PGE

Portland General Electric Company
121 SW Salmon Street • Portland, Oregon 97204
PortlandGeneral.com

August 31, 2007

US Postal Mail

Oregon Public Utility Commission
Attention: Filing Center
550 Capitol Street, N.E., Ste 215
Salem, OR 97301-2551

RE: **Report on the Feasibility of Using Stochastic Modeling in the Annual Update**

Ms. Bailey-Goggins:

In its UE 180 Order, the Commission included a discussion of the potential benefits of stochastic power cost modeling and directed PGE to "submit a report on the feasibility of using stochastic modeling in the Annual Update by September 1, 2007." (Order No. 07-015, Page 12) The requested report is attached. It discusses work done to date and issues related to further development of a stochastic modeling approach. It also requests Commission direction on further work and the incurrence of incremental costs.

If you have questions on the report, please contact me at (503) 464-7580.

Sincerely,

Patrick G. Hager
Manager, Regulatory Affairs

cc: Ed Durrenberger – OPUC
UE 180 Service List

Encls:

g:\ratecase\opuc\projects\report_08.31.07\cover letter.doc

REPORT TO OPUC ON THE FEASIBILITY OF STOCHASTIC POWER COST MODELING

I. Introduction:

In PGE's last general rate case, Docket UE 180, parties expressed a wide range of views on the merits of stochastic modeling for net variable power costs (NVPC). Given this discussion, the Commission stated that "we urge PGE to develop stochastic modeling to develop its NVPC forecast. PGE should submit a report on the feasibility of using stochastic modeling in the Annual Update by September 1, 2007." (Order No. 07-015 at 12) This report is PGE's submittal pursuant to the Commission's request.

The report is organized as follows:

- Section II discusses the history of PGE's investigation of stochastic power cost modeling prior to Order No. 07-015.
- Section III summarizes the work PGE has done since Order No. 07-015.
- Section IV discusses issues related to the implementation of stochastic power cost modeling.
- Section V notes the Commission direction that PGE would need to proceed with implementation of stochastic power cost modeling in its Annual Update process.

II. History Prior to Order No. 07-015:

In Docket UE 165, PGE and Staff signed a Stipulation supporting a hydro-specific power cost adjustment mechanism (PCAM). Although the Commission ruled against this particular PCAM proposal, PGE decided to pursue part of the Stipulation, specifically Section 12, which reads:

> 12. PGE agrees to obtain appropriate consultation services for the purpose of evaluating the statistical distribution of net power costs, at a cost of up to $100,000. The analysis will consider the volatility of hydro generation, electricity prices, natural gas prices, system load, forced outages, and any correlations between these variables. Staff and PGE will work together to formulate a work statement to guide the work of the consultant. PGE will schedule quarterly public workshops to provide progress reports and receive input from interested parties. Staff and PGE reserve the ability to accept or reject the opinion or work product of the consultant for use in ratemaking, including in PGE's next general rate case. The consultant will report results by December 31, 2005, unless Staff and PGE agree to a different date. PGE will not seek recovery of the cost of these consultation services from customers.

Through a targeted Request for Proposal process, PGE selected the PA Consulting Group (PA) to develop and report on preliminary stochastic modeling of PGE's NVPC. We also engaged Marty Howard, a local consultant, to act as a project manager and interface between PGE and PA. In addition to overall management of PA's work, Mr. Howard worked with PA regarding the provision of data and modeling algorithms and the discussion of statistical modeling issues. PGE spent more than $260,000 (incremental costs) on this project.

PA visited with PGE personnel twice and produced a report, "Portland General Electric, Hourly Power Cost Simulation," which was dated July 10, 2006. This report became PGE Exhibit 1803 in docket UE 180. It is also Attachment A to this report. PA developed statistical models with cross-correlations for loads, hydro output, gas prices, electric prices, and plant outages. These stochastic inputs were then used to run 1,000 simulations, each simulation producing an annual PGE NVPC figure. PA's modeling construct produced base total NVPC that were almost the same as those from a Monet run with base inputs. However, there was significant variation between some major components.

PA analyzed the simulation results and provided preliminary answers to two primary questions:

1. Are the mean, or average, NVPC from the simulation study higher or lower than a deterministic Monet forecast based on expected loads, hydro output, gas and electric prices, and plant outages?

2. How might the distribution of simulation outcomes inform the structure of a PCAM mechanism, particularly the size and (a)symmetry of a possible deadband?

In its final report, PA presented a number of preliminary conclusions, based on PGE's plants at that time and the data set used for stochastic input development. These conclusions included:

- The mean NVPC in the simulation study were approximately $10 million greater than in the base Monet forecast.

- The standard deviation of NVPC was at least $55 million.

- Given the large size of the standard deviation relative to the mean vs. base differential, it would be "very difficult to use the numerical results of this prototype for ratemaking or to determine a 'risk adder.'"

- NVPC can increase much more than they can decrease (from the mean).

III. Work Performed After Order No. 07-015:

After the Commission issued Order No. 07-015, PGE consulted with PA to determine the scope of work that would be required to implement stochastic power cost modeling for rate making purposes.

As a first step, PGE requested that PA produce a formal statement of the modeling performed and discussed in the report, "Portland General Electric, Hourly Power Cost Simulation." This statement included two primary elements:

- Formal specifications of how stochastic inputs – loads, hydro output, gas and electric prices, and plant outages – were modeled.

- Formal specification of the mathematical problem – how to meet load requirements at the lowest cost – solved in each iteration of the simulation study.

PA discussed the formal problem statement and noted that one disadvantage of the preliminary modeling completed for the July 2006 report was that the calibration with Monet was not exact. PA suggested that the actual implementation of stochastic power cost modeling for rate making purposes should directly use the Monet model. Each iteration of a simulation study would then be a Monet run with a set of stochastic input variables. The Monet runs per se would then differ from those currently used for rate making only in that the input variables would be stochastic and the dispatch logic for PGE's Coyote Springs and Port Westward plants would be simplified to decrease run time. PGE and PA agreed that the modeling structures used in the preliminary work did not need major revision, although recalibration with new data would be necessary. PGE incurred approximately $15,000 of incremental costs (from PA and Marty Howard) to develop this post-Order No. 07-015 scope of work.

PA then submitted a formal proposal to attempt to implement stochastic power cost modeling. Under this proposal, PA would complete the following five tasks:

1. Recalibrate and revise input models. (Note that no major revisions are foreseen.)

2. Simulate input variables. This would include running several iterations of the models completed in Step 1 and then collecting the results.

3. Develop a Monet "wrapper." This software structure would use the results of Step 2 to run the simulated input data through Monet and keep track of results, iteration by iteration.

4. Report and analyze Monet results. The results from Step 3 would be transferred to an easy to access and analyze format (specifically an Excel add-on called "@Risk"). This would facilitate statistical, graphical, and other forms of analysis.

5. Help write a report on overall simulation results. PA and PGE would jointly complete this step.

PGE would again need to retain Marty Howard as a project manager. We estimate that the total incremental costs for PA and Marty Howard would be approximately $300,000. As noted above, PA has submitted a formal proposal, which is included as Attachment B. However, we believe that this proposal likely underestimates the amount of work that we would need from PA and others to fully implement stochastic power cost modeling. Attachment B is confidential and subject to the protective order in Docket UE 180 (Order No. 06-111). In addition to incremental costs for PA and Marty Howard, PGE would have to allocate significant labor time to the project.

IV. Issues Related to Further Work:

If PGE proceeded with the stochastic power cost modeling work described in Attachment B, several issues would need to be addressed. These include the following:

- Parties would need to agree on an appropriate data set for use in (re)calibrating the stochastic input models. This data set selection could be controversial, as it would reflect various views as to which past data are most reflective of a future test year. For example, should the data set include the West Coast energy crisis period and, if so, how much weight should it be given?

- Parties would need to agree to an appropriate interface between the data used to (re)calibrate the stochastic input models and data more specific to the test year. The latter includes the most recent forward curves and gas and electric positions already taken.

- A simulation study based on one calibration of the stochastic input models would produce a distribution of NVPC. If the input models were stable, we theoretically would expect that, say 100 years of actual NVPC results would roughly follow the theoretical distribution. However, the input models are <u>not</u> stable; they would require frequent recalibration as underlying real world conditions change. Therefore, we would only experience a very few outcomes of a theoretical distribution before that distribution itself would change, due to changes in the underlying stochastic input models. This also refutes the idea that "results will be fair over time."

- One possible use of stochastic power cost modeling is to inform the process of setting parameters of a PCAM. However, the Commission has already set the parameters of PGE's PCAM in Order No. 07-015.

- The increased modeling complexity would raise a number of issues which include:

    - Longer run times, which could require the use of multiple computers over several hours or overnight.

    - Decreased ability of other parties to run the model themselves.

- There would be a need to maintain two models, one with the simplified dispatch logic for Coyote Springs and Port Westward. The simplified logic in the version used for stochastic modeling would have less precision.

- All of the above issues would make stochastic power cost modeling controversial. It would complicate and be a source of conflict among parties in PGE's Annual Update Tariff proceedings.

V. Commission Guidance:

PGE is interested in continuing to study the potential of stochastic power cost modeling. However, as discussed in Sections III and IV of this report, there are a number of significant issues outstanding. A more definitive study could take considerable time and a final approach agreed to by parties for PGE's rate cases would likely not be ready for use in estimating 2009 net variable power costs. Also, given their complexity, we might not reach consensus on all outstanding issues. If PGE continued to develop stochastic power cost modeling, we would first hold a workshop to discuss with parties the modeling approach described in Section III. We would then work with PA to complete that approach, as potentially modified by input from parties at the workshop.

PGE has incurred incremental costs of approximately $275,000 on stochastic power cost modeling to date. If the Commission agrees that we should continue development of this modeling approach, PGE requests an order allowing us to defer the associated future incremental costs.

g:\ratecase\opuc\projects\report_08.31.07\finalreporttoopuc.doc

ATTACHMENT A

PA CONSULTING JULY 10, 2006 REPORT

ATTACHMENT B

PA CONSULTING PROPOSAL FOR FURTHER MODELING

Confidential – Subject to Protective Order No. 06-111

Portland General Electric

Hourly Power Cost Simulation

July 10, 2006

Portland General Electric

Hourly Power Cost Simulation

July 10, 2006

Prepared by: Jonathan Jacobs

PA Consulting Group
390 Interlocken Crescent
Suite 410
Broomfield
Colorado 80021
Tel: +1 720 566 9920
Fax: +1 720 566 9680
www.paconsulting.com

Version: 1.0

TABLE OF CONTENTS

1. Introduction **1-1**

2. Underlying assumptions **2-2**
2.1 Nature of the solution sought 2-2
2.2 14-month time horizon 2-2
2.3 No uncertainty about retirements or new capacity 2-2
2.4 Spot markets are fully liquid 2-2
2.5 Decision vs. valuation timeframes 2-3

3. Data issues **3-4**
3.1 Data by component 3-4
3.2 Summary of data requirements 3-8
3.3 Estimation issues for uncertainty data 3-14

4. Model structure **4-21**

5. Input components – Prototype **5-23**
5.1 Load (and temperature) 5-23
5.2 Hydro availability 5-26
5.3 Gas forward prices 5-26
5.4 Gas spot prices 5-27
5.5 Power spot and forward prices 5-29

6. Resource models – Prototype **6-33**
6.1 Load 6-33
6.2 Beaver power plant 6-34
6.3 Coyote power plant 6-34
6.4 Boardman power plant 6-35
6.5 Colstrip power plant 6-35
6.6 Mid-C hydro contracts 6-35
6.7 Portland General hydro resources 6-36
6.8 Forward hedges 6-36

7. Prototype results **7-38**

8. Summary and Conclusion **8-45**

Appendices

APPENDIX A: Components of Prototype Cost Simulation Model A-1

1. INTRODUCTION

PA Consulting Group has been retained by Portland General Electric to define a "cost simulation model". The basic simulation model would simulate the net variable power cost over a period of time subject to certain assumptions about loads, market prices, hedges in place and hydro conditions. That model could then be run over a large sample of potential realizations of those assumptions in order to estimate the statistical properties of the distribution of net variable power costs.

In the course of PA's work on this assignment it became clear that an important factor limiting the precision of any probabilistic cost simulation is the availability of data describing the distributions and dependencies of its uncertain inputs. PA produced a report explaining the level of detail at which data might be desired, the difficulties in assembling such data, and stopgaps or proxies that might be used on an interim basis. The recommendation from that report was that PGE proceed with the definition of a flexible prototype simulation structure, which could be used to test different data relationships and resource modeling choices. PA has described this flexible structure as a "sandbox".

This document is the Final Report from PA's assignment. It is organized as follows:

- List of major assumptions that define the problem being modeled
- Significant sections of the Data Issues Report
- Overall structure of the prototype model
- Input data modeling in the prototype. This refers to the representation of variables such as load, gas price, power price and hydro conditions, and the relationships between them.
- Resource modeling in the prototype. A "resource" is any source or sink of the power distributed by PGE, or any hedge on PGE's net variable power cost.

2. UNDERLYING ASSUMPTIONS

2.1 NATURE OF THE SOLUTION SOUGHT

PGE was seeking a simulation model that would produce information about the shape of the distribution of net variable power costs. The "location" of that distribution is represented by the Resource Valuation Mechanism (RVM) forecast, that is, a base case Monet run.[1] Therefore the function of the probabilistic simulation is to perturb the Monet input parameters in a way that is consistent with historic distributions of those parameters, and report the shape of the ensuing distribution of net variable power costs (as well as any shifts in the mean).

From this assignment, PGE was seeking a description of a model that it could easily implement if need be. It might be desirable to incorporate the model into Monet but at the very least it should be compatible with Monet. PA interpreted this to mean a simple Excel-based model, which could use, when appropriate, the same logic as Monet (even reusing code from Monet or DLLs it calls).

2.2 14-MONTH TIME HORIZON

The question addressed by this simulation model is the extent by which Portland General's actual costs for a year can differ from the cost forecast used for setting the revenue requirement in the RVM. That revenue requirement is set in November and covers the following year. Therefore the model has a 14-month time horizon, from Nov. 1 through Dec. 31 of the following year (costs during the first two months are not accumulated).

2.3 NO UNCERTAINTY ABOUT RETIREMENTS OR NEW CAPACITY

Because the model horizon is only about a year and a quarter, we assume that PGE's resource base is known. It may change during that period – resources may be retired or added – but the schedule of retirements and additions during the model horizon is known with certainty. Therefore, there is no need for estimation of the parameters of new candidate resources or for capacity expansion decision-making. Furthermore, the technological parameters (heat rates, capacities, outage rates) of those resources are already known. "Disruptive technological change", for example, would not affect the intra-year cost uncertainty, but rather would move the entire distribution of costs including the RVM forecast.

2.4 SPOT MARKETS ARE FULLY LIQUID

The first consequence of this assumption is that each resource may be modeled as if it were dispatched based solely on the price, rather than to meet load. PGE's Monet model accounts for load-following, to the extent that it refers to shaping supply to loads that change within the hour but achieve an anticipated average value, by reserving Mid-C hydro capacity and

[1] The specific Monet run to which the prototype was constrained was based on the file M606PUC05-105-06.xls that was provided to us.

assuming the net energy impact is nil. If the anticipated average value is not achieved it represents a real-time load excursion described in the next paragraph.

The second consequence is that all real-time load excursions beyond those anticipated as "load following" – in other words, all differences between the real-time integrated load for an hour and the day-ahead forecast – can be covered in the spot market. It is not necessary to reserve hydro capability to cover those excursions, although available hydro capability could be used to respond to real-time price fluctuations. It is explicitly assumed that there will always be sufficient spot liquidity to meet all demands or sink excess supply at the spot price (even though that price may be very high).

2.5 DECISION VS. VALUATION TIMEFRAMES

Because the real-time spot market is assumed to be perfectly liquid, and because any net imbalances are settled in real time, we assume that all energy – generated or purchased by PGE, or delivered to loads – is valued at the real-time spot price. However, most resources need to be scheduled in advance. Therefore "advance" prices, e.g., day-ahead prices, have to be used for dispatch decision-making. In some cases, a basic schedule may have to be set based on the advance prices subject to a limited amount of flexibility to respond to real time spot prices.

3. *DATA ISSUES*

3.1 DATA BY COMPONENT

In this section we will review the different components we expect the simulation model to contain, and list the data required by each one. As noted above, we are assuming that PGE's power dispatch decomposes, so that energy resources are scheduled independently. The value of the energy from each resource is computed based on spot power prices; the net value of the resource is its energy value minus the generating cost (fuel and O&M). Similarly the cost to serve load is computed based on spot prices. The net variable power cost is the difference between the cost to serve load and the total net value of resources. Each different resource type is a component of the model.

3.1.1 Load

The load component determines the cost to serve load. It requires two different variables, namely hourly load and the real-time hourly spot price, both of which are uncertain. Forecasting hourly load may require a "hidden variable" such as temperature. Given that the simulation model is completely decomposed and that we assume the real-time market is completely liquid no "load forecast" variables are necessary except for determining the reserve requirement (see the description of the hydro component below).

3.1.2 Gas-fired plants

The gas-fired power plant component includes PGE's gas-fired power plants – Coyote, Beaver and Port Westward (when operational). These plants are scheduled based on the gas price and an advanced power price, and "paid" based on spot power prices. In addition, the Beaver plant may have some ability to modify its schedule to respond to real-time or near-real-time price movements. We understand that there is no regime of tradable NO_x permits in the Northwest (when one is implemented the model would have to be modified to account for NO_x costs, which are likely to covary with load and gas prices).

A. DATA REQUIREMENTS – CERTAIN

The gas-fired plant component requires the following data for each plant:

- Definition of operating states. These can be used to represent dynamic constraints on startup and shutdown, ramping, or a non-constant heat rate.
- State transition matrix. Note that the state detail may be ignored or approximated for this simulation model, relative to the detail in Monet.
- Capacities
- Measure of flexibility for Beaver, which is the one plant whose dispatch apparently can respond in near-real-time
- Heat rates
- Outage rates (may vary by operating state)

- O&M cost
- Gas tariffs
- Maintenance schedules

B. DATA REQUIREMENTS – UNCERTAIN

The gas-fired plant component requires the following uncertainty data:

- Gas prices at Sumas and Stanfield. AECO prices, Rockies (e.g., Opal) prices or even Henry Hub prices could be used as “hidden” variables
- Day-ahead hourly Mid-C power prices, for use in dispatching
- Mid-C power prices with a timing appropriate to Beaver’s scheduling flexibility (probably 4 hours ahead of real time)
- Real-time Mid-C power prices

3.1.3 Coal-fired plants

The coal-fired power plant component includes PGE’s coal-fired power plants – Boardman and its share of Colstrip. These plants are scheduled based on the coal price and an advanced power price, and “paid” based on spot power prices. Although coal-fired plants incur costs associated with sulfur emissions we understand that there is no regime of tradable NO_x permits in the Northwest (when one is implemented the model would have to be modified to account for NO_x costs, which are likely to covary with load and gas prices).

A. DATA REQUIREMENTS – CERTAIN

The coal-fired plant component requires the following data for each plant:

- Definition of operating states. These can be used to represent dynamic constraints on startup and shutdown, ramping, or a non-constant heat rate. Monet actually does not define operating states for these plants, which assumes that they will run pretty much baseloaded and satisfy all operating constraints.
- State transition matrix.
- Capacities
- Heat rates
- Outage rates (may vary by operating state), which includes transmission outages
- O&M cost
- Coal prices – Colstrip is a mine-mouth plant so its coal cost is likely to be known; Boardman currently has a long-term coal contract and it is reasonable to expect that even after that contract runs out it will be supplied on contracts of at least annual duration (so the coal price is not uncertain relative to this model’s 14-month horizon)
- SO_2 emissions rates

- SO_2 allowance prices – we assume that SO_2 allowance prices are certain because the sulfur market is national and associated with baseload generation
- Maintenance schedules
- Transmission costs

B. DATA REQUIREMENTS – UNCERTAIN

The coal-fired plant component requires the following uncertainty data:

- Day-ahead hourly Mid-C power prices, for use in dispatching
- Real-time Mid-C power prices

3.1.4 Hydro plants

The hydro power plant component includes PGE's hydro plants and Mid-C contract. These plants are scheduled with at most intra-month flexibility, that is, the total energy for each month is known at the scheduling time horizon (although it is uncertain at the 14-month horizon).

A. DATA REQUIREMENTS – CERTAIN

The hydro plant component requires the following data for each plant:

- Capacity (for some plants this will be uncertain, that is, dependent on hydro conditions)
- Overmonth storage, that is, the amount of that may be carried from month to month
- Amount of capacity that must be withheld for load-following or reserves (may be expressed relative to load or total fossil generation)
- Amount of capacity that must be scheduled to allow for downward load-following or downward regulation (again, may be expressed relative to load)
- Outage rates (may be 0)
- O&M cost
- Maintenance schedules
- Measure of rescheduling flexibility, which represents the ability to change the schedule to respond to day-ahead or real-time prices (and the way in which changes in energy usage are redistributed within the rest of the month)

B. DATA REQUIREMENTS – UNCERTAIN

The hydro plant component requires the following uncertainty data:

- Monthly energy available by plant.
- Monthly capacity by plant, for those plants whose capacity varies with hydro conditions

- Monthly run-of-river (minimum dispatch level) energy by plant
- Monthly forecasted hourly Mid-C power prices, for use in dispatching
- Day-ahead Mid-C power prices, for rescheduling
- Mid-C power prices with a timing appropriate to Beaver's scheduling flexibility (probably 4 hours ahead of real time), for rescheduling
- Real-time Mid-C power prices

The final model design may not include all the layers of rescheduling described above.

Note that "hydro condition" or "hydro availability" may be a hidden variable that influences energy, capacity and must-run energy. The Monet model that PGE currently uses for revenue requirement forecasting assumes that all hydro units other than Mid-C are inflexible – the hourly dispatch for every hour is fixed relative to the annual average. The Mid-C contracts are assumed to be flexible within the month but not month-to-month. Essentially, over-month storage is disregarded because energy has already been allocated to each month by the NWPP hydro model.

The NWPP model may use rule curves but does not explicitly account for conditional hydro probabilities. An example of a conditional hydro probability would be the probability that inflows in February would be consistent with historical hydro year 1955 given that January was consistent with historical hydro year 1937. Assessing non-uniform conditional probabilities is a difficult task and would be facilitated by a historical forecast database. We believe that the NWPP model deterministically applies the conditions of a specific historical hydro year in each month but have not yet seen model documentation.

3.1.5 Hedges and term power purchases and sales

This component describes PGE's hedges and term power contracts, in place as of November for the following year and modified as improved load and hydro forecasts become available. Forward power contracts don't really affect the distribution of net variable power costs except for moving its mean, but option contracts will affect it. We have not examined PGE's hedge book but we have been led to believe that it is dominantly forwards, fixed-for-float swaps and "vanilla" options priced either at Mid-C or COB.

A. DATA REQUIREMENTS – CERTAIN

The term transaction component requires the following data:

- List of hedges and transactions
- Transaction volumes by month
- Fixed prices for transactions, by month
- Identification of appropriate price indexes for transactions, by month

B. DATA REQUIREMENTS – UNCERTAIN

The hydro plant component requires the following uncertainty data:

- Monthly peak and offpeak average power prices, at Mid-C and COB
- Daily peak and offpeak power prices, at Mid-C and COB

3.1.6 Spot price model

The spot price is not a model component in the same sense as the others. We have mentioned it here merely in order to be able to list some additional "hidden variables" that may be used in forecasting the spot price:

- Total Northwest hydro availability (as opposed to PGE hydro availability)
- Total Northwest regional load
- California load-resource balance
- California intertie capacity

3.2 SUMMARY OF DATA REQUIREMENTS

3.2.1 Certain data (Technological coefficients)

The following table lists the "certain data" identified above. We also label these "technological coefficients" since for the most part they are measurable properties of the technologies implemented in PGE's resource portfolio. For some of these we have initial or stopgap data sources (to be used for the prototype model), and the second column of the table identifies them. Blanks in the second column indicate that a source of initial data has not yet been identified.

Table 1. "Certain" data and preliminary sources

Data requirement	Source
Definition of operating states, Beaver/Coyote	Current Monet spreadsheet
Definition of operating states, Port Westward	
Definition of operating states, coal plants	
State transition matrix, Beaver/Coyote	Current Monet spreadsheet
State transition matrix, Port Westward	
State transition matrix, coal plants	
Capacities, existing fossil-fired and hydro plants	Current Monet spreadsheet
Capacity, Port Westward	
Measure of flexibility for Beaver	Current Monet spreadsheet

Table 1. "Certain" data and preliminary sources

Data requirement	Source
Heat rates, existing fossil-fired plants	Current Monet spreadsheet
Heat rate, Port Westward	
Outage rates, existing fossil-fired and hydro plants	Current Monet spreadsheet
Outage rate, Port Westward	
O&M cost, existing fossil-fired and hydro plants	Current Monet spreadsheet
O&M cost, Port Westward	
Maintenance schedules, existing fossil-fired and hydro plants	Current Monet spreadsheet
Maintenance schedule, Port Westward	
Gas tariffs	
Coal prices	Current Monet spreadsheet
Colstrip transmission cost	
Overmonth storage for hydro plants	
Amount of hydro capacity that must be withheld for load-following or reserves	
Amount of hydro capacity that must be scheduled to allow for downward load-following or downward regulation	
Measure of rescheduling flexibility for hydro plants	
List of hedges and transactions	
Transaction volumes by month	
Fixed prices for transactions, by month	
Identification of appropriate price indexes for transactions, by month	

3.2.2 Description of uncertainty data

A complete description of any individual data series would cover five topics:

- Name or general description of the attribute of interest
- Numerical information (data element) to be used to represent the attribute of interest
- Substitute or proxy data element to be used if needed

- Historical values available for use in modeling the distribution of this data element and its relationship to others. Note that the data element described by the historical values may not be the same at the one to be forecasted

- Data that have been identified or located representing either the historical values or a proxy

The following table describes both the "uncertainty" data used by the model, as well as other "hidden" variables that might prove useful in modeling their distributions.

This table does not describe the other data with which each data element may covary; there were just too many possibilities, given all the different similar variables (such as power prices). Following the table we will list some generic covariances (generic because we refer to "power prices" rather than a specific power price data element) that may exist. Time will not permit all potential covariances to be tested.

Covariance does not imply causation; two correlated variables can have a common "causative" variable, but causation is not really a statistical concept. For example suppose that there is a plausible reason to assume that the value of variable *y* is determined by variable *x*, and in fact the "true" dependence relationship is $y = Ax + b + \varepsilon$ where ε is a normally distributed error. This can equivalently be written $x = A^{-1}y - A^{-1}b + \varepsilon'$ ($\varepsilon' = -\varepsilon$ is a normally distributed error) even though there is no "plausible" model under which *y* "causes" *x*.

Table 2. "Uncertainty" data, covariances and related historical data

Attribute	Numerical values	Proxy	Related Historical data	Identified data
"Uncertainty" data				
Day-ahead power price at Mid-C	Hourly values in $/MWh	Day-ahead price at COB	Daily on and offpeak prices, hourly "scalers" based on ratio of Dow-Jones hourly indices	Dow-Jones daily (day-ahead) prices 1998-2004, hourly prices 2003-2004
Day-ahead power price at COB	Hourly values in $/MWh	Day-ahead price at Mid-C	Daily on and offpeak prices, hourly "scalers" based on ratio of Dow-Jones hourly indices	Dow-Jones daily (day-ahead) prices 1997-2004, hourly prices 2003-2004
Daily peak and offpeak subperiod prices at Mid-C	Average (index) values in $/MWh	Daily subperiod price at COB	Daily on and offpeak prices	Dow-Jones daily (day-ahead) prices 1998-2004
Daily peak and offpeak subperiod prices at COB	Average (index) values in $/MWh	Daily subperiod price at Mid-C	Daily on and offpeak prices	Dow-Jones daily (day-ahead) prices 1997-2004
Monthly peak and offpeak subperiod prices at Mid-C	Average (index) values in $/MWh	Monthly subperiod price at COB, daily subperiod price at Mid-C	Monthly price indices at close of trading of forward contract	
Monthly peak and offpeak subperiod prices at COB	Average (index) values in $/MWh	Monthly subperiod price at Mid-C, daily subperiod price at COB	Monthly price indices at close of trading of forward contract	
Short-term power price at Mid-C	Hourly prices in $/MWh projected four hours ahead	Real-time price at Mid-C		
Real-time power price at Mid-C	Hourly values in $/MWh		Actual historical hourly prices preferably from PGE trading floor	Dow Jones hourly prices 2003-2004

Table 2. "Uncertainty" data, covariances and related historical data

Attribute	Numerical values	Proxy	Related Historical data	Identified data
Monthly projected Mid-C prices, for hydro dispatch	Forecast of hourly prices for the entire month	Day-ahead prices at Mid-C		
Sumas gas price	Daily values	Stanfield gas price	Daily historical prices	Stanfield gas price 1994-2005 (PA data)
Stanfield gas price	Daily values		Daily historical prices	Stanfield gas price 1994-2005 (PA data)
Hydro energy	Available hydro energy by plant		Not necessarily needed as BPA forecasts energy	BPA white book
Hydro capacity	Monthly capacity by plant		Not necessarily needed as BPA forecasts energy	BPA white book
Run-of-river hydro	Minimum hourly hydro dispatch, by plant			
PGE cost-of-service load	Hourly load in MW		Historical hourly loads	Hourly PGE total load 2000-2004 and NCOS load for 2004; historical load from FERC form 714, 1993-2004

Hidden data

AECO gas price	Daily values		Daily historical prices	
Rockies gas price	Daily values		Daily historical prices	
Henry Hub gas price	Daily values		Daily historical prices	
Hydro condition (PGE)	Monthly energy		Historical hourly or daily	PGE hourly plant

Table 2. "Uncertainty" data, covariances and related historical data

Attribute	Numerical values	Proxy	Related Historical data	Identified data
	available		*energy dispatch* by plant	dispatch 1993-2005.
Hydro condition (NW)	Monthly energy available	Hydro condition (PGE)	Historical hourly or daily *energy dispatch* by plant	Federal system daily plant dispatch 1994-2002. Regional total monthly hydro energy 1997-July 2005 from NWPP.
Northwest regional load	Total hourly load in MWh	PGE control area load	Historical hourly load	Historical hourly load by control area from FERC form 714, 1993-2004
California load-resource balance				
California intertie capacity				Hourly intertie capacity data from 1998
Temperature	Daily max/min temperatures		Historical max/min temperature at a particular weather station, e.g., PDX.	

A. POTENTIAL RELATIONSHIPS (COVARIANCES) AMONG UNCERTAIN VARIABLES

- *Power prices* may be related to (depend on) gas prices, regional load, hydro conditions, other power price variables, California load-resource balance, intertie capacity.
- *Gas prices* may be related to (depend on) California load-resource balance, regional loads, or other gas price variables.
- *PGE load* may be related to (depend on) regional load, temperature.
- *Temperature* may be related to (depend on) hydro conditions.
- Historical hydro dispatch (historical data used to estimate hydro conditions) may be related to (depend on) load.

3.3 ESTIMATION ISSUES FOR UNCERTAINTY DATA

As noted above, the goal of the project for which PA was retained is to estimate the statistical properties of the distribution of net variable power costs. This distribution is estimated by simulating costs over a distribution of the values for "uncertainty data". It is therefore important to estimate the distribution of those data. Furthermore the distribution of uncertainty data may be joint and nonseparable, in other words, the different uncertain variables may be mutually dependent. In this section we will address several of the key issues around that estimation. Most of the general issues of modeling dependencies are in section 3.3.6 while the general issues of single-variable modeling are in 3.3.7.

3.3.1 Descriptive vs. prescriptive modeling

The most important issue in estimating the distribution of uncertainty data is to understand the precision with which that distribution must be estimated. That depends on the use to which the end product – the estimated distribution of costs – is to be put.

The cost simulation model at issue here may be used for ratemaking. If the distributional outputs are to be used to set rates based on some kind of "risk-adjusted cost" it would be appropriate to invest considerable effort into the estimation of the underlying variables. We can call this a *prescriptive* analysis, where the model is used to determine a "once and for all" value. On the other hand, if there is an opportunity to "true up" the revenue requirement and the model is used to understand the likely size of the true up, the estimate can be less precise. We can call this, by contrast, a *descriptive* analysis. We will also use the term *descriptive* for a model used to determine a value subject to correction or true up, because it does not prescribe the value once and for all. It is important to clarify which kind of analysis is desired for this project.

As an analogy, consider the use of the Black-Scholes formula for option pricing. The Black-Scholes formula yields the price of a stock option based on two parameters, the risk-free interest rate and the "volatility" of the stock's price. There is an underlying assumption that the stock's price evolves under geometric Brownian motion. If the evolution assumption or the volatility parameter is wrong, the Black-Scholes formula will give the wrong value – and usually neither assumption is wholly correct.

Yet the Black-Scholes formula is used every day to value billions of dollars in options. The key is that the valuation occurs every day. Every day these options are revalued and the user, observing the market's reaction to the previous day's trading, is able to retune the parameters. Furthermore, because portfolios are adjusted every day the exposure to pricing errors is controllable (one can exit a position taken in error). If the formula were used to set the price for illiquid long-term options, with no opportunity to recover from error, it would be appropriate to invest considerably more time and effort into improving the modeling of the underlying random variables.

3.3.2 Estimating the impact of specification error

We believe that an important use for the cost simulation will be to determine the importance of precision in various inputs, and therefore the effort that ought to be invested in precisely estimating different relationships. In this report, we describe a number of questions about data relationships. Resolving all of them, and obtaining correct specifications for the joint distribution of the uncertainty variables, may be prohibitive. It is therefore important to determine the value of information: for each choice of models, how important is the difference?

A prototype cost simulation model can provide a tool for answering that question. A simulation model has three components: a deterministic simulation of dispatch and transactions; sampling from a distribution of input data; and a component that supervises the operation and reports results. Dependence and distribution issues impact the second of those components. Even if the deterministic simulation is just a prototype – not a complete or accurate representation of PGE's operations – it should still provide good *relative* information, as to which errors in input specification would have the greatest impact on results.

Again, the common use of the Black-Scholes model provides an analogy. That model is descriptive rather than prescriptive, in two ways. First, the model is used in an ongoing decision process, rather than to prescribe values "once and for all". Second the model is commonly used as a consistency check between "implied volatilities" of different options. In this usage the model provides not just a numerical result, but more important, a validation of its inputs.

3.3.3 Data availability

Our ability to estimate distributions is limited by the data available. Data unavailability could also affect the model design.

When data are not available we seek proxies that can be used as close substitutes. For example, as a proxy for "hourly forward prices" one often uses daily or monthly forwards to scale observations of hourly spot prices. The choice of a proxy is usually based on theoretical considerations. The underlying assumption is that the proxy is a good statistical predictor of the unavailable variable. The sensitivity of the model results to bias in the proxy should be tested using a descriptive approach.

There are several reasons why consistent historical datasets may not be available:

- Historical data may not exist because a particular data item has no historical analogue – for example, non-Cost of Service load.
- Trading or dispatch decisions may be based on data that is not archived. While actual (metered) load is surely archived, daily forecasts of future loads may not be.
- While historical values for some data may be archived, e.g., historical forward curves, the form in which the data is saved may make it costly to retrieve or organize for forecasting (e.g., daily reports that are not written to a standardized format).
- Data may be commercially sensitive, or its use might compromise commercially sensitive information. Current forward curves are generally considered sensitive. Historical forward curves may not be as obviously sensitive, but the combination of the historical curves and historical trading data might reveal an organization's trading strategies. Similarly, one might be able to infer position or trading limits from historical transaction data.

3.3.4 Impact of operation on recorded data

Distributions of random variables ought to be forecast based on observations untainted by human activity that could serve to mask or mitigate the variability in the underlying variable. This is particularly a problem when modeling power prices and hydro availability. In particular, the input variables to the cost simulation (the "uncertainty data") describe the flexibility or range of generation available from hydro under dispatch control but the historical data for a given year describe an actual realization of hydro generation.

- Power prices can depend on other exogenous variables, such as gas prices. However, they can also be affected by actions taken as a response to observed power prices, such as demand response or changes to the dispatch. By "dispatch" in this context we should understand only the dispatch of PGE units and contracts, which the simulation model represents as a function of prices; the dispatch of non-PGE units is an underlying uncertainty. PGE generation could be used as an explanatory variable in a statistical model of power prices, but the model estimation would need to account for the mutual dependence by using, for example, a multi-stage estimation.
- The situation with hydro condition estimation is even more complex. The data that are generally available describing historical hydro conditions are historical hydro *generation* values, which combine the effects of the hydrological state and dispatch decisions. If one is modeling the dependence of price on hydro one has to account for the fact that hydro generation is partly dependent on price. However, the form of that dependence in the past may not be the same in the future. The physical layout of the hydro system (impairments) changes slowly over time; the rule curves and environmental restrictions change more frequently. Historical hydro variables don't necessarily jibe with the values available as input to a cost simulation model. A separate model that derives monthly hydro capacity and available energy from hydrologic data such as precipitation represents the complexity of the Northwest hydro system.

3.3.5 Relationship between similar variables

Some generic data items, such as “power prices”, are represented by multiple similar but non-identical specific variables. These can represent prices with different degrees of granularity (e.g., hourly, subperiod (peak/offpeak), daily or monthly), forward tenor (difference between the time or date at with the price is observed and the time of spot delivery) or delivery location. The various prices are all different and the various aspects of PGE decision-making rely on different prices. In fact, the differences between some of these prices are important to cost components such as the profit or loss from daily redispatch (trading). The relationships between such similar variables are often quite hard to model; for one thing, the differences between the “true” values of the respective variable can be of the same order of magnitude as the precision with which they are reported, or the bid-ask spreads.

In general theorists have paid the greatest attention to the modeling of forward tenor because it can be analogized to interest rate modeling, which is well studied. For example, the Clewlow-Strickland forward curve evolution model assumes that forward prices of different tenors are correlated but changes to long-dated prices are much smaller (attenuated) than changes to short-dated prices. Fitting such a model can be difficult. During periods of extreme spot-price volatility one may question whether a forward-curve model still holds – if the volatility is clearly due to short-term effects the attenuation should be greater. It is particularly difficult to answer this in the case of the 2000-2001 Western price shocks because liquidity in the forward markets basically dried up and publicly available historical long-dated prices are scarce or nonexistent.

The relationship between prices at geographically separate locations may require a regime-switching model. When transmission capacity is available the price difference is often a constant related to transmission costs (sign depends on the direction of flow) but when transmission capacity is all being used the prices are really uncoupled. Historical data about flows and capacities on interfaces is often hard to get. Analysts often use a single averaged “locational basis”.

The most complex relationship, though, may be between prices of different granularities. For example, there is generally no “hourly forward price” or hourly forecast price. It is customary to create a profile of “scalers” that represent the ratio between an hour’s actual load and that day’s average load, and average the profiles over a period of time. This leads to a number of other questions, often ignored: what is an appropriate measure of the “standard deviation” of a scaler profile? Do forward prices scale the same as spot prices? Does the actual variation in profiles over time really represent variability in some hours’ prices while others stay relatively constant (in which case the whole concept of a “scaler” profile would have to be rethought)? The final question goes to the issue of the appropriate definition of averaging periods; commercially-defined peak and offpeak periods are convenient for contract standardization but not necessarily for hourly price forecasting.

Fortunately, this complex relationship does not need to be modeled precisely. For, what is important in modeling PGE dispatch is not applying a “good” hourly detail to forward indexes; rather it is applying an hourly detail that is faithful to PGE’s operational practice. In this case, statistical modeling is not as important as continued dialog with PGE staff to ensure that their forecasting methods – right or wrong – are replicated in the model.

3.3.6 Relationships between dissimilar variables

In section 3.2.2a above, we provided a list of possible covariances or relationships among variables. To date, PA has focused its attention on possible relationships between power prices and gas prices, regional loads and hydro generation. We did not yet have good data representing such variables as PGE load, California load-resource balance, or intertie capacity. This analysis has involved most of the important issues in modeling relationships among variables:

- Form of the relationship. PA has modeled linear relationships between the input variables, between some of the variables and logarithms of the others, and where some of the variables are replaced by differences. We have not yet tried to neither model more complex transformations such as a logistic, nor have we modeled any form of regime switching. The simple linear model is appropriate for determining the existence and sense of a relationship, and can capture much of the variability. We have usually found log-linear models useful for both "descriptive" and "prescriptive" analysis but we recommend a separate computational test of the potential impact of specification errors. Furthermore, an R-squared value may not be a good measure of the applicability of the data model, as noted below.

- Comparability of variables. Available historical data series may not be appropriately comparable. For example we believe there are good fundamental reasons that power prices should be related to "net thermal load", that is, load minus hydro generation, and the coefficient of interest should be the common absolute value of the respective coefficients. For that relationship to hold, the net thermal load should be the load in a region without many transmission constraints and with substantial internal liquidity, minus the total hydro available in the same region. For load, we had the total load from FERC forms 714 of control areas identifies as being in the Northwest; for hydro generation we had the total generation for large Federal hydro plants, normalized to NWPP monthly hydro energy. Neither of those really corresponds to the region that determines Mid-C prices (the locational price we modeled). Although their coefficients were not of about the same absolute value the difference could be attributable only to their incomparability.

- Statistical independence. Dissimilar variables may not be statistically independent. One example, described earlier, is the relationship between price and hydro generation, since hydro generation responds to price expectations. Relationships that involve such mutual dependence are often estimated using multistage regression. However, in the extreme it may be necessary to employ a structural model of the relationship. In the case of hydro conditions and prices, PGE already uses a regional model to determine monthly hydro availability. A similar approach could be used to determine base case prices in each hydro condition, with the dependence on variables such as gas prices assessed separately. This in turn depends on the assumption that gas prices and hydro conditions affect power prices separable (otherwise the regional price model must be run for combinations of power prices and hydro conditions which is probably prohibitive in time and effort).

- Measure of goodness of fit. We typically assess the relationships between variables using some form of linear least squares. The common measure of goodness of fit is R-squared. While choosing the model with the highest R-squared, or even using any model with "sufficiently high" R-squared, may be reasonable for a cost simulation to be used in a descriptive mode, it does not assure the precision one would want in a prescriptive model. In that case goodness of fit should be measured by the absolute size of the residuals rather than relative to the total variation in the dependent variable, and the residuals themselves should preferably come from holdout samples.

- Stability of the relationship. Any algorithmic approach based on historical observations assumes that data relationships are stable; it is hard to imagine an objective approach based on anything else. (Delphi approaches are not sufficiently quantitative, especially for prescriptive modeling.) Although markets do change, the stability assumption is a reasonable one. Major market transformations will play out over periods of several years; the horizon of this model is only fourteen months, so data relationships derived with an emphasis on the last two years' worth of data should not be too far off the mark. Furthermore, it should be possible to provide manual override, where PGE or the OPUC staff can specify relationships to be imposed on the uncertainty data; however, when input data relationships are specified ad hoc the results should not be used prescriptively.

3.3.7 Modeling distributions of variables and errors

The uncertainty data we have described all involve independent uncertainties. In other words, while one variable may be related to several others it is not completely determined by them. Even though an uncertain variable may covary with many others, it involves an additional underlying uncertainty or error whose distribution must be simulated to provide inputs to the cost model. An independent variable, unrelated to any others, must also have its distribution assessed. The following issues are associated with distribution assessment:

- Relationship between historical data and variables to be forecast or estimated. Data about variables not particular to PGE will probably come from public sources. Even for variables directly related to PGE's dispatch we may have to fall back on public sources out of confidentiality concerns. Public data are often subject to various forms of processing that distort their relation to the variable being described, or to the variable whose future distribution they are used to estimate. An extreme example is the "system lambda" values filed with FERC, which often do not represent system marginal cost but rather a particular rule-based computation. Load data may represent control area load rather than retail deliveries; PGE meter data we have seen so far only separate out NCOS load for 2004.

- Choice of distribution. It is convenient to sample from an analytically defined distribution function, such as a normal, lognormal, exponential or Weibull distribution, rather than to rely on historical data. Extreme values are rare in historical data, but a large sample ought to include a number of extremes. Analytic distributions are used to extrapolate as well as fill in the historical record. However, it is often very difficult to distinguish between various choices of distribution without a theoretical model of the error process to guide the decision. With no other information the most reasonable way to choose the error distribution would be to rely on the Central Limit Theorem, assuming a large number of independent sources of error. If their effects add, the error distribution should be assumed normal; if their effects are multiplicative the error distribution should be assumed lognormal.

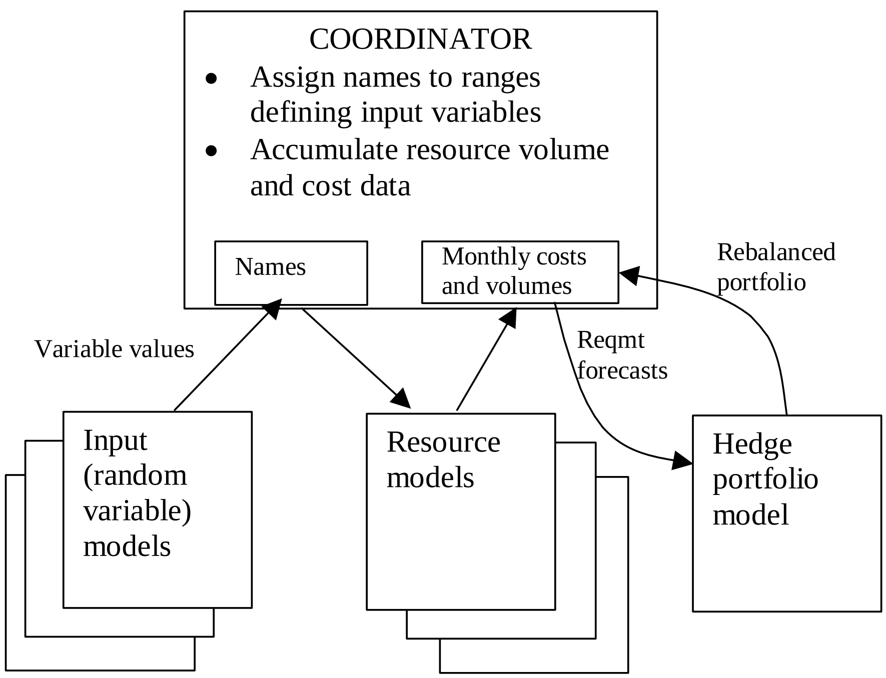
- Closed-from distributions vs. sampling of historical residuals. If there is no reasonable basis to choose an error distribution or if the distribution of historical residual shows clear non-normal behavior (e.g., multiple modes) one often has to fall back on using a discrete distribution based on historical residuals. This approach is more reasonable when there are a large number of historical residuals from which to choose – more than the number of samples to be taken – but it may be the only option when there is no obvious model for the data, even if the historical sample is small. This is the case in modeling hydro conditions. Hydrological conditions are represented by a large set of variables, reduced by a dispatch or regulation model to a set of energy and capacity values. The analytic structure of the energy and capacity data is too complex to model. Therefore one usually samples from a dataset associated with simultaneous historical observations of the hydrological variables, each historical year representing a distinct observation or item in the distribution.

4. MODEL STRUCTURE

We concluded that a cost simulation model had to be flexible and easily modified to test different ways of representing both the interdependencies of input variables and those resources for which unique "best models" cannot be defined – such as the forward hedge portfolio. We referred to this as a "sandbox" model because it provided a simple architecture for testing different approaches to individual resources.

The overall structure of the cost simulation is a Monte Carlo model. Different components of the simulation – different input models and resource models – are contained in separate workbooks (files), to make it easier to test different versions or configurations of the components and "mix and match". The names of the workbooks to be used for a particular run are specified in the Coordinator workbook. We settled on @Risk, a product of Palisade Corporation, as a platform for prototyping the cost simulation because it is Excel-compatible, can run models spread over several workbooks, is easily manipulated, and includes a good functionality for defining and extracting inputs and outputs.[2] @Risk manages the Monte Carlo simulation process.

The overall architecture of the cost simulation is illustrated below:



[2] We did come across one sporadic problem with @Risk. Some resource models, such as Beaver and Mid-C (which used DLLs), had to be calculated with VBA routines rather than just Excel workbook functions. This had to occur after the @Risk "recalculation", in which all the random variables are simulation. But then to get a correct calculation of the summary outputs, which used spreadsheet functions, we had to recalculate certain sheets programmatically (from the VBA code). @Risk would sporadically hang on the recalculation reporting an unknown error during the calculation.

The building blocks of the cost simulation are input models, resource models (of which the hedge portfolio model is a special case) and a Coordinator.

The *input variable models* simulate the various random drivers for each iteration of the Monte Carlo simulation. The input variable models are based on Excel worksheet functions, and possibly VBA functions, but not VBA subroutines, in order to ensure that they are computed as part of the basic @Risk "recalculation" step prior to any programmatic resource models. Each resource model has as its output one or more arrays of random variables such as monthly hydro availability, daily gas price or hourly load. The input variable model "exposes" those arrays to the Coordinator by listing where each array is in the workbook and giving it a name. The Coordinator will assign a workbook a name to that array so that it can be referenced by other input models and the resource model – it looks like a named range in the Coordinator workbook.

The resource models simulate the operation of various resources (load, generators, physical power contracts and hedges). The resource models can obtain the values of input variables from the ranges defined by the names in the Coordinator. Each resource model can have up to four outputs: hourly MWh volume (produced or, in the case of load, consumed); non-power cost to produce those MWh; value of production at spot prices (or cost of load at spot power prices); and net profit. Each resource model includes a reserved range that tells the Coordinator where to find its output.

The resource models can also include VBA routines to be run prior to the simulation (to set up parameters), after the @Risk recalculation (in order to run VBA routines or DLLs such as for dynamic programming models of Beaver or the Mid-C contracts), or after the main group of resource models. The last option is for the use of the forward portfolio model: the forward portfolio is rebalanced over time as information emerges about the net position to be realized, but that requires the net short position (load minus physical resources) already to have been computed.

The Coordinator has three main functions. First, it assigns names to all the outputs of the input variable models, as if those models were storing their outputs in an area identified with the Coordinator. Second, it defines macros that call the VBA routines defined by the resource models, in the right order. Finally it accumulates monthly values of the volume, cost, revenue and profit associated with the resource models and defines them as @Risk output variables.

This architecture does not itself involve a mathematical specification. Mathematical specifications will be given for some of the input and resource models in the next two sections.

5. INPUT COMPONENTS – PROTOTYPE

In this section we describe the various input variables or drivers. There are five "input models" associated with the prototype:

- Temperature
- Load
- Hydro energy
- Gas forward and spot prices
- Power forward and spot prices

We will describe load and temperature together but for expository purposes separate gas forward prices and spot prices into separate sections. In each section we will describe the modeling we did to determine an underlying distribution for each variable, or its relation to other variables. We will then describe how we simulate each variable; the simulation should be consistent with the modeling.

The variables' distributions are set up so that their expectations equal the variable values used in a Monet run to support the RVM. These are the "Monet base case" or "Nov. 1" values.

5.1 LOAD (AND TEMPERATURE)

The first important factor influencing net variable power costs is load. Load obviously influences the total power cost; it also influences the unit cost of power, because (all things being equal) higher loads are served on the margin by higher-cost units, and the higher the load the more load has to be served by non-PGE resources. There is generally a very strong correlation between load and temperature, and load and seasonality. Therefore we modeled load as a function of temperature and time in the year, and also temperature as a function of time in the year.

5.1.1 Temperature

We fit temperatures for five years (2000-2004) to a mathematical model; however the model itself was not used as part of the simulation. Rather, only the distribution of errors around the fitted model was used, as a description of the random fluctuations of temperature around the daily normals used in Monet. The mathematical model was of the form:

(1) $$T_t = T_0 + mt + \alpha_1 \cos 2\pi t + \beta_1 \sin 2\pi t + \alpha_2 \cos 4\pi t + \beta_2 \sin 4\pi t + \alpha_3 \cos 6\pi t + \beta_3 \sin 6\pi t + \varepsilon_t$$

where T is the temperature in degrees Fahrenheit, t is the time (in years) since Jan. 1, 2000, and T_0 is an intercept. The m coefficient allows for any recent temperature trends while the α and β coefficients express seasonality (with a yearly period). The model was fit over five years of data assuming all the ε_t were identically distributed. The fitted coefficients were as follows:

T_0	53.18 (0.23)
m	0.59 (0.08)
α_1	-13.42 (0.16)
β_1	-4.80 (0.16)
α_2	-0.47 (0.16)
β_2	2.42 (0.16)
α_3	0.11 (0.16)
β_3	0.04 (0.16)

Having derived the structural model, we computed the standard deviations of the residuals for each month i.e., σ_1 would be the standard deviation of the set of ε_t for all the days in January for 2000-2004, σ_2 would be the standard deviation of the set of ε_t for all the days in February, etc.

Although not all the coefficients in this model are statistically significant (e.g., α_3) the R-squared value of the regression, 82%, appears to be appropriate for a regression. Therefore it is reasonable to use the monthly σ_m values as measurements of the error in a date based temperature forecast – e.g., daily normals – even if we don't use the specific forecasting model above. In order to simulate temperatures around the PGE forecast (taken from a file supplied to us), we simulated the equation:

(2)
$$T_t = \hat{T}_t + \varepsilon_t, \quad \varepsilon_t \sim Normal(0, \sigma_{m(t)})$$

where $\hat{T}_t$ is the forecast temperature and *m(t)* is the month associated with day *t*. The structural model derived above was not actually used.

5.1.2 Load

To model load as a function of season and temperature we started with a similar, but somewhat more complex model:

(3)
$$L_t = L_0 + mt + \alpha_1 \cos 2\pi t + \beta_1 \sin 2\pi t + \alpha_2 \cos 4\pi t + \beta_2 \sin 4\pi t + \alpha_3 \cos 6\pi t + \beta_3 \sin 6\pi t + A_1 T_t + A_2 T_t^2 + A_3 T_t^3 + d \cdot Wkday + \varepsilon_t$$

where *L* is the daily total PGE load, *t* is the time (in years) since Jan. 1, 2000, *T* is the temperature in degrees Fahrenheit, *Wkday* is a dummy variable that is 1 for weekdays and 0 for Saturdays and Sundays, and L_0 is an intercept. The *m* coefficient allows for load growth over time, the α and β coefficients express seasonality (with a yearly period), the *A* coefficients express temperature dependence, and *d* encapsulates the difference between weekday and weekend loads. This model was fit to five years of data (2000-2004). The fitted coefficients were:

L_0	80590.65 (4110.18)
m	-1115.99 (32.77)
α_1	2623.96 (156.30)
β_1	188.27 (84.60)
α_2	581.91 (73.03)
β_2	217.69 (73.49)
α_3	-66.05 (65.86)
β_3	-428.09 (67.36)
A_1	11.19 (234.30)
A_2	-25.90 (4.34)
A_3	0.29 (0.03)
d	5918.09 (101.51)

As with the temperature model, we computed monthly standard deviations of the residuals, $\sigma_m^{(L)}$.

Note that not all of the coefficients in the above table are significantly different from zero. The most glaring example is α_3. As it happens, and as will be explained below, we actually did not use the time-based (*t*-dependent) part of the structural model.

The fitted load model did not replicate the forecasts $\hat{L}_t$ in the Monet base case, i.e., if we were to forecast load for day *t* based on the normal temperature $\hat{T}_t$ we did not get $\hat{L}_t$:

(4) $$\hat{L}_t \neq L_0 + mt + \alpha_1 \cos 2\pi t + \beta_1 \sin 2\pi t + \alpha_2 \cos 4\pi t + \beta_2 \sin 4\pi t + \alpha_3 \cos 6\pi t + \beta_3 \sin 6\pi t + A_1 T_t + A_2 T_t^{\,2} + A_3 T_t^{\,3} + d \cdot Wkday$$

The goal of this exercise was to estimate the potential variation of net variable power cost around the Monet forecast, assuming that the Monet base case used unbiased forecasts of those uncertain variables. Therefore we normalized the load forecast to the Monet base case:

(5) $$L_t = \left(\hat{L}_t - A_1 \hat{T}_t - A_2 \hat{T}_t^{\,2} - A_3 \hat{T}_t^{\,3}\right) + A_1 T_t + A_2 T_t^{\,2} + A_3 T_t^{\,3} + \varepsilon_t^{(L)}, \quad \varepsilon_t^{(L)} \sim Normal(0, \sigma_{m(t)}^{\ (L)})$$

In other words, the fluctuation of load around its forecast is explained by fluctuations in temperature and an additional normal error. The additional advantage is that the cost simulation will be estimating the cost to serve the same kind of load as Monet, which we believe to be an estimate of the cost-of-service load.

The daily load is converted to 24 hourly loads using a set of scale factors or "scalers". The raw data was again five years' of loads. Each hour was assigned the ratio of its load to the daily total (scaler). Each day was characterized by the month (January to December) and whether it was a weekday (Monday-Friday) or weekend. The scalers were averaged for each hour by month and weekday/weekend indicator. Thus there are 576 scalers (24 hours for 12 months and two daytypes) subject to 24 conditions (for each month and daytype the scalers add to 1). The load for any hour is the simulated load for the day in which the hour occurs, times the appropriate scaler.

5.2 HYDRO AVAILABILITY

We did not have available a consistent set of historical and forecast hydro data. We had three different sets of historical hydro data available – hourly generation from PGE-owned units, daily generation from a set of Federal units, and monthly generation data from the NWPP (including Canada). We used the daily Federal data for the model of power prices described below.

On the other had, we had limited hydro forecast data – basically the information in the 2004 BPA White Book. We constructed a distribution of monthly hydro energy based on two tables from that reference. As a base we used the Water Year (WY) 1937 "Total Hydro Resources" and "Total Surplus/Deficit" for 2006 (Technical Appendix 1, pp. 94-95). To construct a set of variations we used the "Surplus / Deficit by Water Year" from Technical Appendix 1 pp. 122-123. For each water year and month, the monthly energy equaled the WY1937 energy, minus the WY 1937 surplus/deficit, plus the surplus/deficit for the water year under consideration. The fifty water years in the table yield a 50-point discrete distribution for each month, which is converted to a distribution of ratios by dividing by the average over all fifty years (not by the WY1937 value).

The underlying random variable is a vector of 12 monthly energy ratios. Thus, for each iteration of the simulation one of the fifty water years is chosen at random and the ratios for each month of that water year are used. We have no basis to assume any correlation structure across the Northwest, so we assume that every hydro variable is governed by the same ratio. In other words, if the ratio chosen for January is 0.85 then we assume that in January every hydro resource has 85% of its Monet base case energy.

5.3 GAS FORWARD PRICES

To simulate operation costs we would need a model of gas spot prices, but not necessarily forwards. The bulk of PGE's energy comes from hydro and purchases rather than gas so (at least initially) we felt we could ignore gas hedges. However, we did not feel we were able to ignore power hedges. As noted in the next section, we did not have enough historical data on power forwards to construct a model of Mid-C forward prices. We did, however, have enough data on gas prices to construct a model of gas forwards, which we could combine with a model of the relationship between gas and power spot prices to produce an indicative model of the power forward curve. It was therefore important to model the gas forward curve.

Since price simulation is a dynamic process that evolves over time, it is important to maintain consistency in modeling the spot and forward price processes. In Clewlow and Strickland's[3] 1999 paper, they established a consistent model for the entire forward curve. Their price model assumes the forward price is of the form:

$$\frac{dF(t,T)}{F(t,T)} = \sigma e^{-\alpha(T-t)} dz(t) \tag{6}$$

Here $F(t,T)$ is the forward price on date t for delivery on date T. This equation has two volatility parameters; σ determines the level of volatility for spot and forward price returns, while α determines the rate at which the volatility of increasing-maturity forward prices decline, as well as the speed of mean reversion of the shortest-term price. These two parameters can be estimated from the prices of options on the spot price of energy, or forward contracts.

We estimated σ and α based on a relatively short series of gas prices for delivery at Malin, in order to have a consistent set of forwards. We obtained the annualized values σ=0.553 and α=0.482 ("annualized" means t is measured in years). These parameters were used to simulate the evolution of the forward curve from the Monet base case.

In other words, the Monet base case contained a set of forward prices, which we assumed were as of Nov. 1. The model required a simulated gas forward curve for each day of the following calendar year. The forward curve was simulated according to the formula:

$$F(t,T) = F(t-1,T) \cdot \exp(\sigma(t,T)\varepsilon_t \sqrt{\Delta t}) \cdot e^{-\sigma(t,T)^2 \Delta t/2}, \quad \varepsilon_t \sim Normal(0,1) \tag{7}$$

where $\sigma(t,T) = \sigma e^{-\alpha(T-t)}$, Δt=1/365.25 (one day). The last term in (3) corrects for the bias introduced when one exponentiates a random variable. It is important to note that this is a "curve" model, because ε_t depends only on t, not T. Furthermore, even though the parameters of the price process were based on Malin prices, which may be biased relative to the prices seen by PGE, the forward curve is evolved from Monet base case prices that should correct for any bias.

5.4 GAS SPOT PRICES

There are several methods one can take in developing a spot price process. The most prominent one is the Geometric Brownian Motion model with mean reversion. One can also add a jump term to the mean reversion process if historical data shows spikes with meaningful frequency, or when the jump term significantly affects the valuation.

The other question of interest is the definition of mean (as to which the mean spot price should revert). One argument is that the spot simulation should preserve the value of forward price, so one can use the forward price as the expected mean. However, the forward price for a given month (the "current month") is frozen at the start of the month; fundamental market

[3] "Valuing Energy Options in a One Factor Model Fitted to Forward Prices", Les Clewlow and Chris Strickland, April 1999.

information about medium-term effects can arrive during the month and should impact the value toward which spot prices revert. A compromise is to use the daily values of the forward price for the following month (the "prompt month") normalized using the forward for the current month.

The prototype model implements this model for spot gas prices. Let t_0 represent the day on which the current month's forward contract closes, T_0 the current month, and T_1 the prompt month. The prototype implements a simulation of the following model for the spot price S(t):

$$
\begin{aligned}
&S(t) = \frac{F(t,T_0)}{F(t_0,T_1)} F(t_0,T_1) \cdot e^{p(t)} \\
&p(t) = (1-\alpha^s)p(t-1) + \varepsilon_t^s + \zeta_t\left((1+v)e^{\delta_t - \gamma^2/2} - 1\right) - B \\
(8) \qquad &\varepsilon_t^s \sim Normal(0,\sigma^s) \\
&\delta_t \sim Normal(0,\gamma) \\
&\zeta_t \sim Discrete(1\,with\ probability\ \Lambda \cdot \Delta t,\ 0\ otherwise)
\end{aligned}
$$

In the equation for p(t), the first term represents persistence with some reversion to 0, the second is a standard diffusion (geometric Brownian motion) and the third is a jump. The jump probability is $\Lambda \cdot \Delta t$ and the jump amplitude is lognormal. The fourth term (B) is the bias introduced by the reversion coefficient and exponentiation; it has a rather complex form:

$$
(9) \qquad B = \Lambda \cdot v \cdot \Delta t + 1/2\left({\sigma^s}^2 + (1+v)^2 \Lambda \cdot \Delta t\left(e^{\gamma^2} - 1\right) + v^2 \Lambda \cdot \Delta t(1 - \Lambda \cdot \Delta t)\right)\left(1 - \alpha^s\left(1 - \alpha^s\right)\right)
$$

The superscript *s* on σ^s and α^s is to distinguish them from the similar parameters in the forward price process.

σ^s and α^s were estimated using approximately one year of *Gas Daily* prices at Malin. They were actually taken from a simpler version of (8):

$$
\begin{aligned}
&y_t = -\hat{\alpha}x_t + \varepsilon,\ \varepsilon \sim Normal(0,\hat{\sigma}) \\
(8a) \qquad &where\ x_t = \ln\left(\frac{S(t)}{F(t,T_1)}\right)\ and\ y_t = x_t - x_{t-1}
\end{aligned}
$$

We fit model (8a) using an available set for prices spanning the period from Dec. 13, 2004 to Nov. 30, 2005. The dataset contained 253 observations at an average separation of 1.39 days, i.e., Δt=0.00381 years. The estimate for $\hat{\alpha}$ was 0.0341 with a standard error of 0.048. We annualized it by dividing out Δt: $\alpha^s = \hat{\alpha}/\Delta t$ = 89.47. The standard deviation of the residuals was $\hat{\sigma}$ =4.98%; we annualized it (in this case dividing by $\sqrt{\Delta t}$) to get $\sigma^s = \hat{\sigma}/\sqrt{\Delta t}$ =80.6%. Again, the fact that the prices are based on a simulated PGE forward curve should correct for locational biases.

Calibrating a jump model is quite difficult, so we did not do so for this analysis. We used values we consider to be representative: Λ=6, v=0.8 and γ=0.05. We believe the use of these "typical" values is sufficient for prototyping the cost simulation model.

Note that by using a single gas price, and modeling it based on forwards, we are representing less detail Monet, which uses separate locational prices for gas for Beaver and Coyote called GAS_PGE_1 and GAS_PGE_CS respectively, with a basis differential of from $0.08 to $0.33. We based the gas prices in the prototype on GAS_PGE_1, that is, we used the monthly values of that price in Monet as the Nov. 1 forward curve, allowed it evolve under the models described in this section, and used the result for both Beaver and Coyote. This is acceptable for prototyping the simulation; a production version ought to use different locational gas prices for the two plants.

5.5 POWER SPOT AND FORWARD PRICES

A key issue in the determination of net variable power costs is the relationship between power costs and other underlying cost drivers. This is particularly important because while we had a good historical dataset for Mid-C spot power prices, including the "crisis period" of 2000-2001, we did not have a good dataset for Mid-C forward prices. We felt it was important to include the crisis period in our model of power prices, more so than for gas prices – the crisis had to do with the relationship between power and other prices, since currently gas prices have been high without power prices spiking as much as in 2000-2001. If the model of the relationship between power prices and other prices, especially gas prices, did not include the crisis period we felt it would "over-fit", that is, the uncontrolled and independent influences on power prices would be understated.

5.5.1 Spot price model

We identified three key drivers of power prices to test: gas prices, hydro energy availability and load. The power price of interest is the price at Mid-C, which is a regional hub, so the price should represent regional conditions. Therefore we felt we should consider regional hydro availability and regional load as explanatory variables.

- Regional hydro energy was obtained from a dataset obtained from PGE containing 8 years' worth of daily generation from major plants in the Northwest, which we referred to as the "Federal" dataset. Unfortunately this dataset ended in 2002 and no continuation was available.
- Regional load was obtained from the EIA database of FERC form 714 responses. This database contains hourly loads and we summed the loads for all the reporting utilities in NWPA.
- For gas prices we used historical daily prices at Stanfield OR as reported by Bloomberg.
- The independent variable (power prices) was represented by historical daily prices at Mid-C as reported by Bloomberg. We had data on both peak and offpeak prices.

We tested several different models for peak power prices, selecting the "best fit" based on R^2. The models were distinguished by whether they used raw values of the variables, or their logarithms. We further tested several different error models beginning with ordinary least squares (OLS). A Durbin-Watson test indicated the presence of autocorrelation so we tried an autoregressive error model with one lag (AR(1)), autoregressive with two lags (AR(2)), and an AR(2)-GARCH(1,2) model. In PA's judgment the best of these is the AR(2) model with all variables represented by their logarithms:

(10)
$$\ln PP_d = LPP_0 + a\ln GP_d + b\ln PL_d + c\ln H_d + \varphi_d$$
$$\varphi_d = A_1\varphi_{d-1} + A_2\varphi_{d-2} + \varepsilon_d$$

In this specification, PP_d is the peak power price on day *d*, GP_d is the gas price, PL_d is the total load over peak hours, H_d is the hydro energy, and φ_d is the error; the actual random shock is ε_d. All variables are significant at a 95% confidence level (even 99%). The model coefficients are:

LL_0	-4.736 (0.592)
a	0.501 (0.070)
b	1.850 (0.141)
c	-0.578 (0.073)
A_1	-0.738 (0.024)
A_2	-0.231 (0.024)

The standard deviation of the shock ε_t is 0.200. The R^2 of the autoregressive model is quite high at 95.9%.

When we performed a similar regression for offpeak power prices, the coefficients for offpeak load and hydro energy were not significantly different from zero; however, if we used the onpeak load as an explanatory variable, its coefficient was significant. Since the onpeak load influences onpeak price, we tested a regression of offpeak prices on onpeak prices, which seems to work best:

(11)
$$\ln OP_d = LOP_0 + a'\ln PP_d + \delta_d$$

The model coefficients are:

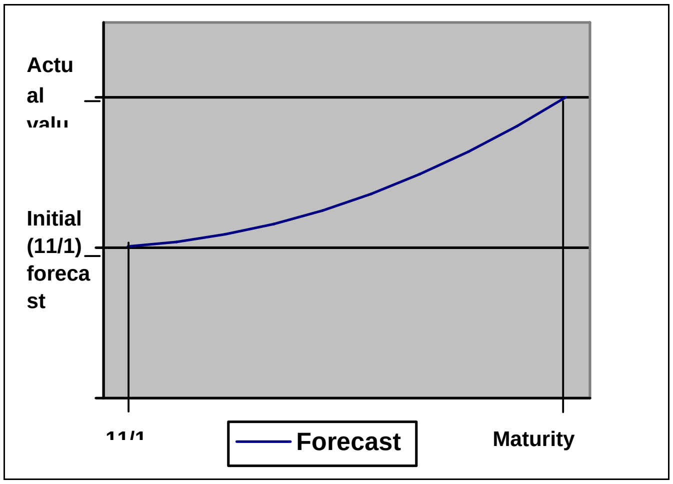
LOP_0	-0.232 (0.029)
a'	0.974 (0.008)

The standard deviation of δ_d is 0.302. The R^2 of this model is 91.1%|. Given that high value of R^2, and the fact that the independent variable (ln PP_d) is already autoregressive, we did not use estimate (11) with an autoregressive error model.

5.5.2 Use of spot price model to construct forward curves

We assumed that forward prices depend on forward versions of these same drivers, in the same way as spot prices. Thus, for example, the price for August power as of March 15 should depend on the August gas forward as of March 15 and the load and hydro energy expected, on March 15, to be realized in August (at the "maturity date").

Unfortunately, the model only simulates the values load and hydro energy and not the accretion of information about hydro conditions or the evolution of load forecasts. In order to construct power forward curves as well as to model hedge rebalancing (discussed below) we require a model of the *information arrival process*. In order to drive the prototype we use a quadratic model of information arrival. The graph below shows how a forecast would increase over time to meet the actual value:



Essentially at any date t between the date of the base case forecast (which is Nov. 1, denoted t_0) and the maturity date T, the state of knowledge of variable v_T is assumed to be

(12) $$v_T^{(t)} = v_T^0 + \left(\frac{t-t}{T-t}\right)^2 \left(v_T - v_T^0\right) = k(v,T,t)$$

We use the following formula as an estimate of the onpeak power forward price as of day *d* for delivery on day *D* (denoted PF(d,D); PF(0,D) is the monthly forward price from the Monet base case – as of Nov. 1 –GF(d,D) is the gas forward and PL_D^0, H_D^0 are the base-case expectations of load and hydro availability respectively):

(13) $$\begin{aligned} &PF(d,D) = PF(0,D)\cdot\left(GF(d,D)/GF(0,D)\right)^a \left(k(PL,D,d)/PL_D^0\right)^b \left(k(H,D,d)/H_D^0\right)^c e^{\varphi_d - B} \\ &\varphi_d = A_1\varphi_{d-1} + A_2\varphi_{d-2} + \varepsilon_d \end{aligned}$$

Here B is a bias due to all the exponentiation. We were not able to come up with a closed-form representation for B analogous to (9) so we estimated it by simulating (13) setting B=0, and then letting B be the ratio of its expectation to PF(0,D). Actually, rather than using a table of B values for all possible parameters *d* and *D*, we fit a linear function of the form

(14) $$B(d,D) = \mu\left(\frac{D-d}{D-t_0}\right) + \beta = \mu' d + \beta'$$

The offpeak power forward price, OF, was simulated using:

(15) $$OF(d,D) = OF(0,D)\cdot\left(PF(d,D)/PF(o,D)\right)^{a'} e^{(\delta_d - V\delta/2)}$$

where Vδ is the variance of δ_d.

Finally, daily average peak and offpeak power prices were simulated using similar formulas but based on the actual gas price, hydro availability and load rather than forwards and "partial knowledge" values (PP_d^0, OP_d^0, GP_d^0 are the values used in the Monet base-case forecast):

(16) $$PP_d = PP_d^0 \cdot \left(GP_d / GP_d^0\right)^a \left(PL_d / PL_d^0\right)^b \left(PL_d / H_d^0\right)^c e^{\varphi_d - B}$$

(17) $$OP_d = OP_d^0 \cdot \left(PP_d / PP_d^0\right)^{a'} e^{\delta_d}$$

Here φ_d and δ_d are as in equations (13) and (15). In equation (16) *B* is a bias similar to the bias in (13) and, similarly, was assessed empirically.

Hourly spot prices were computed from the peak and offpeak averages using scalers, similar to the load scalers described above.

6. RESOURCE MODELS – PROTOTYPE

In this section we describe the various resources. A "resource" is anything that contributes positively or negatively to the net variable power costs. Load is a resource, as are generators; the market-based cost associated with serving load is netted out against the profits obtained by selling generated energy into the market. In other words, the cost simulation is a "mark to market" model.

There are eight "input models" associated with the prototype:

- Load
- Beaver power plant
- Coyote power plant
- Boardman power plant
- Colstrip power plant
- Mid-C hydro contracts
- Portland General hydro resources
- Forward hedges

Some of these resources were modeled using the same DLLs as in Monet, although in one case (Coyote) we used a much simpler model because the DLL took too long to run. We will explain the simplification used as well as the tests carried out to check the impact of the substitution.

For every iteration the model records 180 items of information for each resource. For each month it records:

- MWh (generation or load) – peak, offpeak and total
- Average MW – peak, offpeak and total
- Total fuel or contract cost – peak, offpeak and total. This applies to power plants, which obviously have fuel costs, but also forward hedges, for example, the cost of a swap is the fixed price swapped for the market price. It does not apply to load, which only has a cost based on spot power.
- Total revenue or cost at spot power prices – peak, offpeak and total.
- Net profit (revenue minus costs) – peak, offpeak and total. The net variable power costs is the (negative of the) sum of the monthly net profit figures.

6.1 LOAD

The resource model for load is quite straightforward. The simulated load (described in 5.1) is multiplied by the simulated Mid-C price for each hour, which represents a negative value for "revenue at spot power price".

6.2 BEAVER POWER PLANT

The operation of the Beaver plant is simulated by dynamic programming using the same DLL (bcdispatch.dll) and subroutine (dll_bcd_optimize) as in the Monet model. The operating states, transition matrix, capacities, heat rates and O&M costs were taken directly from the Monet base case. This demonstrates the ability to use Monet components directly.

The DLL is used to dispatch the plant based on simulated gas prices and Mid-C power prices. The "fuel cost" in each hour is actually the sum of the fuel and VOM costs.

6.3 COYOTE POWER PLANT

We initially modeled the Coyote plant as for Beaver, using the same dynamic programming routine as in the Monet model. However, it became clear that because of the size of the Coyote model – 492 states based on 6 separate operating states and 72-hour minimum uptime – it took too much time to run. It would not be reasonable to run a Monte Carlo simulation involving such a complicated operational model.

We constructed a grossly simplified Coyote model, where each day the plant either runs or not, and if it runs then in each hour the model can freely choose any of the six operating states (minimum load; mid load; maximum steam turbine load (full load); full load plus misting; full load plus misting and duct burner; and full load plus misting and duct burner with no steam extraction). One might argue that in optimal daily cycling the plant would not run in offpeak hours, but since its minimum uptime is 72 hours it does have to run in offpeak hours for most of its operating days. We call this the "daily commit" model.

We also constructed a somewhat less grossly simplified model based on weekly commitment: the commitment decision would be made for a week at a time, and for each week the plant would follow one of four patterns: on all week, on from 7AM Monday through 10PM Friday, on from 7AM Monday through 10PM Saturday, or off all week. The first few hours could be adjusted if needed to transition from the previous week's end state, respecting the plant's startup and cooldown ramp rates. In each operating hour the model would freely choose any of the six operating states. We call this the "weekly commit" model.

We tested this simplified models against 100 iterations of the prototype input models. In other words we ran a Monte Carlo simulation of the input variables for 100 iterations (actually these were early versions of the input models) and dispatched each Coyote model against each set of inputs. We then compared the annual total "net revenue" from each of the simplified models against the original dynamic programming (DP) version using linear regression. In other words, we estimated the following two models by regression:

(18) $$NetRevenue(DP) = a \cdot NetRevenue(DailyCommit) + b + \varepsilon$$

(19) $$NetRevenue(DP) = a' \cdot NetRevenue(WeeklyCommit) + b' + \varepsilon'$$

ε, ε' are normally distributed errors. The coefficient values (with standard errors) and R-squareds were:

	Daily Commit	Weekly Commit
Intercept (b)	-4070(150)	1649(72)
Coefficient (a)	1.055(0.0045)	0.9799(0.0024)
R^2	99.8%	99.9%

Both simplified models appear to be quite good predictors of, and therefore acceptable substitutes for, the full dynamic programming model. What is most important is that they are much more efficient: each ran in about 1.5% the time of the DP model. For the results reported below we used the "Daily Commit" model, which is simpler to implement and easier to understand.

6.4 BOARDMAN POWER PLANT

The Boardman plant was modeled assuming daily cycling: it could start any hour from 1 to 7, and shut down any hour from 22 to 24, or not run at all (or be on forced outage). While running it was allowed to run either at minimum or maximum loading. The ownership fraction (65%), operating characteristics, forced outage rate and coal prices were taken from the Monet base case. The only random variable that impacts Boardman is its outage status (available or on outage).

6.5 COLSTRIP POWER PLANT

The Colstrip plant was modeled as two units (3 and 4) with slightly different characteristics (Colstrip 4 had more capacity in the summer and a lower coal cost in the second half of the year). Those characteristics, as well as the forced outage rates, coal prices and ownership fractions (20% of each) were taken from the Monet base case. Both units were allowed actually to cycle hourly, which appears to be how Monet models them. The only random variables that impact the Colstrip units are their outage statuses (available or on outage).

6.6 PORT WESTWARD POWER PLANT

The Port Westward plant was not modeled, since it was not included in the M606PUC05-105-06.xls Monet run we were using as a comparative, and it does not contribute to 2006 net variable power costs. The modeling of the Beaver, Coyote, Boardman and Colstrip should have been sufficient to demonstrate the ability of this modeling approach to represent fossil-fired generators.

6.7 MID-C HYDRO CONTRACTS

The four Mid-C contracts – Priest Rapids, Rocky Reach, Wanapum and Wells – were modeled as a single dispatchable hydro plant, with energy specified monthly and the obligation to provided reserves to cover PGE's other generation. This dispatch modeling was done by dynamic programming using the same DLL (midccomp.dll) and subroutine (dll_midc_optimize) as in the Monet model. All the parameters describing the contracts, and the base-case monthly energy and capacity, came from the Monet base case; however, for each iteration of the simulation model the monthly values were multiplied by the vector of hydro availabilities (see 5.2). Note that both energy and capacity are multiplied by the same scaling value.

6.8 PORTLAND GENERAL HYDRO RESOURCES

There are nine other PGE-owned or –contracted hydro resources: Round Butte, Pelton, Oak Grove, North Fork, Faraday, River Mill, Bull Run, Sullivan, Portland Hydro Project. The Monet model represents each of them with a set pattern of hourly releases based on weekly, monthly and hourly factors; they do not respond to load or prices. We modeled them as depending only on hydro availability, using the same hydro availability as for Mid-C: for each simulation iteration, each hourly generation value is multiplied by the hydro availability for that month.

6.9 FORWARD HEDGES

From our discussions with PGE staff as well as regulators it became clear that it would be important to allow the model to represent a hedge rebalancing process, whereby PGE modifies its hedge portfolio through the year. As it was described to us, PGE does not (by policy) rebalance its hedge portfolio in response to price movements but in response to its evolving view of its short position. This in turn can depend on price movements – increases in the spark spread should encourage greater operation of Beaver, Coyote, Boardman and Colstrip, and reduce the short position.

Specifically it is our understanding that by Nov. 1 PGE is 90% hedged for the coming year. The hedges are a mix of forwards and options; for simplicity the prototype model represents only options. During the year, PGE rebalances the hedges and increases the coverage to 100%. The rebalancing is based on PGE's forecast of its short position and therefore requires the same kind of information arrival modeling as does the power forward curve model (see 5.5.2). In this case we also allow for an error in PGE's forecast of its short position; that is, the arriving information can be incorrect.

Basically we assume that as of Nov. 1 PGE fills 90% of its expected short position L^0. Then on day d, PGE purchases an additional amount of energy for each forward month *M* equal to *k(L', M, d)*. Here L' represents a "perturbed" or unreliable estimate of the net short position and k() is the same quadratic information arrival process as in 5.5.2.

Specifically, for each iteration the net short position (load minus all generation) in month M could be exactly computed as L_M. We assume there is a random error in forecasting L_M; the forecast F_M is normally distributed with mean L_M and standard deviation equal to 2% of L_M. The 2% figure is an assumption made just to drive uncertainty into the modeling. On day d, then, enough additional forward contracts for month M are bought to bring the total position to

$$\left\{0.9+0.1*\left(\frac{d-d_0}{M-d_0}\right)\right\}\left(L_M^0+\left[\left(\frac{d-d_0}{M-d_0}\right)^2\right]\left(F_M-L_M^0\right)\right) \tag{20}$$

The term in curly brackets {} is the fraction of the load forecast that is expected to be filled by day *d*; it grows linearly from 0.9 to 1.0. The term in square brackets [] is the information arrival coefficient.

Note that the forward rebalancing for each iteration needs to be evaluated *after* the input variables have been simulated and the other resource models computed, because it depends on the short position ("filtered" using the information arrival coefficients). The forecast short position is a clouded view of the actual short position in each iteration, which depends on the energy produced by each resource. Therefore the resource models have to be executed, to provide their actual energy values, prior to computing the forward rebalancing.

7. PROTOTYPE RESULTS

PA constructed a prototype of the cost simulation model in order to demonstrate the behavior one can expect from such a model and the range of analysis possible with it. Two advantages of the modular "sandbox" construction were apparent early on, namely the ability to isolate and correct individual model components easily, and the ability to test and substitute different or simpler versions of a component model (such as the Coyote model noted above).

We first tried to compare the results of the cost simulation model with the Monet base case run. To do so we substituted the Monet gas and power prices for the price models in the simulation model and set relative hydro energy for each month to 1 (expected value). The prototype results were not the same as Monet's. This could have been due to errors in the simulation model, which was only a proof of concept rather than a polished model.

The simulation model reported nonzero energy costs associated with the PGE hydro units. These nonzero costs were variable O&M reflecting the O&M costs found in the Monet model, ranging from $0.19/MWh for North Fork and Faraday to $5.02/MWh for Round Butte. Monet reports no O&M costs for any plant, and no costs at all for any PGE hydro resources or contracts except Portland Hydro Project and the Mid-C plants. The costs reported for those plants are fixed contract costs.

The following table summarizes the differences between Monet and the prototype's "base case":

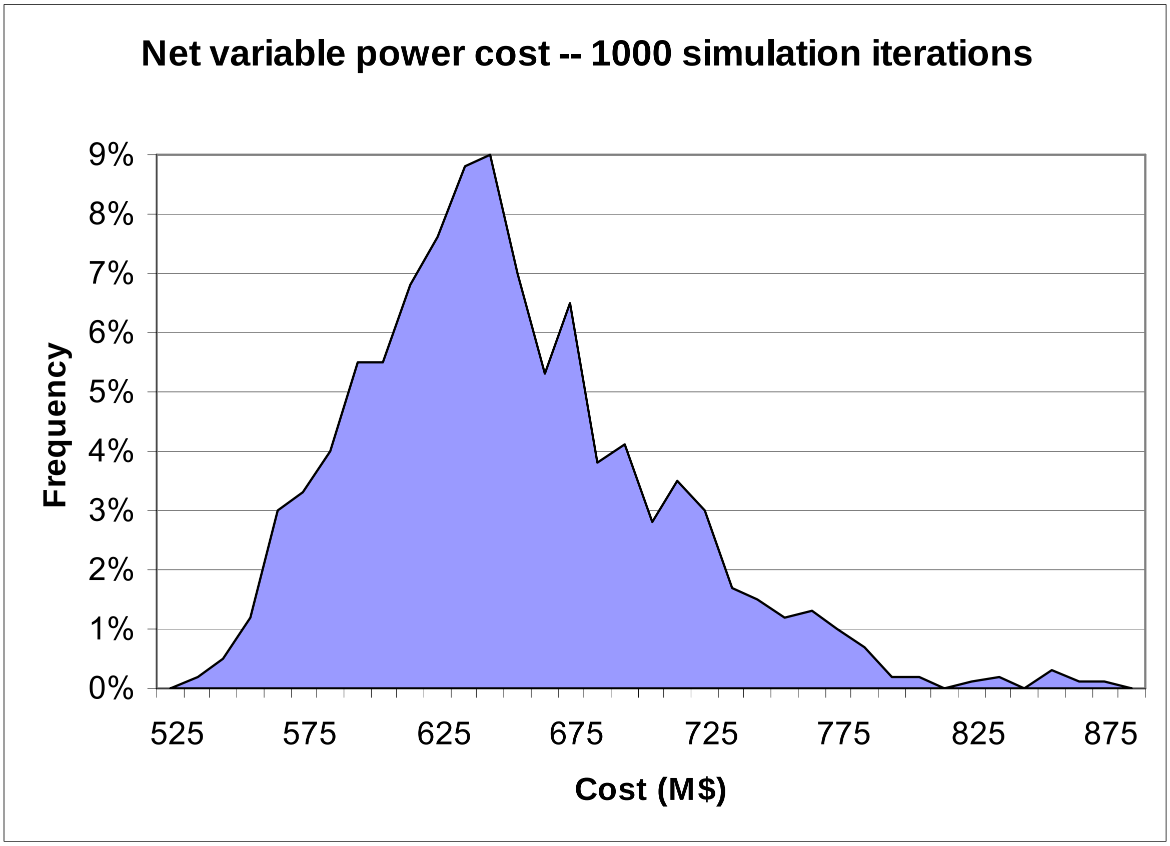
	Monet		Cost simulation prototype with Monet prices	
Resource	Cost (K$)	Energy(GWh)	Cost (K$)	Energy(GWh)
Coyote	66,645	1,183	64,667	1,073
Beaver	9,867	128	-	-
Mid-C	*	2,848	8,978	2,993
Boardman	35,484	2,867	14,223	2,760
Colstrip	14,133	2,087	20,181	2,278
PGE Hydros	4,293	1,992	4,788	2,056
Total market	509,347	8,528	528,105	9,104
Spot	93,256	1,454	528,105	9,104
Hydro contracts	38,759*	N/A	*	
Fwd/other	377,332	7,073	**	
Total	639,770	19,633	640,943	19,633

*-Monet does not really report a variable cost for the Mid-C plants; it reports only a fixed contract charge. That charge has been incorporated into the "market" number below. The simulation prototype, on the hand, reports the variable O&M costs for the Mid-C plants (and in fact it includes variable O&M in each plant's costs). The simulation does not report the fixed costs of the Mid-C contracts, or any other long-term contracts already in the Monet file. On the other hand, Monet does not report variable O&M.

**-Since there was no evolution of forward prices in the base case, all market prices are spot.

Some of the differences between the prototype simulation model and the Monet base case are due to modeling differences, for example the simplified Coyote dispatch. Others are due to its "prototype" nature; we had a limited amount of time to ensure a perfect match between the model and the base case.

Despite these differences, it is useful to examine the prototype results in more detail to see what kind of insights a model of this type might eventually be able to provide, keeping in mind its simplified and prototype nature. The following figure gives the histogram of results from a 1000-iteration run of the cost simulation prototype:



Visually the distribution appears quite skewed. Estimates of its statistical parameters are:

Median (M$)	643.73
Mean (M$)	650.86
Standard deviation (M$)	55.10
Skewness	0.697
(Excess) kurtosis	0.669

The last two items are dimensionless measures of the shape of the distribution. The positive skewness indicates that the distribution is asymmetric, with much more significant outliers to the right (higher costs). Thus its mean is to the right of (greater than) its median. The positive kurtosis indicates the distribution has a higher peak and "fatter tails" than a normal distribution, in other words, that the departures from the mean are likely to be larger than in a normal distribution.

It is often tempting, when facing a skewed distribution like the one above, to assume that it fits a lognormal rather than normal distribution (both the skewness and kurtosis of a normal distribution are zero). In fact this distribution is significantly more skewed and has fatter tails than a lognormal distribution: the corresponding parameters of a lognormal distribution with this mean and standard deviation are 0.255 (skewness) and 0.115 (kurtosis).

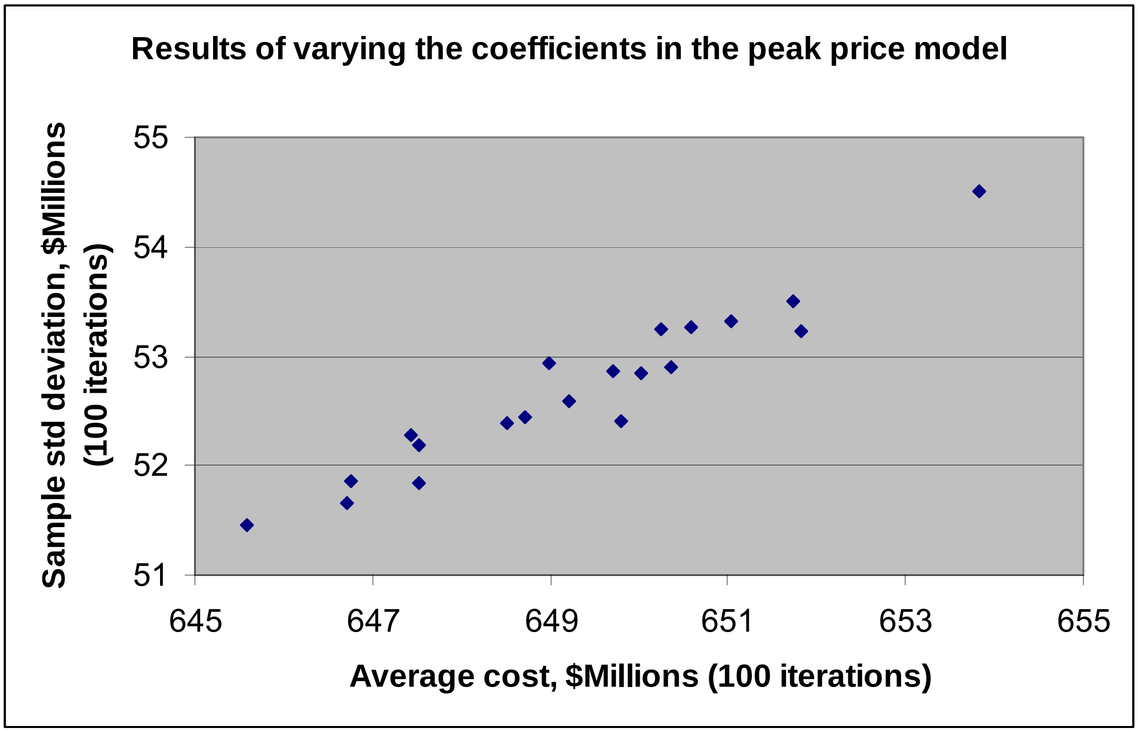
One must use these numbers with care. As noted earlier, many of the parameters underlying the simulation are only imprecisely estimated; in many cases the data used in computing the estimates are only proxies for the values whose properties were being estimated. Furthermore the model itself is a prototype constructed in such a way as to make it easier to test different estimation or approximation techniques. Section 3.3.1 drew a distinction between descriptive and prescriptive models; this prototype is of the descriptive class.

The numbers themselves are estimates of the statistics of an underlying cost distribution, not the statistics themselves, and they may be subject to bias or error. For example, the $55.10 million "standard deviation" is really the square root of the sample variance. The sample variance is generally an unbiased estimator of the variance – that is, tends neither to under nor overestimate it – but its square root is actually a biased estimator of the standard deviation, so that the standard deviation is probably somewhat more than $55.10 million.[4]

The $55.10 million also understates the variability in power costs because the simulation parameters are themselves uncertain. Recall that several of the input variables were represented using parametric statistical models to capture their interrelationships. The parameters were estimated from historical data, necessarily with some uncertainty. (The uncertainty in the estimation is also a proxy for the possibility that the wrong family of statistical models was used.)

The flexibility of the simulation prototype allows us to get a feel for the additional variability introduced by that parameter uncertainly. We simulated the impact of uncertainty in just one of the input models, the model of peak power prices (equation (10) on page 5-30). The table of coefficients for that equation indicates the standard error of each estimate. We ran twenty separate 100-iteration simulations using the cost simulation, with different values of those coefficients: for each run the coefficient values were chosen from normal distributions with the associated mean and variance. A scatter plot of the results follows:

[4] See, e.g., S. L. Sclove, "Concerning the Sample Standard Deviation," University of Illinois – Chicago, http://www.uic.edu/classes/idsc/ids571/samplvar.pdf as of 5/16/06. Specifically, if the sample variance is denoted S^2, so that its square root is S, and the (underlying) standard deviation is σ, then $E(S^2)=\sigma^2$ but $(E(S))^2 = E(S^2) - Var(S^2) < E(S^2)$.



The simulated costs from these twenty runs ranged from $645.6 million to $653.8 million. The uncertainty attributable to this parameter model, as well as the uncertainty in other parameter models, contributes positively to the uncertainty in net variable power cost.

This cost simulation model can *qualitatively* indicate the degree to which there is a greater-than-normal risk of bad outcomes (high costs). Here “normal” really means both “in a normal distribution” and “anticipated in the normal course of life”. Without using the specific numerical values produced by this simulation it is clear that there is significant cost risk: the distribution is quite positively skewed and leptokurtic. The prototype simulation model indicates that PGE’s risks are magnified relative to an estimate based on normal distributions.

The standard deviation of the distribution is $55 million, not a trivial sum even though it is only about 8.5% of the expected costs. Given the amount of hedging assumed in the model, 8.5% is a quite significant variation. It is almost as large as PGE’s total net income for 2005 ($64 million) or 2003 ($60 million as restated) and 60% of PGE’s net income for 2004 ($92 million).[5] Furthermore, the difference between the expected value of $650.86 million and the base case value of $640.94 is positive and statistically significant (p>.999). We had expected that the relationship between hydro conditions and price (in poor hydro conditions more load is exposed to high power prices) would move the mean, and we had been quite surprised when the effect did not show up in an earlier version of the prototype. That serves to demonstrate the dangers of drawing definitive conclusions from early versions of a mathematical model.

---

[5] Net income figures are from Portland General Electric’s Annual Report on Form 10-K filed March 16, 2006.

Although the power price model described in section 5.5 included gas and load as explanatory variables as well as hydro conditions, on a monthly or annual basis the only variable significantly correlated with power prices is hydro energy. The following table gives the correlations observed between the simulated series of various input variables. Recall that the prototype included some complex models for and dependencies among the input variables (sections 5.3-5.5).

CORRELATIONS OF MONTHLY AND ANNUAL AVERAGES OF UNCERTAIN VARIABLES

Month	Load & Peak price	Load & Gas price	Load & Hydro	Peak price & Gas price	Peak price & Hydro	Gas price & Hydro
January	18%	3%	-3%	19%	-85%	0%
February	11%	-1%	-2%	11%	-90%	-1%
March	14%	-1%	-3%	18%	-89%	-4%
April	13%	3%	-2%	17%	-89%	-1%
May	8%	4%	2%	15%	-90%	-1%
June	11%	0%	-3%	20%	-87%	-5%
July	18%	1%	0%	19%	-83%	-4%
August	22%	-2%	-3%	23%	-72%	1%
September	20%	-2%	-6%	15%	-74%	3%
October	9%	-3%	4%	17%	-64%	3%
November	25%	-3%	-2%	20%	-52%	4%
December	16%	-2%	1%	14%	-76%	1%
Total	10%	2%	-4%	14%	-96%	-2%

In order to understand the influence of various variables on the results of the cost simulation, we performed linear regressions of the simulated cost for each month on the average load in that month, average peak period power cost, average gas cost, and relative hydro energy. To put all variables on a common footing they were normalized to have mean zero, unit standard deviation (the means were subtracted and the result divided by the standard deviations). We also included interaction terms. Thus, for each month of the simulation we fit the model:

$$\begin{aligned}\text{Cost} = {} & a^*\text{Load} + b^*(\text{Peak Price}) + c^*(\text{Gas Price}) + d^*\text{Hydro} + e^*\text{Load}^*(\text{Peak price}) \\ & + f^*\text{Load}^*(\text{Gas price}) + g^*\text{Load}^*\text{Hydro} + h^*(\text{Peak price})^*(\text{Gas price}) + \\ & k^*(\text{Peak price})^*\text{Hydro} + m^*(\text{Gas price})^*\text{Hydro} + C_0 + \varepsilon\end{aligned}$$

where a, b, c, d, e, f, g, h, k and m are the model coefficients, C_0 is an intercept and ε is a normally distributed error. The following table gives the coefficient values and standard errors (in parentheses), where coefficients in **bold** have a statistically significant difference from zero (at the 95% confidence level) and coefficients in normal font do not:

Month	Avg Load	Peak Price	Gas Price	Hydro	Load * Peak price	Load * Gas price	Load * Hydro	Peak price * Gas price	Peak price * Hydro	Gas price * Hydro	Intercept
January	**2.03** **(0.14)**	**-0.85** **(0.31)**	**4.71** **(0.15)**	**-9.88** **(0.29)**	0.06 (0.27)	**0.66** **(0.15)**	-0.24 (0.28)	**-1.19** **(0.24)**	**-1.15** **(0.13)**	**-2.99** **(0.24)**	**68.30** **(0.18)**
February	**1.31** **(0.07)**	**-0.50** **(0.19)**	**3.07** **(0.07)**	**-9.72** **(0.18)**	-0.07 (0.17)	0.13 (0.07)	-0.06 (0.18)	-0.13 (0.10)	**-0.70** **(0.08)**	**-1.36** **(0.10)**	**52.56** **(0.09)**
March	**1.15** **(0.05)**	0.13 (0.12)	**1.51** **(0.05)**	**-5.44** **(0.11)**	-0.07 (0.12)	-0.06 (0.05)	-0.05 (0.11)	0.08 (0.06)	**-0.29** **(0.06)**	**-0.59** **(0.06)**	**55.58** **(0.06)**
April	**0.85** **(0.04)**	-0.08 (0.10)	**1.38** **(0.04)**	**-5.12** **(0.09)**	0.07 (0.09)	0.00 (0.04)	0.03 (0.09)	0.02 (0.05)	**-0.23** **(0.04)**	**-0.63** **(0.05)**	**42.90** **(0.04)**
May	**0.55** **(0.04)**	-0.05 (0.09)	**0.91** **(0.04)**	**-4.04** **(0.08)**	0.14 (0.08)	-0.01 (0.04)	0.06 (0.08)	0.00 (0.04)	**-0.13** **(0.03)**	**-0.43** **(0.04)**	**44.91** **(0.04)**
June	**0.58** **(0.03)**	-0.08 (0.07)	**0.82** **(0.03)**	**-3.66** **(0.06)**	0.08 (0.06)	**-0.10** **(0.03)**	0.03 (0.06)	0.06 (0.03)	**-0.19** **(0.03)**	**-0.32** **(0.03)**	**37.05** **(0.03)**
July	**1.40** **(0.05)**	**-0.52** **(0.10)**	**3.00** **(0.05)**	**-4.49** **(0.10)**	0.15 (0.10)	-0.08 (0.05)	0.09 (0.10)	0.03 (0.06)	-0.03 (0.05)	**-0.55** **(0.06)**	**54.54** **(0.06)**
August	**1.55** **(0.07)**	**-0.72** **(0.10)**	**4.26** **(0.07)**	**-2.95** **(0.10)**	-0.17 (0.10)	**-0.19** **(0.06)**	-0.11 (0.10)	**0.17** **(0.07)**	**-0.19** **(0.06)**	-0.11 (0.06)	**59.62** **(0.07)**
September	**0.94** **(0.06)**	**-0.19** **(0.09)**	**2.39** **(0.06)**	**-2.66** **(0.09)**	0.02 (0.08)	**-0.21** **(0.06)**	0.03 (0.09)	0.07 (0.05)	0.01 (0.06)	**-0.31** **(0.05)**	**55.69** **(0.06)**
October	**0.74** **(0.06)**	**-0.29** **(0.08)**	**3.62** **(0.07)**	**-2.31** **(0.09)**	0.13 (0.08)	-0.04 (0.07)	**0.23** **(0.08)**	**-0.17** **(0.07)**	-0.11 (0.08)	**-0.59** **(0.06)**	**54.69** **(0.07)**
November	**1.47** **(0.08)**	-0.02 (0.09)	**3.93** **(0.08)**	**-2.46** **(0.09)**	-0.06 (0.09)	0.01 (0.08)	0.10 (0.09)	-0.15 (0.09)	0.02 (0.08)	**-0.88** **(0.08)**	**56.32** **(0.08)**
December	**1.70** **(0.09)**	0.19 (0.15)	**4.83** **(0.09)**	**-5.44** **(0.15)**	0.06 (0.15)	**0.25** **(0.10)**	0.19 (0.14)	0.10 (0.13)	-0.05 (0.10)	**-1.07** **(0.13)**	**66.94** **(0.11)**
Total	**3.98** **(0.51)**	2.56 (2.05)	**29.56** **(0.54)**	**-39.20** **(1.98)**	1.81 (1.89)	-0.01 (0.53)	0.93 (1.90)	-0.99 (0.72)	**-1.29** **(0.48)**	**-7.01** **(0.70)**	**649.82** **(0.61)**

Note: values in parentheses are standard errors, not p values.

On a monthly basis, hydro condition is the most important determinant of net variable power cost. Although the coefficient on the interaction term between peak price and hydro energy is significant, it is still quite small compared with the intercept, so that the interaction will not have a large effect. In fact, the interaction between gas price and hydro conditions appears to be more important, especially in January and on an annual-average basis. Especially during the first half of the year, in low hydro conditions it appears that increased use of gas-fired plants may be limiting the impact of power price increases.

8. USES OF A COST SIMULATION MODEL TO SUPPORT THE REGULATORY PROCESS

During the course of this project, Portland General Electric staff explained to PA some of the context in which the project was initiated. Apparently there have been discussions between Portland General and the Public Utility Commission of Oregon about the variability in PGE's power costs, and whether it is appropriate for ratepayers to cover that variability, at least in part, through an annual true-up. If there were no true-up then PGE shareholders would bear that risk, providing cost insurance to ratepayers.

In addition to the public policy question of appropriateness, there are several analytic questions that one might try to address with a cost simulation model, e.g., how large is the risk and whether one can compute a risk adjustment to the revenue requirement to compensate the utility for bearing it. This project was initiated to determine how to structure a model that could answer the first of those questions, and whether such a model could also answer the second.

Utilities are generally given the opportunity to earn a return, but not a guarantee that the return will be earned. The return is put at risk to the utility's operational performance and to factors under the control of utility management. Whether fuel price risk, for example, is appropriately placed on the utility may depend on the tools the utility has or has not been given with which to mitigate it. Certain risks may just be too large for the utility reasonably to mitigate. In that case ratepayers, with greater overall financial resources, may appropriately be asked to bear the risk. A simulation model can help characterize the size of the risk.

Using our prototype model we have estimated the standard deviation, that is, the typical range of variation, in net variable power costs. We arrived at a figure of $55.1 million. This figure may not be accurate – it could easily be off by, say, $10 million either way. But we can still make a qualitative statement that the risk is quite sizable. At $55 million, the standard deviation is over half the company's net income in any of the last three years. Suppose the standard deviation of net variable power costs actually were $50 million. If the net variable power costs were normally distributed, there would be a 10% chance that the costs would exceed the net income in two of those three years. Because the cost distribution is positively skewed and fat-tailed the probability is actually greater than 10%.

Uncertainty in the estimate of standard deviation would make it very difficult to use the numerical results of this prototype for ratemaking or to determine a "risk adder". The same problem might apply to other simulation models. As with the Black-Scholes model it is actually quite difficult to calibrate a model that depends on distributional inputs. Therefore if a simulation should only be used for ratesetting if it is in an environment that permits rapid and frequent recalibration of the model and adjustment of the rates based on it.

9. SUMMARY AND CONCLUSION

PA Consulting Group developed a set of assumptions for a probabilistic cost simulation model. The role of the model would be to characterize the uncertainty in PGE's annual net variable power cost. There are many factors that influence net variable power cost, some of which are related to yet subtly different from others. To properly identify all those factors, and model their variability, would be an immense effort; furthermore, information about many of them is not easily available. PA produced a "data issues report" that described the data that would be desirable for such a model; the availability or unavailability of some data; and potential substitutes or proxy data as well as weaknesses in some of those substitutes.

PA concluded that a cost simulation model had to be flexible and easily modified to allow easy substitution of different submodels for uncertain inputs and resource dispatch. PA specified an architecture for such a "sandbox" model that relies on the @Risk add-in to run a set of spreadsheets in a Monte Carlo fashion. A central "coordinator" manages variable names to allow different component spreadsheets to be "plugged" in or out.

We proceeded to develop a prototype of the simulation model. In principle the "base case" of such a prototype, with all random variables set to nominal or expected levels, should produce answers identical to a reference model (in this case a specific Monet run). That did not happen here, partly because the prototype involved simpler dispatch logic for some resources and partly because of the unfinished nature of prototypes.

Even if the base case doesn't line up exactly with the reference model, a cost simulation model can still provide useful information about the distribution of costs. The base case values depend on specific input levels and a good match between specific numerical results of different models can be difficult to achieve. If the inputs are inaccurate, the specific numerical outputs – the locational rather than shape parameters of the distribution – will be undependable. A simulation model can confirm one's intuition about the cost impacts of the relationships among inputs, as well as the approximate magnitude of that impact – e.g., the fact that the expected cost exceeded the "base case" cost by approximately 18% of the standard deviation in costs.

A cost simulation can provide valuable qualitative information about the distribution of net variable power costs. The shape of the distribution of outputs depends on the shapes of the distributions of the inputs and the relationships between the inputs and outputs, that is, the mathematical properties of the model. In other words, distributional shape data encapsulates information about the assumed relationships between inputs and outputs, and those relationships should look the same even if the inputs themselves are inaccurate.

Under the assumptions of the prototype model we conclude that the distribution of net variable power costs is positively skewed (the mean is larger than the median) and leptokurtic (exhibits "fat tails", that is, somewhat elevated chances of extreme values). It is both more skewed and more leptokurtic than a parametric distribution often used to model costs, the lognormal distribution. On the other hand, the mean of the distribution does not seem to depart far from the base case cost. That means that it is not possible to "risk-adjust" the cost distribution (for instance to set a revenue requirement) just by moving the mean to account for correlation of inputs. The precision afforded by a descriptive model such as this is not fine enough to permit one to estimate a "risk adder" but we can say that there is significant variability in the costs.

The approach to cost simulation that PA has prototyped promises to help in the understanding of the way costs vary with uncertain inputs such as hydro conditions and market prices. It is particularly valuable because of its simplicity and flexibility. PA is willing to help PGE implement such an approach or to talk about ways to incorporate it into the Monet architecture.

APPENDIX A: COMPONENTS OF PROTOTYPE COST SIMULATION MODEL

This appendix briefly lists the components of the prototype cost simulation, Monet (base case) data used, other parameters, inputs (from other components) and outputs.

A.1 TEMPERATURE SIMULATION

A.1.1 Base Case data used

- Expected (normal) temperatures

A.1.2 Other parameters

- Description of randomness in temperature distribution

A.1.3 Inputs from other components

- None

A.1.4 Outputs

- Daily temperature and expected temperature

A.2 LOAD SIMULATION

A.2.1 Base Case data used

- Base case daily loads

A.2.2 Other parameters

- Description of randomness in load distribution
- Hourly load scaling factors

A.2.3 Inputs from other components

- Daily temperature

A.2.4 Outputs

- Hourly PGE load

A.3 GAS PRICE SIMULATION

A.3.1 Base Case data used

- Initial forward curves

A.3.2 Other parameters

- Parameters of price models

A.3.3 Inputs from other components

- None

A.3.4 Outputs

- Daily gas forward curve
- Daily gas spot price

A.4 HYDRO SIMULATION

A.4.1 Base Case data used

- None

A.4.2 Other parameters

- Historical distribution of hydro conditions

A.4.3 Inputs from other components

- None

A.4.4 Outputs

- Monthly hydro energy relative to average
- Monthly hydro capacity relative to average

A.5 MID-C POWER PRICE SIMULATION

A.5.1 Base Case data used

- Initial forward curves

A.5.2 Other parameters

- Parameters of price model

A.5.3 Inputs from other components

- Daily gas forward curve
- Daily gas prices
- Daily peak subperiod loads
- Hydro conditions

A.5.4 Outputs

- Daily on/offpeak forward power curve
- Daily on/offpeak spot power price

A.6 COST TO SERVE LOAD

A.6.1 Base Case data used

- None

A.6.2 Other parameters

- None

A.6.3 Inputs from other components

- Hourly loads
- Mid-C power prices

A.6.4 Outputs

- Cost to serve load from market (net variable power costs = cost to serve load from market minus value of production from other resources, plus costs of other resources)

A.7 PGE HYDRO SIMULATION

A.7.1 Base Case data used

- Base case energy
- Monthly, daily, hourly allocation factors
- VOM costs

A.7.2 Other parameters

- None

A.7.3 Inputs from other components

- Hydro energy relative to base case
- Mid-C spot power prices

A.7.4 Outputs

- PGE hydro plants' production in MWh
- Dollar value and cost of production (VOM) from PGE hydro plants

A.8 MID-C HYDRO SIMULATION

A.8.1 Base Case data used

- Base case energy
- VOM costs

- DLL code to optimize Mid-C dispatch
- Parameters of Mid-C optimization DLL routine

A.8.2 Other parameters

- None

A.8.3 Inputs from other components

- Hydro energy relative to base case
- Mid-C spot power prices
- MWh outputs of other PGE plants (for spinning reserve requirement)

A.8.4 Outputs

- Mid-C hydro plants' production in MWh
- Dollar value and cost of production (VOM) from Mid-C hydro plants

A.9 COLSTRIP SIMULATION

A.9.1 Base Case data used

- Monthly capacity, heat rate, maintenance schedule and forced outage rate by unit
- PGE ownership share
- Monthly coal prices
- VOM costs

A.9.2 Other parameters

- None

A.9.3 Inputs from other components

- Mid-C spot power prices

A.9.4 Outputs

- Colstrip production in MWh
- Dollar value and cost of production (fuel + VOM) from Colstrip plant

A.10 BOARDMAN SIMULATION

A.10.1 Base Case data used

- Monthly capacity and heat rate by state (min load / full load)
- Monthly maintenance schedule and forced outage rate

- PGE ownership share
- Monthly coal prices
- VOM costs

A.10.2 Other parameters

- None

A.10.3 Inputs from other components

- Mid-C spot power prices

A.10.4 Outputs

- Boardman production in MWh
- Dollar value and cost of production (fuel + VOM) from Boardman plant

A.11 BEAVER SIMULATION

A.11.1 Base Case data used

- List of operating states and allowable transitions
- Monthly capacity and heat rate by state
- VOM costs by state
- DLL code to optimize dispatch by dynamic programming

A.11.2 Other parameters

- None

A.11.3 Inputs from other components

- Spot gas prices
- Mid-C spot power prices

A.11.4 Outputs

- Beaver production in MWh
- Dollar value and cost of production (fuel + VOM) from Beaver plant

A.12 COYOTE SIMULATION

A.12.1 Base Case data used

- Simplified list of operating states and allowable transitions
- Monthly capacity and heat rate by state

- VOM costs by state
- DLL code to optimize dispatch by dynamic programming

A.12.2 Other parameters

- None

A.12.3 Inputs from other components

- Spot gas prices
- Mid-C spot power prices

A.12.4 Outputs

- Coyote production in MWh
- Dollar value and cost of production (fuel + VOM) from Coyote plant

A.13 FORWARD POWER CONTRACTING SIMULATION

A.13.1 Base Case data used

- None

A.13.2 Other parameters

- Description of purchase strategy

A.13.3 Inputs from other components

- Expected load, expected production from all resources (used to compute expected purchase requirement)
- Mid-C forward power prices

A.13.4 Outputs

- Daily forward purchases by tenor in MWh
- Dollar value and MTM of forward portfolio

A.14 SPOT MARKET SIMULATION

There is no spot market component in the design of the cost simulation as described in section 4. All energy produced or delivered is valued at the simulated spot price, that is, "marked to market". To the extent that simulated loads are greater than the generation simulated from all PGE's resources, power is implicitly bought on the spot market and priced at the simulated Mid-C spot power price.

I hearby certify that I have this day caused PGE's **REPORT ON THE FEASIBILITY OF USING STOCHASTIC MODELING IN THE ANNUAL UPDATE** to be served by electronic mail to those parties whose email addresses appear on the attached service list, and by First Class US Postal Mail, prepaid postage and properly addressed, to those parties on the attached service list who have not waived paper.

Dated at Portland, Oregon, this 31st day of August.

Patrick G. Hager
Manager, Regulatory Affairs

UE-180/UE-181/UE-184 Service List

Attorney

Jim Deason
Attorney at Law
1 SW Columbia St.
Suite 1600
Portland OR 97258-2014
Phone: 503-223-4335
Fax:
Email Add: jimdeason@comcast.net
Service: US Mail
Confidentiality: SPO

Boehm, Kurtz & Lowry

Kurt J. Boehm
Boehm Kurtz & Lowry
36 E. Seventh St., Ste 1510
Cincinnati OH 45202
Phone:
Fax:
Email Add: kboehm@bkllawfirm.com
Service: US Mail
Confidentiality: GPO

Michael L. Kurtz
Boehm, Kurtz & Lowry
36 East Seventh Street
Suite 1510
Cincinnati OH 45202
Phone: (513) 421-2255
Fax: (513) 421-2764
Email Add: mkurtz@bklawfirm.com
Service: US Mail
Confidentiality: GPO

BOMA Association of Building Owners and Managers

Ann Fisher
PO Box 25302
Portland OR 97298-0302
Phone: (503) 721-0181
Fax: (503) 223-2305
Email Add: energlaw@aol.com
Service: Electronic
Confidentiality:

BPA Bonneville Power Administration

Geoffrey M. Kronick
BPA
P.O. Box 3621
LC-7
Portland OR 97208-3621
Phone: (503) 230-4201
Fax: (503) 230-7405
Email Add: gmkronick@bpa.gov
Service: Electronic
Confidentiality: No

Craig Smith
Bonneville Power Administration
P.O. Box 3621-L7
Portland OR 97208-3821
Phone:
Fax:
Email Add: cmsmith@bpa.gov
Service: Electronic
Confidentiality: No

Brubaker & Associates, Inc.

James T. Selecky
Brubaker & Associates, Inc.
1215 Fern Ridge Pkwy
Suite 208
St. Louis MO 63141
Phone: (314) 275-7007
Fax: (314) 275-7036
Email Add: jtselecky@consultbai.com
Service: US Mail
Confidentiality: SPO

Cable Huston Benedict Haagensen & Lloyd LLP

Tamara Faucette
Cable Huston Benedict Haagensen & Lloyd LLP
1001 SW 5th Ave. Ste 2000
Portland OR 97204-1136
Phone: (503) 24-3092
Fax: (503) 224-3176
Email Add: tfaucette@chbh.com
Service: US Mail
Confidentiality: No

Chad M. Stokes
Cable Huston Benedict Haagensen & Lloyd LLP
1001 SW 5th Ave. Ste 2000
Portland OR 97204
Phone: (503) 224-3092
Fax: (503) 224-3176
Email Add: cstokes@chbh.com
Service: US Mail
Confidentiality: No

CADO Community Action Directors of Oregon

Jim Abrahamson
Oregon Energy Partnership/CADO
P.O. Box 7964
Salem OR 97301
Phone: 503-316-3951
Fax:
Email Add: jim@cado-oregon.org
Service: US Mail
Confidentiality: SPO

Constellation Newenergy, Inc.

William H. Chen
Constellation Newenergy, Inc.
2175 North California Blvd
Ste. 300
Walnut Creek CA 94596-4587
Phone: (925) 287-4500
Fax: (925) 287-4501
Email Add: bill.chen@constellation.com
Service: US Mail
Confidentiality: No

CUB Citizens' Utility Board of Oregon

Lowrey Brown
CUB
610 SW Broadway
Suite 308
Portland OR 97205
Phone:
Fax:
Email Add: lowrey@oregoncub.org
Service: US Mail
Confidentiality: No

Jason Eisdorfer
Citizens' Utility Board of Oregon
610 SW Broadway
Suite 308
Portland OR 97205
Phone: (503) 227-1984
Fax: (503) 274-2956
Email Add: jason@oregoncub.org
Service: US Mail
Confidentiality: GPO

Daniel W. Meek, Attorney at Law

Daniel W. Meek
Attorney at Law
Utility Reform Project
10949 SW Fourth Avenue
Portland OR 97219
Phone: (503) 293-9021
Fax: (503) 293-9099
Email Add: meek@meek.net dan@meek.net
Service: US Mail
Confidentiality: SPO

DOJ Department of Justice

Stephanie S. Andrus
Asst. Atty Gen.
Department of Justice
1162 Court St. NE
Salem OR 97310
Phone: (503) 378-4620
Fax: (503) 378-5300
Email Add: stephanie.andrus@state.or.us
Service: US Mail
Confidentiality: GPO

EPCOR Merchant and Capital (US) Inc.

Lorne Whittles
EMC - EPCOR Merchant and Capital Inc.
1161 W. River Street
Suite 250
Boise ID 83702
Phone: (208) 336-9733
Fax: (208) 247-0425
Email Add: lwhittles@epcor.ca
Service: US Mail
Confidentiality: No

Gresham (City of)

John Harris
Transportation Operations Superintendent
City of Gresham
1333 NW Eastman Pkwy
Gresham OR 97030
Phone: (503) 618-2907
Fax: (503) 667-6869
Email Add: john.harris@ci.gresham.or.us
Service: US Mail
Confidentiality: SPO

David Ris
Senior Assistant City Attorney
Gresham City Attorney's Office
1333 NW Eastman Pkwy
Gresham OR 97030
Phone: (503) 618-2507
Fax: (503) 667-3031
Email Add: david.ris@ci.gresham.or.us
Service: US Mail
Confidentiality: No

Hunt Technologies, Cellnet Technology, & Elster]

Scott Debroff
Smigel, Anderson & Sacks
River Chase Office Center
4431 North Front St.
Harrisburg PA 17110
Phone:
Fax:
Email Add: sdebroff@sallp.com
Service: US Mail
Confidentiality: No

ICNU Industrial Customers of NW Utilities

Magdalena Ackenhausen
Brubaker & Associates, Inc.
1215 Fern Ridge Pkwy
Suite 208
St. Louis MO 63141
Phone:
Fax:
Email Add:
Service: US Mail
Confidentiality: SPO

Brad Van Cleve
Attorney
Davison Van Cleve LP
333 SW Taylor, Ste. 400
Portland OR 97204
Phone: (503) 241-7242
Fax: (503) 241-8160
Email Add: mail@dvclaw.com
Service: US Mail
Confidentiality: GPO

Kafoury & McDougal

Linda K. Williams
Kafoury & McDougal
10266 SW Lancaster Rd
Portland OR 97219
Phone: (503) 293-0399
Fax: (503) 245-2772
Email Add: linda@lindawilliams.net
Service: US Mail
Confidentiality: SPO

League of Oregon Cities

Andrea Fogue
League of Oregon Cities
P.O. Box 928
Salem OR 97308
Phone: (503) 588-6550
Fax: (503) 399-4863
Email Add: afogue@orcities.org
Service: US Mail
Confidentiality: SPO

McDowell & Associates PC

Katherine A. McDowell
Attorney
McDowell & Associates PC
520 SW 6th Ave Ste 830
Portland OR 97204
Phone: (503) 595-3922
Fax:
Email Add: katherine@mcd-law.com
Service: US Mail
Confidentiality: No

David Tooze
City of Portland Office of Sustainable Development
721 NW 9th Ave.
Room 350
Portland OR 97209-5311
Phone: (503) 823-5311
Fax: (503)
Email Add: dtooze@ci.portland.or.us
Service: Electronic
Confidentiality: No

Benjamin Walters
Deputy City Attorney
City of Portland- Office of City Attorneys
1220 SW 5th Ave.
Room 315
Portland OR 97204
Phone:
Fax: (503) 823-3089
Email Add: bwalters@ci.portland.or.us
Service: Electronic
Confidentiality: GPO

Preston Gates & Ellis LLP

Harvey P. Spigal
Preston Gates & Ellis LLP
222 SW Columbia Street
Suite 1400
Portland OR 97201-6632
Phone: (503) 228-5788
Fax: (503) 248-9085
Email Add: hspigal@prestongates.com
Service: US Mail
Confidentiality: No

Sempra Energy Solutions

Theodore E. Roberts
Sempra Energy
101 Ash Street
HQ 13D
San Diego CA 92101-3017
Phone: (619) 699-5111
Fax: (619) 699-5027
Email Add: troberts@sempra.com
Service: Electronic
Confidentiality: No

Linda Wrazen
Regulatory Policy Manager
Sempra Global
101 Ash Street
HQ 8C
San Diego CA 92101-3017
Phone: (619) 696-4411
Fax: (619) 696-2500
Email Add: lwrazen@sempraglobal.com
Service: Electronic
Confidentiality: No