EXH. CD-6
DOCKETS UE-22___/UG-22___
2022 PSE GENERAL RATE CASE
WITNESS: DR. CHHANDITA DAS

BEFORE THE
WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION, **Complainant,** **v.** **PUGET SOUND ENERGY,** **Respondent.**	**Docket UE-22____** **Docket UG-22____**

FIFTH EXHIBIT (NONCONFIDENTIAL) TO THE PREFILED DIRECT TESTIMONY OF

DR. CHHANDITA DAS

ON BEHALF OF PUGET SOUND ENERGY

JANUARY 31, 2022

2017 Net Metering Load Research Sample Design and Deployment

Puget Sound Energy (PSE) selected and deployed its net metering customer class load research samples in December 2017. The number of net metering customers has been growing continuously and 5,676 net metering accounts were served by PSE by the end of April 2017. Among those accounts, only 4,878 accounts had a full 12-month billing history. Some of the net metering customers have a single account for energy delivered and returned but multiple accounts for production meters. Therefore, the number of customers eligible for sampling was reduced further to 4,681 after those multiple accounts were aggregated for the same customer. Among them, there were 65 non-core net metering customers classified into Solar/Wind Hybrid, Wind, and Community Solar Generation segments. The samples for each non-core segment were selected manually based on individual customers' billing data so that the resulting sample distribution would be similar to the segment population distribution. Excluding those 65 non-core net metering customers, 397 customers were sampled from the population of 4,616 net metering customers to collect their 15-minute interval load data for the amount of electric energy produced by customer, the amount delivered by PSE to customer and the amount returned by customer to PSE. 186 of those 397 sampled customers were selected on the basis of annual amount of energy delivered, while the remaining 211 customers were sampled on the basis of annual amount returned. The new samples have been producing 15-minute interval load data since January 1, 2018 and will be monitored regularly for their reliability.

The following sections discuss the customer and billing data analyzed for sampling, the statistical methods and analyses performed for sample design, and the initial and final sampling results.

Data

The monthly billing data of net metering customers for the twelve-month period ending April 2017 were analyzed and evaluated to develop a sample design. The purpose of sample design was to draw unbiased samples to collect their 15-minute interval load data for the net metering customer class load profiling.

Statistical Methods and Analyses

A net metering customer is viewed as not only a consumer PSE is obligated to serve but also PSE's potential power supplier. Therefore, the sample design for net metering load research was developed in two ways, using two different criteria: (1) annual total energy delivered by PSE and (2) annual total energy returned to PSE. The net metering class was segmented into nine different customer groups:

1) Residential (Rate Schedule 07) Solar with Battery Storage,
2) Small General Service (Rate Schedule 24) Solar with Battery Storage,
3) Residential Solar with no Battery Storage,
4) Small General Service Solar with no Battery Storage,
5) Medium General Service (Rate Schedule 25) Solar,
6) Large General Service (Rate Schedule 26) Solar,
7) Residential Wind,
8) Residential Solar/Wind Hybrid, and
9) Community Solar.

Since the samples for Customer Segments #5 through #9 were selected manually by analyzing the customers' billing data to simulate the segment population distribution, sample designs were developed only for Customer Segments #1 through #4.[1]

A major obstacle in developing a sample design for the net metering customer class load research is that no actual hourly load data is available for the targeted customers. The class load research sample design produced in this study was developed by analyzing the population statistics of electric billing data of the net metering customers for their annual delivered and returned kWh's for a twelve-month period ending April 30, 2017. The population statistics were calculated only for the customers who were active at the end of April 2017 and had a full 12-month history of billings. A stratified Mean per Unit (MPU) estimation approach was adopted for the net metering load research sample design. The MPU approach assumes a strong positive correlation between hourly loads and annual kWh's across the net metering customers.

To perform a stratified sampling for each of the four customer segments, population data of customers and their annual delivered (or returned) kWh volumes in each segment were sorted by their relative sizes of annual delivered (or returned) kWh volumes. In each case of sample design, a frequency table was created by assigning the customers into appropriate kWh brackets. The number and ranges of kWh brackets were pre-defined on the basis of population size and density of its kWh distribution. The frequency table was then stratified by 3 to 4 groups by combining the kWh brackets and their frequencies. For efficiency and practicality, the most popular numbers of strata being used for a stratified sampling are 2 to 5. For the number of strata set for each of the customer segments, strata boundaries were determined following the Dalenius-Hodges (DH) procedure, also known as the cumulative square root of (uf) procedure where "u" denotes kWh width of a given usage bracket and "f" stands for frequency of the usage bracket. In the DH procedure, strata boundaries are set to have each of the strata yield a similar sum of square-rooted kWh values of (uf).

The Neyman Optimum Allocation formula was then used to calculate total number of sample units for each customer segment and the strata sample sizes. The Neyman Allocation formula estimates total and strata sample sizes required for a given error margin and confidence limit on the basis of the population statistics of number of customers and their kWh use. For each customer segment, total and strata sample sizes were determined by using the following formulae:

$$n = (\sum W_h S_h)^2 / ((d/t)^2 + 1/N * \sum W_h S_h^2)$$

$$n_h = n * (W_h * S_h) / \sum (W_h * S_h)$$

Where N = Total Population size
N_h = Stratum population size
n = Total sample size
n_h = Stratum sample size
$W_h = N_h / N$
S_h = Stratum standard deviation of annual kWh use
d = Tolerable error margin; and

[1] *There were 12 solar customers under Rate Schedule 25, 1 solar customer under Rate Schedule 26, 24 residential wind customers, 11 residential solar/wind hybrid customers, and 17 community solar customers. 9 Rate Schedule 25 customers, 1 Rate Schedule 26 customer, 8 Wind customers, 5 Solar/Wind Hybrid customers, and 15 Community Solar customers were sampled to produce 15-minute interval load readings. However, eight of those customers had meter performance issues and had to be dropped from the final list of the net metering customers selected for 15-minute interval load reading. Therefore, the final sample sizes are 6 for Rate Schedule 25, none for Rate Schedule 26, 6 for Wind, 4 for Hybrid, and 13 for Community Solar.*

t = t-statistic value for pre-set confidence limit

The examples presented below illustrate how the stratified sample sizes were determined for the "Delivered kWh" case of the "Residential Solar with no Battery Storage" customer segment through the two-step procedure explained above. The first step was to set strata boundaries through the DH procedure. At the beginning, all of the net metering customers in the targeted customer segment were sorted by sizes of their annual kWh delivered. A frequency table was created by counting the number of usage points (f) sorted to each of the kWh brackets ascending with an increment (u) of 2,500 kWh. The sum of square rooted value of (uf) for all of the usage brackets was 10,866 kWh. The number of strata set for sampling was 4. The strata boundaries were then determined by making the cut-offs so that all of the four strata yield a sum of square rooted (uf) close to one fourth of 10,866 kWh. The strata boundaries set for the "Residential Solar with no Battery Storage" sample design were 0 – 7,500 kWh for the first stratum, 7,501 – 12,500 kWh for the second, 12,501 – 25,000 kWh for the third, and all of the customers with annual usage over 25,000 kWh assigned to the fourth stratum.

Step 1

Dalenius-Hodges Procedure

Load Profile	Bracket Lower	Bracket Upper	Frequency (f)	u	uf	sqrt(uf)	Cum sqrt(uf)	% Distribution by Stratum
Residential Solar PV No Storage	0	2500	95	2,500	237,500	487	487	
Residential Solar PV No Storage	2501	5000	655	2,500	1,637,500	1,280	1,767	
Residential Solar PV No Storage	5001	7500	824	2,500	2,060,000	1,435	3,202	37.2%
Residential Solar PV No Storage	7501	10000	693	2,500	1,732,500	1,316	4,519	
Residential Solar PV No Storage	10001	12500	550	2,500	1,375,000	1,173	5,691	29.4%
Residential Solar PV No Storage	12501	15000	381	2,500	952,500	976	6,667	
Residential Solar PV No Storage	15001	20000	531	2,500	1,327,500	1,152	7,819	
Residential Solar PV No Storage	20001	25000	239	2,500	597,500	773	8,592	27.2%
Residential Solar PV No Storage	25001	30000	112	2,500	280,000	529	9,121	
Residential Solar PV No Storage	30001	35000	58	2,500	145,000	381	9,502	
Residential Solar PV No Storage	35001	40000	31	2,500	77,500	278	9,781	
Residential Solar PV No Storage	40001	45000	14	2,500	35,000	187	9,968	
Residential Solar PV No Storage	45001	50000	10	2,500	25,000	158	10,126	
Residential Solar PV No Storage	50001	55000	9	2,500	22,500	150	10,276	
Residential Solar PV No Storage	55001	60000	5	2,500	12,500	112	10,388	
Residential Solar PV No Storage	60001	65000	8	2,500	20,000	141	10,529	
Residential Solar PV No Storage	65001	70000	4	2,500	10,000	100	10,629	
Residential Solar PV No Storage	70001	75000	3	2,500	7,500	87	10,716	
Residential Solar PV No Storage	75001	80000	2	2,500	5,000	71	10,786	
Residential Solar PV No Storage	80001	85000	0	2,500	0	0	10,786	
Residential Solar PV No Storage	85001	90000	0	2,500	0	0	10,786	
Residential Solar PV No Storage	90001	100000	1	2,500	2,500	50	10,836	
Residential Solar PV No Storage	100001	110000	0	2,500	0	0	10,836	
Residential Solar PV No Storage	110001	120000	0	2,500	0	0	10,836	
Residential Solar PV No Storage	120001	130000	0	2,500	0	0	10,836	
Residential Solar PV No Storage	130001	140000	0	2,500	0	0	10,836	
Residential Solar PV No Storage	140001	150000	1	2,500	2,500	50	10,886	6.1%
Residential Solar PV No Storage	Total		4,226					100%

2,722	4	=Number of Strata
2,722	1	Stratum 1
5,443	2	Stratum 2
8,165	3	Stratum 3
10,886	4	Stratum 4

The second step was to calculate the minimum number of samples required to meet the sample design criteria of 5% error margin with 95% confidence limit for each residential rate schedule and to allocate the total number of

samples to each stratum. The example provided below illustrates how the population statistics of customers and their annual kWh delivered by PSE were used in the Neyman Optimum Allocation formula to determine the customer segment total and strata sample sizes. The components of the formula calculated with the population statistics are shown in columns of the tables provided below. The final results for the customer segment total and strata sample sizes are presented in column "n(h)."

Step 2

Neyman Allocation for Stratified Sampling
(Total Sample Size Required for 5% Error Margin @ 95% Confidence Limit)

RC *7- Solar PV No Storage*

Min	Max	N(h)	W(h)	Y(h)	S(h)	W(h)*S(h)	n(h)	S(h)^2	W(h)*S(h)^2
0	7500	1,574	0.3725	5,004	1,539	573	17	2,369,549	882,553
7501	12500	1,243	0.2941	9,792	1,452	427	13	2,107,470	619,873
12501	25000	1,151	0.2724	17,082	3,351	913	28	11,226,977	3,057,797
25001	150000	258	0.0611	35,870	13,850	846	26	191,824,709	11,711,021
Total		4,226	1.0000	11,586	8,805	2,758	84		16,271,245

Total sample size required = 83
(@ 5% error margin w/ 95% confidence limit)

Note: While the statistically-required minimum sample size is 83, the final sample size is set to be 84 due to rounding.

Sample Design

As illustrated above for the "Residential Solar with no Battery Storage" sample design, use of the Dalenius-Hodges procedure and the Neyman Optimum Allocation method produced the following sample designs with 5% error tolerance at 95% confidence limit for each of the delivered kWh and the returned kWh cases by customer segment:

Delivered kWh Case

RC *7- Solar PV No Storage*

Min kWh	Max kWh	N(h)	n(h)
0	7,500	1,574	17
7,501	12,500	1,243	13
12,501	25,000	1,151	28
25,001	150,000	258	26
Total		4,226	84

RC *7- Solar PV Yes Storage*

Min kWh	Max kWh	N(h)	n(h)
0	7,500	30	3
7,501	20,000	61	14
20,001	170,000	21	21
Total		112	38

RC *24 - Solar PV No Storage*

Min kWh	Max kWh	N(h)	n(h)
0	12,500	70	6
12,501	30,000	99	13
30,001	60,000	57	13
60,001	220,000	37	35
Total		263	67

RC *24- Solar PV Yes Storage*

Min kWh	Max kWh	N(h)	n(h)
0	7,500	5	2
7,501	25,000	6	2
25,001	200,000	4	1
Total		15	5

Returned kWh Case

RC *7- Solar PV No Storage*

Min kWh	Max kWh	N(h)	n(h)
0	2,000	1,394	22
2,001	4,000	1,358	21
4,001	7,000	1,191	26
7,001	32,000	283	34
Total		4,226	103

RC *7- Solar PV Yes Storage*

Min kWh	Max kWh	N(h)	n(h)
0	1,000	47	6
1,001	4,000	52	20
4,001	14,000	13	13
Total		112	39

RC *24 - Solar PV No Storage*

Min kWh	Max kWh	N(h)	n(h)
0	1,250	77	4
1,251	4,500	80	12
4,501	9,500	73	17
9,501	38,500	33	33
Total		263	66

RC *24- Solar PV Yes Storage*

Min kWh	Max kWh	N(h)	n(h)
0	1,500	5	4
1,501	4,000	5	4
4,001	8,000	5	4
Total		15	12

Sample Drawing

A systematic sampling technique was used to draw the samples for each stratum of the targeted customer segment. The following procedures were followed to select the samples from the segment population list of customers and their annual billing data:

1) Sort the population in ascending order of annual delivered (or returned) kWh and stratify the sorted data by the same kWh strata as was determined in the sample design.
2) Count the number of customers in each kWh bracket (or stratum).
3) Calculate $K = N_h/n_h$ where N_h = population number of customers in stratum h and n_h = number of samples as required by the sample design for stratum h.
4) Draw stratum samples by taking the $(K/2)^{th}$ customer from the top of the stratum population list as the first sample and every k^{th} customer from the first and the samples drawn in sequence until the total number of samples drawn reaches the number of sampling units determined in the sample design.
5) Repeat 3) and 4) to draw the samples for the remaining strata.

Taking the "Delivered kWh" case of the "Residential Solar with no Battery Storage" customer segment as an example, the table below illustrates how the samples were drawn:

Sample Selection

RC 7- Solar PV No Storage

Min	Max	N(h)	n(h)	N(h)/n(h)	1st Sample	Obs. # for 1st Sample	Obs. # for Last Sample
0	7500	1574	17	93	46	46	1,527
7501	12500	1243	13	96	48	1,622	2,769
12501	25000	1151	28	41	21	2,838	3,948
25001	150000	258	26	10	5	3,973	4,221
Total		4,226	84				

Sample Validation

The samples drawn for each of the customer segment were checked against any expected problems with their meter reading and data communication capabilities by evaluating the quality of their previous daily meter readings. When a significant problem was detected, the corresponding sample was replaced with a customer in the same customer segment whose annual amount of delivered (or returned) kWh is similar to the original sample customer's. When an appropriate replacement is not available, the original sample with meter reading problems was simply dropped out of the final sample list. Once the sample list was revised, statistical representation of the final samples was validated by evaluating the percentage difference of sample mean from population mean. The tables provided below list the final sample size and the sample mean percentage difference by kWh bracket and customer segment for the delivered kWh and the returned kWh cases:

Validation of Samples Selected for Delivered kWh Case

Residential Return No Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-2000	1,574	17	5,004	5,002	-0.04%
2001-4000	1,243	13	9,792	9,788	-0.04%
4001-7000	1,151	27	17,082	17,234	0.89%
7001-32000	258	26	35,870	35,351	-1.45%
Total	4,226	83	11,586	11,594	0.07%

Residential Return Yes Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-1000	30	3	5,360	5,351	-0.16%
1001-4000	61	13	12,655	11,983	-5.31%
4001-14000	21	18	45,165	42,250	-6.45%
Total	112	34	16,797	15,882	-5.45%

Commercial Return No Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-1250	70	6	5,789	5,652	-2.36%
1251-4500	99	13	20,854	20,706	-0.71%
4501-9500	57	13	41,324	40,792	-1.29%
9501-38500	37	32	96,174	97,552	1.43%
Total	263	64	31,877	31,863	-0.04%

Commercial Return Yes Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-1500	5	2	4,812	5,051	4.98%
1501-4000	6	2	15,866	15,831	-0.22%
4001-8000	4	1	79,371	70,360	-11.35%
Total	15	5	29,116	26,779	-8.03%

Total Number of Samples for Delivered kWh Case **186**

Validation of Samples Selected for Returned kWh Case

Residential Return No Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-2000	1,394	22	978	976	-0.18%
2001-4000	1,358	21	2,962	2,958	-0.11%
4001-7000	1,191	26	5,240	5,235	-0.09%
7001-32000	283	34	10,235	10,043	-1.87%
Total	4,226	103	3,437	3,421	-0.46%

Residential Return Yes Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-1000	47	6	369	365	-1.20%
1001-4000	52	20	2,068	1,977	-4.39%
4001-14000	13	11	6,207	6,020	-3.01%
Total	112	37	1,835	1,770	-3.58%

Commercial Return No Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-1250	77	4	293	278	-5.20%
1251-4500	80	12	2,864	2,800	-2.25%
4501-9500	73	17	6,635	6,552	-1.25%
9501-38500	33	27	16,954	16,293	-3.90%
Total	263	60	4,926	4,796	-2.64%

Commercial Return Yes Battery	Population Count	Sample Count	Pop. Mean kWh	Sample Mean kWh	% Difference
0-1500	5	4	384	480	25.00%
1501-4000	5	3	2,346	2,569	9.51%
4001-8000	5	4	6,030	5,678	-5.83%
Total	15	11	2,920	2,909	-0.37%

Total Number of Samples for Returned kWh Case **211**

For two customer segments in the delivered kWh case where the population size is small and the kWh variances are high, the total sample mean difference from the population mean is greater than ±5% of the error tolerance limit set initially for sample design. Since the error margins calculated for most of the customer segments and the kWh brackets are well below the ±5% limit and there is no easy solution to remove the sampling bias remained in those few brackets, the sample lists were finalized and were sent to Landis & Gyr and PSE's Meter Data Warehouse (MDW) for deployment. In the future, it may be considered to take the whole population, instead of sampling, for some of those problem kWh brackets where the population size is so small and is not much larger than the required sample size.