



# CASCADE NATURAL GAS LOAD STUDY ANALYSIS

---

DESIGN DOCUMENT

SEPTEMBER 2022

This document describes the data sources and analysis techniques that Cascade is using for its load study along with the results of the load study.

## TABLE OF CONTENTS

Overview .....	3
Data and Data Sources .....	3
Customer Usage Data .....	3
Customer Count Data .....	3
Weather Data .....	4
Sample Size .....	4
Estimation Techniques .....	9
Regression .....	9
Current Analysis and Final Results .....	10

## OVERVIEW

The purpose of this document is to discuss the data gathering, estimation techniques, supporting documentation, and the current load study results for Cascade Natural Gas Corporation (“Cascade” or “Company”). The objective of the load study is to quantify and assign peak and annual responsibility by comparing individual rate schedule demands to system demands. Once Cascade has performed its load study for all customer groups, Cascade will be able to assign service costs for any class or customer based on their contribution to the system peak.

## DATA AND DATA SOURCES

The data inputs for Cascade’s load study include customer usage data, customer count, and weather data.

### CUSTOMER USAGE DATA

Cascade analyzed customer usage data in its load study analysis to determine the appropriate allocation for each rate schedule’s usage annually and on the peak day.

Cascade’s customer usage data for the load study is being captured at the per customer per day by weather area level and will be gathered through Cascade’s online database called ThoughtSpot. Since the usage data is captured at that level, the rate schedule will be easily identifiable for this analysis.

### CUSTOMER COUNT DATA

Cascade analyzed customer count data to determine the representative sample for the load analysis.

For each set of customer usage data, there is one customer behind that data. The remaining number of customers will be determined by utilizing Cascade’s customer count file maintained by Cascade’s accounting department. If forecasted customer

counts are needed, Cascade will utilize the forecasted customer counts from the Integrated Resource Plan (IRP).

## WEATHER DATA

Cascade utilized National Oceanic and Atmospheric Administration (NOAA) to gather daily weather data information. This data is gathered at seven weather locations: four in Washington and three in Oregon. The four in Washington are Bellingham, Bremerton, Yakima, and Walla Walla. The three in Oregon are Baker City, Redmond, and Pendleton. The data obtained from NOAA were actual and normal wind and actual and normal weather. Normal wind is defined as the average daily wind speed and normal weather is defined as the average daily temperatures of the most recent 30 years of historical data which results in the average annual temperatures as well.

The Company uses a heating degree day (HDD) as the unit of measure for temperature. HDD is calculated by taking the average temperature from a day and subtracting it from a reference temperature. If the reference temperature less HDD is negative, then the Company gives that day a 0 value for HDD. The Company uses 60°F as the reference temperature. For example, a 50°F day will result in 10 HDDs (60-50).

## SAMPLE SIZE

Cascade has completed the analysis to determine the representative sample size for each class and weather location. Determining the representative sample at the weather location level ensures the Company there is a representative sample of customers across the different geographic locations. Cascade utilized Confidence Intervals and Relative Precision to determine the appropriate sample size. PURPA suggests that electrical utilities should maintain a minimum 90 percent confidence interval with a 10 percent accuracy at time of class and system peaks.<sup>1</sup> Cascade

---

<sup>1</sup> Load Research Manual, 3<sup>rd</sup> edition, 2017, page 61: "...The desired accuracy should be determined for the study. The desired accuracy is usually expressed as a relative precision with a given confidence level. A design accuracy of +/- 10% at the 90% confidence level (often abbreviated "90/10") at the system and class peak time was specified in

believes these values can be utilized for natural gas as well. In early 2021, Cascade ran an initial sample size count where Cascade pulled a week’s worth of data during the coldest week of 2021, February 9 through February 15. Table 1 provides the known population (actual customer count) and the sample count, which represents the number of Fixed Networks (FN) reads, during the period listed above.

**Table 1: Known Population and FN Sample Count**

			9-Feb-21	10-Feb-21	11-Feb-21	12-Feb-21	13-Feb-21	14-Feb-21	15-Feb-21
<b>Weather Location</b>	<b>RS_CD</b>	<b>Known Population (February 2021)</b>	Sample Count						
Bellingham	CNGWA503	85954	12666	13063	13292	13062	12287	11564	11495
Bellingham	CNGWA504	10390	2070	2129	2149	2093	1990	1854	1864
Bellingham	CNGWA505	196	31	33	34	33	31	25	24
Bellingham	CNGWA511	26	6	6	6	6	6	6	6
Bremerton	CNGWA503	40730	4664	4735	4885	5029	4828	4513	4429
Bremerton	CNGWA504	5062	1166	1166	1174	1209	1200	1157	1131
Bremerton	CNGWA505	58	17	16	17	17	17	17	17
Bremerton	CNGWA511	19	4	4	4	4	3	3	3
Walla Walla	CNGWA503	42784	3611	3942	3815	3779	3990	3634	3269
Walla Walla	CNGWA504	5408	1522	1574	1518	1491	1525	1435	1406
Walla Walla	CNGWA505	54	12	12	12	12	12	11	10
Walla Walla	CNGWA511	17	5	5	4	3	2	1	1
Yakima	CNGWA503	28578	5503	5760	5684	5865	5906	5942	5819
Yakima	CNGWA504	6221	2155	2200	2174	2210	2210	2200	2167
Yakima	CNGWA505	177	43	43	43	43	40	45	42
Yakima	CNGWA511	31	8	8	8	8	7	8	8

Cascade has since re-analyzed the sample count and pulled data from February 12, 2022. February 12 was chosen so the Company can compare the sample counts exactly one year apart. Table 2 shows Cascade has increased the number of sample counts for every rate schedule across all four climate zones in Washington.

---

the Public Utility Regulatory Policy Act (PURPA) for all major rate classes (see Chapter 1 for additional background). This means that the sample will be designed to achieve a 90% confidence interval that is no more than 10% above and below the estimate of the mean.<sup>1</sup> Although these federal standards were withdrawn in 1992, the PURPA specification remains load research common practice, particularly for samples that will be used to support rate cases or other regulatory requirements...”

Table 2: Updated FN Sample Count

RS_CD	Weather Location	Sample Count
CNGWA503	Bellingham	24323
CNGWA504	Bellingham	3613
CNGWA505	Bellingham	70
CNGWA511	Bellingham	14
CNGWA503	Bremerton	7308
CNGWA504	Bremerton	1864
CNGWA505	Bremerton	23
CNGWA511	Bremerton	6
CNGWA503	Walla Walla	11240
CNGWA504	Walla Walla	2468
CNGWA505	Walla Walla	22
CNGWA511	Walla Walla	8
CNGWA503	Yakima	10638
CNGWA504	Yakima	3543
CNGWA505	Yakima	106
CNGWA511	Yakima	18

With the information gathered, Cascade utilized the average, standard deviation, and variance of the actual data along with the z-score for the 90 percent confidence intervals and the 10 percent precision of accuracy, which enabled the Company to calculate a simple random sample-mean-per-unit estimation. The simple random sample-mean-per-unit estimation is a representative sample count estimate based on known average, variance, confidence interval, and precision of accuracy. The formula is:

$$n^0 = \frac{Z^2 * s_y^2}{D^2 * \bar{y}^2}$$

Model Notes:

- $n^0$  = Simple Random Sample – mean – per – unit estimation
- $Z$  = Z – score based on 90 percent confidence interval
- $s$  = variance of  $y$
- $D$  = desired relative precision
- $\bar{y}$  = average use per customer by rate schedule and weather location

Table 3 shows the average, standard deviation, variance, and the simple random sample-mean-per-unit estimation for each weather location and rate schedule for February 9, 2021.

**Table 3: Simple Random Sample-Mean-Per-Unit Estimation**

		9-Feb-21			
Weather Location	RS_CD	Average of Sample	Std Dev of Sample	Variance of Sample	Simple Random Sample-Mean-Per-Unit Estimation
Bellingham	CNGWA503	4.091887467	2.298500747	5.283105682	85
Bellingham	CNGWA504	15.91251765	30.39203721	923.6759258	987
Bellingham	CNGWA505	84.2498571	114.1713456	13035.09616	196
Bellingham	CNGWA511	257.6603383	160.3683081	25717.99423	26
Bremerton	CNGWA503	3.607587832	2.289821685	5.243283349	109
Bremerton	CNGWA504	19.42605161	37.11893037	1377.814992	988
Bremerton	CNGWA505	52.24799059	48.30301843	2333.18159	58
Bremerton	CNGWA511	205.1663875	87.99347814	7742.852195	19
Walla Walla	CNGWA503	3.831241626	2.335362947	5.453920093	101
Walla Walla	CNGWA504	20.81029619	40.51222152	1641.240092	1026
Walla Walla	CNGWA505	74.1282725	111.2028688	12366.07802	54
Walla Walla	CNGWA511	398.0032	242.3198873	58718.92776	17
Yakima	CNGWA503	3.726891392	3.174017324	10.07438597	196
Yakima	CNGWA504	21.60580361	40.65904151	1653.157657	958
Yakima	CNGWA505	110.8799049	114.4932253	13108.69865	177
Yakima	CNGWA511	369.5092738	179.1723757	32102.74021	31

Comparing the simple random sample-mean-per-unit estimation to the actual sample count allows the Company to determine if there is a representative sample for that rate schedule and weather location. If the actual count is greater than the simple random sample-mean-per-unit estimation, then it is determined there is a representative sample. Cascade calculated the average sample count from the February 9 through February 15 data and compared it to the average simple random sample-mean-per-unit estimation for each rate schedule and weather location. Table 4 shows those results.

Table 4: Representative Sample Size Results

Weather Location	RS_CD	Known Population (February 2021)	Average Sample Count	Average Simple Random Sample-Mean-Per-Unit Estimation	Do we have a representative sample?
Bellingham	CNGWA503	85954	12490	87	Yes
Bellingham	CNGWA504	10390	2021	1003	Yes
Bellingham	CNGWA505	196	30	196	No
Bellingham	CNGWA511	26	6	26	No
Bremerton	CNGWA503	40730	4726	117	Yes
Bremerton	CNGWA504	5062	1172	837	Yes
Bremerton	CNGWA505	58	17	58	No
Bremerton	CNGWA511	19	4	19	No
Walla Walla	CNGWA503	42784	3720	107	Yes
Walla Walla	CNGWA504	5408	1496	832	Yes
Walla Walla	CNGWA505	54	12	54	No
Walla Walla	CNGWA511	17	3	12	No
Yakima	CNGWA503	28578	5783	178	Yes
Yakima	CNGWA504	6221	2188	879	Yes
Yakima	CNGWA505	177	43	177	No
Yakima	CNGWA511	31	8	30	No

As seen in Table 4, the results from the 2021 load analysis show that rate schedules 503 and 504 currently have a representative sample size to complete a load study. The results also show that rate schedules 505 and 511 do not have a representative sample. Referring back to Table 3, Cascade has only increased the number of fixed networking so rate schedules 503 and 504 are still high enough to have a representative sample. However, for rate schedules 505 and 511 the rate schedules are still below the necessary representative sample size which will be discussed further in the results section.

The equipment needed to complete the Cascade fixed network installations has been significantly delayed for over 13 months due to ongoing supply shortages. The delay in equipment has moved the estimated completion of the Cascade fixed network in Washington State to the end of 2023. Cascade will continue to work through supply chain issues to install more of the fixed network devices in the service territory to increase coverage of the fixed network and to capture more data for the rate schedule 505 and 511 customers. Although rate schedules 505 and 511 do not have a representative sample, a load study can still be complete with the caveat that the results

are not within the 90 percent confidence intervals and the 10 percent precision of accuracy.

## ESTIMATION TECHNIQUES

Cascade had several techniques to choose from when analyzing a sample set of data. Each technique provides a different analysis with the same end goal result which is to determine the peak allocator within Cascade's cost of service study. However, each technique provides varying results so deciding which technique to use was important. Cascade tested regression models, mean-per-unit models, and ratio models. Cascade ultimately chose the regression model as it is a more robust model that uses statistical analysis to build a relationship between weather and usage. The regression technique is described below:

### REGRESSION

Regression models are most commonly used to describe a relationship between explanatory variables and response variables.

Model:

$$\frac{\text{Therms}}{C_{\text{class}}} = \alpha_0 + \alpha_1 \text{HDD}^D + \text{Fourier}(k) + \text{ARIMA}\epsilon(p, d, q)$$

Model Notes:

- $\frac{\text{Therms}}{C_{\text{class}}} = \text{Usage by customer class.}$
- $\text{HDD}^D = \text{Daily Heating Degree Days from Weather Location.}$
- $\text{ARIMA}\epsilon(p, d, q) = \text{Indicates that the model has } p \text{ autoregressive terms, } d \text{ difference terms, and } q \text{ moving average terms.}$
- $\text{Fourier Term} = \text{Indicates } k \text{ number of seasonalities within dataset}$

Cascade begins by running a simple linear model regressing on HDDs and wind by month. Then, the residuals are analyzed using the Durbin-Watson test to check for autocorrelation. If found, the model then adds an ARIMA term and a Fourier term. In

order to run Fourier terms alongside ARIMA terms, the ARIMA term must be forced to not difference the data. Cascade determined that each model did not have autocorrelation issues, so ARIMA and Fourier terms were not utilized. Cascade pursues this process for each rate schedule and weather zone.

## CURRENT ANALYSIS AND FINAL RESULTS

Cascade gathered daily meter reads from November 17, 2020, to June 18, 2022. Due to data size issues, Cascade had to lump certain customers together who had the same town code. The data was then analyzed at the rate schedule level and regressed against HDDs utilizing NOAA weather data. Below is a sample of the data gathered. Data was gathered at the daily level, by rate schedule, by town (then allocated to weather station), by usage (in therms), and total meters. Cascade was able to then calculate a therm per meter value. This data can be found in Table 5 as well as Exhibit 1.

**Table 5: Load Study Data**

GAS_DAY	RS_CD	DISTRICT_TOWN_ID	Weather Location	Weather Data	Total Therms	Total # of Meters	Therms/Meter
11/17/2020	CNGWA503	47531	Bremerton	8.5	349.799	196	1.784688776
11/17/2020	CNGWA503	47488	Bremerton	8.5	96.495	49	1.969285714
11/17/2020	CNGWA503	47482	Bremerton	8.5	2.978	1	2.978
11/18/2020	CNGWA503	47531	Bremerton	14	408.585	198	2.063560606
11/18/2020	CNGWA503	47488	Bremerton	14	115.497	50	2.30994
11/18/2020	CNGWA503	47482	Bremerton	14	6.453	1	6.453
11/19/2020	CNGWA503	47488	Bremerton	15	88.656	45	1.970133333
11/19/2020	CNGWA503	47531	Bremerton	15	391.558	183	2.139661202
11/19/2020	CNGWA503	47482	Bremerton	15	10.742	2	5.371

After the data was gathered, Cascade utilized R, a statistical analysis software, to run linear regression models using HDDs to forecast each rate schedule by weather zone. Each weather zone and rate schedule provide an intercept coefficient estimate as well as a weather coefficient estimate. In Cascade's results, each p-value shows that it is statistically significant. Table 6, as well as Exhibit 2, shows the regression results for the weather locations and rate schedules.

**Table 6: Load Study Regression Results**

Zone and Rateclass	Estimate	Std. Error	t value	Pr(> t )	R.Squared	ME	RMSE	MAE	MAPE	MASE	1 Year Total Demand Forecast
Bremerton.CNGWA503.(Intercept)	0.420939476	0.016862848	24.96253807	1.65E-129	0.717243422	3.67E-17	0.68302791	0.462998471	Inf	0.442403383	27,666,027
Bremerton.CNGWA503.Weather.Data	0.148955212	0.001341562	111.0312076	0	0.475863199	0	0	0	0	0	-
Bremerton.CNGWA504.(Intercept)	2.878985069	0.120282294	23.93523575	3.48E-119	0.288886893	5.26E-16	4.676270565	2.837425994	86.17794299	0.563887581	15,930,061
Bremerton.CNGWA504.Weather.Data	0.613160063	0.009657476	63.49071504	0	0.410324145	0	0	0	0	0	-
Bremerton.CNGWA505.(Intercept)	8.389037796	0.794820367	10.55463366	1.53E-25	0.752566748	-1.27E-14	23.81006444	16.16176457	Inf	0.777782894	574,441
Bremerton.CNGWA505.Weather.Data	2.09847558	0.063824424	32.87887993	2.90E-199	0.138631141	0	0	0	0	0	-
Bremerton.CNGWA511.(Intercept)	80.73276369	2.737434042	29.49213112	5.35E-145	0.019101903	-1.44E-15	54.9375551	49.41492251	57.99897294	0.798735185	976,870
Bremerton.CNGWA511.Weather.Data	6.384058345	0.218305181	29.24373269	3.91E-143	0.39802445	0	0	0	0	0	-
Bellingham.CNGWA503.(Intercept)	0.590233634	0.012411802	47.55422645	0	0.627950637	-7.19E-17	0.651552316	0.406435568	Inf	0.390393602	57,173,075
Bellingham.CNGWA503.Weather.Data	0.124519602	0.000859423	144.8875131	0	0.600064659	0	0	0	0	0	-
Bellingham.CNGWA504.(Intercept)	1.827156534	0.278979155	6.549437473	6.23E-11	0.378335412	9.85E-16	14.08196716	5.087397659	Inf	0.769731993	30,169,404
Bellingham.CNGWA504.Weather.Data	0.62138526	0.01947732	31.90301629	3.15E-207	0.024188359	0	0	0	0	0	-

To verify results, Cascade applied the actual customer counts, see Exhibit 3, from June 2022 along with normal NOAA HDDs, see Exhibit 4, for an entire year’s worth of data to calculate the estimated daily load by rate schedule by weather zone rolled up to an annual total. Once the results were estimated, Cascade then compared the annual load study results to the average of the three previous years’ actuals. For rate schedules 503 and 504, the results were within a reasonable range from actuals. As for rate schedules 505 and 511, the results were well below the past 3-years of actuals. This is due to the high variance between the customers within those rate schedules, requiring a larger sample size of customers to have a representative sample. Cascade applied current customer counts along with normal HDDs for an entire calendar year to the regression models to calculate annual usage. Table 7 provides the load study results as well as a comparison to 2019-2021 actuals.

**Table 7: Load Study Annual Results**

Weather Zone	Rate Schedule	2019	2020	2021	Average	Load Study Results (therms)
Bellingham	CNGWA503	58,951,560	58,965,281	58,315,686	58,744,176	57,173,075
Bellingham	CNGWA504	29,611,710	27,978,823	28,677,722	28,756,085	30,169,404
Bellingham	CNGWA505	4,163,636	4,004,601	3,542,179	3,903,472	2,458,477
Bellingham	CNGWA511	3,886,734	3,722,799	3,682,732	3,764,088	1,258,040
Bremerton	CNGWA503	26,838,665	26,243,199	26,617,183	26,566,349	27,666,027
Bremerton	CNGWA504	18,242,941	16,725,935	17,459,700	17,476,192	15,930,061
Bremerton	CNGWA505	1,550,749	1,485,041	1,450,144	1,495,311	574,441
Bremerton	CNGWA511	2,782,889	2,779,275	2,421,589	2,661,251	976,870
Walla Walla	CNGWA503	26,244,218	23,516,988	22,852,138	24,204,448	25,691,511
Walla Walla	CNGWA504	21,816,226	17,912,896	18,405,043	19,378,055	17,992,344
Walla Walla	CNGWA505	1,232,062	946,887	930,259	1,036,403	786,680
Walla Walla	CNGWA511	3,464,647	3,183,626	3,015,806	3,221,360	1,041,646

Yakima	CNGWA503	18,673,362	16,471,441	15,893,273	17,012,692	14,251,465
Yakima	CNGWA504	24,567,571	20,606,072	20,986,534	22,053,392	22,342,573
Yakima	CNGWA505	6,207,542	5,151,807	5,330,374	5,563,241	2,939,760
Yakima	CNGWA511	4,989,951	4,795,853	5,063,320	4,949,708	2,340,785

For peak day results, Cascade utilized December 21, 1990, HDDs as it is the coldest day in Cascade's recorded history. Below are the NOAA HDD values for that day by weather station. Table 8 provides the peak HDD for each weather location.

**Table 8: Peak Day HDD**

Location	HDD60
Yakima	58
Walla Walla	66
Bremerton	41.5
Bellingham	45.5

Similar to the annual results, Cascade applied current customer counts along with the peak HDDs to the regression models to calculate what a peak day event would look like if peak temperatures happened under current customer counts. Table 9 provides the Load Study Peak Day Results. The peak day results reflect what the Company would expect usage to be on a single day if a peak day event occurred under current estimated use-per-customer coefficients and actual customer counts. Peak day usage is utilized in Cascade's Integrated Resource Plan, mainly for upstream pipeline modeling with Gas Supply and downstream pipeline modeling with Engineering, to determine the resource needs to serve customers during the highest usage day. Estimating the class core responsibilities on peak day is important as the Company's pipeline system is built to meet peak day loads.

**Table 9: Load Study Peak Day Results**

Weather Zone	Rate Schedule	Peak Day (therms)
Bellingham	CNGWA503	542,197
Bellingham	CNGWA504	315,029
Bellingham	CNGWA505	13,932
Bellingham	CNGWA511	11,806
Bremerton	CNGWA503	274,997
Bremerton	CNGWA504	143,070
Bremerton	CNGWA505	5,347
Bremerton	CNGWA511	6,568
Walla Walla	CNGWA503	391,754
Walla Walla	CNGWA504	242,125
Walla Walla	CNGWA505	9,796
Walla Walla	CNGWA511	5,418
Yakima	CNGWA503	156,583
Yakima	CNGWA504	228,481
Yakima	CNGWA505	21,268
Yakima	CNGWA511	18,795

Cascade believes the annual rate schedules 503 and 504 results fall within the 90 percent confidence and the 10 percent precision of accuracy. Although the rate schedules 505 and 511 fall outside the 90 percent confidence and the 10 percent precision of accuracy, Cascade believes the Company can achieve these levels once more fixed networking is applied to the remaining customers who do not have fixed network meters under the schedule 505 and 511 rates. Cascade is also confident that as the Company gathers more data over time, the confidence and precision of accuracy of these regression results will only continue to improve.