Interfaces

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

The Energy Authority Optimizes Water Routing and Hydroelectric Generation on the Columbia River

Zengjian Hu, John Putz, Yohan Sutjandra, Abel Chan, Edmond Mount, Kristen Baker

To cite this article:
Zengjian Hu, John Putz, Yohan Sutjandra, Abel Chan, Edmond Mount, Kristen Baker (2015) The Energy Authority Optimizes Water Routing and Hydroelectric Generation on the Columbia River. Interfaces 45(1):43-57. http://dx.doi.org/10.1287/inte.2014.0768

Interfaces
Vol. 45, No. 1, January–February 2015, pp. 43–57
ISSN 0092-2102 (print) | ISSN 1526-551X (online)

http://dx.doi.org/10.1287/inte.2014.0768

THE FRANZ EDELMAN AWARD
Achievement in Operations Research

The Energy Authority Optimizes Water Routing and Hydroelectric Generation on the Columbia River

Zengjian Hu, John Putz, Yohan Sutjandra, Abel Chan, Edmond Mount, Kristen Baker
The Energy Authority, Jacksonville, Florida 32202
{zhu@teainc.org, jputz@teainc.org, ysutjandra@teainc.org, achan@teainc.org, emount@teainc.org, kbaker@teainc.org}

We describe a software application that enables owners of generation output from a virtualized Federal Columbia River Power System to safely operate the system, while also shaping the generation to meet their energy and economic needs. The application, the Optimizer, employs modern operations research techniques to convert a highly nonlinear problem into a linear one to create a robust solution for the six-dam system on an hourly basis, over a 10-day time horizon, and within a few minutes. The Optimizer helps to simultaneously manage and optimize the generation portfolios for 13 utilities in a stringent time frame around the clock, and enables hydroelectric (i.e., hydro) planners to ensure that the operation of the river meets all the requirements for flood control, fish management, electrical reliability, safe dam operations, and recreation under high degrees of uncertainty. The Optimizer allows utilities to integrate renewable, environmentally friendly wind and solar generation into their resource portfolio with hydro generation, and empowers these utilities to rapidly make decisions and adapt to changing conditions. We estimate that this project will reap benefits of $765–$952 million between 2011 and 2028.

Keywords: Columbia River; hydro optimization; mixed-integer programming.

Global climate change has necessitated a major shift in the way that the U.S. power industry creates electricity. Power generators are transitioning from traditional fossil fuels such as coal toward carbon-free, renewable hydroelectric (hydro) generation such as wind and solar. This shift creates significant challenges to maintain a reliable energy grid, because the system's reliability depends on moment-by-moment changes in electricity demand being met by equal and offsetting changes in electricity supply. Wind and solar, the two most viable forms of renewable generation at present, hinge on weather conditions and daylight cycles; therefore, they cannot be controlled to match changes in electricity demand.

Until large-scale electricity storage becomes feasible, the only two practical solutions to fill the gaps between these new intermittent supply resources and the variable demand for electricity are gas-fired generators and hydro generation. Gas-fired generators have the capability to start quickly and adjust their output in real time to respond to the imbalances on the grid, but are inefficient and carbon intensive. Hydro generation provides the electric grid with the needed flexibility to integrate intermittent resources into the system, and does so without emitting CO_2 or harmful by-products.

Unfortunately, the ability to build new hydrogeneration capacity is limited by geography and ecological concerns. In the Western United States, these limitations on new hydrogeneration development are compounded by the fact that the existing hydro system at times operates at its maximum ability to absorb rapid changes in supply and demand (i.e., integration capacity).

Increased variability from renewable resources and the inability to add new, clean hydro resources stresses the importance of optimizing the use of existing hydro resources to provide the maximum amount of flexibility to the grid. The Energy Authority, Inc. (TEA) has developed a hydrogeneration and water-routing optimizer, which we refer to as the Optimizer throughout this paper. The Optimizer allows 13 small-to-mid-size public utilities in the Pacific Northwest to harness the power and flexibility of the vast Federal Columbia River Power System (FCRPS). They use this resource to integrate renewable generation into their power supplies and to meet their demands in an environmentally responsible manner, while maintaining low electricity rates for their customer owners.

A number of criteria drove the development of the Optimizer. First, the FCRPS is complex; it consists of many large dams on the Columbia River and Snake River (see Figure 1), and is subject to a wide variety of operational constraints. For a utility to effectively use this resource, cooperation and coordination throughout the organization are required. Second, the interplay between trading in wholesale markets and the need to procure and move power from where it is produced to where it is consumed on an hour-by-hour basis requires rapid decision making around the clock. The Optimizer addresses these requirements by deploying an array of innovative operations research techniques.

The Optimizer creates economic benefits in a number of ways. It maximizes the utility's capability to supply renewable integration and its flexibility to respond to unexpected events, and greatly reduces the number of people required to operate this highly complex FCRPS resource.

The Energy Authority

TEA is a nonprofit corporation established in 1997 by three public power utilities—the South Carolina Public Service Authority (Santee Cooper), the Municipal Electric Authority of Georgia (MEAG), and JEA (formerly the Jacksonville Electric Authority)—to actively participate in the deregulated energy marketplace and take advantage of economies of scale. Its 187 employees in Jacksonville, Florida and Bellevue, Washington provide a full range of power-supply management services to eight member utilities and 41 partner utilities in 23 states across the United States.

Background

Federal Columbia River Power System and Bonneville Power Administration

The FCRPS consists of 31 dams on the Columbia River and Snake River in Washington, Oregon, and Idaho. These dams were constructed by the U.S. Army Corps of Engineers and U.S. Bureau of Reclamation between 1909 and 1975 to harness the power and volume of these rivers to support agriculture, energy production,

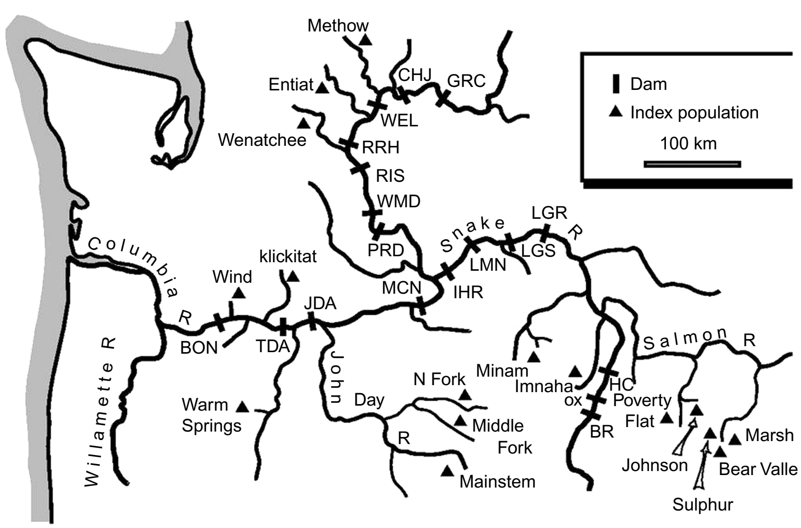

Figure 1: This plot shows the major dams in the FCRPS. The dams discussed in this paper are Grand Coulee (GRC), Chief Joseph (CHJ), McNary (MCN), John Day (JDA), The Dalles (TDA), and Bonneville (BON).

and flood control. Their total generation capacity of 20,460 megawatts represents one-third of 2013's peak summer demand in the region, making the FCRPS the region's predominant source of energy. The Bonneville Power Administration (BPA) was created by the federal government in 1937 to manage the electricity output of the FCRPS. Its primary objective is to supply low-cost power to the public. BPA provides the bulk of this power to community-owned, not-for-profit public utility districts (utilities) in the region.

Columbia River Basin

In addition to power generation, the Columbia River and its tributaries are managed for recreation, irrigation, fish and wildlife habitat, and flood control. The interplay of these interests introduces a wide set of complex, sometimes conflicting, operating requirements for the FCRPS. BPA models these complex requirements as operational constraints on the hydro system. The reservoirs behind each dam form large lakes. Lake elevations must be managed to support recreational needs, including camping, using beaches, boating, swimming, fishing, and engaging in hydroplane races and wind-surfing competitions; however, rapid changes in lake elevations can erode river banks.

In the spring growing season and the hot summer months, water is diverted out of the river to support the multibillion-dollar agricultural industry to irrigate the farms that produce apples, cherries, grapes, potatoes, wheat, and a wide variety of other crops. When this water is diverted, inflows artificially decrease and less water is available for power generation.

The riverine ecosystem provides miles of habitat for numerous fish and wildlife species. Many endangered salmon species, for example, migrate to the Pacific Ocean as juveniles and return as adults to spawn. The FCRPS operates in several specific ways to support this life cycle. During spawning season, flows are established to keep salmon nests at low elevations so they are more likely to stay covered until the eggs hatch. Juvenile salmon are encouraged to migrate with spillwater that falls over the top of a dam rather than passing through generation turbines. However, spill must be carefully limited so dissolved gases are within acceptable bounds and do not kill fish. During spill season, generating units operate within one percent of their peak efficiency to assist the survival of fish swimming through turbines.

Cold and wet weather conditions build winter snowpack in the Cascade and Rocky Mountain ranges and establish the yearly volume of water that flows through the Columbia River Basin. During May, June, and July, the ability to control runoff at storage dams prevents the flooding of downstream cities such as Portland, Oregon. Over the course of the water year, the Army Corps of Engineers determines storage levels at each dam to ensure sufficient capacity to store the runoff when it arrives. Creating too much room may endanger the ability to maintain necessary flows for fish and irrigation requirements in the late summer and dry autumn seasons.

Regional Dialogue Power Sales Agreements

BPA sells the bulk of its power to public utilities in accordance with provisions of the Pacific Northwest Electric Power Planning and Conservation Act of 1980 and the Bonneville Project Act of 1937. These utilities purchase their power through long-term power sales agreements with BPA. The current contracts are known as the regional dialogue power sales agreements; energy deliveries began on October 1, 2011 and extend through September 30, 2028. Beyond 2011, the first contract year, the utility is responsible for meeting any additional load growth and integrating any nonfederal energy resources to meet such loads. There are two regional dialogue contract types.

- The first contract is a full-requirements power sales agreement by which the utility pays a set rate and BPA ensures that the utility's load is supplied with electricity on an hour-by-hour basis (Bonnerville Power Association 2003a).
- The second contract is a slice-block power sales agreement (slice) by which the utility receives a fixed percentage of the actual output of the FCRPS (Bonnerville Power Association 2003b). A slice utility pays only for its slice of the actual FCRPS costs, and receives an equal share of the benefits. Slice guarantees the utility a minimum critical amount of power plus any generation surplus; however, because FCRPS generation varies significantly by season and year, depending on precipitation, snowpack, and weather, the utility also has the responsibility to prudently

manage this flexibility and the corresponding risk that energy and capacity requirements may not be met during any given hourly, daily, or monthly period. Under the slice contract, 16 utilities currently purchase power from BPA; of these, 13 choose to do so via the Optimizer.

Slice Water-Routing Simulator

A slice utility (i.e., a utility that participates in the slice contract) determines its hourly entitled amount of energy using the slice water-routing simulator that BPA built in coordination with all the slice utilities. In this simulator, each utility operates a virtual river, which is a reasonable representation of the FCRPS and is based upon actual river conditions and constraints. This simulated system is composed of six hydraulically linked dams on the Columbia River (see Figure 1). These six dispatchable dams produce up to 85 percent of the total FCRPS generation output.

Each slice utility runs its individual river system independently, and both the historical and future state of its virtual river depends crucially on how that utility chooses to operate the river. Slice utilities route water through the six dams in different ways to make specific generation requests and receive their entitled amount of energy hourly. A slice utility's simulated river does not directly affect how BPA chooses to operate the real river; however, BPA has the obligation to honor the utility's generation requests as long as they are valid. Consequently, the decisions made by all slice utilities have a collective impact on the river's real-time operation.

All requirements that the physical system faces are translated and replicated into 25 constraints types. In addition to these constraints, the system is subject to regulated and unregulated inflows coming into the system, which are forecasted values. To determine the entitled energy, a slice utility submits a simulation request that indicates how it wishes to route water through each of the six dams for the next 10 days without violating a constraint. If it violates a constraint, BPA will not honor the slice utility's energy request, and the utility must make up the difference in the real-time wholesale market at a much higher cost. Stated succinctly, BPA provides the feasibility region as defined by a constraint set, and the slice utility's responsibility is to route the water optimally to suit its needs.

TEA Optimizer Solution

Design Objectives: Speed, Flexibility, and Control

TEA created the Optimizer for the slice utilities to help them (1) address the virtual six-dam dispatchable hydro optimization within defined operating constraints, (2) do so rapidly and robustly, while preserving the full flexibility of the system to respond to changes in variable supply, demand, and electricity prices, and (3) do so in a manner that allows the power traders and schedulers, who are agents trading and scheduling electricity for the utilities, to quickly make optimal decisions. In short, we faced three overarching design criteria: speed, flexibility, and control. Not surprisingly, because these criteria are often contradictory, we need to strike a careful balance between them.

Speed: The overriding requirement for managing the slice contract is to enable the real-time power traders and schedulers to make good decisions within the tight time frames of the wholesale power-trading business. In TEA's case, any single power trader must simultaneously make decisions for as many as five utilities. This led us to the decision to build a globally linearized model with judicious use of mixed-integer programming (MIP).

Flexibility: The second requirement is to preserve as much of the solution space as possible to allow the user to access the full range of generation output. As we mention previously, the ability to adjust generation to respond to changes in both loads and variable supply is of great value in today's energy landscape. A solution that provides fast answers by excluding large parts of the solution space would be a poor trade-off. This requirement led us to apply great care and creativity to the linearization process.

Control: The drawback of a fast linearized solution is that it is deterministic. In reality, supply, demand, hydro conditions, river operations, and the market are all uncertain and vary rapidly. To allow a system to fully automate the dispatch process is not practicable. The solution needs to permit users to account for these uncertainties within strict time windows. They need to be able to manage the solution space such that the Optimizer provides solutions that help users react to constantly changing conditions. This led us to design a system that allows hydro planners, whose role is to ensure all slice utilities operate in compliance with BPA constraints, to (1) modify the solution space with their own constraints, (2) be able to weight those constraints based on the risks involved, and (3) distill the entire complex set of variables to two that are meaningful for a trader—total generation and price.

We know of no other commercially available system that meets all three of these requirements.

Related Work

Other solutions to hydro optimization problems are available (Jacobs et al. 1995, Fleten and Kristoffersen 2008, Mo et al. 2001, Labadie 2004, De Ladurantaye et al. 2009). In particular, the deterministic model described in Hydro-Québec's paper (De Ladurantaye et al. 2009) is similar to the Optimizer model, although it is on a smaller scale (560 megawatts versus 14,000 megawatts) and has a shorter horizon (24 hours versus 240 hours). Moreover, our modeling challenge is more complicated. For example, the Hydro-Québec model's constraints exist for just a single hour and apply only to flow and generation. The constraints facing slice utilities can be formulaic, apply to more variables, such as forebay (the reservoir elevation) and tailwater (the level of the water immediately downstream), and can span multiple hours, a whole day, or even an entire week. The slice model also has more complex nonlinear generation, volume, and tailwater functions.

For slice, the highly complex nonlinearities of the FCRPS constraints, the limited required solve time, obligation to concurrently solve models for 13 unique participants, the contractual requirement to submit 10-day schedules for each run, the validation oversight by BPA for feasibility, and the need to provide robust flexibility to react to real-time market changes present a technological challenge of a greater scale.

Solution Overview

The TEA Optimizer solution consists of load-balanced and high-availability servers that collect the BPA constraints and run optimizations, and a user-facing client application that allows users to set penalties and user-defined constraints. Each optimization server, equipped with Gurobi as the underlying MIP solver, solves the Optimizer model for up to 240 hours within one to three minutes depending on the river conditions and constraint regime. Furthermore, it preserves system flexibility and provides excellent controls for the hydro planners, power schedulers, and electricity traders to drive the system to produce their desired results. To achieve these goals, we developed a novel MIP optimization model to approximate the original slice hydro system and implemented several unique techniques to enhance the solver performance, while maintaining a high degree of accuracy. A typical 10-day model has about 50,000 variables (about 3,000 are binary), 40,000 constraints, and 300,000 nonzeros. Since the system went live on October 3, 2012, it has been running consistently to produce optimized water-routing schedules that address BPA's constraints and observe various user-defined objectives. In so doing, the Optimizer allows our utility clients to fully utilize this highly complex and valuable hydro resource.

The Slice Model

The slice model consists of six dispatchable dams. Figure 2 shows a graphical representation of a single dam. The variables in the figure can be divided into three categories: elevation, water, and electricity (see Table 1). These variables are tightly coupled by both linear and nonlinear equalities, as we describe next.

- Total discharge of each dam is the sum of turbine, spill, and bypass flow.
- Total generation is equal to turbine flow multiplied by the turbine efficiency factor.
- Turbine efficiency factor is a linear function of the head, which is the height difference between forebay and tailwater of a dam.
- Volume of a particular reservoir can be modeled as a nonlinear function of forebay.
- Tailwater is a complex nonlinear function of total discharge and forebay elevation of the downstream dam.

In the slice model, all the dams are linked together both chronologically and geologically. The system state at a given hour depends on the system state and the water-routing decisions made in previous hours. Moreover, all the dams are hydraulically linked together. The inflow of a dam depends on the outflow of its immediate upstream dam a few hours earlier.

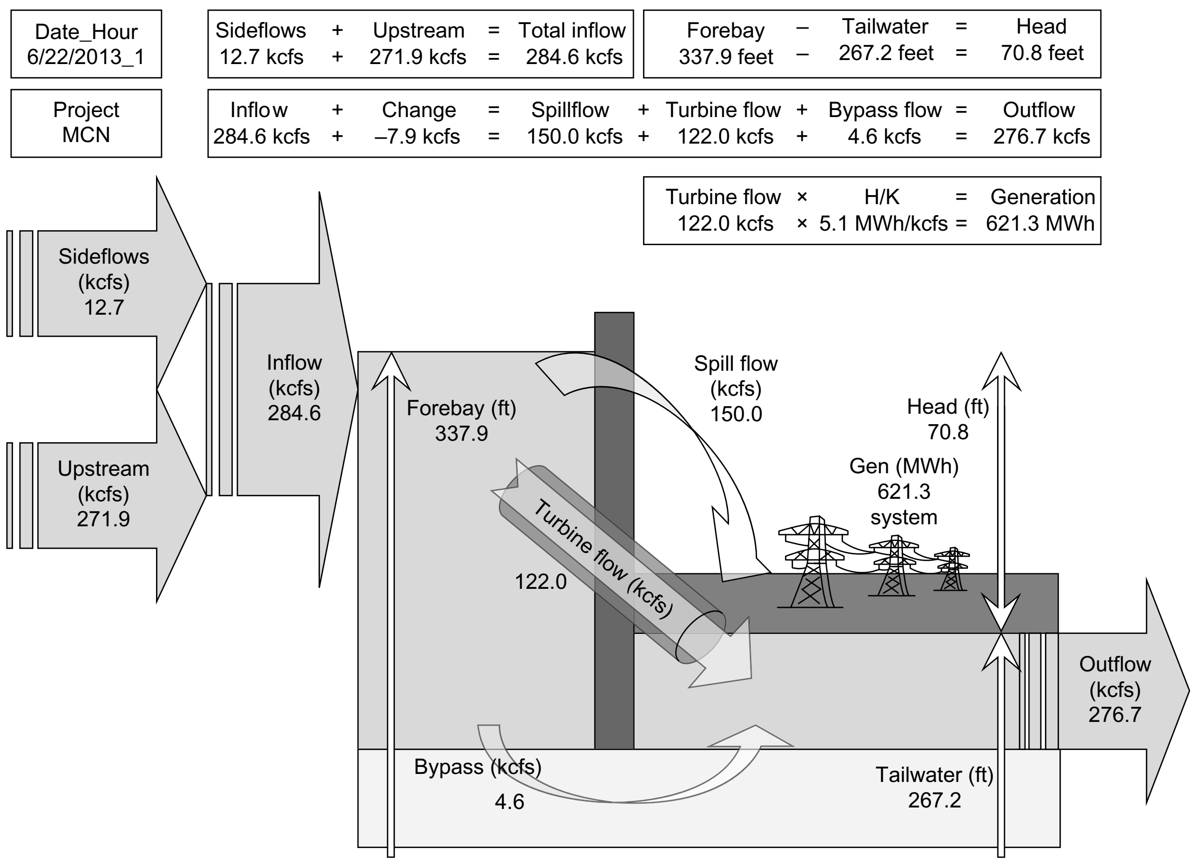

Figure 2: This figure shows a single dam with associated variables and flows at a given hour (Meeker 2013).

At any given hour, the operation of each dam is subject to many operational constraints, which can apply to a hydro variable, its change rate, or its average over a specific period. The constraints may also span multiple hours, days, or even a week.

Dimension Reduction

The challenge to solve the slice model is a result of complex nonlinearities, very limited time frames for solving the model, the requirement to solve it hourly for a 10-day period, while being subject to validation and penalties by a third party (BPA) for infeasibility, and the requirement to provide enough flexibility to react to real-time market changes.

The first step in addressing these challenges is to reduce the model's dimensionality. As we mention previously, the variables in the system are tightly coupled by several complex nonlinear equalities. Maintaining these nonlinear equalities requires introducing a multitude of nonlinear constraints; hence, it greatly slows down the solver.

However, note that the system has only two degrees of freedom: each hour, the utility must determine how much water to discharge downstream and how much water to release through the turbine as opposed to spilling it over the dam. The key decision variables are discharge flow and turbine flow, both of

Category (measure)	Key hydro variable
Elevation (feet)	Forebay, tailwater, head
Water (kilo-cubic feet per second)	Inflow, volume, discharge, spill flow, turbine flow
Electricity (megawatt hour)	Generation, incremental generation reserve, decremental generation reserve

Table 1: This table shows the taxonomy of hydro variables.

which belong to the water category. Consequently, we decided to simplify the model by consolidating all variables into the water category. Because we no longer needed to maintain the complex nonlinear equalities between different variables, the run times improved substantially. For example, a simple two-day prototype model that previously took more than two minutes to solve can now be solved in several seconds; however, dimension reduction also has a cost: this transformation must happen immediately. Whenever BPA adds a new constraint or changes an existing constraint, we must dynamically translate the associated variables into the water-based variables. Although the challenge around nonlinear equalities may be addressed sufficiently, it still produces many nonlinear constraints after the transformation; these must be linearized to enter them into the MIP solver.

Nonlinearities

From a technical perspective, the most challenging operational constraints are those that are nonlinear. The two variables that present the greatest difficulties to solve are the tailwater and the generation.

Tailwater is a complex nonlinear function of the amount of total discharge from the dam and the downstream forebay level; see the graph in Figure 3.

The tailwater is constrained for various reasons. It may be constrained to provide habitat for fish, prevent erosion of river banks, or support navigation and recreational needs. Constraints include tailwater minimum constraints and rate-of-change constraints from hour to hour and over the course of a day, and constraints that have tailwater as a parameter. Given the complexity of the tailwater function and the nonlinear nature of some tailwater constraints, we devoted much work to linearizing these constraints.

For the minimum tailwater constraint, we can imagine transecting the three-dimensional tailwater curve shown in Figure 3 with a horizontal plane at the minimum value of tailwater. Figure 4 shows a plot of the curve that represents the intersection of the constant (minimum) tailwater plane and the tailwater function. The shaded region outside this curve represents feasible regions in the downstream forebay-discharge parameter space. Therefore, if we can dynamically linearize the constraint in terms of downstream forebay and discharge, we can place linear constraints on those variables. We also want to preserve as much of the solution parameter space as possible to help achieve our objective of preserving the flexibility of the generation output.

We employ the standard piecewise linear approximation algorithm as follows. Begin on the leftmost end of the Figure 4 curve and move along the curve to the right, drawing the tangent line at each point. When the distance between the tangent line and the starting point reaches a maximum allowed deviation, add that tangent line to the linear approximation. Then, to determine the stopping point for that tangent

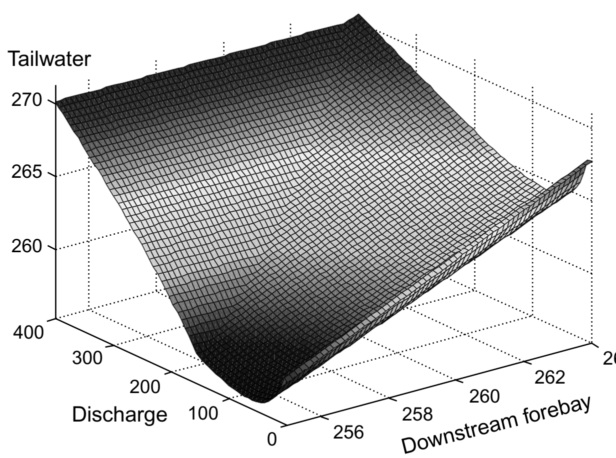

Figure 3: This plot shows the tailwater function of the McNary dam.

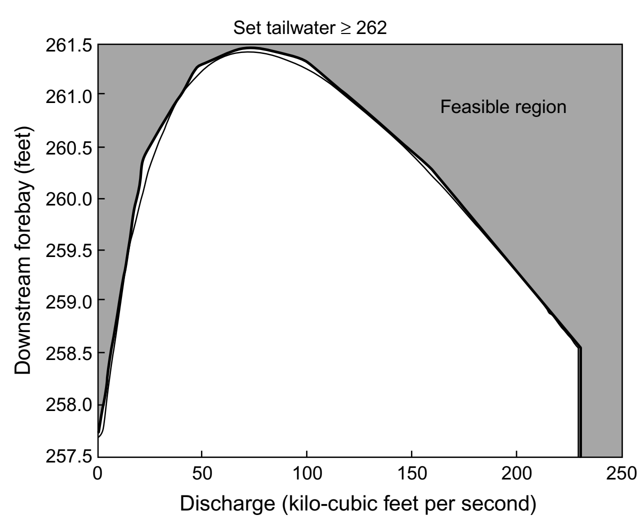

Figure 4: This curve shows the minimum tailwater in the downstream forebay discharge plane with linear approximation.

line in the other (rightward) direction, extend that tangent line to the right until the distance between the tangent line and the curve again reaches the maximum allowed deviation. Repeat this process using the end of one tangent line as the starting point for the next one and add segments until the linear approximation is complete.

Generation is nonlinear because it is a product of turbine flow and the turbine efficiency factor, which is proportional to the head of the reservoir. Therefore, generation is proportional to forebay minus tailwater times turbine flow.

In deciding how to tackle this nonlinearity, we took two approaches that we alternated depending upon the conditions of the water system. In the first approach, we examined historical data and determined that the head typically remains reasonably constant over the 10-day horizon; to be more precise, the 24-hour levels of the head remain fairly constant, because the reservoirs typically fill up throughout the night and empty throughout the day. Therefore, in this approach, we keep the head constant throughout the optimization, but preserve some generation buffer (approximately five percent) within the feasible range to allow for estimation error. Once the optimization is complete and we know the forebays, tailwaters, and turbine flows, we can calculate the correct generation.

This approach generally works well when the range of generation is wide and the system is not overly generation constrained. It also has the virtue of being fast to solve; however, the range of allowed generation is sometimes narrow and can be less than the set-aside buffer to account for the inaccurate efficiency estimation. In some situations, typically during spring runoff when the system is running full bore, the system is highly generation constrained. For these circumstances, we developed the second approach—a more precise approximation to calculate the generation. The approach uses a Taylor series expansion of the generation function around the most recent values of the generation function input variables (see Appendix B). With this approximation, we reduce the generation buffer from approximately five percent to 0.1 percent with little speed trade-off.

Measuring System Flexibility

Throughout the global linearization process, the Optimizer must preserve the solution space as much as possible and measure the system's generation flexibility. The task is harder than one might think; because the dams are all serially coupled to each other, any constraint upon a dam or a constraint at another serially coupled dam can potentially be the binding constraint upon minimum- or maximum-generation output from that dam.

The solution is to add feasible minimum- and maximum-generation variables to the model, in addition to the hydro and generation variables. These represent the minimum or maximum generation allowed for each dam. Each constraint will include one or the other of these variables in the constraint equation. For example, for a maximum turbine flow constraint, which can potentially limit the maximum amount of generation, the feasible maximum generation divided by the turbine efficiency factor must be less than the maximum turbine flow and greater than or equal to the turbine flow. With this approach, we can determine the maximum and minimum generation by the most binding constraints for each dam.

Figure 5 shows the feasible minimum, maximum, and optimal system generation and energy prices. Figure 6 shows the breakdown of the total system generation to individual dams combined and the overall feasible minimum and maximum system generations.

Two major benefits accrue from measuring system flexibility. First, BPA requires slice utilities to give back a certain amount of flexibility, which is specified by imposing minimum incremental and decremental capacity constraints. Therefore, the Optimizer can provide this flexibility in the most cost-effective way. Figure 7 shows how the incremental capacity is allocated among the different dams. The shades of gray represent allocations to individual dams. As the figure shows, it is a fairly complex allocation that changes from hour to hour. It would be exceedingly difficult for a manual solution to allocate the capacity as efficiently. Second, knowing the system's capability allows power traders and schedulers to determine how to manage customer portfolios by calculating the available capacity to absorb real-time wholesale market volatility and the uncertainty of various renewable resources (e.g., wind and solar generations).

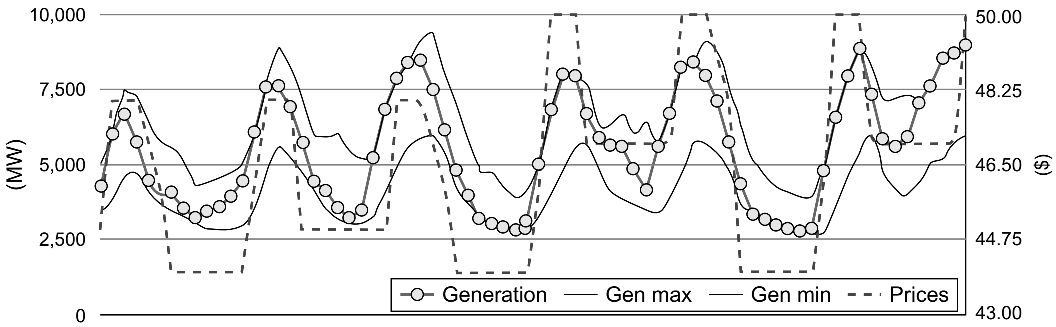

Figure 5: This chart shows the total generation of the six dispatchable dams (circled line), feasible minimum and maximum total generations (solid lines), and prices (dashed line).

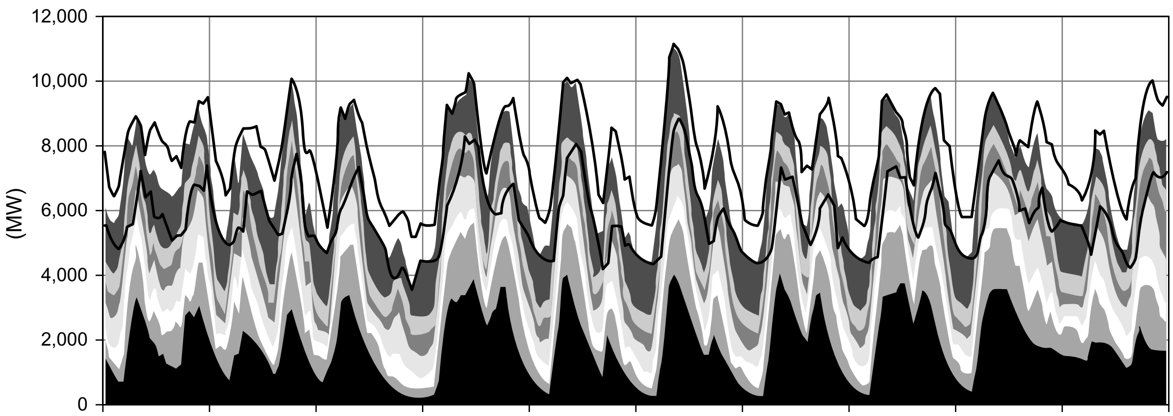

Figure 6: This chart shows the total system generation (in megawatts) with contribution for each dam, and the system-wide feasible minimum and maximum generations (solid lines) over a 10-day period.

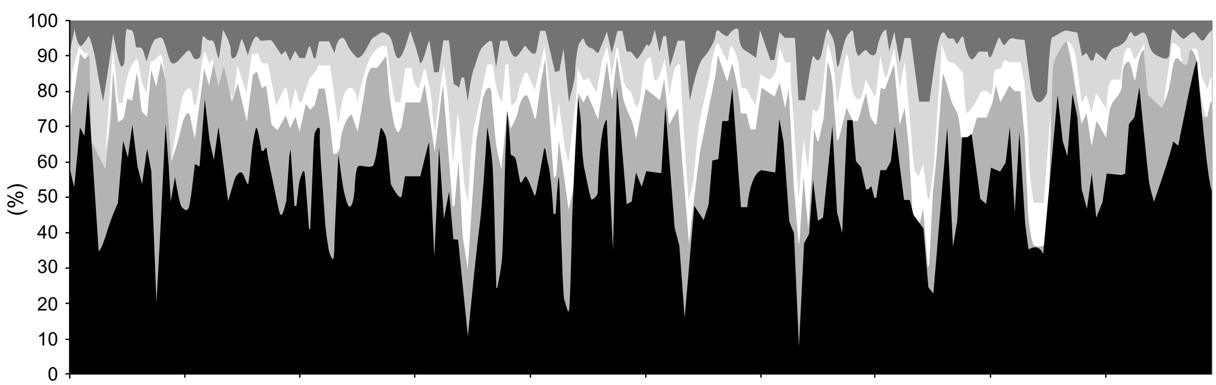

Figure 7: This chart shows the incremental capacity by dam allocated by the Optimizer as a percentage of the total incremental capacity of all six dams over a 10-day period.

Handling Uncertainties

Because the Optimizer is a deterministic tool, hydro planners are tasked with managing the uncertainties that affect the system's operation, particularly inflow uncertainty and unexpected changes to the system constraints. The Optimizer allows planners to accomplish both objectives by adding user-defined constraints to the system.

Planners determine when to draw down or fill the storage reservoirs. They can add their own versions of any type of constraints that BPA imposes. They can also set the levels and the penalties associated with these constraints. No constraint within the Optimizer is a hard constraint; all constraints are implemented with slack variables, which allow any constraint to potentially be violated should the Optimizer find a more optimal (i.e., lower-cost) solution by violating it. Penalties applied to constraints are all denominated in dollars per megawatt hour within the Optimizer. These are the units in which power is priced in the market; for example, by converting all penalties from dollars per feet for elevations, or dollars per unit of flow, to dollars per megawatt hour, planners can set varying penalties in their preferred units of measure for different constraints, and the Optimizer automatically converts these penalty values into dollars per megawatt hour.

TEA planners typically will set guidance constraints to protect the system against unexpected changes in both inflows and BPA constraints. These guidance constraints are intended to be enforced in all cases, unless doing so would force violation of a BPA constraint; the penalties for violating guidance constraints are therefore high, but less than BPA constraint penalties. Figure 8 shows that the forebay curve output from the Optimizer observes both the BPA constraints and TEA guidance constraints.

The Fast Pace of Wholesale Power Trading

Although we designed the Optimizer to preserve as much generation flexibility as possible, power traders and schedulers will ultimately manage the slice system in a fast-paced environment. Therefore, the Optimizer must be easy to use and understand, and must quickly solve the model.

The modern power industry runs on an hourly clock. In regions such as the Northwest that rely on bilateral power markets, the role that humans play is based upon a repeated hourly process that occurs every hour of the year. Within each hourly cycle, the power scheduler for each utility must do the following:

- Determine the expected electricity demand for the next hour (and subsequent hours).
- Determine which generators are available to meet that demand.
- Determine the hourly prices at which the utility buys and sells electricity.
- Decide how much of the utility's own generation to use and how much to buy or sell in the wholesale markets.
- Execute any trades required.
- Schedule the flow of power from generators to loads over particular transmission paths.
- Communicate information required by BPA.

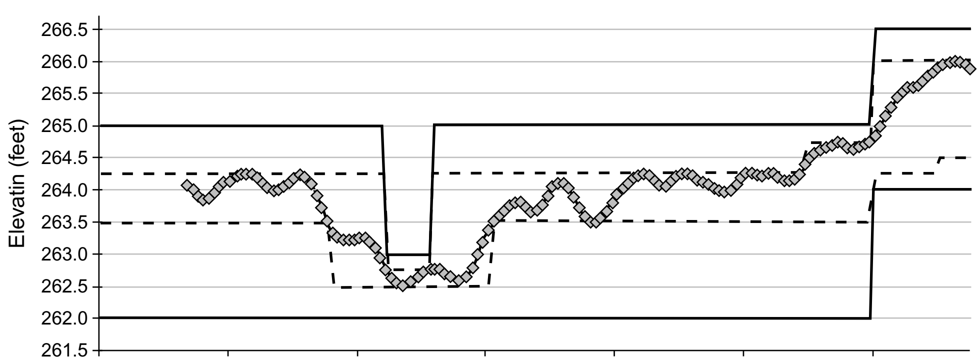

Figure 8: This chart shows the BPA constraint (solid lines), the TEA planner constraint (dashed lines), and the actual forebay elevation (curve in the middle) at the Dallas dam.

If renewable generators are in a utility's portfolio, the trader and (or) power scheduler must also forecast that generation and how much it is changing. Typically, the trader plans the system several hours in advance to avoid having to buy or sell large amounts of electricity at inopportune times.

Power schedulers have a small window of five to 10 minutes in which to make decisions and act on them. During this time, they must determine how to dispatch their slices of the system for the next hour and submit that information to BPA. The slice contract also includes additional requirements that intensify the demands upon the slice utilities within this limited time frame.

The slice contract requires utilities to manage their slice of the six large dispatchable federal dams on the main stem of the Columbia River under the same constraints as those imposed on the actual river. All schedules for hydro and other types of generation must be feasible for each hour that they are submitted to BPA. If they are not, the utility is penalized through forfeiture of the power generated in that hour from one or more of the dams (up to all six), depending on the extent of the violation.

In addition to producing and submitting feasible next-hour schedules to BPA each hour, the utility is required to create and submit at least once a day a fully feasible set of schedules for each hour of the next 10 days. Because of this requirement and the nature of a hydro system (i.e., decisions made in one hour will affect the system in all subsequent hours), the Optimizer solution solves the full system each hour for the next 10 days, and usually several times an hour, for each TEA slice client.

From a process-flow perspective, two innovations facilitate operations within the real-time framework. Automatic runs of the entire workflow are initiated each hour at the beginning of the hour when the updated data are available from BPA. For all 13 utility clients, updated data are automatically downloaded, full 10-day optimizations are run, and the results are uploaded to BPA. This automation creates a feasible set of schedules should the trader choose (or be forced to choose because of contingencies on the trading desk) to accept that solution for that hour.

More typically, the traders will want to adjust the schedule to balance changes in load or wind-generation forecasts or to take advantage of current market conditions. Therefore, they will want to rerun the optimization; however, although a full 10-day optimization is usually performed within two to three minutes, if the trader must do that for five utilities, the time required is a significant portion of the approximately 20-minute window that trader has to finalize the schedules for the next hour.

The solution was to introduce a process called stitching. It allows the traders to adjust generation for the next few hours, which will only affect the hydro dispatch within the first day or two of the optimization period. Therefore, we use the automatically generated 10-day fully feasible solution as a starting point; we then effectively freeze the solution for the final eight or nine days. We allow only the first one or two days to vary, but still force the solution to be fully feasible over the entire 10 days. By doing so, we reduce the optimization time by a factor of 10 or so, to 10–20 seconds.

Model Run Performance

Production Statistics

The Optimizer has been operational since October 2012. Table 2 shows its performance during that period. In more than 120,000 auto runs, we have

Year	Month	Simulation count	Feasibility percentage	10-day opt. (mins:secs)	Stitching count	Stitching opt. (mins:secs)
2012	Oct	6,264	96	1:45	9,460	0:11
2012	Nov	6,489	93	1:54	8,019	0:13
2012	Dec	6,696	93	1:18	7,794	0:12
2013	Jan	6,678	96	0:57	7,210	0:12
2013	Feb	6,191	93	1:34	6,774	0:12
2013	Mar	6,605	98	1:21	7,550	0:12
2013	Apr	6,472	94	3:50	7,159	0:16
2013	May	6,643	91	3:06	4,923	0:18
2013	Jun	6,733	92	3:17	5,756	0:24
2013	Jul	8,527	92	4:10	3,941	0:28
2013	Aug	8,388	98	3:30	5,162	0:22
2013	Sep	9,264	100	2:56	6,449	0:16
2013	Oct	9,668	100	2:42	6,806	0:24
2013	Nov	9,339	99	2:47	7,285	0:28
2013	Dec	9,628	99	3:10	6,461	0:25
2014	Jan	9,604	98	2:51	6,263	0:29
Average		7,699	96	2:34	6,688	0:18

Table 2: This table shows the Optimizer performance since it became operational.

achieved a fully feasible rate of 96 percent and an average solve time of two minutes and 34 seconds. Recall that all constraints have slack variables allowing them to be violated if the violation results in a lower-cost solution; by fully feasible, we mean that all the BPA constraint slack variables have a value of zero; that is, in a fully feasible solution, all the BPA constraints are satisfied as if they were hard constraints. We have also achieved more than 100,000 stitching optimizations with an average solve time of 18 seconds.

Lazy Constraints

To further enhance system performance, we utilize a simple, yet effective, approach. We identify a priori a set of lazy constraints that are less likely to be binding, but we do not initially include them in the model. Instead, we enter them into the solver later through a callback during the branch-and-bound step. The MIP solver rapidly solves the simplified model without the lazy constraints; meanwhile, it can also quickly adjust the solution upon receiving these lazy constraints, because they are usually not binding. The Optimizer can operate in both lazy and nonlazy modes. For a group of 10 benchmarks, the average optimization time on the lazy and nonlazy mode is 64.1 seconds and 219.1 seconds, respectively. Moreover, the run time is much more stable in lazy mode (see Figure 9).

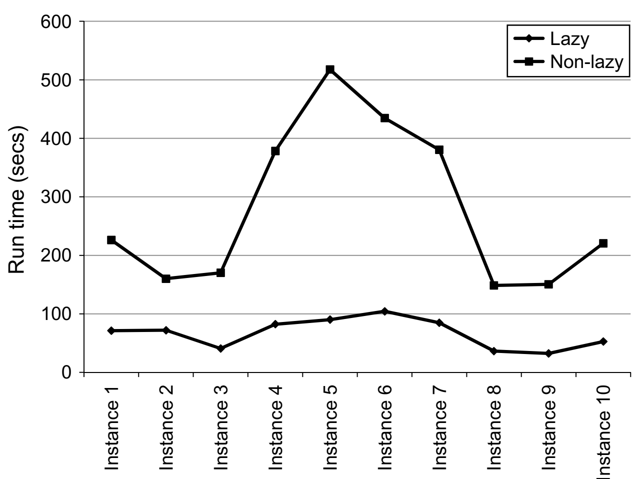

Figure 9: This plot compares the run times of lazy mode (diamond curve) and nonlazy (square curve) mode.

Estimated Economic Benefits

As we mention previously, BPA provides two types of power-delivery contracts for the utilities. Under the BPA full-requirements contract, integration of nonfederal power resources has an additional cost. Under the slice contract, the utility is not obligated to use resource integration, but must invest in the power-supply management infrastructure. Each utility has three options: (1) choose the BPA full-requirements product and pay the additional integration cost; (2) choose the BPA slice product, but implement the slice product manually; or (3) choose the BPA slice product and the TEA Optimizer solution. The benefits of option (3) are its reduced costs versus options (1) or (2) and the additional increased revenue it provides as a result of optimization, which is only available in option (3). We thus look at the following three components of the benefits of option (3).

1. Reduced costs of renewable resource integration, particularly wind, in comparison to option (1).
2. Reduced costs of the power-supply management infrastructure in comparison to option (2).
3. Increased revenue through optimized hydro output for sale in the wholesale power market.

Note that the reduced costs are not additive. We thus calculate the reduced costs individually and take the minimum of benefits (1) and (2) and then add benefit (3) to estimate the total benefits of using the Optimizer.

Reduced Costs of the Power-Supply Management Infrastructure

An individual utility would require a full-time equivalent (FTE) staff of between 14 and 19 people to manually plan and execute a power-supply portfolio using the BPA slice product. It would need at least two shifts of 24-7 staff of five to seven employees per shift; one shift would be dedicated to marketing and scheduling power in the hourly markets, and one shift dedicated solely to manually calculating a feasible hydro routing schedule every hour as per the BPA slice contract. A utility that uses the Optimizer does not require these 14 to 19 people. Instead, TEA employs 46 shared FTE employees and maintains the Optimizer for all 13 utilities. The total cost of a FTE is assumed to be $200,000, and the annual cost for Optimizer licensing and maintenance is $1,000,000.

For each utility, the annual infrastructure cost (without the Optimizer) is between $2,800,000 ($14 \times$ $200,000) and $3,800,000 ($19 \times$ $200,000), and the annual shared infrastructure costs with the Optimizer is $784,615 [$(46 \times \$200{,}000 + \$1{,}000{,}000)/13$]. Thus, the net annual reduced infrastructure cost for the 13 utilities is:

- Low estimate = $(\$2{,}800{,}000 - \$784{,}615) \times 13 =$ $26,200,000 per year;
- High estimate = $(\$3{,}800{,}000 - \$784{,}615) \times 13 =$ $39,200,000 per year.

Reduced Costs of Resource Integration

BPA's full-requirements customers integrate resources through BPA's diurnal flattening service (DFS) and the resource-shaping charge (RSC). For TEA's nine full-service utility clients, the expected annual energy output of the 406-megawatt total installed nonfederal resources is 1,478,338 megawatt hours. The BPA DFS and RSC calculated unit cost is $18.34 per megawatt hour (Bonneville Power Association 2009). We estimate the savings through reduced costs of resource integration as follows.

- Annual resource integration costs per utility = $(\$1{,}478{,}338 \times 18.34)/9 =$ $3,012,524.
- Shared annual slice implementation costs per utility = $784,615.
- Net avoided annual resource integration costs for 13 utilities = $(\$3{,}012{,}524 - \$784{,}615) \times 13 =$ $28,962,817 per year.

Optimization Result: Increased Revenue in Wholesale Power Market

Estimated increased revenues of the Optimizer are based on a prior analytical study that TEA conducted over a defined study period; in this study, TEA compared an optimized and unoptimized water-routing solution. The optimized solution (average price: $38.40 per megawatt hour) is from the Optimizer. The unoptimized solution (average price: $34.88 per megawatt hour) is based on the commonly applied strategy of pass inflows (i.e., setting outflows always equal to inflows). The estimated benefit is about 10 percent of the expected realized value and is then halved to five percent because of the speculative nature of how a utility that does not use the Optimizer would manage the river. The annual estimated revenue increase for the total 13 utilities is as follows.

- Low = 63,072 gigawatt hours × 4.5877 percent × $40 per megawatt hour × 5 percent = $18,402,000 per year.
- Average = 83,220 gigawatt hours × 14.5877 percent × $40 per megawatt hour × 5 percent = $24,280,000 per year.
- High = 91,980 gigawatt hours × 14.5877 percent × $40 per megawatt hour × 5 percent = $26,836,000 per year.

The previous calculations are based on the following: The expected total annual hydro system generation is 63,072 gigawatt hours (low), 83,220 gigawatt hours (average), or 91,980 gigawatt hours (high), depending on the water conditions. The total percentage of all 13 utilities is 14.5877 percent. Assumed electricity price is $40 per megawatt hour based on the Mid-Columbia wholesale power price curve as of November 2013.

Total Combined Estimated Benefits

Finally, taking the minimum of the reduced costs and then adding the estimated increased revenue, we determine the annual total estimated benefits of the Optimizer as follows.

- Low estimate = $\min(\$28{,}962{,}817, \$26{,}200{,}000) +$ $\$18{,}402{,}000 \approx \$45{,}000{,}000$ per year.
- High estimate = $\min(\$28{,}962{,}817, \$39{,}200{,}000) +$ $\$26{,}836{,}000 \approx \$56{,}000{,}000$ per year.

Thus, through the entire 17-year period of the slice contract from October 1, 2011 through September 30, 2028, the total combined estimated benefits of the Optimizer are

- Low estimate = $\$45{,}000{,}000 \times 17 = \$765{,}000{,}000$.
- High estimate = $\$56{,}000{,}000 \times 17 = \$952{,}000{,}000$.

Acknowledgments

We sincerely thank the following public utilities: Benton County Public Utility District, Clark County Public Utilities, Cowlitz County Public Utility District, Emerald People's Utility District, Franklin County Public Utility District, Grays Harbor County Public Utility District, Klickitat County Public Utility District, Lewis County Public Utility District, and Pacific County Public Utility District for investing in TEA to create the Optimizer, and the City of Idaho Falls, Okanogan County Public Utility District, the Eugene Water Electric Board, and Clatskanie People's Utility District for adopting the Optimizer in their power trading. We also thank the Slice Group of the Bonneville Power Administration and Gurobi Optimization for their unparalleled support and assistance.

Appendix A. Objective Function

The total objective function of the Optimizer is the sum of the four terms in Table A.1.

Notation:

- t: hour 0–240.
- d: one of the six dams.
- c: a constraint.
- $sys\_gen(t)$: total system generation at hour t in megawatt hours.
- $load(t)$: total system load at hour t in megawatt hours.
- $buy(t)$: total purchases of electricity at hour t in megawatt hours.
- $sell(t)$: total sales of electricity at hour t in megawatt hours.
- $trans\_cost(t)$: total transaction cost at hour t in dollar per megawatt hour.
- $violation(t, d, c)$: violation amount of constraint c for dam d at hour t (normalized to per megawatt hour).
- $penalty\_violate\_constraint(t, d, c)$: penalty violating unit amount of constraint c for dam d at hour t in dollars per megawatt hour.
- $target(t)$: total user requested generation amount in megawatt hours.
- $penalty\_miss\_target(t)$: penalty of missing target at hour t in dollars per megawatt hour.

Term (in dollars)	Expression
Total revenue of electricity sale	$\sum_t \{sys\_gen(t) \cdot price(t)\}$
Total transaction cost	$-\sum_t \{\lvert sys\_gen(t) - load(t) + buy(t) - sell(t)\rvert \cdot trans\_cost(t)\}$
Total penalty violating constraints	$-\sum_t \sum_d \sum_c \{violation(t, d, c) \cdot penalty\_violate\_constraint(t, d, c)\}$
Total penalty violating generation targets	$-\sum_t \{\lvert sys\_gen(t) - target(t)\rvert \cdot penalty\_miss\_target(t)\}$

Table A.1: This table shows the breakdown of the objective function.

Appendix B. Multivariate Taylor Expansion for Generation Constraints

Let G, Q, H, and η be the simulated generation, turbine flow, water head, and turbine efficiency factor, respectively. In the slice model, $G = Q \cdot \eta = Q \cdot (H \cdot \eta^*/H^*)$, where η^*, H^* stand for the average turbine efficiency factor and average water head in the previous day, respectively. Consequently, transforming a maximum generation constraint $G \leq c$ in terms of turbine flow, we get $Q \leq (c \cdot H^*/\eta^*)/H$. Note that to linearize the aforementioned constraint, it suffices to approximate $1/H$ as $(c \cdot H^*/\eta^*)$, which is a constant.

Note that water head, which is equal to forebay minus tailwater, is a highly complex nonlinear function. Let $H = H(v, u, d)$ be the simulated water head, where v, u, and d stand for volume, volume of the downstream dam, and total discharge, respectively. Moreover, let v^*, u^*, and d^* be appropriate constant values of v, u, and d. Let $x = (v, u, d)$ and $x^* = (v^*, u^*, d^*)$. We apply the following multivariate Taylor expansion to estimate $1/H$:

$$H(x)^{-1} \approx H(x^*)^{-1} - H(x^*)^{-2} \cdot \nabla H(x^*) \cdot (x - x^*),$$

where $\nabla H(x^*) = \left(\frac{\partial H}{\partial v}(x^*), \frac{\partial H}{\partial u}(x^*), \frac{\partial H}{\partial d}(x^*)\right)$. The high-order terms are discarded to ensure the linearity of the approximation. Applying this approximation, we can employ a generation buffer as low as 0.1 percent as opposed to five percent if we assume $\eta = \eta^*$. This helps to preserve the generation flexibility of the hydro system.

References

Bonneville Power Association (2003a) Conformed load following template. Accessed January 4, 2014, http://www.bpa.gov/power/pl/regionaldialogue/implementation/d%ocuments/docs/Conformed_LF_Template_12_13_13.docx.

Bonneville Power Association (2003b) Conformed slice/block contract template. Accessed January 4, 2014, http://www.bpa.gov/power/pl/regionaldialogue/implementation/documents/docs/Conformed_Slice-Block_Template_12_13_13.docx.

Bonneville Power Association (2009) Illustrative example for DFS for a wind resource in power PUD's contract—includes transmission steps. Accessed January 4, 2014, http://www.bpa.gov/power/pl/regionaldialogue/implementation/previous-meeting-materials/documents/2009/2009-06-22_DFSand-TransmisionExample.pdf.

De Ladurantaye D, Gendreau M, Potvin J-Y (2009) Optimizing profits from hydroelectricity production. *Comp. Oper. Res.* 36(2):499–529.

Fleten SE, Kristoffersen TK (2008) Short-term hydropower production planning by stochastic programming. *Comp. Oper. Res.* 35(8):2656–2671.

Jacobs J, Freeman G, Grygier J, Morton D, Schultz G, Staschus K, Stedinger J (1995) Socrates: A system for scheduling hydroelectric generation under uncertainty. *Ann. Oper. Res.* 59(1):99–133.

Labadie JW (2004) Optimal operation of multireservoir systems: State-of-the-art review. *J. Water Resource Planning Management* 130(2):93–111.

Meeker C (2013) Slice water routing simulator. *Bonneville Power Administration* (U.S. Department of Energy, Washington, DC).

Mo B, Gjelsvik A, Grundt A (2001) Integrated risk management of hydro power scheduling and contract management. *IEEE Trans. Power Systems* 16(2):216–221.

Zengjian Hu is the chief architect of the Slice optimization engine. He is responsible for mathematical modeling, system design, and implementation. Prior to joining The Energy Authority, Zengjian has held several industry research and development positions and has published over a dozen papers for major science conferences and journals. Zengjian received his PhD in computer science from Simon Fraser University of Canada and a master's degree in computer science from the Institute of Computing Technology at the Chinese Academy of Sciences in Beijing.

John Putz is a 14-year veteran of the energy, energy software, and trading industries. John has experience in a wide variety of areas including risk management, trading, energy software, asset and portfolio optimization, trade and asset valuation, and hydrogeneration optimization. John received his PhD in experimental particle physics from the University of Washington in 2000.

Yohan Sutjandra is responsible for fundamental modeling of the electricity markets. He has worked on a variety of projects, including the application of utility theory to evaluate risk-averse strategies, development of a prototype to the Slice Optimizer, and developing FTR bidding strategies. Before his career at The Energy Authority, Sutjandra worked as a software developer for Cadence Design Systems, an electronic design automation company in California. He earned his master's and bachelor's degrees in electrical engineering from the University of Texas at Austin.

Abel Chan is responsible for leading all software development and business intelligence initiatives for The Energy Authority. Additionally, he manages software engineering services related to strategic planning, development, deployment, and maintenance of applications. Before joining The Energy Authority, Chan was a software developer for Power Resource Managers, Inc., and BSquare Corporation. Chan earned a master's degree in civil engineering and a bachelor's degree in computer engineering from the University of Washington.

Edmond Mount is responsible for delivering power supply planning and operations services to The Energy Authority's clients in the Pacific Northwest. He also served as the overall program manager for The Energy Authority's Slice Application Development Project. Mount has more than 20 years of experience in the Pacific Northwest electric utility industry, spanning activities in resource planning and management, trading, risk management, and contract negotiations. He has a master's degree in economics from the University of Washington, and a bachelor's degree in economics from Whitman College.

Kristen Baker is responsible for performing the daily operational functions required for the delivery of power supply planning and operations services to The Energy Authority's clients in Pacific Northwest. She is the primary user and tester for the Slice Optimization Tool. Baker joined The Energy Authority in 2006 as an applications systems specialist and programmer, designing and maintaining load forecast models for residential, commercial, industrial, and irrigation load types. She earned a bachelor's degree in computer science from The University of Puget Sound in Tacoma, Washington.