EXHIBIT NO. ___(MDR-3C)
DOCKET NO. UE-13____
2013 PSE PCORC
WITNESS: MATTHEW D. RARITY

BEFORE THE
WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION, **Complainant,** **v.** **PUGET SOUND ENERGY, INC.,** **Respondent.**	**Docket No. UE-13____**

SECOND EXHIBIT (CONFIDENTIAL) TO THE
PREFILED DIRECT TESTIMONY OF
MATTHEW D. RARITY
ON BEHALF OF PUGET SOUND ENERGY, INC.

REDACTED
VERSION

APRIL 25, 2013

Hour-Ahead Balancing Model Methodology

Overview

The Hour-Ahead Balancing Model ("HABM") is a SAS-based program designed to estimate the cost of providing balancing reserve capacity to manage load and wind volatility. The model is iterative and scenario driven, capable of utilizing an unlimited number of unique economic dispatch and price scenarios from production cost models, such as AURORA, and then iteratively adjusting unit dispatch to provide capacity to balance the net fluctuations in load and wind over each scheduling hour. Additionally, the model allows for the evaluation of alternative portfolio scenarios, either through the addition or subtraction of wind, thermal, or hydro assets.

Balancing Reserves

Balancing reserve capacity is defined as the amount of generating capacity required to be held in reserve to balance the differences between real-time load and wind generation from their hourly fixed schedule. Balancing capacity can be separated into two types: "INC" capacity, which is unloaded capacity that can be deployed as incremental energy, and "DEC" capacity, which is the available capacity on a generating resource that can be utilized to decrease energy output. INC capacity can be further sub-divided into spinning and non-spinning capacity. Spinning capacity comes from unloaded capacity on a resource that is online and synchronized to the system. Non-spinning capacity is unloaded capacity that is not online or synchronized to the system, but can be available within 10 minutes.

Balancing reserves can be calculated for wind assets alone, for load alone, and for net load, which allows deviations in load and wind to net against each other. For example, if wind generation is 10 MW above schedule and load is simultaneously 20 MW greater than schedule, on net the system requires 10 MW of incremental capacity to be deployed, rather than holding and deploying two separate amounts of reserves for wind and load individually. This effect is not confined to the relationship between wind and load, and can exist between any number of generating resources or loads.

In addition to providing balancing reserve capacity to balance hourly load and wind, PSE must hold capacity as contingency reserves, as part of its obligation as a Balancing Authority. As opposed to balancing capacity which is continuously deployed to balance load and wind, contingency reserves can only deployed in specific circumstances, called qualifying events. These include events such as sudden losses of generation or the loss of non-interruptible imports. Contingency reserves require additional spinning and non-spinning INC capacity to be set aside each hour, beyond the amount set aside for hourly balancing reserve needs.

Dispatch Scenarios

To simulate the cost of procuring balancing reserve capacity under various market and hydro conditions, the HABM incorporates hourly dispatch output from AURORA. Utilizing the same 70 AURORA scenarios in PSE's 2013 PCORC, each scenario reflects a unique set of dispatch conditions regarding power prices and hydro volatility, which in turn impact economic unit dispatch and the subsequent cost of re-dispatching assets to provide balancing capacity. Model inputs include:

- Hourly power and gas prices;
- Hourly dispatch for PSE gas-fired assets;
- Hourly generation from PSE's Mid-Columbia ("Mid-C") projects;
- Asset operating characteristics (heat rate, capacity, etc); and
- Hourly balancing reserve capacity values

The initial dispatch state from the unaltered AURORA dispatch is referred to as "economic" dispatch. From this dispatch, adjustments are made to ensure, at every hour, enough INC and DEC capacity is available to meet hourly balancing capacity requirements. These adjustments potentially lead to units being run below their economically efficient generation point, being turned on when market prices dictate doing so is otherwise uneconomical, or shifting hydro generation across the day. The incremental cost (or at times, revenue) of altering the economic dispatch is computed at an hourly level, allocated to load and wind (see 'Allocating HABM Costs' section), and aggregated to the annual level. The expected cost of providing hour-ahead balancing reserve capacity for Wild Horse across the 70 AURORA simulations is reported as the hour-ahead wind integration cost in the 2013 PCORC.

Model Methodology

The HABM takes an iterative approach to adjusting the economic dispatch to provide balancing reserve capacity rather than a full-fledged optimization. Each iteration is layered to reflect a least-cost approach to procuring reserves, as well as operational preferences and constraints when modifying unit output. As a first step, the hourly hydro and gas-fired dispatch are scanned to ascertain the existing availability of INC and DEC capacity from the economic dispatch. For gas-fired units, this includes further distinction based on their maintenance schedules and other outages. In hours where the economic dispatch does not provide sufficient capacity the model will first pass through the hydro dispatch to make adjustments.

Hydro Dispatch

When altering the hydro dispatch, changes are subject to three constraints: dispatch cannot exceed PSE's contractual capacity on the Mid-C, dispatch cannot dip below the plant's minimum generation level, and total hydro generation must be maintained on a daily level. The daily generation requirement is a compromise between fully modeling PSE's hydro storage, yet recognizing real operational constraints prevent hydro generation from constantly being pushed in a given direction, potentially depleting or exceeding available hydro storage. The daily requirement allows for hydro generation to be flexibly shifted between hours, while ensuring the net daily change to dispatch, and therefore storage, is zero. Additionally, the daily generation constraint controls for spring runoff periods when continuously high generation levels cannot be shifted and Mid-C flexibility is severely limited.

For a given hour with a shortage of INC or DEC capacity, the hydro dispatch will be shifted upward or downward in an effort to provide the necessary capacity. For example, in an hour where more DEC capacity is needed, the model will attempt to shift hydro dispatch up, subject to not violating the INC capacity requirement and PSE's maximum generation limit on the Mid-C. In addition, to meet the daily generation requirement, the increase in hydro dispatch in the example hour must be balanced with a decrease in dispatch in another hour within the day. The model will look to do this in the most economical hour, subject to not violating any hour's reserve requirement.

Gas-Fired Generation Dispatch

If there are still hours with unmet INC or DEC capacity requirements after making adjustments to the hydro dispatch, the remaining unmet capacity requirement will be shifted to PSE's gas-fired dispatch. Here, the model will make asset-by-asset passes through the dispatch, based on a user-defined unit stack. Presently, units are ordered based on their economic merits and fast-start capabilities. As the stack is user-defined, the model allows for controlling which units can provide balancing reserves, the order they provide reserves, and offers the ability to insert or subtract units into the stack to assess their impact on providing balancing capacity.

In adjusting the gas-fired dispatch, the primary constraint is the operating range of each unit. Dispatch cannot exceed the plant's maximum generation level ("baseload") and dispatch cannot be below the plant's minimum generation level. The minimum and maximum generation levels offer the two extreme dispatch points for the model, and a third possible dispatch point is located at the midpoint of the two extremes. The model will only dispatch a unit to one of these three generation levels, or leave the unit offline.

While these dispatch limitations apply to all gas-fired assets, combined-cycle combustion turbines ("CCCT") face additional dispatch constraints. CCCTs require lengthy start-up and shutdown times which are generally planned well in advance of the operating hour, and also must meet minimum runtime and minimum downtime requirements. As the

HABM is designed to reflect the procurement of reserve capacity in the hour-ahead time frame for use in the following operating hour, to reflect these operational constraints the model will only make CCCT dispatch changes if the unit is already economically dispatched. The model will not start a CCCT if it is offline.

This limitation is not placed on PSE's simple cycle combustion turbines ("SCCT"). As every unit in PSE's SCCT fleet is available to start in less than one hour, the option to start an offline SCCT to meet balancing capacity requirements is available to the model. Potential SCCT starts are limited by outage events.

With these constraints in mind, the model will move through each gas-fired unit's economic dispatch, making adjustments downward if INC capacity is required, dispatching units at higher levels if additional DEC capacity is required, or starting SCCTs, if necessary.

At present, PSE's modeled gas-fired resource stack is listed below in order of increasing heat rate. [redacted] are placed at the bottom of the stack as they are not 10-minute ready and cannot respond as quickly to balancing capacity needs.

1. [redacted]
2. [redacted]
3. [redacted]
4. [redacted]
5. [redacted]
6. [redacted]
7. [redacted]
8. [redacted]
9. [redacted]

REDACTED

Calculating Hour-Ahead Wind Integration Costs

The total cost of providing balancing reserve capacity for the PSE system is a combination of opportunity and real costs. For the gas-fired units, there are real fuel costs associated with changes in dispatch and where that dispatch level lies on the unit's heat rate curve. Opportunity costs arise as hydro generation is shifted throughout the day, from on-peak to off-peak hours, and gas-fired units are withheld from the market when market prices dictate they should be committed. To maintain load-resource balance, the net change in PSE resource generation for each hour is offset by an equal and opposite market purchase or sale.

To reflect PSE's obligation to hold both contingency reserves and balancing reserve capacity, two HABM models are run. The first uses only PSE's contingency reserve obligation as the capacity reserve requirement, and the second uses the sum of both the contingency reserves and balancing reserves for load and wind. The total cost from the first run is subtracted from the total cost of the second run, to remove the costs associated with providing contingency reserves for the PSE Balancing Authority Area from the costs of providing balancing reserve capacity for hourly load and wind integration.

Once reduced to only the costs to provide balancing reserve capacity, a portion of this cost is allocated to wind resources based on wind's contribution to total volatility (see 'Allocating HABM Costs' section). The wind allocation of balancing reserve costs is further allocated to each wind resource, based each wind resource's share of total wind capacity, measured as the total installed nameplate capacity of wind facilities balanced in the PSE Balancing Authority Area ("BAA"). As Wild Horse is the only PSE-owned wind facility in the PSE BAA, the Wild Horse share is the hour-ahead wind integration cost recovered in the 2013 PCORC.

Allocating Wind Balancing Reserve Costs

Load and wind each contribute to net system volatility and the aggregate need for balancing reserve capacity. However, while each in isolation may have its own balancing reserve requirements, PSE does not balance each resource individually. Rather PSE balances the net deviations in load and wind from their respective schedules. At the system level, the balancing reserve requirement is not equal to the sum of each resource's reserve requirement. At times load may be extremely volatile while wind resources are generating at constant output, vice-versa, or they move in an offsetting manner.

This diversity in moment-to-moment fluctuations is beneficial in that less balancing capacity is required when aggregating both resources. To allocate the costs equitably, a methodology inspired by Brendan Kirby and Eric Hirst (2000)[1] is used and modified using mathematical properties of variance and covariance. The results of this methodology are identical to the methodology employed by the Bonneville Power Administration ("BPA") to allocate resources their share of the net balancing reserve capacity held by BPA, termed Incremental Standard Deviation ("ISD").

Equation 1: Composite Variance of Net Load

$$\sigma_N^2 = \sigma_W^2 + \sigma_L^2 + 2cov(W, L) \qquad (1)$$

[1] B. Kirby, E. Hirst, "Customer-Specific Metrics for the Regulation and Load-Following Ancillary Services" (Oct. 2000)

Where the subscripts 'N', 'W', and 'L' represent 'net load', 'wind', and 'load' respectively, σ is the standard deviation of the 10-minute level forecast errors, and $cov(W, L)$ is the covariance of the wind and load forecast errors.

Equations 2&3: Covariance Properties

$$cov(X + Y, Z) = cov(X, Z) + cov(Y, Z) \quad (2)$$

$$cov(X, X) = \sigma_X^2 \quad (3)$$

Using these properties and substituting with equation 1, a resource's share of net cost becomes a relationship involving the resource's covariance with the net load and the net load error variance.

Equations 4&5: Wind & Load Shares of Net Volatility

$$\text{Wind share} = \frac{cov(W, N)}{\sigma_N^2} \quad (4) \qquad \text{Load share} = \frac{cov(L, N)}{\sigma_N^2} \quad (5)$$

While more computationally intensive than several alternative methods, this methodology thrives in properly allocating balancing capacity costs based on each resource's contribution to net system volatility. Most critically, this methodology will produce identical results regardless of how resources are aggregated together. In a hypothetical scenario with three wind facilities and a single load, this methodology will assign the same share to load, regardless of whether the calculation is done as 4 shares (for each wind facility and load), or one aggregate wind share and one load share.

Model Output

The model provides summary data for each scenario on an annual level. The total cost is broken down by asset type: hydro, CCCT, and SCCT. The total cost is further broken down to the share for providing reserves for wind, allocated based on the source's contribution to net system volatility. The model also reports the amount of unmet capacity and hours, if any. A unit breakdown is also provided for the gas-fired generating resources, reporting the number of starts and run hours in the economic dispatch and the final dispatch. These results are further broken down by on- and off-peak hours.