The Western Energy Market: Inherent Risks and Market Solutions

by Jeffrey D. Roark
with sidebar information
by Frederick H. Pickel, Ph.D

WPTF
Western Power Trading Forum

The Western Energy Market:
Inherent Risks and Market Solutions

by Jeffrey D. Roark

Introduction

It is not uncommon to see in the daily news another article about the death of electricity restructuring, alongside stories about the problems and alleged misdeeds of the companies that pioneered the effort. For those who saw so much promise for consumers from restructuring, this is distressing. The ideal of a competitive electricity industry is neither wrong nor unattainable, but it has been mutilated by mismanagement and imprudence, and even tarnished by potentially criminal acts.

The Federal Energy Regulatory Commission (FERC) seems to be the last bastion of resolve for going forward instead of backward, but the FERC is enchanted by the success of markets in the Eastern United States. In the West, we have participated in the most notorious market meltdown in the short history of electricity restructuring. Industry officials worldwide list their "Lessons from California" that will help them avoid the same fate.[1] And yet, even today there is still significant disagreement about what really happened in the West in 2000 and 2001. Without really knowing what went wrong, we cannot determine how best to go forward, nor can we pretend to know how to forestall such problems in the future. To understand what went wrong, we have to understand how hydro dominance really makes the western electric system different from its eastern counterpart, and how that led the West into shortage.

While the FERC insists that its reforms are aimed at avoiding further market meltdowns, its prescriptions appear to come directly from markets that haven't (yet?[2]) experienced the stresses that the West has suffered. However, in some ways the West had created a model to follow before California created the model to avoid. The Eastern markets have not enjoyed as long a history of benefits from liquid trading in competitive wholesale markets as the West. Liquid trading hubs arose naturally in the West, long before the Pennsylvania-Jersey-Maryland (PJM) established an ISO in the East, and well ahead of California's restructuring. Wholesale market pricing data extends as far back as 1994 for hubs such as Palo Verde (in Arizona) and the California-Oregon Border (COB), hubs that have helped the West optimize the use of its natural energy resources in spite of their wide geographic differences. (Reportedly, wholesale trading has been active at these locations since the 1970s.) In addition, since 2000 there has been more merchant capacity built in Arizona and Nevada than in California, despite the fact that Arizona and Nevada have not restructured. Apparently then, some things were working rather well in the West.

[1] A recent IEEE Power Engineering Review, for example, included the following articles: "Will the California Crisis Perturb Spain's Liberalization Process?", "California Crisis Influences Further Reforms in Latin America," and "Can Brazil Learn From California?" This author has personally addressed utility executives from China, Germany, The Philippines, and Turkey, all curious and concerned about what happened in California.

[2] The market designs in the Eastern U.S. have been successful in managing operational problems, and in guiding the proper location of new capacity. It is a valid question, however, whether the East is even susceptible to the kind of shortage that the West experienced in 2000-2001, as this paper will discuss further.

The Western crisis was not a failure of wholesale markets or restructuring; it was a physical situation at its root. However, many of the difficulties encountered in the crisis, from arcane trading strategies to FERC's difficulties bringing the situation under control, can be traced to inconsistencies in rules, jurisdictions, and market structures. Worse, perhaps, was California's intentional limitation on risk management by the utilities, which left them in the spot market and pushed them into insolvency.

Energy resources vary widely among regions of the West

Energy Generated by Fuel Type
Calendar 2000

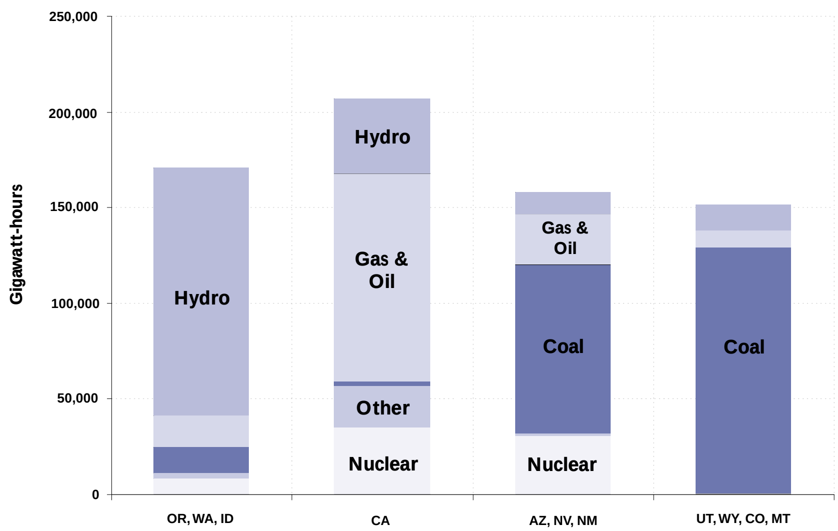


Figure 1 – The wide geographic diversity of the Western system's resources creates a natural need for wholesale liquidity to rationalize the operation of these resources for maximum efficiency. Two regions are strongly dominated by baseload technologies, while the Northwest is dominated by hydro. California holds the lion's share of the gas-fired capacity.

Source: Derived from EIA Electric Power Monthly

California is not an independent electric system; it is a dependent part of a hydro-dominant Western system unlike all others in the United States. Restructured hydro-dominant systems in other countries have proven their abilities to manage stresses such as in 2000-2001, but in these systems rules are consistent system-wide and risk management is practiced. FERC's vision for markets is one of widespread rule consistency, supporting wholesale liquidity, risk management, and reliability. The West must continue to assert its uniqueness so that the implementation of this vision is sufficiently robust, as the problems faced only by the West are substantial. However, consistency of rules and market structures across the West is an important objective to support.

The West Really Is Different: **Physical Characteristics**

The West was a natural system to develop wholesale electricity market, but there were some structural features created along the way that really helped establish wholesale liquidity. (See sidebar) The West has its own natural resources: mighty rivers, abundant coal and natural gas, even without considering western Canada with its plentiful hydro and natural gas. The problem, of course, has always been that the resources are not located where the population is, and the distances between them are staggering in some instances. Each region of the West naturally developed its native resources for electricity, as the local economics dictated. This approach to developing local electricity resources made simple economic sense, but it resulted in sharp differences in fuel mix among the various corners of the system. California, with the West's largest concentration of population, developed its natural hydro and gas resources, too, but as its population grew it imported energy, primarily through gas pipelines and high-voltage transmission lines from neighboring states.

Wholesale Liquidity Offers Efficiency Without Centralized Control

For many years it was thought that for a large power system to run as efficiently as possible required a central operator making it happen, running computer programs to make all of the coordination decisions. The West is testament to the fact that this isn't necessary. With enough transmission and wholesale liquidity, a decentralized market takes the place of a central coordinator over the whole electric system.[3] In fact, this may even do a better job of optimizing than a single central operator could do. A central operator has to make approximations and judgments far from the action, requiring intensive communication between the central controller and all of the various parts of the system. Decentralized market players, on the other hand, are like hard-working ants: crawling everywhere, close to the action, each in command of the details of its own little surroundings, communicating with others nearby through multiple transactions. It isn't necessary for anybody to grasp the whole and optimize it; it happens automatically through the combined actions of all of these separate "agents." The West has saved many dollars this way, with decentralized trader activity setting up major power flows between regions in the day-ahead and monthly term market. Then on each operating day, the various system operators physically dispatch the generators in their smaller areas, incorporating the inter-control-area schedules as imports or exports.

Essentials for Liquidity Arose Naturally in the West

The last major debacle of the era of regulated utilities, nuclear plant development, actually provided the bridge to the future in the West. California had the size, the financial strength, the need, and the economic engine to attempt its own nuclear plants, but the utilities in the Desert Southwest were not big enough to do this alone. Rather, a group of utilities in several southwestern states, California included, combined resources and built the gargantuan Palo Verde Nuclear plant and its associated transmission system. Perhaps the co-developers didn't realize what they had created when they completed the Palo Verde switchyard, but it became a special place.

Over most of the country, when you see transmission lines or substations, you typically see the property of only one utility, maybe two if you happen to be standing at a boundary between systems. Utilities in those days had interconnection agreements with all of the utilities that they were physically connected to, and these were the only utilities with which they could directly transact. To transact with anybody else involved transacting with whatever utility stood in between on the transmission system. Without open access for merchant players, there was no ability for anybody other than the interconnected utilities themselves to transact in power. Palo Verde switchyard was special because the property of 10 different utilities came together in that one spot. They were interconnected in a point rather than at seams between their retail service territories; therefore they could transact directly with each other. In other words, by bringing together a number of buyers and sellers with physical rights at a single location, Palo Verde's joint-ownership arrangement on its transmission system created a marketplace in the Desert Southwest. A second jointly owned project, the California-Oregon Intertie, created a second market in the north, effectively capping California at both ends with wholesale electricity marketplaces. It was a small step to liquidity when the National Energy Policy Act of 1992 allowed non-utility Power Marketers into the mix. It is no accident that Palo Verde and COB were the very first liquid trading points listed by the new energy-trading press, and the first points selected for futures contracts. In a way these are accidental market structures, but they are market structures nonetheless.

It is also no coincidence that PJM was the major "point" of liquidity in the east. PJM's member companies each had rights to move power to or from each interface on its 500kV network, and each had a portion of each interface with PJM's neighbors. PJM's 500kV network was therefore a market point as well, where multiple companies had physical rights to take or deliver power at multiple locations. This fact alone was enough to promote wholesale liquidity in and around PJM well before PJM established itself as an independent grid operator.

J. Roark

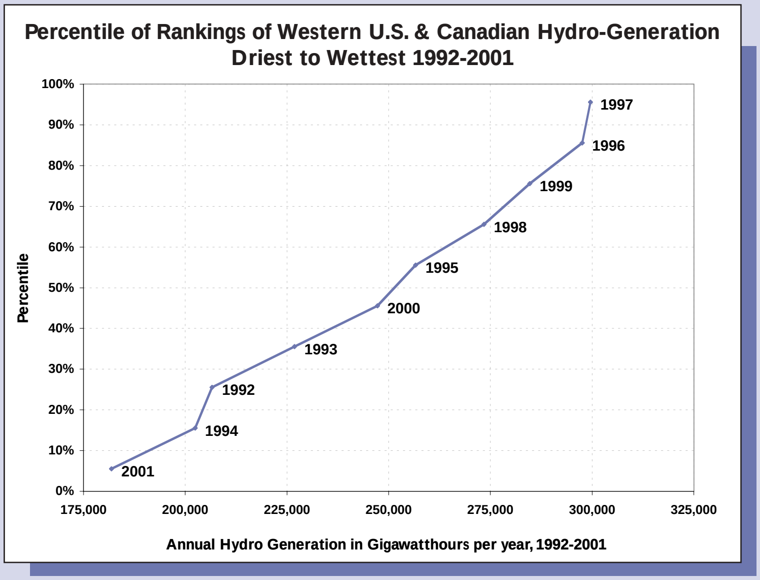
When hydroelectric resources are plentiful, the mechanism of wholesale trading spreads the surplus[4] energy throughout the Western system. In particular, surplus hydro in the West has one major effect on other generators in the system – it reduces the output of fossil-fired generators. This simple fact reflects the relative economics of nuclear energy, fossil energy, and renewable sources. Economics dictate that a system will run its nuclear plants, its cogenerators, and its other renewable sources as much of the time as it can. Like hydro, these other sources displace the burning of fossil fuels, which is literally an economic last resort. Among fossil fuels, a coal plant is generally less expensive to run than a gas plant, so coal plants generally run as much as they can practically run. Gas- and oil-fired generators are truly the last generators in line; they usually generate what is needed after all other generator types have done all they can do.[5] This doesn't happen automatically, of course, as if a central operator were dictating the details for the whole market. Rather, it happens through wholesale trading, where each generator, trader, or utility trades energy with others before the operating day, according to its needs and available resources. When all the trading is done, the coal plants in the West are often still scheduled to run almost "flat out," but the gas plants are lined up to "swing." That is, they run only enough to satisfy the next day's energy and reliability needs.[6] Economically, this is exactly what should happen.

These benefits of wholesale liquidity are evident in public data[7] that reveal the operating details of the Western system. Coal-fired generating plants, which produce the grand majority energy generated in the Rocky Mountain regions, run very nearly at full output[8] year-round, despite the fact that local demand couldn't possibly consume it all, and local demand doesn't have a flat load shape in any case. Gas-fired generators in California, Arizona, and Nevada do track local demand variations, but they also behave as the high-cost units for the entire West, responding to variations in energy available from hydroelectric resources in the Western system. In other words, the Western Interconnection in many ways operates efficiently as if it were an integrated system, implemented through the West's long-standing tradition of wholesale liquidity.

How big is the swing in annual hydro production? The 100,000 gigawatt-hour hydro gap.

Hydro produced electricity averages one third of the electricity produced in the western US and Canada. The average and median are about 250,000 gigawatt-hours out of a total western generation demand of about 750,000 gigawatt-hours per year. Hydro production, however, swings over a very wide range – as illustrated in this chart showing the percentile rankings of hydro production for the US and Canadian West over a ten year period ending 2001.

Note: This chart shows actual hydro generation for the year. Forward prices – and because of the opportunity costs of hydro generation, spot prices in a hydro linked system – reflect expectations of the future. In the early spring of 2001, based on actual precipitation and snowfall, forecasts of hydro runoff were anticipating a short-fall of 100,000 to 125,000 gigawatt-hours rather than the 70,000 gigawatt-hour shortfall that actually occurred. Following the above normal years of 1995-1999, which averaged 283,000, the threatened drop in hydro production from above very long-term average 250,000 gigawatt-hours to possibly below 150,000 gigawatt-hours triggered severe shortage fears and high price expectations.



Data Sources: Hydro data compiled by Economic Insight from US EIA and Canadian statistics

F. Pickel

[3] This is not to suggest that there is no need for system operators or market facilitators, or even optimizers for daily dispatch decisions. The issue here is one of inter-control-area trades from short-term to long-term that provide economical energy or "insurance" of the type discussed in subsequent sections of this paper.

[4] The word "surplus" in this case means simply that the hydro output is greater than local demand. In a sense, electric energy is never in surplus, but there are times when water energy can be lost, literally spilt over the dam, if it is not used to generate electricity. Hydro operators try to avoid spillage, and will sell electricity to optimize the value of their resource.

Did the Wholesale Market Bring California Down?

While we can look at the West and see the benefits of decentralized trading on operations at the macro scale, we are quickly reminded that the California market design was an exercise of extreme faith in this very same concept. The physical dispatch of generators in California was constructed through multiple trades between buyers and sellers in the now-defunct Power Exchange (PX); the central operator couldn't improve the efficiency of the dispatch even if it could see how to do it. It is a legitimate question whether California took this concept too far and too literally, and it is a legitimate question whether California's market design led to its ultimate meltdown in 2000 and 2001. We now know that the physical roots of the crisis had nothing to do with California's market design. At the same time, however, the constraints that California placed on the regulated utilities limited the ability of the market to manage the crisis through economic means, and produced the financial meltdown that followed. Consumers were given a price guarantee, while the utilities required to serve them were not allowed to protect themselves against increases in wholesale prices. This incomplete market design left the utilities dangerously exposed, turning a major power problem into a major financial disaster. That the power problem – originating as it did in natural physical occurrence – was neither anticipated nor planned for is the opening act to this tragedy.

Monthly Hydroelectric Output

in the Western U.S, 1996 - 2002

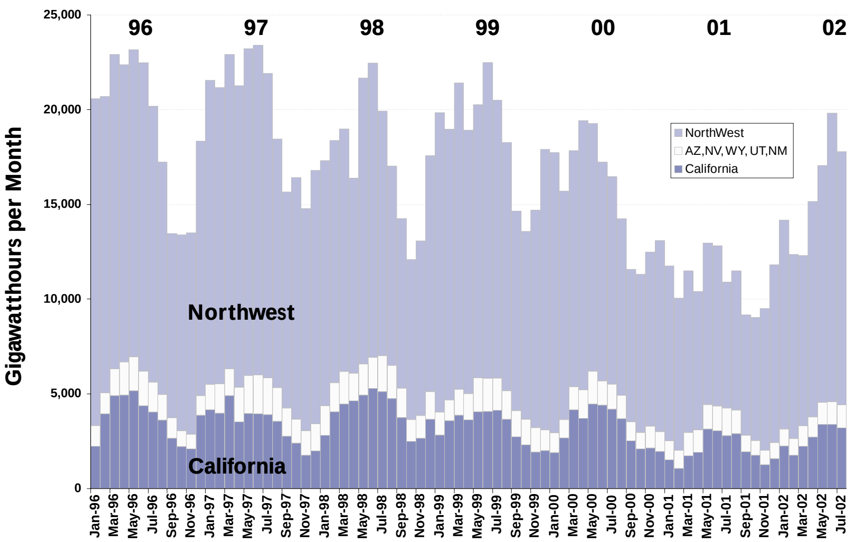


Figure 2 – Hydro output fell off rapidly after Spring of 2000 and did not recover for two years. It is important to understand that the hydro resource in the West is very large, usually producing more than twice as much hydro energy as the entire rest of the country. It is also highly variable, ranging from producing 40% of the total Western supply in 1997 to about 21% in 2001, a very large amount of change.

Source: EIA Electric Power Monthly

[5]All gas-fired plants cannot be simply turned off because some are absolutely necessary to maintain the stability and reliability of the electrical transmission grid. These units became visible in California when they were designated Reliability Must-Run (RMR) units prior to their being sold to competitive suppliers. The RMR contracts gives the system operator the ability to call these units into service at any time they are needed for this reliability service, irrespective of market prices. The operator reimburses the generators for the cost of providing the service. These contracts are invoked routinely, but they would be most important in the event of a hydro surplus, when market prices would likely sink so low as to otherwise force these RMR generators offline. Figure 5 includes periods such as this, such as May of 1998 when California prices were zero much of the time. The RMR contracts also mitigate the "locational market power" that these generators could otherwise gain from foreknowledge that their capacity is needed to maintain grid reliability.

[6] Oil-fired generation fills this role after gas-fired units, but there is so little oil-fired capacity in the West that it would complicate the issue to refer to it repeatedly. While oil-fired units definitely increased their outputs by multiples in response to the crisis, it is clear from that oil nevertheless plays a small part on a West-wide scale.

[7] Hourly megawatt output, heat input, and emissions for all major fossil-fired generating units in the country are available from the Environmental Protection Agency (EPA). In addition, the federal Energy Information Administration (EIA) provides monthly generation figures for most power plants in the country.

[8] Practical full-output levels accommodate periodic maintenance outages, in addition to random forced outages. EPA data shows that coal plants in the West maintain high availability on the whole, but there is a clear pattern of maintenance outages in the off-peak seasons.

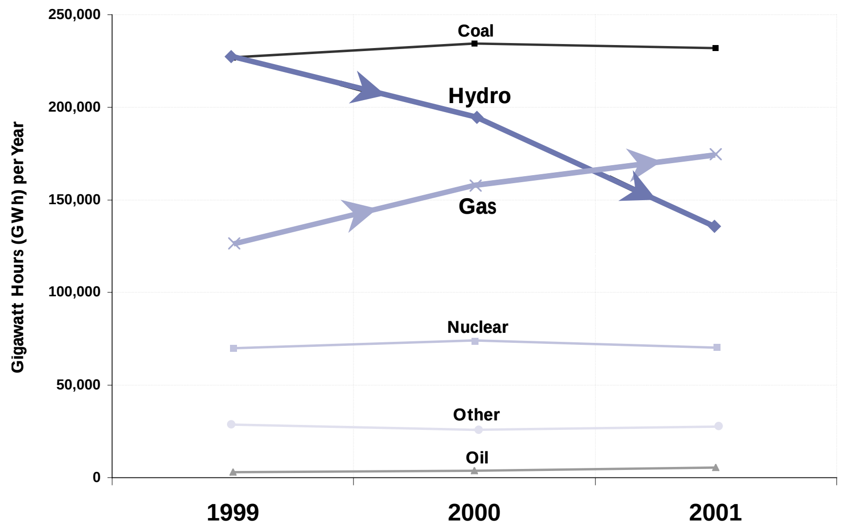
Critical Analysis of the Events of 2000-2001

The root causes of the shortage in the West are frequently listed and dispensed with in a few sentences, without showing their extent. The most publicly visible aspects of the crisis were high prices, bankruptcies, and the concomitant political aftermath, and these issues usually dominate. In this paper we will concentrate on the physical roots of the crisis, examining some of the public data that exposes the extremity of the situation. The data raise the question of whether the shortage happened simply because there wasn't enough capacity to gracefully weather the physical shocks that occurred. And, if such physical shocks are expected periodically, then we need to consider how to make the Western system sufficiently robust.

The Western Energy Crisis of 2000-2001 was a series of physical events that accumulated problems as it progressed. Physically, the crisis had a clear beginning and a clear end, both quite independent from strategies of unscrupulous traders, price caps set by regulators, or state-purchased long-term contracts.

Energy Generated by Fuel Type in the Western U.S.

1999 - 2002



Source: EIA Electric Power Monthly

Figure 3 – As hydro energy declined in 2000, it was replaced by gas-fired generators, mostly in California. Coal generators increased their output slightly, but coal, nuclear, and "other" generators had very little slack available. The gas category shown here is total, including gas-fired cogeneration. Oil generation increased quite a bit in percentage terms, but there is very little oil-fired capacity on the system.

Hydro is Huge in the West

Hydroelectric energy is dominant in the entire Western market, even though the majority of it is generated in the Northwest. In 1997, hydroelectric dams contributed over 40% of the energy generated in the Western U.S., generating twice as much hydro energy as enjoyed by the rest of the country in that year.[9] In this sense the West has more in common with New Zealand (about two-thirds hydro in 2000) or Argentina (about one-fourth hydro in 1999) than with the Eastern United States (5% hydro or less).[10] The curse of hydro, of course, is its variability, as illustrated in Figure 2. From its high point in 1997, hydro provided about 33% of the total energy in 2000, and only about 21% in 2001.

In 2000 the Western system encountered an unusual wide-spread heat wave that produced surprisingly high demands all around the West in May and June.[11] Hydro generation, which had appeared normal until late May, began declining rapidly soon thereafter.[12] As Figure 3 shows on an annual basis, gas-fired generators were essentially the only resource with any slack capacity that could step up production to replace what was lost when hydro declined. For gas-fired generators it is unimportant whether demand is high or hydro is low; the effect is the same.

How does this 100,000 gigawatt-hour hydro generation gap compare to the overall market scale?

The 2001 hydro shortfall looked like it could substantially exceed 100,000 gigawatt-hours in comparison to the 1995-1999 averages. Just how much conservation or additional generation might be needed to compensate for such a shortage?

Since this shortage was an annual energy generation shortage, just saving or generating energy on "peak days" is nearly meaningless. The new generation production or conservation savings have to come over the full year.

For perspective on the scale of the conservation or replacement generation challenge:

- **The entire electricity consumption for Washington State, all consumers, was 94,000 gigawatt-hours during 1999.**
- **California consumed 211,000 gigawatt-hours in 1999.**
- **All industrial consumers in Washington, Oregon, and California combined consumed 99,000 gigawatt-hours in 1999.**
- **All consumers in Southern California Edison's territory consume about 77,000 gigawatt-hours in 1999.**
- **The average US nuclear unit produces about 5,600 gigawatt-hours per year, so 100,000 gigawatt-hours is equivalent to the average annual output from almost 18 nuclear generation units.**

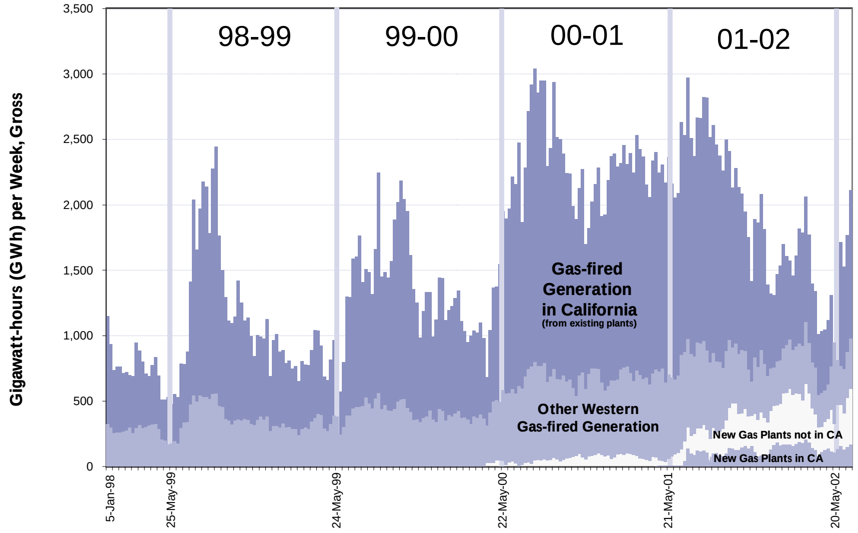
Given the scale of the hydro shortage, replacement generation from just a limited group of thermal plants or conservation by one state or one sector of the economy would not be enough to address the shortfall.

Data Sources: EEI Statistical Yearbook, NERC generation availability reports.

F. Pickel

Beginning abruptly in late May of 2000, the production of energy by gas-fired generators, the high-cost generators in the West, rose to unprecedented levels, as illustrated in Figure 4. As hydro declined further, emergency conditions in July and August of 2000 coincided with the highest output days, when most of the coal and gas capacity was online and generating at high output levels.[13] There was precious little slack generating capacity remaining anywhere in the West.

Electric Production by New and Existing Gas-Fired Generators in the West



Source: EPA Acid Rain Program Hourly Emissions Data

Figure 4 – Energy output from gas-fired generators in the West rose quickly to extreme levels in Summer of 2000 because of high heat-driven demands, but remained at high levels for 18 months in response to unusually low hydro supplies. By mid-2001 new plants were becoming significant generators of energy, even as the total demand for gas-fired energy was falling to normal levels. This graph includes all major gas-fired generators in the West, but does not include gas-fired co-generators.

[9] From EIA Electric Power Monthly, March, 1998. Other EIA statistics show that 1997 was the peak year for hydro output for the entire U.S..

[10] EIA Country Energy Data Reports

[11] Demand in May of 2000 was over 13,000 MW greater than in May, 1999, as reported in the Western Electricity Coordinating Council 10-Year Coordinated Plan Summaries.

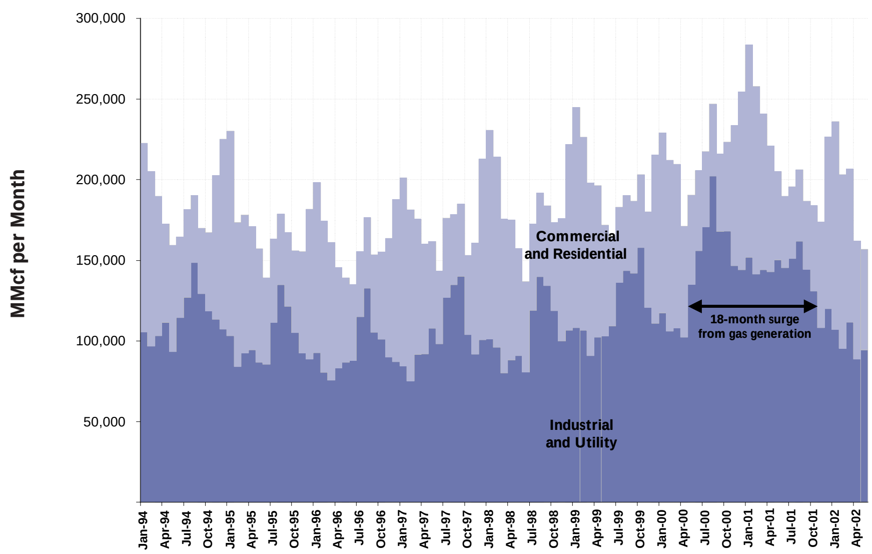
[12] This pattern of hydro runoff was noted in a Northwest Power Planning Council report, indicating that this unusual pattern of runoff probably "fooled" a number of participants in the market. "Study of Western Power Market Prices, Summer 2000, Final Report," October 11, 2000, Northwest Power Planning Council document 2000-18

Albeit difficult for the consuming public to see or understand, production levels for this topmost portion of the "supply stack" barely receded following the summer peak, melding the summer 2000 and 2001 peaks together in an unprecedented 18-month-long super-peak season. The center of this extended production surge, the Winter of 2000-01, coincides with the first "trough" of the hydro slump in Figure 2, book-ended by the two summer peak seasons. This 18-month period came to be known as the California Energy Crisis, even though it really involved the whole Western system. California is where most of the gas-fired power plants are located, usually operating at low utilization rates while the state imports hydropower from the Northwest. When hydro fortunes fell, the California gas-fired plants were suddenly very important, and so was the California wholesale market.

The surge in electric production by gas-fired generators was just the beginning of a chain of events. As summer turned to fall and winter, some of the older units broke down under the strain. Units that could run for only a limited amount of time per year (because of emissions restrictions) used up their allotted time. Transmission lines became congested as the gas-fired electric energy flowed to consumers in a pattern totally unlike prior years. After all, the hydro energy had been greatly reduced in the Northwest; most of the replacement energy was coming from the Desert Southwest and Southern California. This was a large amount of energy movement from one corner of the system to another. In winter, the gas generators that were available in Northern California were operating at maximum output,[14] yet rolling blackouts became necessary. Gas demand from electric generators created an unusual gas consumption peak in August (Figure 5), to be followed only five months later by an even higher natural-gas peak in winter, draining gas in storage to low levels and increasing imports of gas into California to a new all-time high, thereby stressing gas pipeline capacity. To make matters worse, NOx emissions in Southern California had an absolute upper limit, represented by a limited supply of emissions credits that generators needed to purchase in advance in order to run. The supply of emission credits became scarce as the generators burned much more gas than anyone had expected. It is little wonder that prices for gas, credits, and electricity all spiked together.

Gas Consumption in CA, AZ, and NV

By Month, 1994 - 2002



Source: EIA Natural Gas Monthly

Figure 5 – Gas consumption in these states shows the telltale shape of the gas-fired electricity production surge. Note that the Industry and Utility sector increased in 1999, but the mild winter allowed overall gas consumption to remain low for the year. The generation surge pushed August into a new gas-consumption peak, and pushed January consumption to a new peak 15% higher. Additional history provides perspective on the magnitude and duration of the surge in total gas consumption during this period.

[13] This can be observed by aggregating hourly dispatch data obtained from the EPA. While all gas-fired generators are not represented in the data, all of the large ones are. With the exception of gas-fired cogeneration, most of the gas-fired energy is represented. In particular, most of the merchant capacity in California is included and can be observed in hourly dispatch, except for small combustion turbines.

The end of the crisis is marked by two helpful trends, both evident in Figure 4. First, new gas-fired generators came into production beginning in mid-2001, giving the Western system (and particularly Northern California) additional support that greatly increased reliability. By late 2001, Northern California was receiving energy from 1,000MW of new natural-gas-fired combined-cycle units, raising the stock of responsive gas capacity in that constrained part of the system by almost 20%. As the output of new generators continued growing into 2002, conservation by consumers and more typical hydro conditions combined to bring the total demand for gas-fired energy back down to pre-crisis levels. Physically, the system had returned to "normal" by the end of 2001.

How much gas would be needed to generate 100,000 gigawatt-hours and replace the missing hydro supply?

Historically in the coastal west (the NW Power Pool and California), a pattern of north/ south trade had developed since the 1970's. During wet hydro years, surplus power was sold into the California market, and California gas-fired generators reduced their production. During dry hydro years, California's gas-fired generation has increased production, exporting the surplus energy to the Northwest.

Under extreme drought, like the 100,000 gigawatt-hour shortfall feared in 2001, the existing gas-fired generation and pipeline/ storage infrastructure could not handle the increase:

- **Assuming a typical 11,000 Btu/kWh consumption rate for the gas-fired power plant, a 100,000 gigawatt-hour shortage creates an increased 3 billion cubic feet per day (Bcf/d) gas demand for the west.**
- **In 2000, about 80% of the gas-fired generation in coastal west was in California, so about 2.4 Bcf/d gas demand was created in California.**
- **This 2.4 Bcf/d would increase gas usage by 28% in system that was already expected to be flowing at 5.8 Bcf/d under average conditions in 2001 – or attempting to flow 8.2 Bcf/d or 107% of capacity. The total pipeline receipt capacity into California in early 2001 was 6.6 Bcf/d plus about 1.0 Bcf/d California gas production for a total of 7.6 Bcf/d. Remember, these are average ANNUAL flows – so peak days for home heating would exceed capacity.**
- **The existing gas-fired generation in the coastal west would have to run an additional 2900 hours in the year, or 33% of the year MORE than in an average year.**

Data Sources: California Energy Commission, Natural Gas Infrastructure, 2001, and related reports; NERC Electricity Supply/ Demand Balance report and database; Sam Van Vactor and Fred Pickel, "Money, Power and Trade", Public Utilities Fortnightly, May 2001.

F. Pickel

Was it all physical?

This description of the energy crisis is all about physical changes and physical limits, without market manipulation, physical withholding, trading strategies, or even high prices. But if generators withheld supply, then the evidence is obscured by their very high output levels. If plant outages were high, it was following a summer of unprecedented continual stress all the way to the top of the "stack" of units, where, for example, 46-year-old generators struggling through their final years were suddenly pushed as hard as the younger, more dependable ones. This extreme situation was created by physical phenomena, and the physical response by gas-fired generators in the West was greater than even savvy market participants would have imagined.

As California Power Authority Chair, S. David Freeman said as rotating blackouts began in California, "It's not all that complicated — we just have a shortage. All the conspiracy theories — you don't need to go there." [15] The important point it is that it is not necessary to resort to conspiracies to find a reason for a physical shortage. With the data now available it is clear that the crisis had unavoidable physical root causes, culminating in the 18-month surge in demand for energy from gas-fired units around the West that pushed the limits of generation, transmission, pipelines and emissions credits. In such conditions, intentional withholding is not necessary to produce shortage conditions or high prices. Unit outages will accumulate naturally under continual stresses and make conditions worse. This does not rule out the possibility of harmful trading strategies intended to manipulate the market for extra profits on the part of some market participants, of course. But regardless of whether some people tried to manipulate the market or not, they could not have caused the heat wave, the drop in hydro or the surge in gas-fired generation. These were not minor events, and they are not shrouded in mystery.

---

[14] This is extremely unusual operation for gas-fired plants in California, to run at peak output for days on end. These units had a cost advantage over units in Southern California because generators in the north don't have to buy NOx credits, and gas is often cheaper in the north than in the south. Statements in the press indicate anecdotally that congestion was limiting flows from south to north in California often during this winter period, and there were significant outages in Northern California during this period as well.

[15] California in State of Emergency Over Power," New York Times, January 18, 2001.

Financial Fallout

In 2000 and 2001, Californians paid four times as much for power as they had in 1999, according to state officials.[16] Given the physical changes and stresses that the system went through, it is not surprising the prices spiked, and given the duration of the hydro slump, it is not surprising or unfair that Californians would pay more for power during this period than before. After all, normally they don't burn all of that gas in-state, economically preferring to import hydro energy. What is unfortunate, however, is that the California utilities bought most of their power on the spot market at the high spot prices. This was quite unnecessary, but it was quite deliberate. Reportedly intended to prevent the utilities from simply rebuilding their former dominance through contracting, the regulatory limitations on contract purchases prevented prudent risk management from being implemented on behalf of consumers, and to some extent prevented merchant generators from managing their risks as well.

It is basic that market prices go up when demand threatens to exceed supply. Before the Western crisis, electricity markets didn't have regulatory price caps, even after hourly prices in Midwest wholesale markets surpassed $7,500/MWh in 1998.[17] Such prices are consistent with the cost of additional capacity whose duty would be to stand in reserve until all other resources are exhausted.[18] This oft-repeated logic works well in Eastern markets, where peak stress conditions can be measured in hours rather than in years. But with the possibility of two-year slumps in hydro availability, the West does not fit this Eastern model very well. Shortages and prices at shortage levels are not acceptable to the public for such long periods, especially when large blocks of the market are transacting at the spot price.

The West needs to think about reliability differently from the Eastern United States. As a hydro-dominant system, it is prone to random, if well-separated, hydro slumps. If these slumps are to be passed without notice, then the West needs to maintain a stock of fossil-fired capacity that will seem to be a surplus in most years. For example, generation seemed to be plentiful in 1999, a very mild year for gas-fired generation, except for an unusual production period in October. But if the system didn't have enough capacity in 2000, it likely didn't have enough in 1999, either, to be safe from the possibility of a hydro slump. Especially in the West, the system shouldn't need to run all of its capacity in every year. That is, unless there are ways to adjust demand to match supply in real time and to allocate scarce resources among competing uses. This is what markets are meant to do. Markets use the price mechanism as the primary allocating factor. That said, the wide variability of such a large block of hydro energy may produce shortages that are too deep and prices that are too high to be acceptable to the public. Awareness of the risks and positive risk management actions are necessary on the part of public officials and energy consumers.

[16] "Governor 'Outraged' at Stance FERC Took," San Diego Union-Trib, December 13, 2002.

[17] This price was reportedly for a single hour. A FERC staff investigation determined that the market was reacting to unusual circumstances that created temporary shortage conditions. "Staff Report to the Federal Energy Regulatory Commission on the Causes of Wholesale Electric Pricing Abnormalities in the Midwest During June 1998"

[18] For instance, suppose a combustion turbine needs to recover $75/kW per year to meet its fixed costs. This is a reasonable fixed cost level. If it must recover that amount in only 10 hours of operation on the summer peak, then it must receive $7.50/kWh, or $7,500/MWh, in order to meet fixed costs. A regulated utility planning to build such a turbine to avoid 10 hours of outage would be assuming that consumers would place at least this value on the power in those 10 hours. There is even an industry term and acronym for this assumed customer value: the VOLL, or Value of Lost Load.

Risk Management As Insurance

In everyday life, when individuals or companies want to avoid a possible catastrophic loss, they buy insurance. They hope never to suffer the loss, but just in case, they pay periodic premiums to an insurer, who maintains a financial reserve sufficient to reimburse its policyholders in the rare event that such a loss occurs. The insurer protects itself by building a portfolio of customers whose risks of loss are different and unlikely to occur all at once. The concept of insurance is fully appropriate in the context of power markets, except that in power it is more of a process of combining and managing risks that cancel each other out. Consumers and suppliers are said to be on opposite sides of price risk in power. What is a risk for the supplier is often a reward for the consumer, and vice versa. Risk management is a process of meeting in the middle in a way that resolves these issues ahead of time, and eliminating the risks for both.

Spare generating capacity is a form of insurance, perhaps best thought of as physical insurance. Spare capacity in the West is capacity that likely would not be needed at all during years with normal or surplus hydro. Such capacity would be called upon, however, when hydro supplies are low and demand remains high. Spare capacity is generally the older and highest-cost units on a system, those pushed out of the routine dispatch by lower-cost plants. We can think of the lower-cost plants as energy producers, and the spare capacity as insurance, in case the energy producers fail. We know that the West needs spare capacity to generate during hydro shortages. The issue for the West to decide is how it wishes to purchase its insurance and from whom.

Regulated utilities provide insurance as a matter of obligation to serve reliably. Using sophisticated statistical models, central reliability planners determine how much spare, or reserve, capacity is needed, making assumptions about the value of service to consumers. Under the regulatory model, the cost of this reserve capacity is collected from consumers on a regular basis, not just when it is used. In regulation, the insurance against outages and price spikes comes bundled in every-day electricity prices. Consumers are unaware of the risks and have no choice in the matter. Price signals are simple and few, in spite of the complexity behind them. Without competition, and without the ability for consumers of various types to exercise individual choices, these prices are also likely to be higher than they should be.

Competitive entities can build power plants as well, and can provide insurance in many forms and flavors. In fact, competitive entities have demonstrated an ability to build power plants faster and cheaper than regulated entities. The key for competitive entities building power plants, however, is the existence of an active "power insurance" market. Consider how a competitive entity could survive holding a portfolio of spare capacity in the West. Looking at Figure 2, could a competitive supplier of spare capacity have survived financially through the hydro years of 1996 through 1999, waiting without income for a hydro shortage to develop? Not likely, unless there were revenue support for spare capacity, i.e., an insurance-like contract to maintain plant availability over time.

During the energy crisis, California was uninsured by design. Not only was California physically exposed to physical shortage in the event of a hydro deficit in the Northwest, but the California utilities, who were still buying power on behalf of consumers, were unable to buy power through long-term contracts, and therefore unable to ensure themselves against wholesale price increases. This also left the competitive suppliers uninsured as well, unable to contract with the utilities.[19] By design then, suppliers and consumers were completely disconnected until each day's market was established. Everybody was forced into being speculators to some extent, virtually guaranteeing that the market remained unprepared for a sudden shortage.[20] The result noted earlier was that Californians paid four times as much for power as they had the year before. If instead they had bought 80% or 90% of their anticipated needs for 2000 back in 1999 or even 1998, they would have had a much better deal. Spot market prices still would have skyrocketed all around the West when the hydro receded, but financially it would have mattered much less to the utilities or to the public.

Naturally, other hydro-dominated competitive markets in the world have experienced similar physical shortage situations, though none have prohibited risk management. New Zealand (about two-thirds hydro in 2000) suffered a drought and shortage in Winter of 2001, and Norway (almost 100% hydro) suffered one in 2002. In New Zealand, wholesale prices increased ten-fold, initiating load-reduction programs and creating some financial hardships. However, a government review concluded that the market had "worked much as expected during winter 2001, with very high prices signaling an increasingly tight supply situation and record demand."[21] Through the market, the situation was managed without supply interruptions, in spite of record-low water inflows. While some in the market were adequately hedged, others were not.

"Some major retailers and large users were seriously under hedged against dry-year spot prices. Although hedges were available, several years of surplus generating capacity and record low spot prices affected buyers' assessments about investing in hedges. As a consequence of experience in 2001, increased awareness of dry-year risk is likely to result in better risk management. More sophisticated (and liquid) hedge and contracting arrangements are likely to emerge." [22]

While this draws analogies to the Western situation in 2000-2001, it also points to possible problems even when risk management means are available. Several years of continual surplus may lull some energy buyers into a false sense of security, leading them to go without hedges. If hedging behavior is necessary to promote sufficient capacity building, then long periods of surplus may naturally end in rude awakenings, at least for the unhedged portion of the market. At least in that situation it is clear that those who suffer the high prices are the ones who failed to insure themselves. It is, of course, a matter of public policy whether regulatory means should be used to support fossil-fired capacity at a minimum level to avoid periodic shortages. But if such a level were adopted and properly supported through market mechanisms, competitive suppliers can and will supply the needed capacity.[23]

---

[19] Undoubtedly some of the California merchant generators hedged with other utilities in the Western system, as capacity was expected to be short in summer of 2000. Some of them may have also hedged with third parties who were speculators.

[20] Interrupting the contracting activities between suppliers and demand interrupts downstream insurance actions as well. Unhedged generators generally do not buy fuel under long-term contracts, because it is speculative to do so, putting them at risk from the fuel dropping in value. On the other hand, a generator who has sold power under a long-term contract must buy fuel under a matching contract to mitigate the risk that fuel prices might increase and wipe out profits from the power sale. Since spot power prices in California generally at least maintain parity with spot gas prices, a gas-fired generator selling spot electricity assumes no additional risk by deferring gas purchases into the spot market.

[21] Electricity Post Winter Review, www.winterreview.govt.nz

[22] Main conclusions from the Post Winter Review, ibid.

[23] Competitive suppliers will likely build surpluses periodically even if the matter is left entirely up to markets and prices. But as with the slow variation in hydro availability, several continuous years of surplus eliminate continued construction, leaving the system exposed to the rude awakening.

The West is Different -- Can Markets Solve the Problem?

There was never a need for the hydro-driven thermal-generation surge to become a supply crisis, but such a surge was not planned for throughout the West. The West needs to be far more aware of the precariousness of its situation, and to remain highly aware of it in its planning. The kind of planning needed is special, too, because this is a regional issue, not a state issue. Hydro power declines in the Northwest, and as a result more gas flows into Arizona and Southern California. How do state regulators deal with cross-jurisdictional issues like that?

It is natural for people to think of centralized actions to address problems that go beyond our abilities to control locally. If the problem is bigger than state control, then we bring the states together to control things at a higher level, to encourage coordination and cooperation among the states. These things should happen and have happened, but we should be wary of the natural reaction to a situation that got out of control: the tendency to control more than necessary. This is why it is so important for policy makers to understand the true nature of the crisis.

Naturally we would like to suggest that market structures can help sort out the Western issues, as they have done for many years. As we have seen, wholesale liquidity already moves power all around the Western system and converts a far-flung system of highly diverse energy resources into a system that operates as efficiently as it can under the current market structure. Wholesale liquidity moves gas from the supply regions to the places where it is needed for generation at the time it is needed, as well as it can do given the physical transportation system. In addition, most of the new capacity in the West, the real force that brought the market back into control, was provided by merchant energy companies, and much of that was underway before the crisis made the need for it so clear. Markets and competitive suppliers are already a large part of the solution today, and they must be part of the solution in the future.

However, work is necessary at the policy level. The relevant question for policy makers and market participants alike is one of risk. How much risk is acceptable? None? Some? How does it appear? How is that quantified? How does it work its way through the system? Where does the buck stop? The impetus to central control is merely an administrative move to avoid risk. But control is probably an over-reaction to risk that should be managed through financial means. The governor of California famously took this tact at the height of the energy crisis, reasoning that long-term contracts were necessary to get control of the market. This was based on the assumption that the market was at fault for the crisis and was out of control. But the market, as we now know, was missing its usual hydro support, something that was entirely out of his control or anybody else's. Worse, the hydro support was due to return at any time. The policies that appeared in the heat of the moment were designed to impose central control. The risks were not understood, and the solutions simply made a bad situation worse. They did nothing to bring the market under control; the market soon corrected itself. We can only imagine how things might have been different if policy makers had understood the risks before they fell victim to them.

If one believes that the crisis of 2000-2001 was manufactured by traders or by withholding on the part of merchant energy suppliers, then one's attention is focused away from the problem of quantifying risks and figuring out what to do about it. How often will hydroelectric resources recede to 2000-2001 levels? If it happens again in two years, say, where on the system will the energy be replaced next time? Will the transmission system be sufficient to move the power to where it is being consumed? Will we need to ration consumption again if another shortage appears? How can we make these decisions?

The West is different. The risks that the West faces are unlike the risks that the rest of the nation faces, and the solutions must be unique. The West has to deal with a possible swing in hydro output that is as great as the hydro output of the rest of the nation. However, the cycles in hydro variation are long, much longer than market memory. These things demand dispassionate analysis, both forensic and forward. A legitimate role of government in this situation is bringing the analytical resources of the entire region together to analyze the risks and to recommend policies accordingly. The Northwest Power Planning Council is adept at analyzing its hydro situation, but is naturally focused on its 4-state region. Since the entire West is hydro dependent to a certain extent, this capability needs to be extended so that others in the region can benefit from it as well.

Markets can handle the short-term allocation problems as long as policy makers are willing to let price do its work, and willing to let market structures evolve toward greater west-wide consistency. Markets will move fuel to generators and power to consumers. Markets and consumers can jointly ration consumption, too, if that is necessary. Markets may consume surplus capacity rather than preserve it, however. The West may have to be structured to tolerate it and maintain it. Of course, generators must also be paid to maintain it, if maintain it they must. But this is the essence of risk management, once the risks are understood.

How much did demand response contribute to meeting the 100,000 gigawatt-hour gap?

The shortage in overall electric energy supply was so severe that it could not be solved by additional generation for existing "swing" gas generation alone. New generation could not be added quickly enough, or could the pipeline system handle the additional requirement at existing gas fired generation. So a significant reduction in demand was required throughout the West. As illustrated by the "How big is the 100,000 gigawatt-hours?" discussion, the conservation effort could not be handled by one sub-region in the West alone nor by increases in gas-fired generation.

What kind of reduction was needed? Just shaving peak demand wasn't enough. While peak demand reduction is very valuable in electricity markets that run into capacity limits on a few peak days, it does not address a hydro production shortage that might last one or more years. Overall annual electric energy usage must be cut – this is short-term energy conservation of the "ugly kind" involving ceasing using the energy electric for specific uses over one or more years. The clearest example of these reductions is the BPA buyout of electricity usage for the "Direct Service Industries" in the Northwest – cutting aluminum production (and jobs) for a one to two year period.

The Northwest led these reductions, achieving a 2500 MW to 4000 MW average energy reduction by the summer of 2001. California's efforts were delayed, with the primary demand response beginning once it was clear that retail prices were going to be increased. California appears to have reduced demand by 3,000 MW to 6,000 MW. These market-driven buyouts and demand savings, conservatively, achieved a 5,000 MW average electric energy reduction, or about 44,000 gigawatt-hours per year.

These demand reductions occurred more quickly than new production capacity could be added – and were key, along with the additional gas-fired electricity production at existing plants, for closing the 100,000 gigawatt-hour hydro shortage gap and keeping the lights on for more critical uses in the West.

Data Sources: California Energy Commission, Emergency Conservation and Supply Response 2001; Readiness Steering Committee and NWPP Council Staff, Coping with the 2000-2001 Energy Crisis; Sam Van Vactor, Economic Insight, FERC Docket EL02-60-003, October 15, 2002.

F. Pickel

Risk management was the discipline missing in California. If no one knows the risks exist, then they will not bother to manage them. It is safe to say that in the early 2000, "the market" in the West was no more familiar with its price risks than were government officials.[24] The crisis caught everybody by surprise. At this moment in time, however, everybody has received a painful education, even if they may have a variety of ideas about how it all happened. It would seem that the market for risk management in the West is ready to take the next step.

Is the FERC's Standard Market Design (SMD) the answer?

Among the goals of SMD are widespread market rules that are consistent with the physics and economics of generation and transmission of electric energy. As currently envisioned, SMD could enhance the efficiency of daily operations in the West, but would perhaps be most fruitful in providing market guides for the proper location of new generating capacity and transmission lines. The unique characteristics of the West don't invalidate the market design any more than they invalidate physics or economics. SMD may not adequately address the West's unique reliability[25] issues, but the West, and California in particular, have not adequately addressed these issues either. If SMD lacks realism for the West, it is the West's job to bring realism to SMD. It is only convenient, and in no way critical, that the two major interconnections in the U.S. have the exact same rules. However, for the West's ability to manage and allocate its resources in times of stress such as 2000-2001, West-wide consistency and realism in market rules are critical.

Policy makers in the West have some decisions to make, and these decisions should be made together and in full knowledge of the characteristics of the Western system. The hydro resource can vary a lot in a relatively short period of time. A wet, surplus year with low prices can be followed by one with high prices and shortages. Price-responsive demand can avoid service interruptions, and simple hedging can avoid economic dislocations to a large extent. But, since hydro supplies can be low for relatively long periods of time, the public may not tolerate supply tightness as well as it might tolerate supporting the cost of slack capacity. The same logic applies to gas-pipeline capacity as well. Policy makers in the West should know what these risks are and should make a conscious decision about how to either assure adequacy or deal with occasional tightness.

In many states, drivers are required to carry certain insurance coverage, because without the requirement many drivers will go uninsured. The FERC is considering requiring power companies to carry "insurance" as a part of SMD, in the form of a capacity requirement. Whether this or any other "forced insurance" concept is required to assure adequacy over the long term in the West is not known, but it is likely that some form of capacity insurance market will be required to support adequacy in the West. Perhaps a light-handed requirement coupled with sound West-wide adequacy monitoring is all that is required. If capacity is properly valued, it will not be necessary to resort to government or regulated utilities for building and owning capacity. However, these important decisions should be made consciously from a base of knowledge, and should not be made by default.

Acknowledgements

Financial support for the preparation of this paper was provided by the Western Power Trading Forum. The views expressed herein are those of the authors Mr. Jeff Roark, an independent consultant, and Dr. Fred Pickel of the firm Tabors-Caramanis & Associates. As such, those views do not necessarily represent those of the many members of the Western Power Trading Forum, or of any of their respective officers, sponsors, or clients.

[24] "The Market" had never really seen a dry year. The West had seen a dry year as late as 1994, but that year was really before third party had really become involved in Western markets to any great extent. There was no price data available for 1994, for instance. The price expectations of twenty-something traders in the West were all developed in the intervening years, during some very wet years. Everybody learned from the crisis.

[25] The term "reliability" is used here in a broad sense that includes the wide variability of Western system operations in response to large deviations in hydro energy availability, as well as the heightened exposure to transmission outages and/or stability problems resulting from lengthy transmission lines.