when the system load is modeled as a constant impedance (z model)

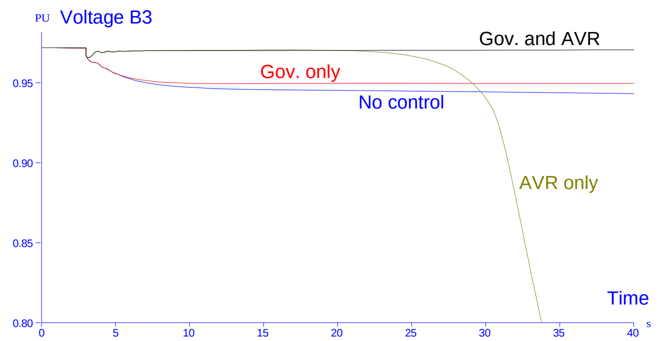

Fig. VII.3. A Voltage in bus 3 as a function of time for the 4 cases simulated (1: no control, 2: governor only, 3: AVR only, 4: governor and AVR) when the system load is modeled as a constant impedance (z model)

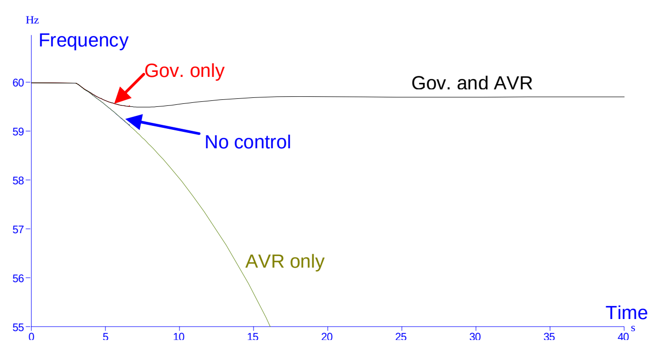

Fig. VII.1. Frequency as a function of time for the 4 cases simulated (1: no control, 2: governor only, 3: AVR only, 4: governor and AVR) when the system load is modeled as a constant power injection (p model)

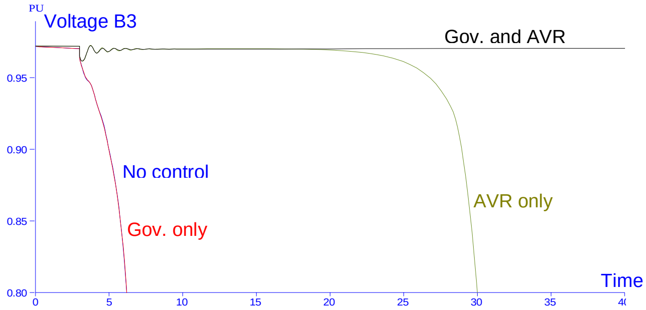

Fig. VII.1. Voltage in bus 3 as a function of time for the 4 cases simulated (1: no control, 2: governor only, 3: AVR only, 4: governor and AVR) when the system load is modeled as a constant power injection (p model)

It can be concluded from references [3] and [4] that the dynamic behavior of the system can be considerable improved by using better control strategies, local and remote information, and new technologies such as FACTS, non-linear controllers, etc.

Two main research questions arise in context of fast dynamics. First, the problem of how to create incentives for using more effective primary controllers and control strategies in the new electricity industry? Second, the problem of how can the performance of these technologies be compared in order to create these incentives? Another important issue is the effect of the protection system design, due to the fact that the characteristics of controllers needed depend on settings and robustness of the protection system.

VII. Conclusions And Open Research Issues

Assessing many restructuring issues in terms of the economies of scope is helpful in order to understand best ways of managing the sharable efforts and inputs which are used at different times and by the different agents throughout the electric power network. For example, some areas in the power network might present peak load demand at different times and the system has to meet this demand by supplying the cheapest and fastest plants available. Meanwhile, the other equipment is being used to supply base load. Another example is the failure of a generator, in which case the other plants must supply more electricity in order to meet its given demand.

We suggest that metrics in terms of economies of scope should be developed for quantifying the effects of unbundling and to further identify which products and services must be re-bundled. Moreover, unique to the fact that there is hardly any storage in current systems, it is important to decompose (aggregate) network users according to their time-of-use (TOU), either for real-time pricing or for longer-term contracts. It is illustrated in this paper how differentiating according to TOU (annual, monthly, hourly) could result in more stable electricity prices. Similarly for accounting who is contributing how much to the delivery bottlenecks (transmission congestion), it is necessary to cluster network users according to their relative contributions to network. Similarly, the end-users requiring reactive power/voltage support should be clustered according to their needs for this service.

Based on the examples presented in this paper we conjecture that it is conceptually impossible to unbundled the electricity service into products and/or services (private or public good types) without a significant loss of related economies of scope.

In the other hand, each sub-product and/or sub-service can be accounted for through systematic aggregation (clustering, decomposition) of its suppliers and users. Contrary to the common belief, it is possible to define decomposable demand and supply functions at the right level of aggregation, with quantifiable accuracy.

Finally, one should aggregate suppliers and users according to the quality of service (broadly characterized through rate of response, rate of interruptions, rates of equipment failures, and alike). Dynamic adaptive aggregation leads to manageable IT-supported multi-layered interactions for re-bundling purposes, therefore recovering economies of scope that might have gotten lost in an entirely decentralized industry.

The design of methods for (1) product differentiation; (2) decomposition of suppliers and consumers associated with similar specifications of these products, and (3) on-line IT networks for multi-layered protocols for re-bundling these services is the key R&D challenge. Without these methods, huge opportunities would be missed.

References

[1] F. Wu, P. Varaiya, P. Spiller and S. Oren, "Folk Theorems on transmission access: proofs and counterexamples" *Journal of Regulatory Economics*, vol. 10, pp. 5–23, July 1996.

[2] Z. Wu, M. D. Ilic "Toward a Systematic Design of Capacity Sub-Markets for Electric Power: Stratum Markets Structure" Qualification paper of Engineering and Public Policy department, Carnegie Mellon University, January 2004

[3] J. W. Chapman, M. D. Ilic, et al "Stabilizing a Multi-machine Power System via Decentralized Feedback Liberalizing Excitation Control" *IEEE Trans. on Power Systems*, vol. 8, no 3, August 1993.

[4] M. Ghandhari, G. Andersson, M. Pavella, D. Ernst, "A control strategy for controllable series capacitor in electric power system", *Automatica*, vol 37, 2001.

[5] D. L. Kaserman, J. W. Mayo, "The measurement of vertical economies and the efficient structure of the electric utility industry," *Journal of Industrial Economies*, vol. XXXIX, pp. 483-502, Sep. 1991

[6] J. E. Kwoka, "Vertical economies in electric power: evidence on integration and its alternative," *International Journal of Industrial Organization* vol. 20, issue 5, pp. 653-671, May 2002

[7] J. W. M. Cheng, F. D. Galiana, D. T. .McGillis, "Studies of bilateral contracts with respect to steady-state security in a deregulated environment," *IEEE Power Industry Computer Applications.*, 1997. 20th International Conference on , 11-16 May 1997, pp. 31-36.

[8] J.C. Panzar, R.D. Willig, "Economies of Scope," *The American Economic Review*, vol. 71, No. 2, pp. 268-267, May 1981

[9] K.E. Train, "Optimal Regulation: The Economic Theory of Natural Monopoly", MIT Press, 1991

[10] C. Helfat, K. Eisenhardt, “Inter-temporal Economies of Scope, Organization Modularity, and the Dynamics of Diversification,” Forthcoming Strategy Management Journal.