Interfaces

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Noninvasive Test Scheduling in Live Electricity Markets at Transpower New Zealand

Steve Batstone, Geoff Pritchard, Golbon Zakeri

Interfaces
Vol. 46, No. 6, November–December 2016, pp. 482–492
ISSN 0092-2102 (print) | ISSN 1526-551X (online)

https://doi.org/10.1287/inte.2016.0865

Noninvasive Test Scheduling in Live Electricity Markets at Transpower New Zealand

Steve Batstone
Whiteboard Energy, Auckland, New Zealand 1143; and Sapere Research Group, Auckland, New Zealand 1140, steve@whiteboardenergy.co.nz

Geoff Pritchard
Department of Statistics, University of Auckland, Auckland, New Zealand 1010, g.pritchard@auckland.ac.nz

Golbon Zakeri
Department of Engineering Science, University of Auckland, Auckland, New Zealand 1010, g.zakeri@auckland.ac.nz

In 2013, Transpower New Zealand commissioned a new high-voltage, direct current link to transfer electrical power between the North and South Islands of New Zealand. This was a substantial and prolonged undertaking, requiring approximately 400 in-situ capability tests. Transpower elected to perform these tests "live," without suspending the normal operation of the wholesale electricity market. Instead, Transpower's trading team attempted to create suitable flow conditions for each test by entering into innovative financial derivative contracts with power generation firms. We created a stochastic dynamic programming model to handle the contingent scheduling of the tests; its most important random variable was the state of water storage available to hydropower plants.

Keywords: stochastic dynamic programming; scheduling; electricity; financial derivatives.
History: This paper was refereed. Published online in *Articles in Advance* October 20, 2016.

The high-voltage, direct current (HVDC) line between the North and South Islands is arguably the most important single piece of electricity infrastructure in New Zealand. It is both the largest-capacity transmission asset in the country and the sole electrical connection between the two main islands (Figure 1). Approximately 6 percent of New Zealand's electricity consumption is transferred across the HVDC link each year and, at full capacity, the link can transfer up to 25 percent of peak demand to either island. This makes it a crucial asset in ensuring the security of the supply.

The HVDC link is bidirectional, incorporating AC-to-DC and DC-to-AC converters at both ends of the DC transmission line so that power can be transferred in either direction. In 2007, the conversion equipment comprised a set of 1960s-era mercury-arc converters (Pole 1) together with a set of thyristor converters added in 1991 (Pole 2). For comprehensive references on high voltage engineering, see Ryan (2001) and Gill (2009). However, the unexpected and (eventually) permanent failure of Pole 1 in late 2007 reduced the transmission capacity of the HVDC by approximately half.

For the next five years, the New Zealand electricity system operated with this limited capacity, increasing the risk of shortages in the event of low inflows to hydropower reservoirs. Although no shortages ever materialized, market participants experienced frequent constraints on flow across the HVDC, with consequential price volatility in New Zealand's location-based wholesale electricity market. The transmission congestion may have cost consumers as much as $17 million per month. (References to dollars in this article refer to New Zealand currency.)

In September 2008, the national grid operator of New Zealand, Transpower, received regulatory approval to invest $672 million in HVDC equipment (Figure 2, Pole 3) to replace Pole 1. Although the restoration of the link's full transmission capacity was

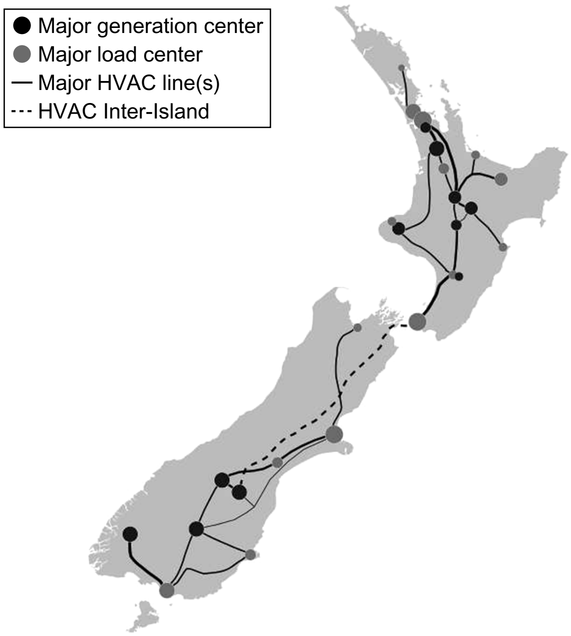

Figure 1: New Zealand's transmission grid contains high-voltage DC lines that connect generation sources to load nodes, which are depicted by the dashed line.

welcome, the design, installation, and commissioning of such a complex piece of electrical infrastructure was far from trivial. Siemens was awarded the contract to build Pole 3. As part of the terms of this contract, Siemens needed to subject the new equipment to over 2,000 tests to prove its capability under a variety of circumstances. Most of these tests could be conducted in Siemens' laboratory in Munich. However, as with many other large HVDC transmission lines, the final tuning and demonstration of capability needed to take place in a live environment, after the equipment had been installed in the field. For Transpower and Siemens, this required approximately 400 tests across the full range of the asset's capability to be completed in 2013.

Although not unique, this commissioning challenge was arguably one of the most complex in the world. Further adding to its complexity was Transpower's choice to create the right flow conditions for each test by using opportunistic scheduling and market instruments, rather than regulatory power—the first time in the world that a national grid operator used this approach. This created a stochastic scheduling

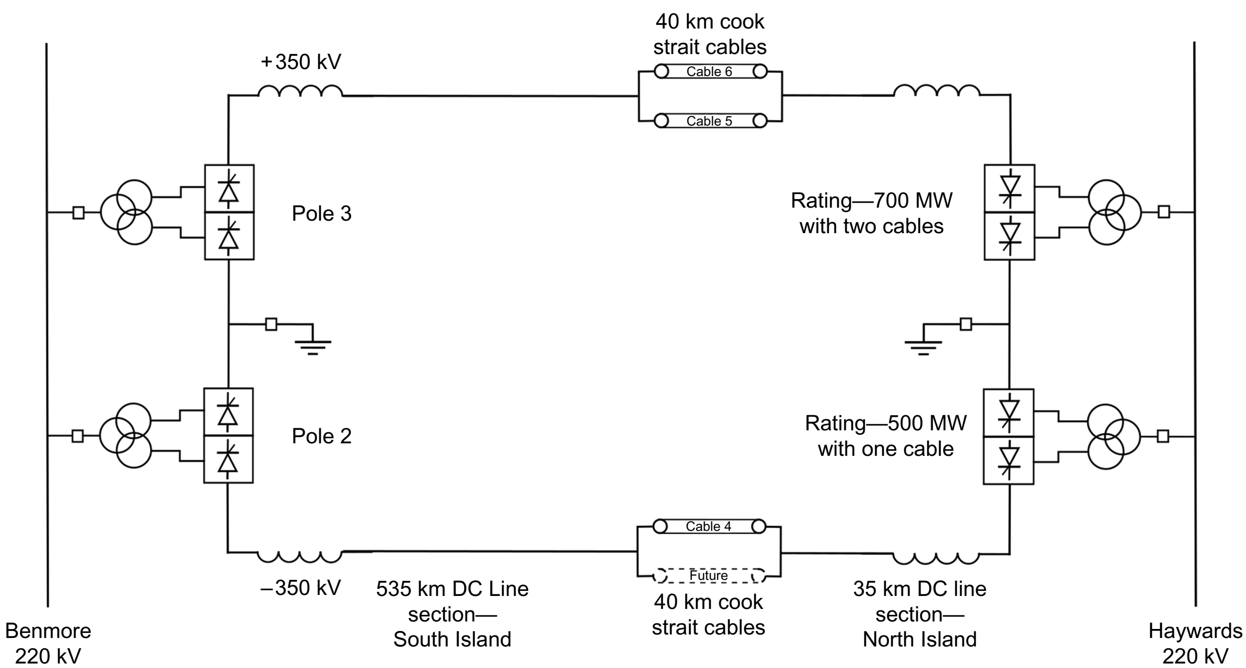

Figure 2: The DC connection consists of two essentially separate circuits that are connected at the common AC nodes at each end. During the period 2007–2013, only Pole 2 was operational. This HVDC figure depicts the connection after the full deployment of Pole 3.

problem: minimize the cost and disruption of the testing program by taking advantage of normal power flows as determined by the wholesale power market.

The implications of this choice should not be underestimated. New Zealand operates a lightly regulated wholesale market for power, with generators making largely unfettered offers of capability on a half-hourly basis. For comprehensive background on electricity market operations, including nodal pricing, see Schweppe et al. (1987), Chao and Huntington (1998), and Alvey et al. (1998). The supply side of the market is dominated by hydropower, which accounts for (on average) 55 percent of all generation. Hydrological conditions (e.g., stream flows and the levels of reservoirs) are thus a primary determinant of power flow on any given day. These conditions, although relatively transparent to the market, are highly unpredictable over the time horizon (a few months) required for the testing program. Generators' offers to the market are also driven by fuel cost and availability, portfolio positions including sales commitments, and expectations of wholesale prices, all of which are relatively opaque. Predicting the offers of the market participants (and hence the expected flow on the HVDC) for any given time period was an enormous challenge, even 24 hours in advance. The market itself offered little help: although tentative offers are provided to the market operator seven days in advance, reliable indications of generation intentions do not emerge until a few hours prior to real time. Gate closure (after which generators are unable to amend offers) occurs two hours prior to real time; however, even after that, variation in demand and wind-power generation can materially influence the final outcome. The uncertainty in HVDC flows, even after gate closure, would often exceed the allowable range of HVDC flows for some tests.

Moreover, as we outline below, the commissioning tests often required the HVDC to be constrained, resulting in price differences across it. Helping market participants manage the resulting price risk was also a key objective. (A financial-transmission rights market, which might have helped to hedge these risks, did not exist in New Zealand in 2013.)

Transpower needed to take an approach to the testing program, which combined a dynamic and contingent schedule of tests, based on uncertain system conditions, with trading mechanisms (i.e., derivative contracts with market participants) to adjust the HVDC flow in situations in which the market clearing process did not naturally provide the right flows. These contracts were devised to provide incentives to generators to adjust their desired output to levels that created the correct flow conditions required for each test. The expected cost of contracting to manufacture the right conditions had to be traded off against the costs of delaying the test until system conditions were more favorable. Because of the highly sequential nature of the testing program, delaying a test had a ripple effect on other remaining tests.

The Tests

Most transmission lines consist only of poles or towers, conductors, and insulators. Embedding a direct-current line within an alternating-current system requires a substantial additional investment in power electronics equipment and control systems at each end. The need to replace this equipment led to a commissioning process whose complexity far exceeded the complexity for an ordinary transmission line. Approximately 400 tests were required in the live environment, primarily to ensure that the new asset could handle the behavior of the New Zealand power system, and did not cause any undue instability.

The test requirements ranged from the relatively simple, such as providing flows on the HVDC that remained between upper and lower bounds for a specified period, to far more complex specifications, requiring very precise flow levels, and (or) ramping up and down power flows at specified rates. Most (70 percent) of the tests were an hour or less in length; however, a modest number exceeded six hours, and the longest was 26 hours.

Consistent with a normal approach to commissioning a complex asset, the testing program required groups of tests, pertaining to different aspects of the HVDC's capability, to be performed repeatedly at gradually increasing power transfer levels. The overall program had two major stages: the commissioning of the new Pole 3 equipment (approximately 200 tests), and then the commissioning of the new control system to ensure that Pole 3 was properly integrated with the existing Pole 2.

The Approach

Transpower's approach to the commissioning challenge incorporated three elements:

- An optimization of the sequence of tests, which created an ideal plan that dynamically responded to emerging hydrology. The primary purpose of this optimization was to develop a suggested plan and be able to simulate different testing scenarios.
- A set of analytical tools and visualizations that allowed a trading team to produce probabilistic scenarios on expected generation and demand, and on the prospect of transmission constraints elsewhere on the grid and (crucially) expected natural HVDC flows should the market be left to its own devices.
- A set of financial contracts, acceptable to the generation community that provided the correct incentives to generators to adjust their production plans with sufficient confidence that the test could be scheduled and completed. This included fine-tuning contracts with those few participants permitted to vary their power levels in real time: wind generators and demand.

Optimal Test Scheduling

Few projects have tackled the challenge of project scheduling in an uncertain environment; for an excellent survey, see Verderame et al. (2010). In principle, however, several standard approaches are applicable to constructing an optimal schedule under uncertainty. These include stochastic programming, dynamic programming, and Markov decision processes. For a comprehensive treatment of these topics, we refer the reader to Birge and Louveaux (2011) and White (1993).

For initial planning purposes, the HVDC test-scheduling team developed a scheduling optimization model, which it named Moshe. It was clear from the start that Moshe needed to be responsive to hydrological conditions (e.g., inflows and hydro energy storage in reservoirs), because these are the most important drivers of normal power flows. However, the hydrological conditions were unforeseeable in advance, and could change significantly during the testing program. Any single, inflexible test schedule, no matter how well optimized for expected conditions, would potentially impose HVDC flows very different from those that would otherwise occur. This would have been unacceptable to industry participants.

The team thus decided that Moshe would be a stochastic dynamic program (SDP), with its stochastic element comprising a climatic and hydrological model. We provide a discussion of SDP models and terminology in the appendix. Moshe's output consisted of state-dependent policies, rather than a single schedule. Stochastic dynamic programs have a long history of application, including hydro-thermal scheduling in electricity systems; examples include Terry et al. (1986), Pereira and Pinto (1991), and Johnson et al. (1998). These models have also been applied in the areas of telecommunications and mining. For an example in telecommunications, we refer the reader to Lesaint et al. (2000), and for an example in mining to Newman et al. (2010). Perhaps the closest SDP model to ours is maintenance scheduling for bridge networks as explored in Frangopol and Liu (2007). However, none of these applications involves an intricate live electricity market.

This design concept suggested an intuitive way to present the results: for any assumed hydrological sequence (e.g., from a recent historical year), the test schedule could be determined and displayed (Figure 3). Naturally, it had to be emphasized that this full schedule would not be visible from the start to anyone implementing it; rather, it would be determined dynamically as the hydrological events played out.

As a simplified example, suppose that we have only to perform the three tests shown in Table 1. The precedence structure is such that the third test (26.001) must be performed last. The costs of performing the tests are known with certainty only for tests begun in the next five hours; after that, a random hydrology change (to a low or high hydrology state, with a 50 percent probability for each) will determine costs, as Table 2 shows. All three tests must be completed within eight hours, and an additional terminal cost depends on the finish time, as Table 3 shows.

At each point in time, the subset of the tests completed so far may be $\{\}$, $\{12.001\}$, $\{65.004\}$, $\{65.004, 12.001\}$, or $\{65.004, 12.001, 26.001\}$. In addition, from Time 6 onward, there are two possible hydrology states. This gives the dynamic program the state space depicted in Figure 4.

In each system state, the possible actions are to do nothing for one time step, or to begin one of the tests not yet performed.

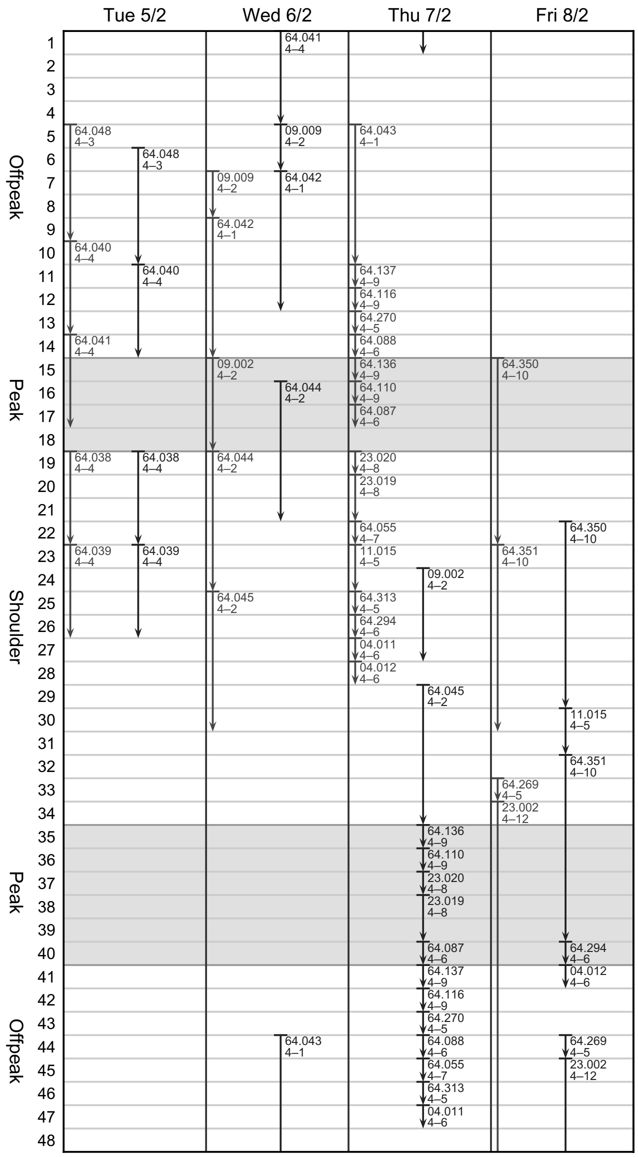

Figure 3: Moshe produced the indicative test schedules for February 5 to February 8, 2013. In the diagram, the vertical axis denotes the day, and the horizontal axis denotes the time (half-hourly) within the day. The text at the left in each box shows the optimal test schedule if the hydrology is as in the summer of 2007–2008; the text at the right within each box shows the prospective schedule under hydrology matching 2010–2011. For example: Test 64.350 (top center) is best scheduled during the morning peak period if the hydrology is as in the summer of 2007–2008; it is best scheduled in the shoulder period if the hydrology matches the 2010–2011 scenario.

The solution to the problem proceeds by the standard stochastic dynamic programming technique (see the appendix). Note that there is no single optimal test plan: the optimal path is contingent on the hydrology state from Time 6 onward. In presenting this solution, it would be possible to present a single schedule only for the first two tests; the exact timing of the third test is not determined at the outset.

The Moshe SDP has one further complication not present in the above example. Tests must be assembled into work shifts of at most 12 hours, with a minimum 11-hour break between shifts. This consideration adds a dimension to the system-state space, meaning that the possible system states s are 4-tuples:

$$s = (t, P, h, w),$$

where
- t is the time, in half-hourly time steps;
- P is the program-completion state (i.e., the subset of the tests already completed);
- h is the hydrological state;
- w is the work-shift state: the number of half-hours of work done so far in the current shift.

The actions available in any given state could include performing a test, ending the current work shift, or simply idling for a half-hour. The durations of the tests varied from a single half-hour up to 26 hours (with very long tests being exempt from the work shift rules).

States in which all the tests have been completed are terminal. The value assigned to such a state depends only on the time of program completion; sooner is better (i.e., a lower value) than later. Also terminal are those states in which the total available time—typically around six months—has run out with the test program still not completed; such states are assigned the value ∞.

A typical instance of Moshe involved approximately 170 tests, each with an associated duration, time window, and other timing constraints. Various precedence requirements constrained the order in which the tests could be performed; these were computationally helpful, because they reduced the size of the problem's state space. A further reduction arose from noting a few small groups of tests with identical scheduling requirements: the tests within such a group could be performed (not necessarily consecutively) in an arbitrary order. The eventual size of the program-completion space, from "nothing done yet" to "all done" and every possibility in between, typically ran to about 600 states.

Test	Description	Duration (hours)	Flow range (Megawatts)	Ramp	Prerequisites
65.004	Simulated filter protection trip, 100 MW	3	80 to 120	No	66.001, 66.003, 66.005
12.001	Reversal of power flow direction, Pole 3, 70 MW	2	−98 to 98	Yes	07.005, 11.006, 14.003
26.001	Reduced voltage Pole 3, 300 MW	1	249 to 300	No	65.004, 12.001

Table 1: The first of the three typical tests shown simply requires northward power flow of between 80 and 120 megawatts to be maintained throughout the test. The second is a more complex test requiring the flow to vary ("ramp"), including reversals of direction. The third must be done after the other two have completed.

We derived Moshe's objective function (the "costs" incurred when tests are performed) from a table giving the expected change in HVDC flow needed to achieve a given flow at a given date, time, and hydrological state; this table was a distillation of a database that records flow patterns on the power grid over the five years since the failure of Pole 1. For tests requiring varying HVDC flows (i.e., ramping), we assumed a cost equal to the greater of the flow changes required to achieve the initial and final flows of the ramp; this rather pessimistic assumption reflects the difficulty of such tests. It soon became apparent that some tests were best done at night; others could be performed most easily under peak load conditions in the morning or evening. For many others, the ideal time of day depended strongly on the hydrology state.

Start time	Hydrology state	Cost to perform 65.004	12.001	26.001
0		3	4	
1		4	3	
2		5	4	
3		5	3	
4		5	2	
5			3	3
6	Low			1
6	High			2
7	Low			0
7	High			2

Table 2: The costs to perform the tests are contingent from Time 6 onward.

Finish time	Terminal cost
6	0
7	2
8	3

Table 3: The terminal cost structure rewards early finishes.

We eventually adopted a very simple hydrological model: a six-state Markov chain, with three energy storage states (low, medium, and high) and two inflow states (low and high). This model used weekly time steps, with state changes occurring notionally between Friday and Saturday—a circumstance that led occasionally to some curious optimal solutions, because Moshe scheduled tests late on Friday nights, or early Saturday mornings, to avoid or exploit the impending hydrology update.

The combination of approximately 600 program-completion states, 24 shift-completion states, and six hydrology states produced a total state space size of approximately 80,000, for each of the half-hourly time steps over several months. The resulting problem could be solved on a standard laptop computer within about 20 minutes.

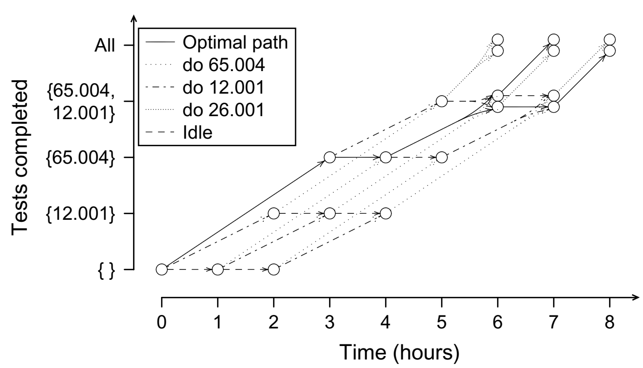

Figure 4: We seek a least-expected-cost path to one of the terminal states on the top row of the diagram. Each arrow has an attached cost, which represents the cost of performing the associated test. Each terminal state also has a cost, representing the cost of completing the test program at that time. There is a 50 percent chance that the costs will all change at Time 6, making each test-completion state from then on into a pair of problem states. The optimal path branches at Time 6: decisions from then on are contingent on costs observed at that time.

While immensely useful for planning, Transpower did not use Moshe in the operational environment of testing. The rapidity with which the testing and system environment changed, often on an hourly basis, did not permit the frequent rerunning and interrogation of a complex mathematical model. Moshe's primary value was as a tool of communication (internally and with the industry), and for testing various timings of testing, especially as delays were experienced in the physical works. Moshe was run on a weekly basis to give an indication of the sequence of tests for each week. Financial contracts, as we describe in the next subsection, would then be procured to obtain proper flow conditions for a test schedule close to what Moshe suggested.

Financial Contracting

To adjust HVDC power flows in the desired direction, the project team used purpose-designed financial instruments. Most often, the counterparties to these contracts were the power generation firms. Several principles governed the nature of the contracts:

- The form of contract had to align with the International Swaps and Derivatives Association (ISDA) arrangement, to minimize the extent of negotiation of the terms and conditions. Generally speaking, all market participants were familiar with the ISDA arrangement, which immeasurably sped up the contracting process.
- The contracting options available needed to provide flexibility in terms of the trade-off between certainty of outcome and cost.
- The contracting options needed to cover a range of different "products" that needed to be secured.

The team developed two forms of primary energy products. The first was a traditional swap or contract for difference (CfD), where a contract quantity and price was negotiated for the period concerned. Such instruments do not guarantee any particular physical outcome from a generator, but rather provide a financial incentive for generators to vary their output. However, the exact degree of variation depended on their fuel cost structure (including water value), other financial commitments, and anticipated influence over the wholesale price. Few of these factors were visible to the trading team; hence, these pure financial CfDs could, at best, be relied on for achieving a degree of shift in the generator's position. They could not be relied on for precision, but were useful for tests where a relatively wide range of HVDC flows was acceptable.

For example, test 26.001 in Table 1 required flow in the range 249–300 MW for one hour; however, if Moshe performed this test in mid-March under medium lake-storage levels, the unmodified HVDC flows would typically be higher than the following: circa 500–700 MW during the day and 350–500 MW at night. Therefore, Moshe would likely schedule the test at night. Once it became apparent that the test would probably occur in a particular one-hour nighttime period, the team would, for example, be able to negotiate a CfD for 100 MW, for that hour only, with a North Island (i.e., downstream) generation firm. The firm could then be expected to generate 100 MW more than it otherwise would to cover its contracted position, reducing the HVDC flow by 100 MW. Additional contracts could be added to refine the position as the time for the test approached.

For tests requiring more precise control, such as 12.001 in Table 1, we developed a more precise form of contract, termed a physical CfD. These contracts mimicked the payoff of a pure CfD, except that the payoff was conditional on the generator meeting a physical requirement; this requirement usually pertained to providing greater than, or less than, a specific output from a specified set of power stations. Physical CfDs have been used in other commodity markets; however, to the best of the authors' knowledge, their use in this project was a first in the context of the electricity market.

The contract had a two-part pricing structure. If the generator met the physical requirement, the contract was settled according to a price that was usually favorable to the generator. If the generator failed to meet the physical requirement, the contract was settled according to a price that was usually favorable to Transpower. Over the course of the 400 tests, less than 10 contracts were knocked out. Generator counterparties became very good at ensuring their physical conditions were met by the way they offered into the wholesale market.

Shifting generators' energy output was only one type of product required. Others include the following:

- Reserve contracts: In New Zealand's wholesale market, the size of the HVDC requires it to be considered a risk for the purposes of scheduling spinning reserve. Spinning reserve is defined as extra generating capacity that is available by increasing the power output of generators that are already dispatched. Hence, for tests that required flow levels implying a high degree of spinning reserve, reserve contracts were developed, which provided incentives to market participants to maximize their reserve offers (which were, in turn, highly dependent on their generation level, introducing a feedback loop for the trading team).

• Capacity contracts: Some tests required the scheduling of large generation units. Some market participants preferred to be compensated for making this capacity available for the duration of the test, and charged a fixed capacity fee (rather than a price-dependent CfD).

• Load-balancing contracts: New Zealand electricity-distribution utilities retain control of ripple-controlled water heating, which allows some degree of control (usually downward) over electricity demand. The team contracted with one network company in each island, so that load curtailments could be used to adjust the HVDC flow in either direction. Up to 100 MW change was achieved at times, dependent on system conditions.

• Wind-balancing contracts: Another innovation in the New Zealand market emerged as a result of the necessary real-time control. No generator in the market over 10 MW is permitted to change its output after the gate closure (i.e., two hours prior to real time), except for intermittent generation (i.e., wind). This meant that wind generation became the only form of generation that could assist the trading team to tweak the HVDC flow during testing if system conditions deviated from expectations. About 300 MW of wind was contracted to reduce output when required by feathering the wind turbine blades. If system conditions reversed, the wind generator could increase output again as long as the wind conditions permitted. Modern control systems allowed these changes to be executed often within three to five minutes of the instruction being given. The contracts in place allowed the wind generator to be compensated according to a calculation of what its generation would otherwise have been.

Analytics

The day- and week-ahead volatility in the wholesale market—weather events and their impact on wind, hydro and demand, generation plant outages, and general trading behavior—needed to be incorporated into the testing process via predictive analytics.

In a system in which half-hourly demand varied between 5,000 MW and 7,000 MW, achieving desired testing ranges of ±100 MW proved challenging. Theoretically, if every generating plant in the system (except the frequency keeper) could be contracted to provide a specific output, the only residual variation would be demand, which—within each island—would be met by the frequency keeper. Such an approach would be immensely costly, time consuming, and difficult. The decision-making process became one of trading off the acceptable variation in HVDC flow for a given test, with the cost and difficulty of contracting (and the implied residual uncertainty in HVDC flow after contracting). Hence, the analytical process needed to highlight which generating plants (or demand regions) were creating the greatest uncertainty, implicitly generating a "merit order" of contracting. This process needed to also account for the changes in HVDC flow (from its natural level) required for the test.

With the copious historical data available on demand and on the half-hourly dispatch of generators, long-term distributions could easily be formed to estimate the uncertainty associated with each generator. However, this was of little use, due to the highly conditional nature of variability: variation in demand and wind output were highly contingent on the weather conditions at the time, while hydro output was, to a large extent, dependent on the inflow patterns of recent weeks. For larger generators, the firm's net financial position, which is not visible to the market, was an equally important driver, and changed on a daily basis. Therefore, the distributions of potential output needed to be cognizant of the current and immediate-past system, participant, and weather behavior.

The analytical problem had a classic Bayesian structure, with the probabilities needing to be constantly updated with real-time data. The team considered the construction of a formal Bayesian network, but did not pursue it due to concerns over solution time.

Team members expected that the analysis would need to be conducted several times daily, as new information emerged, and as contract negotiations continued.

The team developed an analytical tool, which attempted to capture the Bayesian structure by isolating the variation caused in recent HVDC flow by each individual system element (demand, generation), thus establishing current priors. An input could be any of a number of variables, such as electricity demand from irrigation in a particular region or wind generation. We would then examine the historical relationship between this variable and HVDC flows, for different states of the world (e.g., day of the week, time of day). The analyst would then be able to specify the likely state for a future test (e.g., day, time, forecast weather conditions), and the tool would use the historical relationships to generate a distribution of HVDC flows based on the underlying drivers.

It was the role of the trading analyst to establish whether this prior distribution was reflective of the conditions expected for the test (e.g., by seeking weather forecasts or adjusting for plant outages), thus providing posterior adjustments. Finally, if the level and (or) variation in the resulting HVDC flow did not meet the requirements of the test, then contracting with market participants was necessary, which again provided a form of posterior updating.

The contracts developed for real-time flow management through the use of wind generation and demand response had a short lead time (between 5 and 10 minutes); thus, they required HVDC flows during the test to be monitored with a high degree of precision. In practice, the team monitored HVDC flows in real time with a one-minute resolution. This also allowed real-time monitoring of the performance of contracted and other generators.

The Outcome

In May 2013, after 163 contracted tests, Transpower successfully commissioned Pole 3 into the market, in time for the critical winter period. Commissioning of the final Pole 2 interface and control system was achieved on December 9, 2013 after an additional 165 contracted tests. Over the whole six-month commissioning period, the trading team successfully conducted 328 contracted commissioning tests (including 60 retests, mostly due to the new equipment failing engineering tests). This equated to approximately five tests for each day that testing took place—a high level of activity for the team, given the amount of effort required to plan and execute each test. Only three tests needed to be performed again because they did not achieve the right HVDC transfer. Despite some doubts from industry participants about the robustness of the in-market contracting approach, the successful commissioning, with only a one percent failure rate due to contracting, showed it to be entirely practical. In terms of the impact on the market, the contracted tests took 270 hours; the longest individual test was the 26-hour heat run for Pole 3 commissioning. Although a difficult metric to calculate, we estimate that a total of 120 Gigawatt hours (GWh) of market shift (deviation of HVDC flow from its natural position) occurred, equalling only 0.5 percent of national demand over the commissioning period. The heat run alone required approximately 15 GWh of market shift (Figure 5). A greater degree of market shift was observed with Pole 2 commissioning, largely as a result of the much higher flow targets (north and southward) that were required to test the bipole control system. By the time these high-flow tests were required (November 2013), natural HVDC flows were strongly northward as a result of hydrological conditions. This made the high south-flow tests conducted between November 7 and November 21 among the most challenging to contract, because these five tests required 28 GWh of market shift.

Probably the most remarkable statistic is that 1,179 deals were negotiated over the commissioning period. This number does not include standing arrangements for load control. This means that, on average, the trading team was concluding five deals per day for Pole 3. It was also conducting a remarkable 8.3 deals per day for Pole 2. The higher Pole 2 figure partly reflects higher flow tests (because Pole 3 was already commissioned) and the need for more reserve contracts at higher flows. For comparison, the entire industry transacted only 721 over-the-counter contracts in all of 2012. This represents a rate of two per day spread across all market participants.

Acknowledgments

The authors thank Chris Otton, Andrew Gard, David Reeve, and the rest of the project team at Transpower Ltd. and Whiteboard Energy for their invaluable contributions to this

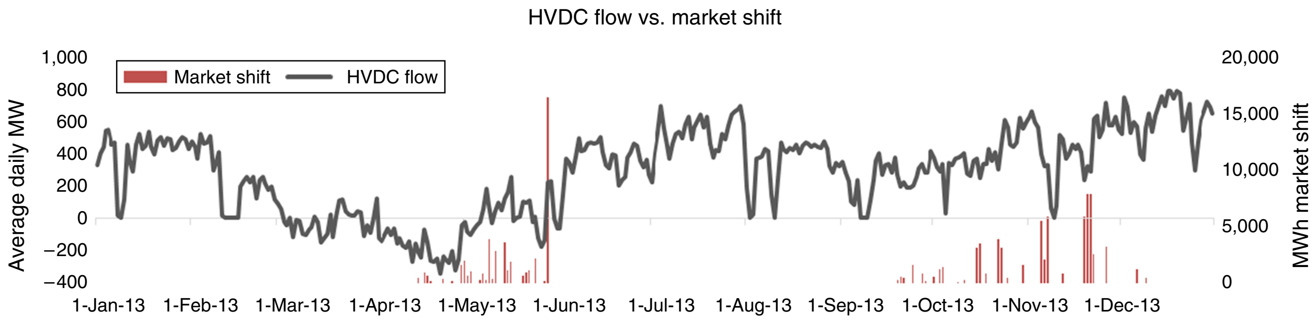

Figure 5: (Color online) The response of HVDC flow to contracted market shift was often clearly visible during Pole 3 commissioning.

research. They also thank the anonymous referees for their highly valuable comments that helped improve this paper.

Appendix. Stochastic Dynamic Programming

The Moshe stochastic dynamic program (SDP) is an example of a class of SDPs, which may be described as follows (White 1993). The system to be optimized may be in any of various *states*; a state s is thought of as being achieved at some specific (integer) time $t(s)$. (System configurations achieved at different times are considered different states, even if otherwise identical.) A state s has an associated set of *actions* $A(s)$, which may be taken when the system is in state s. Taking an action $a \in A(s)$ incurs a cost $c(a)$ and precipitates one of an associated set of *transitions* $T(a)$; the transition $\tau \in T(a)$ occurs with probability $p(\tau)$ independently of all previous random events. Of course, we must have $\sum_{\tau \in T(a)} p(\tau) = 1$ for any action a. When transition τ occurs, the system is transformed to a new state $\sigma(\tau)$. Every transition consumes some time, so that we have

$$t(\sigma(\tau)) > t(s) \quad \text{for each state } s,\ a \in A(s), \text{ and } \tau \in T(a).$$

The *value* $V(s)$ of a state s is the expected cost of all future actions, if the system begins in state s and is managed optimally thereafter. It is easy to see that the value function V satisfies the Bellman equation

$$V(s) = \min_{a \in A(s)} \left(c(a) + \sum_{\tau \in T(a)} p(\tau) V(\sigma(\tau)) \right) \tag{1}$$

for any state s with $A(s) \neq \varnothing$. States s with $A(s) = \varnothing$ are considered terminal: in such a state, there is nothing left to do. Values $V(s)$ for terminal states s are given as part of the problem specification; Equation (1) can then be used to find values for other states.

References

Alvey T, Goodwin D, Ma X, Streiffert D, Sun D (1998) A security-constrained bid-clearing system for the New Zealand wholesale electricity market. *IEEE Trans. Power Systems* 13(2):340–346.
Birge J, Louveaux F (2011) *Introduction to Stochastic Programming* (Springer, New York).
Chao H-P, Huntington HG (1998) *Designing Competitive Electricity Markets* (Kluwer, Boston).
Frangopol D, Liu M (2007) Bridge network maintenance optimization using stochastic dynamic programming. *J. Structural Engrg.* 133(12):1772–1782.
Gill P (2009) *Electrical Power Equipment Maintenance and Testing* (CRC Press, Boca Raton, FL).
Johnson RB, Svoboda AJ, Greif C, Vojdani A, Zhuang F (1998) Positioning for a competitive electric industry with PG&E's hydrothermal optimization model. *Interfaces* 28(1):53–74.
Lesaint D, Voudouris C, Azarmi N (2000) Dynamic workforce scheduling for British telecommunications plc. *Interfaces* 30(1):45–56.
Newman AM, Rubio E, Caro R, Weintraub A, Eurek K (2010) A review of operations research in mine planning. *Interfaces* 40(3):222–245.
Pereira MVF, Pinto LMVG (1991) Multi-stage stochastic optimization applied to energy planning. *Math. Programming* 52(1):359–375.
Ryan HM (2001) *High Voltage Engineering and Testing* (Institution of Electrical Engineers, London).
Schweppe F, Caramanis MC, Tabors RD, Bohn RE (1987) *Spot Pricing of Electricity* (Springer, Boston).
Terry LA, Pereira MVF, Araripe Neto TA, Silva LFCA, Sales PRH (1986) Coordinating the energy generation of the Brazilian national hydrothermal electrical generating system. *Interfaces* 16(1):16–38.
Verderame PM, Elia JA, Li J, Floudas CA (2010) Planning and scheduling under uncertainty: A review across multiple sectors. *Indust. Engrg. Chemistry Res.* 49(9):3993–4017.
White DJ (1993) *Markov Decision Processes* (Wiley, New York).

Verification Letter

Andrew Gard, Project Director, HVDC Pole 3; Chris Otton, Contracts Coordination Manager, HVDC Pole 3, writes:

"Transpower NZ Ltd is happy to support the submission of the abstract 'Non-invasive test scheduling for vital transmission infrastructure over live electricity markets' to INFORMS as a candidate for the Wagner prize in 2014.

"We verify that the approach outlined in the abstract (and subsequent paper) was successfully followed in practice, which, together with the substantial core engineering achievements, allowed us to complete commissioning of the new HVDC equipment in December 2013.

"We note the trading approach adopted was specifically mentioned by a panel of judges when they awarded Transpower the Deloittes' Energy Company of the Year in 2013."

Steve Batstone is a director at Sapere Research Group. He has more than 15 years experience in the electricity sector, including wholesale markets, portfolio and risk management, market design and transmission. His management experience in business planning, operations, and strategy analysis is complemented by being a specialist modeller in the field of decision-making under uncertainty. Stephen holds a PhD in operations research from the University of Canterbury.

Geoff Pritchard is a senior lecturer at the University of Auckland Department of Statistics. He is a principle at the Electric Power Optimization Centre (EPOC). He has more than 20 years of experience in probability and statistics with special application to electricity markets. Geoff holds a PhD in mathematics from University of Wisconsin-Madison.

Golbon Zakeri is a senior lecturer at the University of Auckland in the Department of Engineering Science. She is a director of the Electric Power Optimization Centre, the deputy director of UoA Energy Centre, and the president of the New Zealand Operations Research Society. She has more than 20 years of experience in optimization, with a focus on electricity markets. She holds a PhD in mathematics and computer sciences from University of Wisconsin–Madison.