Southern California Gas Residential End-Use Model

I. Residential End-Use Model Description

Introduction:

SoCalGas used the End Use Forecaster model to generate annual gas demand forecasts for the residential market. The software's market segmentation and end-use modeling framework analyzes the impacts of competitive strategies (gas vs. electricity) and market scenarios on gas demand and market shares. The model separates the residential market into five building types (B-level).

These groups are identified by the premise code classification found in the company billing files. The five residential groups are:

- Single-Family(SF);
- Multi-Family <= 4 units (MF2);
- Multi-Family > 4 units (MF3);
- Master Metered (MM); and
- Sub-Metered (SM).

The residential model identifies eight end-uses (N-level) that are the primary drivers of natural gas demand:

- Space heating;
- Water heating;
- Cooking;
- Drying;
- Pool heating;
- Spa heating;
- Fireplace; and
- Barbeque.

The model assumes two fuel choices (F-level) for end-uses:

- Natural gas; and
- Electricity.

The model assumes up to four efficiency levels (E-level) for the various end-uses. In general, the efficiency levels are:

- Stock;
- Standard;
- High efficiency; and
- Premium efficiency.

See Figure 1 for a classification of the number of efficiency levels for each end-use by customer segment type.

A set of post-model adjustments were applied to the model's annual demand forecast. The first adjustment calibrates to the recorded 2017 weather-adjusted demand. Next, the annual forecast was parceled out to a series of monthly forecasts by a process which involves two steps. These two steps consist of (1) using the fitted equation for customer demand to generate a forecast of use per customer that varies with the number of calendar days and heating degree days in a given month and (2) calculating a series of weights based on the customer's predicted monthly usage share in total annual consumption. The shares obtained from the latter step were then applied to annual totals to derive the stream of monthly forecasts which are conditional on the particular weather design specification for the entire year. An adjustment to the forecast offsets the throughput by the energy efficiency savings. Annual conservation benefits associated with AMI are estimated by SoCalGas to represent 1% of the core gas throughput in the post deployment period. The residential load was reduced by the expected AMI energy savings for customers with non-AMI meters.

Figures 2-5 illustrate the monthly forecasts for each weather scenario.

Data Sources:

The information used to perform the modeling and to generate the forecast includes historical 2017 consumption and customer counts; meter counts, growth, and decay; use per customer by vintage and unit energy consumption (UEC) values; fuel costs and price elasticity; equipment capital costs and availability; building and equipment lives and decay. The historical 2017 data is in Figure 6.

Meter Counts, Growth and Decay:

Regression equations were developed for each of the 5 building types. The meter count forecast is a company-specific forecast based on actual meter counts within the SoCalGas service territory. Data on meter decay rates were obtained from the Energy Information Administration (EIA). See Figure 7 for the meter forecast used as an input to the End-Use Model.

Use Per Customer by Vintage and UEC:

Use per customer and Unit Energy Consumption (UEC) data were based on company marketing data and the California Measurement Advisory Council. See Figure 8 for the appliance UEC's.

Fuel Costs and Price Elasticity:

Average and marginal gas prices ($/therm) were calculated from forecasts of the residential rate components. Residential rates have two consumption tiers. We used the simple average of the second tiers' projected monthly prices for each forecast year as the marginal rate. The marginal rate was used for each housing segment type.

For a given housing segment type, the average gas commodity rate was calculated using a pair of weights for the two consumption tiers applied to the simple average of each tier's monthly rate. The average commodity rate in each forecast year was developed using the same consumption tier weights, but with the forecasts of rates for each residential rate tier. The average gas price each year was then calculated by including the non-volumetric customer charges with the year's average gas commodity price. Figure 9 illustrates the gas price forecasts.

Electric Price Data:

Both average prices (cents/kWh) and marginal prices (cents/kWh) were developed as electricity price inputs. Forecasts for the SCE residential customer class were developed based on the California Energy Commission's December 2017 updated forecast rates for California energy demand (forecast for the SCE planning area, under "Mid-Case" demand for electricity) for the SCE service area through our forecast time horizon.

To impute average electricity prices to each residential housing type, we simply calculated the ratio of the housing type's average gas price to the overall residential gas price for each housing type, then multiplied by the overall average electricity price.

The marginal prices for each residential housing type were calculated by multiplying each year's respective average price by a ratio. These ratios were 1.513 for the SF, MF2 and MF3 housing types, 1.034 for the MM housing type and 1.125 for the SM housing type. These various ratios ware estimated from analyses of SCE Schedule D rate schedule for housing types SF, MF2 and MF3; SCE Schedule DM for housing type MM; and SCE Schedule D as applied to sub-metered buildings for housing type SM. Copies of these rate schedules were obtained from the SCE web-site in March 2006.

Equipment Capital Costs and Availability:

Data on equipment capital costs and availability were from EIA, the Residential Appliance Saturation Survey (RASS), Energy Star (EPA & DOE), and SoCalGas company data. See Figures 11 and 12 for gas and electric appliance equipment cost.

Building and Equipment Lives and Decay:

Building decay rates are based on the building shell lifetimes, where the lifetime is defined as the length of time it takes for either a demolition or a major renovation to occur. For single-family residential buildings, an exponential rate of decay of 0.3% per year was assumed. See Figure 13 for the building decay rates.

Data on equipment lives and decay rates are based on EIA, RASS, Energy Star, and SoCalGas company data. See Figure 14 for the average lifetimes of gas appliances.

Saturations, Fuel and Efficiency Shares:

Saturation values, fuel shares, and efficiency shares were extracted from SoCalGas company data files and the most recent the RASS survey. Please see Figures 15-18 for saturations, fuel, and efficiency shares.

AMI:

Mass deployment of AMI gas modules began in 2011. The conservation benefits estimated by SoCalGas represent approximately 1% of core gas throughput in post-2018 (post deployment year). The conservation benefits were incorporated in the forecast as a post-model adjustment.

II. Residential End-Use Model Data

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 1: Number of Efficiency Levels by End Use by Customer Segment

	Space Heating		Water Heating		Cooking		Drying		Pool		Spa		Fireplace		BBQ	
	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric
Single Family	4	4	4	4	2	2	2	2	1	1	1	1	1	1	1	1
Multi-Family <= 4 Units	4	4	4	4	2	2	2	2	0	0	0	0	0	0	1	1
Multi-Family > 4 Units	4	4	4	4	2	2	2	2	0	0	0	0	0	0	1	1
Master Meter	4	4	4	4	2	2	2	2	0	0	0	0	0	0	1	1
Sub-Meter	4	4	4	4	2	2	2	2	0	0	0	0	0	0	1	1

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 2: Average Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	34,207	30,153	26,004	21,512	15,545	11,980	11,504	11,478	11,348	14,153	22,670	35,925	246,479
2018	33,928	29,907	25,792	21,336	15,419	11,882	11,410	11,384	11,255	14,037	22,485	35,631	244,466
2019	33,522	29,549	25,483	21,081	15,234	11,740	11,273	11,248	11,121	13,869	22,216	35,205	241,541
2020	33,053	29,135	25,126	20,786	15,021	11,576	11,116	11,091	10,965	13,675	21,905	34,712	238,159
2021	32,594	28,731	24,778	20,498	14,812	11,415	10,961	10,937	10,813	13,485	21,601	34,231	234,857
2022	32,044	28,246	24,359	20,151	14,562	11,223	10,776	10,752	10,630	13,258	21,236	33,652	230,889

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 3: Cold Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	38,963	34,218	29,056	23,706	16,425	12,195	11,543	11,512	11,431	14,743	25,105	41,037	269,934
2018	38,697	33,985	28,858	23,545	16,313	12,112	11,465	11,434	11,353	14,642	24,933	40,757	268,092
2019	38,285	33,622	28,550	23,294	16,139	11,982	11,342	11,312	11,232	14,486	24,668	40,323	265,235
2020	37,804	33,200	28,192	23,001	15,936	11,832	11,200	11,170	11,091	14,304	24,358	39,817	261,905
2021	37,340	32,793	27,846	22,719	15,741	11,687	11,063	11,033	10,955	14,129	24,059	39,328	258,692
2022	36,770	32,292	27,420	22,372	15,500	11,508	10,894	10,864	10,788	13,913	23,691	38,727	254,739

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 4: Hot Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	29,452	26,088	22,953	19,318	14,666	11,766	11,464	11,444	11,265	13,563	20,235	30,812	223,023
2018	29,163	25,832	22,728	19,128	14,522	11,650	11,352	11,332	11,154	13,430	20,037	30,511	220,839
2019	28,768	25,483	22,420	18,869	14,326	11,493	11,198	11,178	11,003	13,248	19,765	30,097	217,848
2020	28,315	25,081	22,066	18,572	14,100	11,311	11,021	11,002	10,830	13,039	19,454	29,623	214,414
2021	27,867	24,684	21,717	18,278	13,877	11,133	10,847	10,828	10,658	12,833	19,146	29,154	211,022
2022	27,341	24,218	21,307	17,933	13,615	10,922	10,642	10,623	10,457	12,591	18,785	28,604	207,039

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 5: Base Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	11,306	10,577	11,306	10,941	11,306	10,941	11,306	11,306	10,941	11,306	10,941	11,306	133,483
2018	11,066	10,352	11,066	10,709	11,066	10,709	11,066	11,066	10,709	11,066	10,709	11,066	130,645
2019	10,790	10,094	10,790	10,442	10,790	10,442	10,790	10,790	10,442	10,790	10,442	10,790	127,397
2020	10,483	9,807	10,483	10,145	10,483	10,145	10,483	10,483	10,145	10,483	10,145	10,483	123,766
2021	10,167	9,511	10,167	9,839	10,167	9,839	10,167	10,167	9,839	10,167	9,839	10,167	120,033
2022	9,824	9,191	9,824	9,507	9,824	9,507	9,824	9,824	9,507	9,824	9,507	9,824	115,991

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 6: 2017 Historical Data

Segment	2017 Therm Sales	2017 Meter Count	2017 Meter Count: Pre-1979 Customers	2017 Meter Count: 1979-2016 Customers	2017 Meter Count: 2017 "New" Customers	Avg Annual Consumption: Pre-1979 Customers	Avg Annual Consumption: 1979-2016 Customers	Avg Annual Consumption: 2017 "New" Customers	Price Elasticity
Single Family	1,735,240,312	3,742,704	2,409,585	1,312,290	20,829	464	465	337	-0.1053
Multi-Family <= 4 Units	182,211,284	552,881	405,384	145,121	2,376	334	318	251	-0.11171
Multi-Family > 4 Units	359,351,881	1,203,899	701,151	495,862	6,885	305	290	244	-0.07145
Master Meter	141,169,129	36,731	32,706	3,897	128	3,377	7,334	16,750	-0.0688
Sub-Meter	46,814,430	1,756	1,651	105	0	26,344	31,532	0	-0.1053

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 7: Meter Count Forecast

Year	Single Family	Multi-Family <= 4 Units	Multi-Family > 4 Units	Master Meter	Sub-Meter
2017	3,742,704	552,881	1,203,899	36,731	1,756
2018	3,766,828	558,772	1,216,727	36,731	1,756
2019	3,793,093	565,017	1,230,325	36,731	1,756
2020	3,820,834	571,730	1,244,942	36,731	1,756
2021	3,849,333	578,807	1,260,353	36,731	1,756
2022	3,878,156	586,206	1,276,464	36,731	1,756

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 8: Appliance Unit Energy Consumption (Gas in Therms, Electric in Kwh)

		Single Family		Multi-Family <= 4 Units		Multi-Family > 4 Units		Master Meter		Sub-Meter	
End-Use	**Efficiency**	**Gas**	**Electric**	**Gas**	**Electric**	**Gas**	**Electric**	**Gas**	**Electric**	**Gas**	**Electric**
Space Heating	Stock	270	4,110	150	730	110	730	130	730	250	1,340
	Standard	250	3,730	140	670	100	670	120	670	230	1,210
	High	230	3,450	130	620	100	620	110	620	220	1,120
	Premium	210	3,170	120	570	90	570	100	570	200	1,030
Water Heating	Stock	170	2,440	130	2,440	100	2,440	120	2,440	160	2,010
	Standard	150	2,220	120	2,220	100	2,220	110	2,220	150	1,830
	High	140	2,110	110	2,110	100	2,110	100	2,110	140	1,740
	Premium	140	2,050	110	2,050	90	2,050	100	2,050	140	1,690
Cooking	Stock	28	574	26	465	26	465	26	465	25	514
	Standard	24	488	22	395	22	395	22	395	21	437
Drying	Stock	41	1,442	35	1,442	30	1,442	33	1,442	35	873
	Standard	39	1,370	33	1,370	28	1,370	31	1,370	33	830
Pool	Stock	123	3,431	-	-	-	-	-	-	-	-
Spa	Stock	100	290	-	-	-	-	-	-	-	-
Fireplace	Stock	17	0	-	-	-	-	-	-	-	-
BBQ	Stock	16	0	15	0	13	0	14	0	16	0

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 9: Average and Marginal Gas Prices ($/therm)

Year	Res Price Deflator	R SF Average Price	R SF Marginal Price	R MF2 Average Price	R MF2 Marginal Price	R MF3 Average Price	R MF3 Marginal Price	R MM Average Price	R MM Marginal Price	R SM Average Price	R SM Marginal Price
2017	97.7	1.1188	1.1588	1.1076	1.1588	1.0807	1.1588	1.0840	1.1588	1.1194	1.1588
2018	100.0	1.0997	1.1397	1.0885	1.1397	1.0616	1.1397	1.0649	1.1397	1.1003	1.1397
2019	102.4	1.1186	1.1580	1.1076	1.1580	1.0811	1.1580	1.0843	1.1580	1.1192	1.1580
2020	105.1	1.1502	1.1911	1.1388	1.1911	1.1114	1.1911	1.1147	1.1911	1.1509	1.1911
2021	107.8	1.1673	1.2084	1.1558	1.2084	1.1282	1.2084	1.1316	1.2084	1.1680	1.2084
2022	110.5	1.2253	1.2672	1.2136	1.2672	1.1854	1.2672	1.1889	1.2672	1.2260	1.2672

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 10: Average and Marginal Electricity Prices (Cents/KWh)

Year	R SF Average Price	R SF Marginal Price	R MF2 Average Price	R MF2 Marginal Price	R MF3 Average Price	R MF3 Marginal Price	R MM Average Price	R MM Marginal Price	R SM Average Price	R SM Marginal Price
2017	18.10	27.39	17.92	27.12	17.49	26.46	17.54	18.14	18.11	20.38
2018	19.00	28.75	18.81	28.46	18.34	27.76	18.40	19.03	19.01	21.40
2019	20.03	30.31	19.83	30.01	19.36	29.30	19.42	20.08	20.04	22.56
2020	20.51	31.04	20.31	30.73	19.82	29.99	19.88	20.56	20.52	23.10
2021	21.09	31.92	20.88	31.60	20.39	30.85	20.45	21.14	21.10	23.75
2022	21.74	32.90	21.54	32.59	21.04	31.83	21.10	21.82	21.76	24.48

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 11: Gas Appliance Equipment Cost (Nominal $)

End-Use	Customer Class	Stock Efficiency	Standard Efficiency	High Efficiency	Premium Efficiency
Space Heating	Single Family	4,000	4,600	4,800	5,000
	Multi-Family 2-4 Units	2,000	2,300	2,400	2,500
	Multi-Family > 4 Unit	1,600	1,840	1,920	1,980
	Master Meter	1,000	1,150	1,200	1,250
	Sub-metered	1,600	1,840	1,920	1,980
Water Heating	Single Family	550	650	700	750
	Multi-Family 2-4 Units	330	390	420	450
	Multi-Family > 4 Unit	330	390	420	450
	Master Meter	330	390	420	450
	Sub-metered	330	390	420	450
Cooking	Single Family	500	1,400	-	-
	Multi-Family 2-4 Units	300	1,400	-	-
	Multi-Family > 4 Unit	250	1,400	-	-
	Master Meter	250	1,400	-	-
	Sub-metered	250	1,400	-	-
Drying	Single Family	328	482	-	-
	Multi-Family 2-4 Units	328	482	-	-
	Multi-Family > 4 Unit	328	482	-	-
	Master Meter	328	482	-	-
	Sub-metered	328	482	-	-
Pool	Single Family	1,200	-	-	-
Spa	Single Family	2,000	-	-	-
Fireplace	Single Family	150	-	-	-
Barbecue	Single Family	1,000	-	-	-
	Multi-Family 2-4 Units	600	-	-	-
	Multi-Family > 4 Unit	600	-	-	-
	Master Meter	600	-	-	-
	Sub-metered	600	-	-	-

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 12: Electric Appliance Equipment Cost (Nominal $)

End-Use	Customer Class	Stock Efficiency	Standard Efficiency	High Efficiency	Premium Efficiency
Space Heating	Single Family	4,100	-	-	-
	Multi-Family 2-4 Units	2,050	-	-	-
	Multi-Family > 4 Unit	1,640	-	-	-
	Master Meter	1,025	-	-	-
	Sub-metered	1,640	-	-	-
Water Heating	Single Family	550	650	700	750
	Multi-Family 2-4 Units	330	390	420	450
	Multi-Family > 4 Unit	330	390	420	450
	Master Meter	330	390	420	450
	Sub-metered	330	390	420	450
Cooking	Single Family	500	1,400	-	-
	Multi-Family 2-4 Units	300	1,400	-	-
	Multi-Family > 4 Unit	250	1,400	-	-
	Master Meter	250	1,400	-	-
	Sub-metered	250	1,400	-	-
Drying	Single Family	328	482	-	-
	Multi-Family 2-4 Units	328	482	-	-
	Multi-Family > 4 Unit	328	482	-	-
	Master Meter	328	482	-	-
	Sub-metered	328	482	-	-
Pool	Single Family	1,200	-	-	-
Spa	Single Family	2,000	-	-	-
Fireplace	Single Family	150	-	-	-
Barbecue	Single Family	1,000	-	-	-
	Multi-Family 2-4 Units	600	-	-	-
	Multi-Family > 4 Unit	600	-	-	-
	Master Meter	600	-	-	-
	Sub-metered	600	-	-	-

Southern California Gas Company
Figure 13: Building Lives and Decay Rate

Building Type	Building decay Rate
Single Family	0.003
Multi-Family 2-4 Units	0.006
Multi-Family > 4 Units	0.006
Master Meter	0.008
Sub-Meter	0.008

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 14: Gas Appliance Age (Years)

		Single Family		Multi-Family 2-4 Units		Multi-Family > 4 Units		Master Meter		Sub-metered	
End-Use	**Vintage**	**Average**	**Max**	**Average**	**Max**	**Average**	**Max**	**Average**	**Max**	**Average**	**Max**
Space Heating	Pre-1979	17	17	15	15	15	15	16	16	16	16
	1979-2004	10	17	12	15	11	15	11	16	11	16
	2005-Current	3	17	4	15	4	15	4	16	4	16
Water Heating	Pre-1979	7	7	7	8	6	8	6	8	6	8
	1979-2004	7	7	8	8	8	8	8	8	8	8
	2005-Current	3	7	2	8	4	8	4	8	4	8
Cooking	Pre-1979	12	12	10	10	10	11	14	14	14	14
	1979-2004	10	12	9	10	11	11	11	14	11	14
	2005-Current	2	12	2	10	4	11	3	14	3	14
Drying	Pre-1979	8	8	7	9	6	8	8	8	8	8
	1979-2004	8	8	9	9	8	8	8	8	8	8
	2005-Current	6	8	3	9	3	8	4	8	4	8
Pool	Pre-1979	13	13	-	-	-	-	-	-	-	-
	1979-2004	9	13	-	-	-	-	-	-	-	-
	2005-Current	3	13	-	-	-	-	-	-	-	-
Spa	Pre-1979	11	11	-	-	-	-	-	-	-	-
	1979-2004	8	11	-	-	-	-	-	-	-	-
	2005-Current	3	11	-	-	-	-	-	-	-	-
Fireplace	Pre-1979	15	15	-	-	-	-	-	-	-	-
	1979-2004	15	15	-	-	-	-	-	-	-	-
	2005-Current	15	15	-	-	-	-	-	-	-	-
Barbecue	Pre-1979	7	7	5	6	5	5	5	9	5	9
	1979-2004	7	7	6	6	5	5	9	9	9	9
	2005-Current	5	7	3	6	5	5	2	9	2	9
Other	Pre-1979	15	15	15	15	15	15	15	15	15	15
	1979-2004	15	15	15	15	15	15	15	15	15	15
	2005-Current	15	15	15	15	15	15	15	15	15	15

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 15: End-Use Saturations

End-Use	Vintage	Single Family	Multi-Family 2-4 Units	Multi-Family > 4 Units	Master Meter	Sub-metered
Space Heating	Pre-1979	1.00000	1.00000	1.00000	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	1.00000	0.00000
Water Heating	Pre-1979	1.00000	1.00000	1.00000	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	1.00000	0.00000
Cooking	Pre-1979	1.00000	1.00000	0.99633	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	1.00000	0.00000
Drying	Pre-1979	0.85795	0.63122	0.20040	0.47158	0.47158
	1979-2004	0.89516	0.69314	0.42764	0.57182	0.57182
	2005-Current	0.92508	0.75324	0.74161	0.74768	0.00000
Pool	Pre-1979	0.15644	-	-	-	-
	1979-2004	0.17913	-	-	-	-
	2005-Current	0.16916	-	-	-	-
Spa	Pre-1979	0.12651	-	-	-	-
	1979-2004	0.21695	-	-	-	-
	2005-Current	0.19134	-	-	-	-
Fireplace	Pre-1979	0.22973	-	-	-	-
	1979-2004	0.27252	-	-	-	-
	2005-Current	0.26269	-	-	-	-
Barbecue	Pre-1979	0.13716	0.09015	0.04723	0.07424	0.07424
	1979-2004	0.25180	0.13557	0.06165	0.10179	0.10179
	2005-Current	0.31442	0.23862	0.07818	0.16198	0.00000
Other	Pre-1979	1.00000	1.00000	1.00000	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	1.00000	N/A

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 16: Gas Fuel Shares (average)

End-Use	Single Family	Multi-Family 2-4 Units	Multi-Family > 4 Units	Master Meter	Sub-metered
Space Heating	0.98200	0.94116	0.91179	0.92461	0.92461
Water Heating	0.97630	0.95244	0.89871	0.92997	0.92997
Cooking	0.83890	0.80100	0.82622	0.81058	0.81058
Drying	0.80258	0.74410	0.59654	0.70306	0.70306
Pool	0.49003	-	-	-	-
Spa	0.60804	-	-	-	-
Fireplace	0.56361	-	-	-	-
Barbecue	0.95008	0.90406	0.85803	0.89234	0.89234
Other	1.00000	1.00000	1.00000	1.00000	1.00000

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 17: Gas Efficiency Shares

End-Use	Customer Class	Stock Existing	Stock New	Standard Existing	Standard New	High Existing	High New	Premium Existing	Premium New
Space Heating	Single Family	0.59	0.59	0.34	0.34	0.06	0.06	0.01	0.01
	Multi-Family 2-4 Units	0.70	0.70	0.28	0.28	0.01	0.01	0.01	0.01
	Multi-Family > 4 Units	0.50	0.50	0.48	0.48	0.01	0.01	0.01	0.01
	Master Meter	0.50	0.50	0.48	0.48	0.01	0.01	0.01	0.01
	Sub-metered	0.59	0.59	0.34	0.34	0.06	0.06	0.01	0.01
Water Heating	Single Family	0.10	0.10	0.68	0.68	0.21	0.21	0.01	0.01
	Multi-Family 2-4 Units	0.22	0.22	0.61	0.61	0.16	0.16	0.01	0.01
	Multi-Family > 4 Units	0.13	0.13	0.76	0.76	0.10	0.10	0.01	0.01
	Master Meter	0.13	0.13	0.76	0.76	0.10	0.10	0.01	0.01
	Sub-metered	0.10	0.10	0.68	0.68	0.21	0.21	0.01	0.01
Cooking	Single Family	0.90	0.90	0.10	0.10	-	-	-	-
	Multi-Family 2-4 Units	0.95	0.95	0.05	0.05	-	-	-	-
	Multi-Family > 4 Units	0.95	0.95	0.05	0.05	-	-	-	-
	Master Meter	0.95	0.95	0.05	0.05	-	-	-	-
	Sub-metered	0.95	0.95	0.05	0.05	-	-	-	-
Drying	Single Family	0.75	0.75	0.25	0.25	-	-	-	-
	Multi-Family 2-4 Units	0.75	0.75	0.25	0.25	-	-	-	-
	Multi-Family > 4 Units	0.75	0.75	0.25	0.25	-	-	-	-
	Master Meter	0.75	0.75	0.25	0.25	-	-	-	-
	Sub-metered	0.75	0.75	0.25	0.25	-	-	-	-
Pool	Single Family	1.00	1.00	-	-	-	-	-	-
Spa	Single Family	1.00	1.00	-	-	-	-	-	-
Fireplace	Single Family	1.00	1.00	-	-	-	-	-	-
Barbecue	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family 2-4 Units	1.00	1.00	-	-	-	-	-	-
	Multi-Family > 4 Units	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-
Other	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family 2-4 Units	1.00	1.00	-	-	-	-	-	-
	Multi-Family > 4 Units	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-

Southern California Gas Company
2020 Triennial Cost Allocation Proceeding (TCAP)
Figure 18: Electric Efficiency Shares

End-Use	Customer Class	Stock Existing	Stock New	Standard Existing	Standard New	High Existing	High New	Premium Existing	Premium New
Space Heating	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family 2-4 Units	1.00	1.00	-	-	-	-	-	-
	Multi-Family > 4 Units	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-
Water Heating	Single Family	0.10	0.10	0.68	0.68	0.21	0.21	0.01	0.01
	Multi-Family 2-4 Units	0.22	0.22	0.61	0.61	0.16	0.16	0.01	0.01
	Multi-Family > 4 Units	0.13	0.13	0.76	0.76	0.10	0.10	0.01	0.01
	Master Meter	0.13	0.13	0.76	0.76	0.10	0.10	0.01	0.01
	Sub-metered	0.10	0.10	0.68	0.68	0.21	0.21	0.01	0.01
Cooking	Single Family	0.90	0.90	0.10	0.10	-	-	-	-
	Multi-Family 2-4 Units	0.95	0.95	0.05	0.05	-	-	-	-
	Multi-Family > 4 Units	0.95	0.95	0.05	0.05	-	-	-	-
	Master Meter	0.95	0.95	0.05	0.05	-	-	-	-
	Sub-metered	0.95	0.95	0.05	0.05	-	-	-	-
Drying	Single Family	0.75	0.75	0.25	0.25	-	-	-	-
	Multi-Family 2-4 Units	0.75	0.75	0.25	0.25	-	-	-	-
	Multi-Family > 4 Units	0.75	0.75	0.25	0.25	-	-	-	-
	Master Meter	0.75	0.75	0.25	0.25	-	-	-	-
	Sub-metered	0.75	0.75	0.25	0.25	-	-	-	-
Pool	Single Family	1.00	1.00	-	-	-	-	-	-
Spa	Single Family	1.00	1.00	-	-	-	-	-	-
Fireplace	Single Family	1.00	1.00	-	-	-	-	-	-
Barbecue	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family 2-4 Units	1.00	1.00	-	-	-	-	-	-
	Multi-Family > 4 Units	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-

Weather for SoCalGas: Heating Degree Days – Average and Cold Year Designs; and Winter Peak Day Design Temperatures

I. Overview

Southern California Gas Company's service area extends from Fresno County to the Mexican border. To quantify the overall temperature experienced within this region, SoCalGas aggregates daily temperature recordings from fifteen U.S. Weather Bureau weather stations first into six temperature zones and then into one system average heating degree-day ("HDD") figure. The table below lists weather station locations by temperature zones.

Table 1

Weather Stations by Temperature Zones and Weights

Temperature Zone	Weight	Station (After 10/31/2002)	Station (Before 11/1/2002)
1. High mountain	0.0058	Big Bear Lake	Lake Arrowhead
2. Low desert	0.0385	Palm Springs El Centro	Palm Springs Brawley
3. Coastal	0.1854	Los Angeles Airport Newport Beach Santa Barbara Airport	Los Angeles Airport Newport Beach Harbor Santa Barbara Airport
4. High desert	0.0716	Bakersfield Lancaster Airport Fresno	Bakersfield Airport Palmdale Visalia
5. Interior valleys	0.3831	Burbank Pasadena Ontario Rialto	Burbank Pasadena Pomona Cal Poly Redlands
6. Basin	0.3156	Los Angeles Civic Center Santa Ana	Los Angeles Civic Center/ Downtown-USC Santa Ana

SoCalGas uses 65° Fahrenheit to calculate the number of HDDs. One heating degree day is accumulated for each degree that the daily average is below 65° Fahrenheit. To arrive at the HDD figure for each temperature zone, SoCalGas uses the simple average of the weather station HDDs in that temperature zone. To arrive at the system average HDDs figure for its entire service area, SoCalGas weights the HDD figure for each zone using the proportion of gas customers within each temperature zone based on calendar year 2017 customer counts. These weights have been used in calculating the data shown from January 1998 to December 2017.

Daily weather temperatures are from the National Climatic Data Center or from preliminary data that SoCalGas captures each day and posts on its internal Company web-site at the URL: http://utilinet.sempra.com/departments/massmarkets/weather/default.htm for various individual weather stations as well as for its system average values of HDD. Annual HDDs for the entire service area from 1998 to 2017 are listed in Table 2, below.

Table 2

Calendar Month Heating Degree-Days (Jan. 1998 through Dec. 2017)

	Month												Total
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	"Cal-Year"
1998	269	282	187	185	87	21	0	0	5	43	167	322	1568
1999	265	246	284	235	78	39	1	2	5	8	128	246	1537
2000	247	243	210	81	26	5	2	1	3	64	248	242	1372
2001	379	338	195	208	26	6	4	3	3	21	146	359	1688
2002	334	202	226	149	79	11	2	4	8	77	92	315	1499
2003	141	233	166	180	74	17	1	1	3	16	200	306	1338
2004	292	301	86	85	18	8	3	2	4	73	227	292	1391
2005	287	208	176	116	35	11	4	1	9	44	99	235	1225
2006	272	201	338	163	29	3	0	1	5	36	104	279	1431
2007	347	215	125	117	50	16	1	1	12	37	126	353	1400
2008	347	263	149	124	76	8	1	0	2	23	75	334	1402
2009	196	259	194	134	19	16	3	4	1	44	117	320	1307
2010	254	220	174	164	72	14	8	9	14	42	203	268	1442
2011	251	307	212	105	80	27	3	3	6	40	207	350	1591
2012	224	236	222	118	38	11	6	1	1	16	111	300	1284
2013	330	264	126	66	17	4	1	2	2	44	103	257	1216
2014	142	148	90	76	20	4	0	1	1	5	66	223	776
2015	180	94	64	67	69	5	1	0	1	4	162	316	963
2016	281	111	114	54	46	8	1	1	3	14	111	268	1012
2017	319	208	99	44	50	6	1	0	4	12	50	174	967
20-Yr-Avg (Jan1998-Dec2017)													
Avg.	267.9	229.0	171.9	123.6	49.5	12.0	2.2	1.9	4.6	33.2	137.1	288.0	1320.5
St.Dev.	66.7	61.2	68.5	53.4	25.1	9.0	2.1	2.1	3.6	21.8	55.9	48.8	236.4
Min.	141.0	94.0	64.0	44.0	17.0	3.0	0.0	0.0	1.0	4.0	50.0	174.0	776.0
Max.	379.0	338.0	338.0	235.0	87.0	39.0	8.0	9.0	14.0	77.0	248.0	359.0	1688.0

II. Calculations to Define Our Average-Temperature Year

The simple average of the 20-year period (January 1998 through December 2017) was used to represent the Average Year total and the individual monthly values for HDD. In this TCAP, the standard deviation has been calculated using an approach that compensates for the annual HDD values for the years 2014-2017 in SoCalGas' service territory being dramatically lower than in any preceding year going back to 1950. A regression with a time trend and a dummy variable for the years 2014-2017 has been used to estimate a shift in the level of annual HDD that occurred beginning in 2014. A dummy variable takes the value one for some observations to indicate the presence of an effect or membership in a group and zero for the remaining observations. Estimating the effect of the dummy variable gives an estimate of that effect or the impact of membership in that group. A dummy variable is used here to estimate the average effect on annual HDD of a given year having membership in the group of years 2014-2017. The dataset is SoCalGas system-wide annual HDD for the years 1998-2017. The regression equation is:

$$HDD_t = \alpha + \beta * t + \beta_{2014-2017} * D_{2014-2017} + \varepsilon$$

where $D_{2014-2017}$ is a dummy variable for the years 2014-2017 and $\beta_{2014-2017}$is the corresponding dummy coefficient. This regression equation estimates average HDD over the period 1998-2017 controlling for time trends in HDD and the warm weather regime of years 2014-2017. It's important to note that p-value for the estimate of $\beta_{2014-2017}$ is 0.11% indicating an extremely low probability that membership in the group of years 2014-2017 had no effect on annual HDDs. Please see table 3 below for the full regression output.

Table 3

Dummy Regression for Calculation of Heating Degree-Day Standard Deviation

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.877127127
R Square	0.769351997
Adjusted R Squa	0.742216937
Standard Error	120.0395315
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	817095.635	408547.8175	28.35269274	3.84656E-06
Residual	17	244961.315	14409.48912		
Total	19	1062056.95			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	26248.60272	12961.00191	2.025198584	0.058843735
Regime Dummy	-364.8759058	93.16452578	-3.916468235	0.001111279
YEAR	-12.38115942	6.462711128	-1.915784131	0.072366763

The dummy variable's estimated effect, $\beta_{2014-2017}$, is subtracted from the actual annual HDD data for years 2014-2017 to adjust the data to remove the level shift. The standard deviation has been calculated using this adjusted dataset. This standard deviation has been used to design the two Cold Years based on a "1-in-10" and "1-in-35" chance, ***c***, that the respective annual "Cold Year" $\boldsymbol{hdd_c}$ value would be exceeded.

A probability model for the annual HDD is based on a t-Distribution with N-1 degrees of freedom, where, N is the number of years of HDD data we use, μ is the average of the last 20 years of HDD, and S_{20} is the average of the standard deviations of the 20 most recent 20 year periods:

$U = (HDD_y - \mu)/S_{20}$, has a t-Distribution with N-1 degrees of freedom.

III. Calculating the Cold-Temperature Year Weather Designs

Cold Year HDD Weather Designs

For SoCalGas, cold-temperature-year HDD weather designs are developed with a 1-in-35 annual chance of occurrence. In terms of probabilities this can be expressed as the following for a "1-in-35" cold-year HDD value in equation 1 and a "1-in-10" cold-year HDD value in equation 2, with Annual HDD as the random variable:

(1) Prob { Annual HDD > "1-in-35" Cold-Yr HDD } = 1/35 = 0.0286

(2) Prob { Annual HDD > "1-in-10" Cold-Yr HDD } = 1/10 = 0.1000

An area of 0.0286 under one tail of the T-Distribution translates to 2.025 standard deviations *above* an average-year based on a t-statistic with 19 degrees of freedom. Using the standard deviation calculated as described earlier, 135.1 HDD, these equations yield values of about 1,594 HDD for a "1-in-35" cold year and 1,499 as the number of HDDs for a "1-in-10" cold year (an area of 0.1000 under one tail of the T-Distribution translates to 1.328 standard deviations *above* an average-year based on a t-statistic with 19 degrees of freedom). For example, the "1-in-35" cold-year HDD is calculated as follows:

(3) Cold-year HDD = 1,594 which equals approximately

1,320 average-year HDDs + 2.025 * 135.1

Table 4 shows monthly HDD figures for "1-in-35" cold year, "1-in-10" cold year and, average year temperature designs. The monthly average-temperature-year HDDs are calculated from weighted monthly HDDs from 1998 to 2017, as shown as the bottom of Table 2, above. For example, the average-year December value of 288.0 HDD equals the simple average of the 20 December HDD figures from 1998 to 2017. SoCalGas calculates the cold--temperature-year monthly HDD values using the same distribution of average-year HDDs. For example, since 21.8 percent of average-temperature-year HDDs occurred in December, the estimated number of HDDs during December for a cold-year is equal to 1,594 HDDs multiplied by 21.8 percent, or 347.6 HDDs.

Table 4

Calendar Month Heating Degree-Day Designs

	Cold		Average	Hot	
	1-in-35 Design	1-in-10 Design		1-in-10 Design	1-in-35 Design
January	323.3	304.1	267.8	231.4	212.2
February	276.4	259.9	228.9	197.8	181.4
March	207.5	195.1	171.8	148.5	136.1
April	149.1	140.3	123.5	106.8	97.9
May	59.7	56.1	49.4	42.7	39.2
June	14.5	13.6	12.0	10.4	9.5
July	2.6	2.4	2.1	1.9	1.7
August	2.2	2.1	1.8	1.6	1.5
September	5.6	5.2	4.6	4.0	3.6
October	40.0	37.6	33.1	28.6	26.3
November	165.5	155.6	137.1	118.5	108.6
December	347.6	326.9	287.9	248.8	228.1
	1594	1499	1320	1141	1046

IV. Calculating the Peak-Day Design Temperature

SoCalGas' 1-in-35 Peak-Day design temperature of 40.3 degrees Fahrenheit, denoted "Deg-F," is determined from a statistical analysis of observed annual minimum daily system average temperatures constructed from daily temperature recordings from the fifteen U.S. Weather Bureau weather stations discussed above. Since we have a time series of daily data by year, the following notation will be used for the remainder of this discussion:

(1) $AVG_{y,d}$ = system avg value of temperature for calendar year "y" and day "d".

The calendar year, y, can range from 1950 through 2017, while the day, d, can range from 1 to 365, for non-leap years, or from 1 to 366 for leap years. The "upper" value for the day, d, thus depends on the calendar year, y, and will be denoted by n(y)=365, or 366, respectively, when y is a non-leap year or a leap year.

For each calendar year, we calculate the following statistic from our series of daily system average temperatures defined in equation (1) above:

$$\text{(2)} \qquad MinAVG_y = \min_{d=1}^{n(y)} \{ AVG_{y,d} \}, \text{ for } y=1950, 1951, \ldots, 2017.$$

(The notation used in equation 2 means “For a particular year, y, list all the daily values of system average temperature for that year, then pick the smallest one.”)

The resulting minimum annual temperatures are shown in Tables 5.1 and 5.2, below. Note that most of the minimum temperatures occur in the months of December or January; however, for some calendar years the minimums occurred in other months (the minimum for 2006 was observed in March).

The statistical methods we use to analyze this data employ software developed to fit three generic probability models: the Generalized Extreme Value (GEV) model, the Double-Exponential or GUMBEL (EV1) model and a 2-Parameter Students’ T-Distribution (T-Dist) model. [The GEV and EV1 models have the same mathematical specification as those implemented in a DOS-based executable-only computer code that was developed by Richard L. Lehman and described in a paper published in the Proceedings of the Eighth Conference on Applied Climatology, January 17-22, 1993, Anaheim, California, pp. 270-273, by the American Meteorological Society, Boston, MA., with the title “Two Software Products for Extreme Value Analysis: System Overviews of ANYEX and DDEX.” At the time he wrote the paper, Dr. Lehman was with the Climate Analysis Center, National Weather Service/NOAA in Washington, D.C., zip code 20233.] The Statistical Analysis Software (SAS) procedure for nonlinear statistical model estimation (PROC MODEL) was used to do the calculations. Further, the calculation procedures were implemented to fit the probability models to observed *maxima* of data, like heating degrees. By recognizing that:

$$- MinAVG_y = - \min_{d=1}^{n(y)}\{AVG_{y,d}\} = \max_{d=1}^{n(y)}\{ -AVG_{y,d}\}, \text{ for } y=1950, \ldots, 2017;$$

this same software, when applied to the *negative* of the minimum temperature data, yields appropriate probability model estimation results.

The calculations done to fit any one of the three probability models chooses the parameter values that provide the “best fit” of the parametric probability model’s calculated cumulative distribution function (CDF) to the empirical cumulative distribution function (ECDF). Note that the ECDF is constructed based on the variable “-$MinAVG_y$” (which is a *maximum* over a set of *negative* temperatures) with values of the variable $MinAVG_y$ that are the same as shown in Tables 5.1 and 5.2, below.

In Tables 6.1 and 6.2, the data for -$MinAVG_y$ are shown after they have been sorted from “lowest” to “highest” value. The ascending *ordinal* value is shown in the column labeled “RANK” and the empirical cumulative distribution function is calculated and shown in the next column. The formula used to calculate this function is:

$$ECDF = (RANK – \alpha)/[MaxRANK + (1 – 2\,\alpha)],$$

where the parameter "α" (shown as *alpha* in Table 6.1 and Table 6.2) is a "small" positive value (usually less than ½) that is used to bound the ECDF away from 0 and 1.

Of the three probability models considered (GEV, EV1, and T_Dist) the results obtained for the T_Dist model were selected since the fit to the ECDF was better than that of either the GEV model or the EV1 model. (Although convergence to stable parameter estimates is occasionally a problem with fitting a GEV model to the ECDF, the T_Dist model had no problems with convergence of the iterative procedure to estimate parameters.)

The T_Dist model used here is a three-parameter probability model where the varable $z = (-MinAVG_y - \gamma) / \theta$, for each year, y, is presumed to follow a T_Dist with location parameter, γ, and scale parameter, θ, and a third parameter, ν, that represents the number of degrees of freedom. For a given number of years of data, N, then ν=N-2.

The following mathematical expression specifies the T_Dist model we fit to the data for "$-MinAVG_y$ " shown in Table 6.1 and Table 6.2, below.

(3) $ECDF(-MinAVG_y) = Prob\{ -T < -MinAVG_y \} = T\_Dist\{z; \gamma, \theta, \nu=N-2\}$,

where "T_Dist{ . }" is the cumulative probability distribution function for Student's T-Distribution[1], and

(4) $z = (-MinAVG_y - \gamma) / \theta$, for each year, y, and

the parameters "γ" and "θ" are estimated for this model for given degrees of freedom ν=N-2. The estimated values for γ and θ are shown in Table 6.2 along with the fitted values of the model CDF (the column: "Fitted" Model CDF).

Now, to calculate a *peak-day design temperature*, $TPDD_\delta$, with a specified likelihood, δ, that a value less than $TPDD_\delta$ would be observed, we use the equation below:

(5) $\delta = Prob\{ T \leq TPDD_\delta \}$, which is equivalent to

(6) $\delta = Prob\{ [(-T - \gamma) / \theta] \geq [(-TPDD_\delta - \gamma) / \theta] \}, = Prob\{ [(-T - \gamma) / \theta] \geq [z_\delta] \}$,

where $z_\delta = [(-TPDD_\delta - \gamma) / \theta]$. In terms of our probability model,

(7) $\delta = 1 – T\_Dist\{ z_\delta; \gamma, \theta, \nu=N-2\}$,

which yields the following equation for z_δ,

(7') $z_\delta = \{ TINV\_Dist\{ (1-\delta); \gamma, \theta, \nu=N-2\}$, where "TINV_Dist{ . }" is the inverse function of the T_Dist{ . } function[2]. The implied equation for $TPDD_\delta$ is:

---

[1] A common mathematical expression for Student's T-Distribution is provided at http://en.wikipedia.org/wiki/Student%27s_t-distribution; with a probability density function

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$

such that $T\_Dist\{z; \gamma, \theta, \nu=N-2\} = \int f(t)\, dt$, from t=-∞ to t=z. Also, the notation Γ(.) is known in mathematics as the GAMMA function; see http://www.wikipedia.org/wiki/Gamma_function for a description. Also, see *Statistical Theory*, 3rd Ed., B.W. Lindgren, MacMillian Pub. Inc, 1976, pp. 336-337.

(8) $TPDD_{\delta} = -[\gamma + (z_{\delta})(\theta)]$.

To calculate the minimum daily (system average) temperature to define our extreme weather event, we specify that this COLDEST-Day be one where the temperature would be lower with a "1-in-35" likelihood. This criterion translates into two equations to be solved based on equations (7) and (8) above:

(9) solve for "z_{δ}" from equation (7') above with $(1-\delta) = (1 - 1/35) = 1 - 0.0286$,

(10) solve for "$TPDD_{\delta}$" from $TPDD_{\delta} = -[\gamma + (z_{\delta})(\theta)]$.

The value of $z_{\delta} = 1.938$ and $TPDD_{\delta} = -[\gamma + (z_{\delta})(\theta)] = 40.3$ degrees Fahrenheit, with values for "v=N-2"; along with "γ" and "θ" in Tables 6.1 & 6.2, below.

SoCalGas' 1-in-10 peak-day design temperature of 42.0 degrees Fahrenheit, is calculated in a methodologically similar way as for the 40.3 degree peak day temperature. The criteria specified in equation (9) above for a "1-in-35" likelihood would be replaced by a "1-in-10" likelihood.

(9') solve for "z_{δ}" from equation (7') above with $(1-\delta) = (1 - 1/10) = 1 - 0.1000$,

which yields a "z_{δ}" value of $z_{\delta} = 1.295$ and, $TPDD_{\delta} = -[\gamma + (z_{\delta})(\theta)] = 42.0$ with values for "v=N-2"; along with "γ" and "θ" in Tables 6.1 and 6.2, below.

A plot of the cumulative distribution function for $MinAVG_y$ based on "v=N-2", the fitted model parameters, "γ" and "θ" with values in Tables 6.1 and 6.2, below, is shown in Figure 1.

[2] Computer software packages such as SAS and EXCEL have implemented statistical and mathematical functions to readily calculate values for T_Dist{ . } and TINV_Dist{ . } as defined above.

Table 5.1

YEAR	MINAVG	Month(MinAvg)
1950	40.8634	Jan
1951	44.5730	Dec
1952	43.0734	Jan
1953	45.6758	Feb
1954	45.6884	Dec
1955	45.8342	Dec
1956	44.9053	Feb
1957	39.4961	Jan
1958	46.2669	Nov
1959	48.2571	Feb
1960	42.3300	Jan
1961	47.2153	Dec
1962	43.4200	Jan
1963	42.5981	Jan
1964	45.2353	Nov
1965	44.7995	Jan
1966	46.7088	Jan
1967	40.7616	Dec
1968	40.6340	Dec
1969	44.8417	Jan
1970	46.8281	Dec
1971	42.9888	Jan
1972	41.4200	Dec
1973	45.0642	Jan
1974	42.9862	Jan
1975	44.6346	Jan
1976	44.8440	Jan
1977	48.3449	Jan
1978	41.6562	Dec
1979	41.3803	Jan
1980	50.3474	Jan
1981	49.3422	Jan
1982	45.3450	Jan
1983	48.6797	Jan
1984	46.9101	Dec
1985	45.1244	Feb
1986	48.5959	Feb
1987	43.4583	Dec
1988	43.2855	Dec
1989	40.6087	Feb
1990	39.0053	Dec
1991	48.6689	Mar
1992	47.3482	Dec
1993	46.1072	Jan
1994	47.1497	Nov

Table 5.2

YEAR	MINAVG	Month(MinAvg)
1995	49.8387	Dec
1996	44.9460	Feb
1997	48.3524	Jan
1998	43.6350	Dec
1999	49.0089	Jan
2000	48.7903	Mar
2001	47.1587	Feb
2002	45.8195	Jan
2003	47.0887	Dec
2004	48.2140	Nov
2005	47.2790	Jan
2006	45.7926	Mar
2007	41.5310	Jan
2008	45.8173	Dec
2009	45.2622	Dec
2010	44.7088	Dec
2011	46.7556	Feb
2012	46.7788	Dec
2013	43.9135	Jan
2014	48.0734	Dec
2015	45.6143	Jan
2016	46.7328	Dec
2017	47.5690	Jan

Table 6.1

alpha= 0.375

YEAR	Month(-MinAvg)	Days/Yr	**-MinAvg**	Rank	Empirical CDF	Fitted Model CDF
1980	Jan	366	-50.3474	1	0.0145	-2.232
1995	Dec	365	-49.8387	2	0.0290	-1.929
1981	Jan	365	-49.3422	3	0.0435	-1.738
1999	Jan	365	-49.0089	4	0.0580	-1.593
2000	Mar	366	-48.7903	5	0.0725	-1.475
1983	Jan	365	-48.6797	6	0.0870	-1.375
1991	Mar	365	-48.6689	7	0.1014	-1.286
1986	Feb	365	-48.5959	8	0.1159	-1.207
1997	Jan	365	-48.3524	9	0.1304	-1.134
1977	Jan	365	-48.3449	10	0.1449	-1.067
1959	Feb	365	-48.2571	11	0.1594	-1.004
2004	Nov	366	-48.2140	12	0.1739	-0.946
2014	Dec	365	-48.0734	13	0.1884	-0.890
2017	Jan	365	-47.5690	14	0.2029	-0.837
1992	Dec	366	-47.3482	15	0.2174	-0.786
2005	Jan	365	-47.2790	16	0.2319	-0.737
1961	Dec	365	-47.2153	17	0.2464	-0.690
2001	Feb	365	-47.1587	18	0.2609	-0.644
1994	Nov	365	-47.1497	19	0.2754	-0.600
2003	Dec	365	-47.0887	20	0.2899	-0.557
1984	Dec	366	-46.9101	21	0.3043	-0.514
1970	Dec	365	-46.8281	22	0.3188	-0.473
2012	Dec	366	-46.7788	23	0.3333	-0.433
2011	Feb	365	-46.7556	24	0.3478	-0.393
2016	Dec	366	-46.7328	25	0.3623	-0.354
1966	Jan	365	-46.7088	26	0.3768	-0.315
1958	Nov	365	-46.2669	27	0.3913	-0.277
1993	Jan	365	-46.1072	28	0.4058	-0.239
1955	Dec	365	-45.8342	29	0.4203	-0.202
2002	Jan	365	-45.8195	30	0.4348	-0.165
2008	Dec	366	-45.8173	31	0.4493	-0.128
2006	Mar	365	-45.7926	32	0.4638	-0.091
1954	Dec	365	-45.6884	33	0.4783	-0.055
1953	Feb	365	-45.6758	34	0.4928	-0.018
2015	Jan	365	-45.6143	35	0.5072	0.018
1982	Jan	365	-45.3450	36	0.5217	0.055
2009	Dec	365	-45.2622	37	0.5362	0.091
1964	Nov	366	-45.2353	38	0.5507	0.128
1985	Feb	365	-45.1244	39	0.5652	0.165
1973	Jan	365	-45.0642	40	0.5797	0.202
1996	Feb	366	-44.9460	41	0.5942	0.239
1956	Feb	366	-44.9053	42	0.6087	0.277
1976	Jan	366	-44.8440	43	0.6232	0.315
1969	Jan	365	-44.8417	44	0.6377	0.354
1965	Jan	365	-44.7995	45	0.6522	0.393

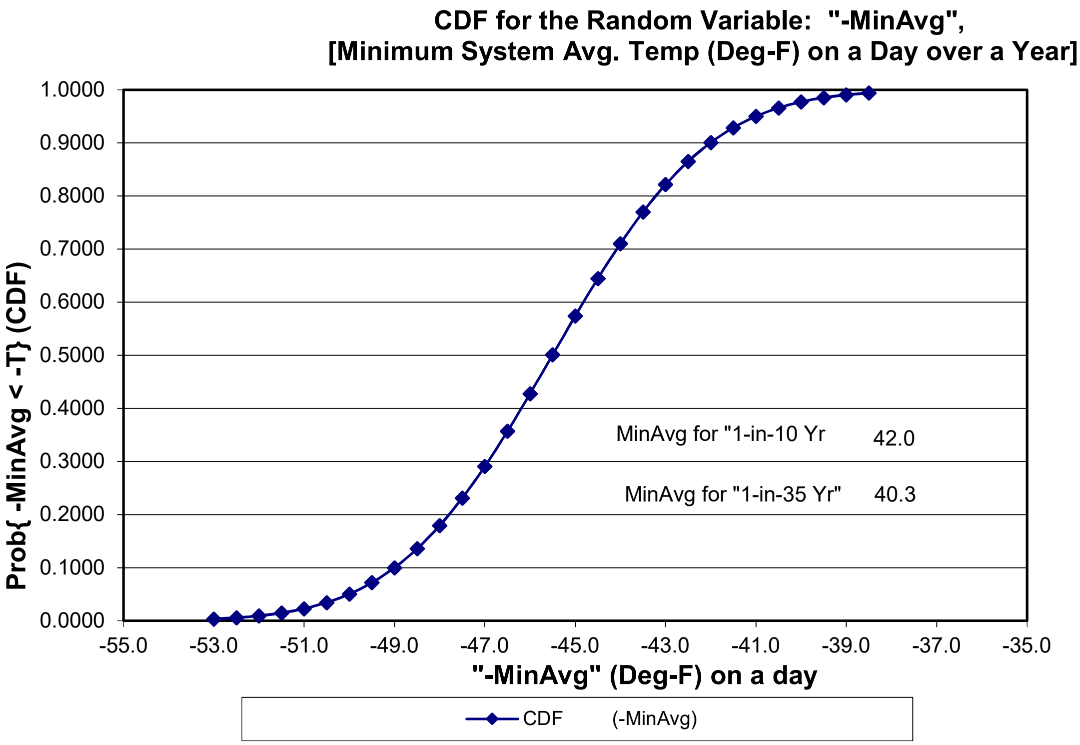
Table 6.2

alpha= 0.375

YEAR	Month(-MinAvg)	Days/Yr	**-MinAvg**	Rank	Empirical CDF	Fitted Model CDF
2010	Dec	365	-44.7088	46	0.6667	0.433
1975	Jan	365	-44.6346	47	0.6812	0.473
1951	Dec	365	-44.5730	48	0.6957	0.514
2013	Jan	365	-43.9135	49	0.7101	0.557
1998	Dec	365	-43.6350	50	0.7246	0.600
1987	Dec	365	-43.4583	51	0.7391	0.644
1962	Jan	365	-43.4200	52	0.7536	0.690
1988	Dec	366	-43.2855	53	0.7681	0.737
1952	Jan	366	-43.0734	54	0.7826	0.786
1971	Jan	365	-42.9888	55	0.7971	0.837
1974	Jan	365	-42.9862	56	0.8116	0.890
1963	Jan	365	-42.5981	57	0.8261	0.946
1960	Jan	366	-42.3300	58	0.8406	1.004
1978	Dec	365	-41.6562	59	0.8551	1.067
2007	Jan	365	-41.5310	60	0.8696	1.134
1972	Dec	366	-41.4200	61	0.8841	1.207
1979	Jan	365	-41.3803	62	0.8986	1.286
1950	Jan	365	-40.8634	63	0.9130	1.375
1967	Dec	365	-40.7616	64	0.9275	1.475
1968	Dec	366	-40.6340	65	0.9420	1.593
1989	Feb	365	-40.6087	66	0.9565	1.738
1957	Jan	365	-39.4961	67	0.9710	1.929
1990	Dec	365	-39.0053	68	0.9855	2.232

"Gamma" (Fitted) =	**-45.51**
"Theta" (Fitted) =	**2.70**
Deg. Freedom=	**66**

Figure 1


CDF for the Random Variable: "-MinAvg",
[Minimum System Avg. Temp (Deg-F) on a Day over a Year]
Prob{ -MinAvg < -T} (CDF)
1.0000
0.9000
0.8000
0.7000
0.6000
0.5000
0.4000
0.3000
0.2000
0.1000
0.0000
MinAvg for "1-in-10 Yr 42.0
MinAvg for "1-in-35 Yr" 40.3
-55.0
-53.0
-51.0
-49.0
-47.0
-45.0
-43.0
-41.0
-39.0
-37.0
-35.0
"-MinAvg" (Deg-F) on a day
CDF (-MinAvg)

V. Estimating the Uncertainty in the Peak-Day Design Temperature

The calculated peak-day design temperatures in section IV above also have a statistical uncertainty associated with them. The estimated measures of uncertainty recommended for our use are calculated from the fitted model for the probability distribution and are believed to be reasonable, although rough, approximations.

The basic approach used the estimated parameters for the probability distribution (see the results provided in Tables 6.1 and 6.2, above) to calculate the fitted temperatures as a function of the empirical CDF listed in Tables 6.1 and 6.2, above. These fitted temperatures are then compared with the observed temperatures by calculating the difference = "observed" – "fitted" values. The full set of differences are then separated into the lower third (L), the middle third (M) and the upper third (U) of the distribution. Finally, values of the root-mean-square error (RMSE) of the differences in each third of the distribution are calculated, along with the RMSE for the entire set of differences overall. The data in Tables 7.1 and 7.2, below, show the temperature data and the resulting RMSE values.

The formula below is used to calculate the RMSE for a specified set of "N" data differences:

$$\text{RMSE} = \text{SQRT}\left\{ \left(\sum_{i=1, \ldots, N} e[\,i\,]^2 \right) / \,(N-2) \right\},$$

where e[i] = *observed* less *fitted* value of temperature, T[i]. The number of estimated paramters (3 for the GEV model, 2 for the T-Dist and EV1 models) is subtracted from the respective number of data differences, N, in the denominator of the RMSE expression.

Since both the "1-in-35" and "1-in-10" peak-day temperature values are in the lower third quantile of the fitted distribution, the calculated standard error for these estimates is 0.5 Deg-F.

Table 7.1

Quantile: (Lower, Middle, Upper 3rd's)	Observed $T_{[i]}$ Temp. Ranked	Fitted Value of $T_{[i]}$	Residual $e_{[i]}$: Obs'd. less Fitted Value of $T_{[i]}$	Square of $e_{[i]}$:
U	50.3474	52.0385	-1.6911	2.859685
U	49.8387	50.9561	-1.1174	1.248662
U	49.3422	50.3582	-1.0160	1.032201
U	49.0089	49.9286	-0.9196	0.845695
U	48.7903	49.5866	-0.7963	0.634079
U	48.6797	49.2990	-0.6193	0.383540
U	48.6689	49.0484	-0.3795	0.144011
U	48.5959	48.8249	-0.2290	0.052433
U	48.3524	48.6219	-0.2694	0.072603
U	48.3449	48.4350	-0.0902	0.008129
U	48.2571	48.2612	-0.0041	0.000017
U	48.2140	48.0981	0.1159	0.013427
U	48.0734	47.9440	0.1293	0.016731
U	47.5690	47.7975	-0.2284	0.052175
U	47.3482	47.6574	-0.3092	0.095602
U	47.2790	47.5229	-0.2439	0.059470
U	47.2153	47.3933	-0.1780	0.031677
U	47.1587	47.2679	-0.1092	0.011933
U	47.1497	47.1462	0.0035	0.000012
U	47.0887	47.0279	0.0608	0.003700
U	46.9101	46.9124	-0.0023	0.000005
U	46.8281	46.7994	0.0287	0.000822
U	46.7788	46.6887	0.0901	0.008123
M	46.7556	46.5799	0.1757	0.030861
M	46.7328	46.4729	0.2599	0.067539
M	46.7088	46.3675	0.3414	0.116539
M	46.2669	46.2633	0.0037	0.000013
M	46.1072	46.1602	-0.0530	0.002814
M	45.8342	46.0582	-0.2240	0.050173
M	45.8195	45.9569	-0.1373	0.018862
M	45.8173	45.8562	-0.0389	0.001516
M	45.7926	45.7561	0.0366	0.001337
M	45.6884	45.6562	0.0322	0.001036
M	45.6758	45.5566	0.1191	0.014193
M	45.6143	45.4571	0.1572	0.024727
M	45.3450	45.3575	-0.0125	0.000155
M	45.2622	45.2577	0.0046	0.000021
M	45.2353	45.1575	0.0778	0.006048
M	45.1244	45.0569	0.0675	0.004559
M	45.0642	44.9556	0.1087	0.011811
M	44.9460	44.8535	0.0926	0.008567
M	44.9053	44.7504	0.1548	0.023966
M	44.8440	44.6463	0.1978	0.039106
M	44.8417	44.5408	0.3009	0.090559
M	44.7995	44.4338	0.3657	0.133757

Table 7.2

Quantile: (Lower, Middle, Upper 3rd's)	Observed $T_{[i]}$ Temp. Ranked	Fitted Value of $T_{[i]}$	Residual $e_{[i]}$: Obs'd. less Fitted Value of $T_{[i]}$	Square of $e_{[i]}$:
L	44.7088	44.3250	0.3837	0.147254
L	44.6346	44.2143	0.4202	0.176600
L	44.5730	44.1014	0.4716	0.222419
L	43.9135	43.9859	-0.0723	0.005234
L	43.6350	43.8675	-0.2325	0.054068
L	43.4583	43.7458	-0.2875	0.082665
L	43.4200	43.6204	-0.2004	0.040172
L	43.2855	43.4908	-0.2054	0.042173
L	43.0734	43.3563	-0.2830	0.080085
L	42.9888	43.2163	-0.2275	0.051736
L	42.9862	43.0697	-0.0835	0.006972
L	42.5981	42.9156	-0.3175	0.100776
L	42.3300	42.7525	-0.4225	0.178531
L	41.6562	42.5787	-0.9225	0.851040
L	41.5310	42.3918	-0.8608	0.741056
L	41.4200	42.1889	-0.7688	0.591100
L	41.3803	41.9653	-0.5850	0.342207
L	40.8634	41.7148	-0.8513	0.724757
L	40.7616	41.4271	-0.6655	0.442919
L	40.6340	41.0852	-0.4512	0.203574
L	40.6087	40.6555	-0.0468	0.002193
L	39.4961	40.0576	-0.5615	0.315316
L	39.0053	38.9753	0.0300	0.000900

Overall RMSE ($e_{[i]}$):	**0.45**	**°F**
Upper 3rd RMSE ($e_{[i]}$):	**0.60**	**°F**
Middle 3rd RMSE ($e_{[i]}$):	**0.18**	**°F**
Lower 3rd RMSE ($e_{[i]}$):	**0.51**	**°F**

VI. The Relationship between Annual Likelihoods for Peak-Day Temperatures and "Expected Return Time"

The event whose probability distribution we've modeled is the likelihood that the minimum daily temperature over a calendar year is less than a specified value. And, in particular, we've used this probability model to infer the value of a temperature, our *peak-day design temperature* ($TPDD_\delta$), that corresponds to a pre-defined likelihood, δ, that the observed minimum temperature is less than or equal to this design temperature.

(1) δ = Prob{ Minimum Daily Temperature over the Year < $TPDD_\delta$ }.

For some applications, it is useful to think of how this specified likelihood (or "risk level" δ) relates to the expected number of years until this Peak-Day event would first occur. This expected number of years is what is meant by the *return period*. The results stated below are found in the book: ***Statistics of Extremes***, E.J. Gumbel, Columbia University Press, 1958, on pages 21-25.

(2) E[#Yrs for Peak-Day Event to Occur] = 1 / δ,

1 / Prob{ Minimum Daily Temperature over the Year < $TPDD_\delta$ }.

For our peak-day design temperature (40.3°F) associated with a 1-in-35 annual likelihood, the return period is 35 years (δ=1/35). For the 42.0°F peak-day design temperature, the return period is 10 years (δ=1/10). Occasionally, a less precise terminology is used. For example, the 40.3°F peak-day design temperature may be referred to as a "1-in-35 year cold day"; and the 42.0°F peak-day design temperature may be referred to as a "1-in-10 year cold day."

The probability model for the *return period*, as a random variable, is a geometric (discrete) distribution with positive integer values for the *return period*. The parameter δ = Prob{ Minimum Daily Temperature over the Year < $TPDD_\delta$ }.

(3) Prob{ *return period* = r } = $(1 - \delta)^{(r-1)}\, \delta$, for r = 1, 2, 3, . . .

The expected value of the *return period* is already given in (2) above; the variance of the *return period* is:

(4) Var[*return period*] = (E[*return period*])^2 x (1- (1 / E [*return period*])),

(4') Var[*return period*] = (E[*return period*]) x (E [*return period*] - 1).

Equations (4) and (4') indicate that the standard deviation (square root of the variance) of the *return period* is nearly equal to its expected value. Thus, there is substantial variability about the expected value—a *return period* is not very precise.

SoCalGas SAS Code #1: SysAvgVar(Daily).sas

```
Title1 "Combine PRELIMINARY weather data (from 'wdaily46.sas7bdat') and combine with previous" ;
Title2 "FINAL NWS data ('dly50_96.sas7bdat') and UPDATE ('Updt4NWS.sas7bdat') from TWO NWS CD-ROM files. " ;

/*******************************************************************************/
/*                                                                             */
/* Two SAS-Formated data sets are created:                                     */
/*     1). "scgwea.DlySys_D" contains Daily System Average values of HDD, CDD  */
/*         and Temperature ("AVG")                                             */
/*     2). "scgwea.DlyStn_D" contains Daily of HDD,CDD and Temperature         */
/*         for each of our 15 weather stations, our 6-Climate Zones, and our   */
/*         "System Averages"                                                   */
/*                                                                             */
/*******************************************************************************/




options mprint ;
options ls=160 ps=90 ; **<<PORTRAIT:   SAS-Monospace w/Roman 6pt. Font >>**;
options date number notes ;
/* %cour8l
   %cour8p */
*options ls=211 ps=69 ; **<<LANDSCAPE:  SAS-Monospace w/Roman 6pt. Font >>**;



*******************************************************************************;

* Data Importation Section;

*libname scgdf '/data/home/scgdf';
libname scgwea '/EDS_RB/Weather/2018 Cgr/SoCalGas';

**** Calculate Hdd, Cdd from PRELIMINARY data set! ****;
****libname InWea  '/data/home/scgdf/weather/wdaily46.sas7bdat';

  **<< Use the copy of "wdaily46.sas7bdat" saved at this folder. >>** ;
  **<< It has data we need through Feb. 27th, 2018 for both SoCalGas and the SDGandE locations. >>** ;



proc contents data=scgwea.wdaily46 ;  **<< File was current through 2/27/2018! >>>*** ;
run ;


**<< New Weights for 6-Climate Zones: (per Avg. Mtr. Cnt for Yr)        >>** ;

**<<   Zone:       <<Dec 2015 Values>>  <<Year 20xx Values>>   >>** ;
**<< 1 High Mtn.     0.005842                 a.aaaa          >>** ;
**<< 2 Low Desert    0.038341                 b.bbbb          >>** ;
**<< 3 Coastal       0.186160                 c.cccc          >>** ;
**<< 4 High Desert   0.071125                 d.dddd          >>** ;
**<< 5 Intr. Valleys 0.385713                 e.eeee          >>** ;
**<< 6 Basin         0.312819                 f.ffff          >>** ;
**<< -------------- ---------                 ------          >>** ;
**<<   TOTAL         1.000000                 1.0000          >>** ;


**<<   Zone:       <<Dec 2014 Values>>  <<Year 20xx Values>>   >>** ;
**<< wtz5 calcd value was 0.3813, but used 0.3812 so total adds to 100.00 exactly. >>** ;
**<< 1 High Mtn.     0.0061              a.aaaa          >>** ;
**<< 2 Low Desert    0.0423              b.bbbb          >>** ;
```

```
**<< 3 Coastal       0.1763               c.cccc          >>** ;
**<< 4 High Desert   0.0747               d.dddd          >>** ;
**<< 5 Intr. Valleys 0.3812  (0.3813)      e.eeee          >>** ;
**<< 6 Basin         0.3194               f.ffff          >>** ;
**<< -------------- -------               -------         >>** ;
**<<   TOTAL         1.0000               1.0000          >>** ;


**<<   Zone:       <<Dec 2013 Values>>  <<Year 20xx Values>>   >>** ;
**<< wtz5 calcd value was 0.3793, but used 0.3792 so total adds to 100.00 exactly. >>** ;
**<< 1 High Mtn.     0.0062               a.aaaa          >>** ;
**<< 2 Low Desert    0.0424               b.bbbb          >>** ;
**<< 3 Coastal       0.1772               c.cccc          >>** ;
**<< 4 High Desert   0.0742               d.dddd          >>** ;
**<< 5 Intr. Valleys 0.3792  (0.3793)      e.eeee          >>** ;
**<< 6 Basin         0.3208               f.ffff          >>** ;
**<< -------------- -------               -------         >>** ;
**<<   TOTAL         1.0000               1.0000          >>** ;


**<< New Weights for 6-Climate Zones: (per Avg. Mtr. Cnt for Yr)        >>** ;

**<<   Zone:       <<Dec 2010 Values>>  <<Year 20xx Values>>   >>** ;
**<< wtz6 calcd value was 0.3199, but used 0.3198 so total adds to 100.00 exactly. >>** ;
**<< 1 High Mtn.     0.0062               a.aaaa          >>** ;
**<< 2 Low Desert    0.0418               b.bbbb          >>** ;
**<< 3 Coastal       0.1774               c.cccc          >>** ;
**<< 4 High Desert   0.0746               d.dddd          >>** ;
**<< 5 Intr. Valleys 0.3802               e.eeee          >>** ;
**<< 6 Basin         0.3198               f.ffff          >>** ;
**<< -------------- -------               -------         >>** ;
**<<   TOTAL         1.0000               1.0000          >>** ;

**<< New Weights for 6-Climate Zones: (per Avg. Mtr. Cnt for Yr)        >>** ;

**<<   Zone:       <<Dec 2009 Values>>  <<Year 20xx Values>>   >>** ;
**<< 1 High Mtn.     0.0062               a.aaaa          >>** ;
**<< 2 Low Desert    0.0417               b.bbbb          >>** ;
**<< 3 Coastal       0.1779               c.cccc          >>** ;
**<< 4 High Desert   0.0745               d.dddd          >>** ;
**<< 5 Intr. Valleys 0.3794               e.eeee          >>** ;
**<< 6 Basin         0.3203               f.ffff          >>** ;
**<< -------------- -------               -------         >>** ;
**<<   TOTAL         1.0000               1.0000          >>** ;

**<< New Weights for 6-Climate Zones: (per Avg. Mtr. Cnt for Yr)        >>** ;

**<<   Zone:       <<Year 2006 Values>>  <<Year 20xx Values>>   >>** ;
**<< 1 High Mtn.     0.0057               a.aaaa          >>** ;
**<< 2 Low Desert    0.0354               b.bbbb          >>** ;
**<< 3 Coastal       0.1888               c.cccc          >>** ;
**<< 4 High Desert   0.0676               d.dddd          >>** ;
**<< 5 Intr. Valleys 0.3854               e.eeee          >>** ;
**<< 6 Basin         0.3171               f.ffff          >>** ;
**<< -------------- -------               -------         >>** ;
**<<   TOTAL         1.0000               1.0000          >>** ;

**<<   Zone:       <<Year 2002 Values>>  <<Year 20xx Values>>   >>** ;
**<< 1 High Mtn.     0.0062               a.aaaa          >>** ;
**<< 2 Low Desert    0.0332               b.bbbb          >>** ;
**<< 3 Coastal       0.1998               c.cccc          >>** ;
**<< 4 High Desert   0.0662               d.dddd          >>** ;
**<< 5 Intr. Valleys 0.3807               e.eeee          >>** ;
**<< 6 Basin         0.3139               f.ffff          >>** ;
**<< -------------- -------               -------         >>** ;
**<<   TOTAL         1.0000               1.0000          >>** ;
```

```
**<<   Zone:      Year 1992 values    <<Year 2000 Values>>  <<Year 2001 Values>>   >>** ;
**<< 1 High Mtn.     0.0065                0.0063                0.0062          >>** ;
**<< 2 Low Desert    0.0282                0.0312                0.0320          >>** ;
**<< 3 Coastal       0.1900                0.2020                0.2021          >>** ;
**<< 4 High Desert   0.0620                0.0655                0.0638          >>** ;
**<< 5 Intr. Valleys 0.3862                0.3799                0.3809          >>** ;
**<< 6 Basin         0.3271                0.3151                0.3150          >>** ;
**<< -------------- -------                -------               -------         >>** ;
**<<   TOTAL         1.0000                1.0000                1.0000          >>** ;



%global wtz1 wtz2 wtz3 wtz4 wtz5 wtz6 ;

**<< Zone Weights a/o December 2017.      >>** ;
**<< (Data kindly provided by Andrew Tung/Idan Enright.) >>** ;
%let wtz1=0.005822 ;
%let wtz2=0.038497 ;
%let wtz3=0.185384 ;
%let wtz4=0.071615 ;
%let wtz5=0.383107 ;
%let wtz6=0.315576 ;



**<< Code inserted to do the daily "system avg." calc. from Six-Zone formulas  >>**;
**<< that are used to calc. the daily "system avg." values for the NWS data.   >>**;

data prelimD ;
   set scgwea.wdaily46(keep=date
       max_1-max_46
       min_1-min_46 );

 ** if date >= '01jan97'd; **<< Previous NWS data w/update goes through Dec. 31, 2000.    >>** ;

 ** if date >= '01nov02'd; **<< Current NWS data w/update NOW goes through Oct. 31, 2002. >>**;

   if date >= mdy(11,01,2002); **<< Current NWS data w/update NOW goes through Oct. 31, 2002. >>**;
                               **<< Use this format to "test" the DATE variable!!             >>**;

   year = year(date);
   month = month(date);

   ARRAY A_MAX(i) S0442MAX S1048MAX S1194MAX S4671MAX
                  S5085MAX S5114MAX S5115MAX S6175MAX
                  S6624MAX S6635MAX S6719MAX S7050MAX
                  S7306MAX S7851MAX S7888MAX S7905MAX S9367MAX;

         ARRAY A_MIN(i) S0442MIN S1048MIN S1194MIN S4671MIN
                  S5085MIN S5114MIN S5115MIN S6175MIN
                  S6624MIN S6635MIN S6719MIN S7050MIN
                  S7306MIN S7851MIN S7888MIN S7905MIN S9367MIN;

         ARRAY A_AVG(i) S0442AVG S1048AVG S1194AVG S4671AVG
                  S5085AVG S5114AVG S5115AVG S6175AVG
                  S6624AVG S6635AVG S6719AVG S7050AVG
                  S7306AVG S7851AVG S7888AVG S7905AVG S9367AVG;

         ARRAY A_hdd(i) S0442hdd S1048hdd S1194hdd S4671hdd
                  S5085hdd S5114hdd S5115hdd S6175hdd
                  S6624hdd S6635hdd S6719hdd S7050hdd
                  S7306hdd S7851hdd S7888hdd S7905hdd S9367hdd;

         ARRAY A_cdd(i) S0442cdd S1048cdd S1194cdd S4671cdd
                  S5085cdd S5114cdd S5115cdd S6175cdd
                  S6624cdd S6635cdd S6719cdd S7050cdd
                  S7306cdd S7851cdd S7888cdd S7905cdd S9367cdd;
```

```
  ** array a_max(i) max_1-max_46;  **<< There are NOW 46 stations on this file! >>**;
  ** array a_min(i) min_1-min_46;
  ** array a_avg(i) avg_1-avg_46;
  ** array a_hdd(i) hdd_1-hdd_46;
  ** array a_cdd(i) cdd_1-cdd_46;

%macro MapStatn(el);
*****************    "NWS" Data-Station         <---  "WeaDaily" Prelim Data-Station ***;
   s5115&el = (&el._3 ) ; *  LA Civic Cntr.     <---  Los Angeles Civic Center         ;
   s5114&el = (&el._4 ) ; *  LA Air Port        <---  Los Angeles International Apt.    ;
   s1194&el = (&el._9 ) ; *  Burbank            <---  Burbank                          ;
   s7050&el = (&el._11) ; *  Pomona             <---  Ontario                          ;
   s6719&el = (&el._12) ; *  Pasadena           <---  Pasadena                         ;
   s7306&el = (&el._46) ; *  Redlands           <---  RIALTO (replaced San Bernardino, Aug-2005)  ;
   s6175&el = (&el._20) ; *  Newport Beach      <---  Newport Beach                    ;
   s7888&el = (&el._21) ; *  Santa Ana          <---  Santa Ana                        ;
   s4671&el = (&el._23) ; *  Lake Arrowhead     <---  Big Bear                         ;
   s6624&el = (&el._24) ; *  Palmdale           <---  Lancaster                        ;
   s1048&el = (&el._26) ; *  Brawley            <---  El Centro                        ;
   s6635&el = (&el._28) ; *  Palm Springs       <---  Palm Springs                     ;
   s7905&el = (&el._31) ; *  Santa Barbara-Apt  <---  Santa Barbara (City?)            ;
   s0442&el = (&el._34) ; *  Bakersfield        <---  Bakersfield                      ;
   s9367&el = (&el._35) ; *  Visalia            <---  Fresno                           ;

   **<< Note: The two NWS stations below are NOT used in the climate zone averages. >>**;
   **<<       The NWS data set("dly50_96") has 17 stations, only 15 are used.       >>**;
   s5085&el = (&el._2 ) ; *  Long Beach(Apt)      <---  Long Beach                     ;
   s7851&el = (&el._32) ; *  San Luis Obispo(@CSU)<---  San Luis Obispo                ;

%mend;


%mapstatn(max);
%mapstatn(min);

*** %mapstatn(avg); **<< Do not need to do these calcs.! >>** ;
*** %mapstatn(hdd);
*** %mapstatn(cdd);

     do over a_max;
       a_avg = (a_max+a_min)/2;
       if a_avg-int(a_avg)=.5 then if mod(int(a_avg),2)=0
       then a_avg=int(a_avg);
       else a_avg=int(a_avg)+1;
       a_hdd = max(0,65-a_avg);
       a_cdd = max(0,a_avg-65);
     end;




  **<< Calc. Climate Zones and "System Average" values for:          >>**;
  **<<          Hdd, Cdd and Avg (temperature)                       >>**;

     do over a_hdd;
       a_hdd = round(a_hdd);
       a_cdd = round(a_cdd);
     end;


     hddz1 =   S4671hdd * &wtz1;
     hddz2 = ((S1048hdd + S6635hdd)/2) * &wtz2;
     hddz3 = ((S5114hdd + S6175hdd + S7905hdd)/3) * &wtz3;
     hddz4 = ((S0442hdd + S6624hdd + S9367hdd)/3) * &wtz4;
     hddz5 = ((S1194hdd + S6719hdd + S7050hdd + S7306hdd)/4) * &wtz5;
     hddz6 = ((S7888hdd + S5115hdd)/2) * &wtz6;
```

```
  **hdd   = round(sum (of hddz1-hddz6),0.1);
    hdd   =       sum (of hddz1-hddz6)     ;  **<< Do not round! >>** ;

    cddz1 =   S4671cdd * &wtz1;
    cddz2 = ((S1048cdd + S6635cdd)/2) * &wtz2;
    cddz3 = ((S5114cdd + S6175cdd + S7905cdd)/3) * &wtz3;
    cddz4 = ((S0442cdd + S6624cdd + S9367cdd)/3) * &wtz4;
    cddz5 = ((S1194cdd + S6719cdd + S7050cdd + S7306cdd)/4) * &wtz5;
    cddz6 = ((S7888cdd + S5115cdd)/2) * &wtz6;
  **cdd   = round(sum (of cddz1-cddz6),0.1);
    cdd   =       sum (of cddz1-cddz6)     ;  **<< Do not round! >>** ;

    avgz1 =   S4671avg * &wtz1;
    avgz2 = ((S1048avg + S6635avg)/2) * &wtz2;
    avgz3 = ((S5114avg + S6175avg + S7905avg)/3) * &wtz3;
    avgz4 = ((S0442avg + S6624avg + S9367avg)/3) * &wtz4;
    avgz5 = ((S1194avg + S6719avg + S7050avg + S7306avg)/4) * &wtz5;
    avgz6 = ((S7888avg + S5115avg)/2) * &wtz6;
  **avg   = round(sum (of avgz1-avgz6),0.1);
    avg   =       sum (of avgz1-avgz6)     ;  **<< Do not round! >>** ;


    drop   max_1-max_46
           min_1-min_46
           avg_1-avg_46
           hdd_1-hdd_46
           cdd_1-cdd_46;     **<< Drop ALL the PRELIMINARY weather data! >>** ;

run;






**** Calculate  HDD and CDD from FINAL NWS data set (dly50_96.sas7bdat) w/update (Updt4NWS.sas7bdat). ****;

* Commented Out by GDT;
***libname in 'S:\Weather\2005Bcap\Updt-NWS\';
***libname in2 'S:\Weather\2005Bcap\Updt-NWS\';

* Commented Out by GDT;
*libname in 'C:\Weather\2005Bcap\Updt-NWS\';
*libname in2 'C:\Weather\2005Bcap\Updt-NWS\';

  **<< This is the built-in drive on the computer BMW uses when telecommuting.  >>**;
  **<< It also has a copy of the same "wdaily46.sd2" file on the "S:\" drive.   >>**;
  **<< Having the file on the built-in drive allows for faster execution.       >>**;




 proc contents data=scgwea.dly50_96 ;
 run ;


 proc contents data=scgwea.Updt4NWS ;
 run ;



/******************************
******************************/


data Updt4NWS ;
    set scgwea.Updt4NWS ;
```

```
  **<< Keep only the Post-1996 observations thorugh YE-2002.              >>** ;
  **<< Note: the file "in2.Updt4NWS" actually goes through Oct. 31st, 2002! >>** ;
  if ( (date > mdy(12,31,1996))
   and (date < mdy(11,01,2002)) ) ;

run ;


data finalD ;
     merge scgwea.dly50_96  Updt4NWS ;
   by date ;

     year  = year(date);
     month = month(date);
     ARRAY A_MAX(I) S0442MAX S1048MAX S1194MAX S4671MAX
              S5085MAX S5114MAX S5115MAX S6175MAX
              S6624MAX S6635MAX S6719MAX S7050MAX
              S7306MAX S7851MAX S7888MAX S7905MAX S9367MAX;
     ARRAY A_MIN(I) S0442MIN S1048MIN S1194MIN S4671MIN
              S5085MIN S5114MIN S5115MIN S6175MIN
              S6624MIN S6635MIN S6719MIN S7050MIN
              S7306MIN S7851MIN S7888MIN S7905MIN S9367MIN;
     ARRAY A_AVG(I) S0442AVG S1048AVG S1194AVG S4671AVG
              S5085AVG S5114AVG S5115AVG S6175AVG
              S6624AVG S6635AVG S6719AVG S7050AVG
              S7306AVG S7851AVG S7888AVG S7905AVG S9367AVG;
     ARRAY A_hdd(I) S0442hdd S1048hdd S1194hdd S4671hdd
              S5085hdd S5114hdd S5115hdd S6175hdd
              S6624hdd S6635hdd S6719hdd S7050hdd
              S7306hdd S7851hdd S7888hdd S7905hdd S9367hdd;
     ARRAY A_cdd(I) S0442cdd S1048cdd S1194cdd S4671cdd
              S5085cdd S5114cdd S5115cdd S6175cdd
              S6624cdd S6635cdd S6719cdd S7050cdd
              S7306cdd S7851cdd S7888cdd S7905cdd S9367cdd;
     DO OVER A_MAX;
       A_AVG = (A_MAX+A_MIN)/2;
      IF A_AVG-INT(A_AVG)=.5 THEN IF MOD(INT(A_AVG),2)=0
      THEN A_AVG=INT(A_AVG);
      ELSE A_AVG=INT(A_AVG)+1;
      a_hdd = max(0,65-a_avg);
      a_cdd = max(0,a_avg-65);
     END;


     hddz1 =   S4671hdd * &wtz1;
     hddz2 = ((S1048hdd + S6635hdd)/2) * &wtz2;
     hddz3 = ((S5114hdd + S6175hdd + S7905hdd)/3) * &wtz3;
     hddz4 = ((S0442hdd + S6624hdd + S9367hdd)/3) * &wtz4;
     hddz5 = ((S1194hdd + S6719hdd + S7050hdd + S7306hdd)/4) * &wtz5;
     hddz6 = ((S7888hdd + S5115hdd)/2) * &wtz6;
  ** hdd   = round(sum (of hddz1-hddz6),0.1);
     hdd   =       sum (of hddz1-hddz6)     ;   **<< Do not round! >>** ;

     cddz1 =   S4671cdd * &wtz1;
     cddz2 = ((S1048cdd + S6635cdd)/2) * &wtz2;
     cddz3 = ((S5114cdd + S6175cdd + S7905cdd)/3) * &wtz3;
     cddz4 = ((S0442cdd + S6624cdd + S9367cdd)/3) * &wtz4;
     cddz5 = ((S1194cdd + S6719cdd + S7050cdd + S7306cdd)/4) * &wtz5;
     cddz6 = ((S7888cdd + S5115cdd)/2) * &wtz6;
  ** cdd   = round(sum (of cddz1-cddz6),0.1);
     cdd   =       sum (of cddz1-cddz6)     ;   **<< Do not round! >>** ;

     avgz1 =   S4671avg * &wtz1;
     avgz2 = ((S1048avg + S6635avg)/2) * &wtz2;
     avgz3 = ((S5114avg + S6175avg + S7905avg)/3) * &wtz3;
     avgz4 = ((S0442avg + S6624avg + S9367avg)/3) * &wtz4;
```

```
     avgz5 = ((S1194avg + S6719avg + S7050avg + S7306avg)/4) * &wtz5;
     avgz6 = ((S7888avg + S5115avg)/2) * &wtz6;
   **avg   = round(sum (of avgz1-avgz6),0.1);
     avg   =       sum (of avgz1-avgz6)     ;  **<< Do not round! >>** ;

  run ;

*;

/***********

libname out2 'S:\Weather\2016Tcap-Phase II\SoCalGas\';


**** Prelim and Final (Daily Data!) HDD, CDD and AVG together. ****;

data out2.DlySys_D ;
     set finalD prelimD ;  **<< Save a copy of DAILY Hdd, Cdd and Avg data for later use. >>** ;

  if (date < mdy(01,01,2015)) ;  **<< Complete DAILY data NOW through month of December 2014! >>**;

  keep date hdd cdd avg ;
run;


proc contents data=out2.DlySys_D ;
run ;


**** Prelim and Final HDD, CDD and AVG (Daily Data!) together--BY Station! ****;

data out2.DlyStn_D ;  **<< Save a copy of the Hdd by Station for later use. >>** ;
   set finalD prelimD ;
  if (date < mdy(01,01,2015)) ;  **<< Complete DAILY data NOW through month of December 2014! >>**;
run ;


proc contents data=out2.DlyStn_D ;
run ;
***********/


/***********
***********/

* Commented Out by GDT;
*libname out3 'C:\Weather\2016Tcap-Phase II\SoCalGas\';   **<< Alternative directory on my PERSONAL Share-Drive! >>** ;

**** Prelim and Final HDD, CDD and AVG (Daily Data!) together--BY Station! ****;
```

```
data scgwea.DlySys_D ;
     set finalD prelimD ;  **<< Save a copy of DAILY Hdd, Cdd and Avg data for later use. >>** ;

  if (date < mdy(01,01,2018)) ;  **<< Complete DAILY data NOW through month of December 2015! >>**;

  keep date hdd cdd avg ;
run;


proc contents data=scgwea.DlySys_D ;
run ;



data scgwea.DlyStn_D ;  **<< Save a copy of the Hdd by Station for later use. >>** ;
   set finalD prelimD ;
  if (date < mdy(01,01,2018)) ;  **<< Complete DAILY data NOW through month of December 2015! >>**;
run ;


proc contents data=scgwea.DlyStn_D ;
run ;




quit ;
```

SoCalGas SAS Code #2: SysAvgVar(Monthly).sas

```
Title1 "Create MONTHLY Summaries of System Average HDD and CDD data from previousl created DAILY" ;
Title2 "system average variables.  Calc. these summaries for System-Wide and by 15 Component Stations." ;

/*********************************************************************************/
/*                                                                 */
/*                                                                 */
/* Current Version of program file for our 2018 CGR work is:                      */
/*                                                                 */
/* Saved at: "/EDS_RB/Weather/2018 Cgr/SoCalGas/SysAvgVar(Monthly).sas"  */
/*                                                                 */
/*            on Feb 28th, 2018.                                          */
/*                                                                 */
/*                                                                 */
/* Origial AUTHOR: Loan Nguyen, on September 12, 2001                    */
/*    from file: "s:\Weather\2003Bcaphdd(LXN).sas"                    */
/*                                                                 */
/*********************************************************************************/




options mprint ;
/* %cour8p
   %cour8l */


 options ls=211 ps=69 ; **<<LANDSCAPE:  SAS-Monospace w/Roman 6pt. Font >>**;
*options ls=160 ps=90 ; **<<PORTRAIT:   SAS-Monospace w/Roman 6pt. Font >>**;

options date number notes ;


**** Calculate Monthly Summaries for Hdd, Cdd from daily data sets! ****;

libname scgdf '/EDS_RB/Weather/2018 Cgr/SoCalGas';


**<< Use copy saved to Built-in drive on "XP-lap top". >>**;
libname scgwea '/EDS_RB/Weather/2018 Cgr/SoCalGas';


proc contents data=scgwea.DlyStn_D ;
run ;


%global wtz1 wtz2 wtz3 wtz4 wtz5 wtz6 ;



**<< Zone Weights a/o December 2017.      >>** ;
**<< (Data kindly provided by Andrew Tung/Idan Enright. >>** ;
%let wtz1=0.005822 ;
%let wtz2=0.038497 ;
%let wtz3=0.185384 ;
%let wtz4=0.071615 ;
%let wtz5=0.383107 ;
%let wtz6=0.315576 ;


/**********************************************

**<< Zone Weights a/o December 2015.      >>** ;
```

```
**<< (Data kindly provided by Idan Enright. >>** ;
%let wtz1=0.005842 ;
%let wtz2=0.038341 ;
%let wtz3=0.186160 ;
%let wtz4=0.071125 ;
%let wtz5=0.385713 ;
%let wtz6=0.312819 ;


**<< Zone Weights a/o December 2014.      >>** ;
**<< (Data kindly provided by Idan Enright. >>** ;
%let wtz1=0.0061 ;
%let wtz2=0.0423 ;
%let wtz3=0.1763 ;
%let wtz4=0.0747 ;
%let wtz5=0.3812 ;  **<< wtz5 calcd value was 0.3813, but used 0.3812 so total adds to 100.00 exactly. >>** ;
%let wtz6=0.3194 ;

**<< Zone Weights a/o December 2013.      >>** ;
**<< (Data kindly provided by Andrew Tung. >>** ;
%let wtz1=0.0062 ;
%let wtz2=0.0424 ;
%let wtz3=0.1772 ;
%let wtz4=0.0742 ;
%let wtz5=0.3792 ;  **<< wtz5 calcd value was 0.3793, but used 0.3792 so total adds to 100.00 exactly. >>** ;
%let wtz6=0.3208 ;

**<< New Weights for 6-Climate Zones: (per Avg. Mtr. Cnt for Yr)        >>** ;
**<< Zone Weights a/o December 2010.      >>** ;
**<< (Data kindly provided by Andrew Tung. >>** ;
%let wtz1=0.0062 ;
%let wtz2=0.0418 ;
%let wtz3=0.1774 ;
%let wtz4=0.0746 ;
%let wtz5=0.3802 ;
%let wtz6=0.3198 ;  **<< wtz6 calcd value was 0.3199, but used 0.3198 so total adds to 100.00 exactly. >>** ;

**<< Zone Weights a/o December 2009.      >>** ;
**<< (Data kindly provided by Steve Tung. >>** ;
%let wtz1=0.0062 ;
%let wtz2=0.0417 ;
%let wtz3=0.1779 ;
%let wtz4=0.0745 ;
%let wtz5=0.3794 ;
%let wtz6=0.3203 ;

**<< Zone Weights a/o Year-2006.          >>** ;
**<< (Data kindly provided by Steve Tung. >>** ;
%let wtz1=0.0057 ;
%let wtz2=0.0354 ;
%let wtz3=0.1888 ;
%let wtz4=0.0676 ;
%let wtz5=0.3854 ;
%let wtz6=0.3171 ;

**<< Zone Weights a/o Year-2002.          >>** ;
**<< (Data kindly provided by Steve Tung. >>** ;
%let wtz1=0.0062 ;
%let wtz2=0.0332 ;
%let wtz3=0.1998 ;
%let wtz4=0.0662 ;
%let wtz5=0.3807 ;
%let wtz6=0.3139 ;

**<< Zone Weights a/o Year-2001.          >>** ;
**<< (Data kindly provided by Steve Tung. >>** ;
%let wtz1=0.0062 ;
```

```
%let wtz2=0.0320 ;
%let wtz3=0.2021 ;
%let wtz4=0.0638 ;
%let wtz5=0.3809 ;
%let wtz6=0.3150 ;

**<< Zone Weights a/o Year-2000.          >>** ;
**<< (Data kindly provided by Steve Tung. >>** ;
%let wtz1=0.0063 ;
%let wtz2=0.0312 ;
%let wtz3=0.2020 ;
%let wtz4=0.0655 ;
%let wtz5=0.3799 ;
%let wtz6=0.3151 ;

**<< Zone Weights a/o Year-1992. >>** ;
%let wtz1=0.0065 ;
%let wtz2=0.0282 ;
%let wtz3=0.1900 ;
%let wtz4=0.0620 ;
%let wtz5=0.3862 ;
%let wtz6=0.3271 ;
**********************************************/




**<< Code inserted to do the daily "system avg." calc. from Six-Zone formulas  >>**;
**<< that are used to calc. the daily "system avg." values for the NWS data.   >>**;

data DlyData ;
    set scgwea.DlyStn_D ;



 ***<< Calc. Climate Zones and "System Average" values for:          >>**;
 ***<<         Hdd, Cdd and Avg (temperature)                       >>**;
 **  hddz1 =   S4671hdd * &wtz1;
 **  hddz2 = ((S1048hdd + S6635hdd)/2) * &wtz2;
 **  hddz3 = ((S5114hdd + S6175hdd + S7905hdd)/3) * &wtz3;
 **  hddz4 = ((S0442hdd + S6624hdd + S9367hdd)/3) * &wtz4;
 **  hddz5 = ((S1194hdd + S6719hdd + S7050hdd + S7306hdd)/4) * &wtz5;
 **  hddz6 = ((S7888hdd + S5115hdd)/2) * &wtz6;
 ****hdd   = round(sum (of hddz1-hddz6),0.1);
 **  hdd   =       sum (of hddz1-hddz6)     ;  **<< Do not round! >>** ;
 ** ;
 **  cddz1 =   S4671cdd * &wtz1;
 **  cddz2 = ((S1048cdd + S6635cdd)/2) * &wtz2;
 **  cddz3 = ((S5114cdd + S6175cdd + S7905cdd)/3) * &wtz3;
 **  cddz4 = ((S0442cdd + S6624cdd + S9367cdd)/3) * &wtz4;
 **  cddz5 = ((S1194cdd + S6719cdd + S7050cdd + S7306cdd)/4) * &wtz5;
 **  cddz6 = ((S7888cdd + S5115cdd)/2) * &wtz6;
 ****cdd   = round(sum (of cddz1-cddz6),0.1);
 **  cdd   =       sum (of cddz1-cddz6)     ;  **<< Do not round! >>** ;
 ** ;
 **  avgz1 =   S4671avg * &wtz1;
 **  avgz2 = ((S1048avg + S6635avg)/2) * &wtz2;
 **  avgz3 = ((S5114avg + S6175avg + S7905avg)/3) * &wtz3;
 **  avgz4 = ((S0442avg + S6624avg + S9367avg)/3) * &wtz4;
 **  avgz5 = ((S1194avg + S6719avg + S7050avg + S7306avg)/4) * &wtz5;
 **  avgz6 = ((S7888avg + S5115avg)/2) * &wtz6;
 ****avg   = round(sum (of avgz1-avgz6),0.1);
 **  avg   =       sum (of avgz1-avgz6)     ;  **<< Do not round! >>** ;
 ** ;

 **<< Recover Avg, Hdd and Cdd by TZone: >>** ;
```

```
   avgz0 = avg ;
   avgz1 = avgz1 / &wtz1 ;
   avgz2 = avgz2 / &wtz2 ;
   avgz3 = avgz3 / &wtz3 ;
   avgz4 = avgz4 / &wtz4 ;
   avgz5 = avgz5 / &wtz5 ;
   avgz6 = avgz6 / &wtz6 ;

   hddz0 = hdd ;
   hddz1 = hddz1 / &wtz1 ;
   hddz2 = hddz2 / &wtz2 ;
   hddz3 = hddz3 / &wtz3 ;
   hddz4 = hddz4 / &wtz4 ;
   hddz5 = hddz5 / &wtz5 ;
   hddz6 = hddz6 / &wtz6 ;

   cddz0 = cdd ;
   cddz1 = cddz1 / &wtz1 ;
   cddz2 = cddz2 / &wtz2 ;
   cddz3 = cddz3 / &wtz3 ;
   cddz4 = cddz4 / &wtz4 ;
   cddz5 = cddz5 / &wtz5 ;
   cddz6 = cddz6 / &wtz6 ;


    keep   year month date
           avg avgz0-avgz6
           hdd hddz0-hddz6
           cdd cddz0-cddz6

         S0442hdd S1048hdd S1194hdd S4671hdd
         S5085hdd S5114hdd S5115hdd S6175hdd
         S6624hdd S6635hdd S6719hdd S7050hdd
         S7306hdd S7851hdd S7888hdd S7905hdd S9367hdd

         S0442cdd S1048cdd S1194cdd S4671cdd
         S5085cdd S5114cdd S5115cdd S6175cdd
         S6624cdd S6635cdd S6719cdd S7050cdd
         S7306cdd S7851cdd S7888cdd S7905cdd S9367cdd   ;  **<< Drop ALL othe variables! >>** ;

run;


*;
proc means data=DlyData nway noprint;
    class year month;
    var hdd hddz0-hddz6
        S0442hdd S1048hdd S1194hdd S4671hdd
        S5085hdd S5114hdd S5115hdd S6175hdd
        S6624hdd S6635hdd S6719hdd S7050hdd
        S7306hdd S7851hdd S7888hdd S7905hdd S9367hdd
        cdd cddz0-cddz6
        S0442cdd S1048cdd S1194cdd S4671cdd
        S5085cdd S5114cdd S5115cdd S6175cdd
        S6624cdd S6635cdd S6719cdd S7050cdd
        S7306cdd S7851cdd S7888cdd S7905cdd S9367cdd;
    output out=smByStn  sum=;
run ;
*;


data smByZone ;
    set smByStn  ;

    hdd = round(hdd);
  ** hddz0 = hddz0 ;  **<< Do not round this variable! >>** ;
```

```
     hddz1 = round(hddz1);
     hddz2 = round(hddz2);
     hddz3 = round(hddz3);
     hddz4 = round(hddz4);
     hddz5 = round(hddz5);
     hddz6 = round(hddz6);

     cdd = round(cdd);
  ** cddz0 = cddz0 ;  **<< Do not round this variable! >>** ;
     cddz1 = round(cddz1);
     cddz2 = round(cddz2);
     cddz3 = round(cddz3);
     cddz4 = round(cddz4);
     cddz5 = round(cddz5);
     cddz6 = round(cddz6);

     keep year month hdd cdd hddz0-hddz6  cddz0-cddz6 ;
run;


data ByStn ;
   set smByStn ;

  dateYYMM = mdy(month,1,year) ;

  if (dateYYMM < mdy(1,1,2018)) ;  **<< Only Pre Jan-2018 observations! >>** ;

  hdd = round(hdd) ;
  cdd = round(cdd) ;

 **<< The "labeling" below is based on the pre-1997 data from NWS! >>** ;
  label s5115hdd = "LA Civic Cntr." ;
  label s5114hdd = "LA Air Port" ;
  label s1194hdd = "Burbank" ;
  label s7050hdd = "Pomona" ;
  label s6719hdd = "Pasadena" ;
  label s7306hdd = "Redlands" ;
  label s6175hdd = "Newport Beach" ;
  label s7888hdd = "Santa Ana" ;
  label s4671hdd = "Lake Arrowhead" ;
  label s6624hdd = "Palmdale" ;
  label s1048hdd = "Brawley" ;
  label s6635hdd = "Palmlabel springs" ;
  label s7905hdd = "Santa Barbara-Apt" ;
  label s0442hdd = "Bakersfield" ;
  label s9367hdd = "Visalia" ;

  label s5085hdd = "Long Beach(Apt)" ;
  label s7851hdd = "San Luis Obispo(@CSU)" ;

  label hdd = "Syst-Avg. Hdd" ;
run ;


proc sort data=ByStn ;
   by year month ;
run ;


proc print data=ByStn  uniform split="/" ;
 id year;
 by year;
```

```
 sumby year;
    var dateYYMM hdd
        S4671hdd
        S1048hdd   S6635hdd
        S5114hdd   S6175hdd   S7905hdd
        S0442hdd   S6624hdd   S9367hdd
        S1194hdd   S6719hdd   S7050hdd   S7306hdd
        S7888hdd   S5115hdd  ;
    **  s5085hdd   s7851hdd   ;
    sum hdd
        S4671hdd
        S1048hdd   S6635hdd
        S5114hdd   S6175hdd   S7905hdd
        S0442hdd   S6624hdd   S9367hdd
        S1194hdd   S6719hdd   S7050hdd   S7306hdd
        S7888hdd   S5115hdd  ;
    **  s5085hdd   s7851hdd   ;
    format  dateYYMM worddate3.  hdd
        S4671hdd
        S1048hdd   S6635hdd
        S5114hdd   S6175hdd   S7905hdd
        S0442hdd   S6624hdd   S9367hdd
        S1194hdd   S6719hdd   S7050hdd   S7306hdd
        S7888hdd   S5115hdd   6. ;
    **  s5085hdd   s7851hdd   6. ;

  label s5115hdd = "Los Angeles/ Civic Cntr/Zone6/--------"
       s5114hdd = "Los Angeles/ (Apt) /Zone3/--------"
       s1194hdd = "Burbank /Zone5/--------"
       s7050hdd = "Pomona /Zone5/--------"
       s6719hdd = "Pasadena /Zone5/--------"
       s7306hdd = "Redlands /Zone5/--------"
       s6175hdd = "Newport /Beach /Zone3/--------"
       s7888hdd = "Santa Ana /Zone6/--------"
       s4671hdd = "Lake Arrow-/ head /Zone1/--------"
       s6624hdd = "Palmdale /Zone4/--------"
       s1048hdd = "Brawley /Zone2/--------"
       s6635hdd = "Palm Springs /Zone2/--------"
       s7905hdd = "Santa Bar-/bara(Apt) /Zone3/--------"
       s0442hdd = "Bakers-/ field /Zone4/--------"
       s9367hdd = "Visalia /Zone4/--------"

       hdd = "Syst-Avg./ Hdd /Zone1-6/--------" ;

 **<< Note: The last two("s5085hdd" and "s7851hdd") are not in the "Sys-Avg" calc. >>** ;
 **<<      The first six(6) sets are used in Climate Zones 1-6, respectively.     >>** ;
 **<<      The variable "hdd" is the System-Average value from Climate Zones 1-6. >>** ;
 **  s5085hdd = "Long Beach/ (Apt) /No-Zone/--------"        ;
 **  s7851hdd = "San Luis /Obispo(@CSU) /No-Zone/--------"   ;

run ;



proc means data=ByStn noprint ;
   class year ;
    var hdd
        S4671hdd
        S1048hdd   S6635hdd
        S5114hdd   S6175hdd   S7905hdd
        S0442hdd   S6624hdd   S9367hdd
        S1194hdd   S6719hdd   S7050hdd   S7306hdd
        S7888hdd   S5115hdd  ;
     ** s5085hdd   s7851hdd   ;
    output out=SumByYr sum=;
run ;
```

```
/**
data SumByYr ;
  set SumByYr ;
  if (year=.) then delete ;
run ;
**/

proc print data=SumByYr  uniform split="/" ;
  where year ne . ;
     var year hdd
        S4671hdd
        S1048hdd   S6635hdd
        S5114hdd   S6175hdd   S7905hdd
        S0442hdd   S6624hdd   S9367hdd
        S1194hdd   S6719hdd   S7050hdd   S7306hdd
        S7888hdd   S5115hdd  ;
      ** s5085hdd   s7851hdd   ;
     format   hdd
        S4671hdd
        S1048hdd   S6635hdd
        S5114hdd   S6175hdd   S7905hdd
        S0442hdd   S6624hdd   S9367hdd
        S1194hdd   S6719hdd   S7050hdd   S7306hdd
        S7888hdd   S5115hdd   6. ;
      ** s5085hdd   s7851hdd   6. ;

  label s5115hdd = "Los Angeles/ Civic Cntr/Zone6/--------"
       s5114hdd = "Los Angeles/ (Apt) /Zone3/--------"
       s1194hdd = "Burbank /Zone5/--------"
       s7050hdd = "Pomona /Zone5/--------"
       s6719hdd = "Pasadena /Zone5/--------"
       s7306hdd = "Redlands /Zone5/--------"
       s6175hdd = "Newport /Beach /Zone3/--------"
       s7888hdd = "Santa Ana /Zone6/--------"
       s4671hdd = "Lake Arrow-/ head /Zone1/--------"
       s6624hdd = "Palmdale /Zone4/--------"
       s1048hdd = "Brawley /Zone2/--------"
       s6635hdd = "Palm Springs /Zone2/--------"
       s7905hdd = "Santa Bar-/bara(Apt) /Zone3/--------"
       s0442hdd = "Bakers-/ field /Zone4/--------"
       s9367hdd = "Visalia /Zone4/--------"

       hdd = "Syst-Avg./ Hdd /Zone1-6/--------" ;

 **<< Note: The last two("s5085hdd" and "s7851hdd") are not in the "Sys-Avg" calc. >>** ;
 **<<       The first six(6) sets are used in Climate Zones 1-6, respectively.     >>** ;
 **<<       The variable "hdd" is the System-Average value from Climate Zones 1-6. >>** ;
 **  s5085hdd = "Long Beach/ (Apt) /No-Zone/--------"        ;
 **  s7851hdd = "San Luis /Obispo(@CSU) /No-Zone/--------"   ;

run ;


**** Prelim and final HDD and CDD together ****;
data all;
    set smByZone ;
  date = mdy(month,01,year) ;  format date monyy7. ;
  if (date < mdy(01,01,2018)) ;  **<< Complete Monthly data only through month of December 2017. >>**;
run;

proc sort data=all out=monthly;
    by year month;
```

```
run;




/***********
***********/
data scgwea.mn50_17 ;  **<< Save a copy of the Monthly Hdd/Cdd data for later use. >>** ;
   set monthly ;
 **  keep date year month hdd cdd ;   **<< Keep all zone data as well as system average data. >>** ;
 **<< Export a copy as ".xls" as well. >>** ;
run ;

data scgwea.ByStn ;  **<< Save a copy of the Hdd by Station for later use. >>** ;
   set ByStn ;
 **<< Export a copy as ".xls" as well. >>** ;
run ;






***<< Print Summary Tables of Hdd and Cdd by month  with Annual Totals >>*** ;

proc transpose data=monthly out=hddSmry prefix=mon ;
   by year;
   id month ;
   var hdd ;
run ;

data hddSmry ;
   set hddSmry ;
 hddTot = sum(of mon1-mon12) ;
run ;

proc print data=hddSmry ;
 id year ;
 var hddTot mon1-mon12 ;
title1 'Monthly Heating Degree-Days from 1950 thru 2017 (Month-to-Date)';
title2 "   " ;
run ;



proc transpose data=monthly out=cddSmry prefix=mon ;
   by year;
   id month ;
   var cdd ;
run ;

data cddSmry ;
   set cddSmry ;
 cddTot = sum(of mon1-mon12) ;
run ;

proc print data=cddSmry ;
 id year ;
 var cddTot mon1-mon12 ;
title1 'Monthly Cooling Degree-Days from 1950 thru 2017 (Month-to-Date)';
run ;
```

```
/********  BEGIN:  Surpress execution of Stat. Analysis **************************
proc print data=monthly;
     where 1950 <= year <= 2014;
     id year month;
     var date hdd cdd ;
title1 'Monthly Heating/Cooling degree days from 1950 thru 2014(Month-to-Date)';
run;


proc plot data=monthly ;
     where 1950 <= year <= 1959;
    plot hdd*date='*' ;
    plot cdd*date='+' ;
title1 'Plot of MONTHLY Heating/Cooling degree days from 1950 thru 1959';
run ;

proc plot data=monthly ;
     where 1960 <= year <= 1969;
    plot hdd*date='*' ;
    plot cdd*date='+' ;
title1 'Plot of MONTHLY Heating/Cooling degree days from 1960 thru 1969';
run ;

proc plot data=monthly ;
     where 1970 <= year <= 1979;
    plot hdd*date='*' ;
    plot cdd*date='+' ;
title1 'Plot of MONTHLY Heating/Cooling degree days from 1970 thru 1979';
run ;

proc plot data=monthly ;
     where 1980 <= year <= 1989;
    plot hdd*date='*' ;
    plot cdd*date='+' ;
title1 'Plot of MONTHLY Heating/Cooling degree days from 1980 thru 1989';
run ;

proc plot data=monthly ;
     where 1990 <= year <= 1999;
    plot hdd*date='*' ;
    plot cdd*date='+' ;
title1 'Plot of MONTHLY Heating/Cooling degree days from 1990 thru 1999';
run ;

proc plot data=monthly ;
     where 2000 <= year <= 2009;
    plot hdd*date='*' ;
    plot cdd*date='+' ;
title1 'Plot of MONTHLY Heating/Cooling degree days from 2000 thru 2009';
run ;

proc plot data=monthly ;
     where 2010 <= year <= 2014;
    plot hdd*date='*' ;
    plot cdd*date='+' ;
title1 'Plot of MONTHLY Heating/Cooling degree days from 2010 thru 2014';
run ;




proc means data=all nway ;   ** noprint;
     class year;
     var hdd cdd ;
     output out=year sum=hdd cdd ;
title1 'Summary Statistics by YEAR:' ;
```

```
run;

proc print data=year;
     where 1950 <= year <= 2014;
     id year;
     var hdd cdd;
     sum hdd cdd;
title1 'Annual Heating/Cooling degree days from 1950 thru 2014';

proc plot data=year ;
     where 1950 <= year <= 2014;
    plot hdd*year='*' ;
    plot cdd*year='+' ;
title1 'Plot of ANNUAL Heating/Cooling degree days from 1950 thru 2014';
run ;


run;

proc means data=monthly nway ;   ** noprint;
     where 1950 <= year <= 2014;
     class month;
     var hdd cdd;
     output out=avg mean=;
title1 'Summary Statistics by MONTH (for years 1950-2014):' ;
run;

proc print;
     id month;
     var hdd cdd ;
     sum hdd cdd ;
title1 '65-Year Average Monthly Heating Degree Days';
run;

**** "65" <<year standard deviation>> ****;
proc means data=year ;   **noprint;
     where 1950 <= year <= 2014;
     var hdd cdd;
     output out=std65 std=Hstd65 Cstd65;
run;

**** "65" <<year "normal" as Average Year>> ****;
proc means data=year ;   **noprint;
     where 1950 <= year <= 2014;
     var hdd cdd ;
     output out=avg65 mean=hdd cdd ;
run;

**** Small sample - Use t-statistics with 65-1 DF ****;
**** 1 in 35 probability for "HOT" and "COLD" years ****;
data finalyr;
     set avg65;
     if _n_ = 1 then set std65;
     p35   = 1-(1/35);
     tstat = tinv(p35,65-1);
*    zstat = probit(p35);
     Havgyr = round(HDD);
     Hcolyr = round(Havgyr + (tstat * Hstd65)); **<< "More" Hdd means LOWER temps.! >>**;
     Hhotyr = round(Havgyr - (tstat * Hstd65);

     Cavgyr = round(CDD);
     Ccolyr = round(Cavgyr - (tstat * Cstd65));
     Chotyr = round(Cavgyr + (tstat * Cstd65)); **<< "More" Cdd means HIGHER temps.! >>**;
run;

data finalmo;
     set avg(rename=(hdd=Havg_mo cdd=Cavg_mo));
```

```
     if _n_ = 1 then set finalyr;

     Hratio  = Havg_mo / Hdd;
     Havg_mo = Hratio * Havgyr;
     Hhot_mo = Hratio * Hhotyr;
     Hcol_mo = Hratio * Hcolyr;

     Cratio  = Cavg_mo / Cdd;
     Cavg_mo = Cratio * Cavgyr;
     Chot_mo = Cratio * Chotyr;
     Ccol_mo = Cratio * Ccolyr;

     drop _type_ _freq_ p35;
run;

proc print;
     id month;
     var Havgyr Hcolyr Hhotyr Havg_mo Hcol_mo Hhot_mo tstat Hstd65;
     format Havg_mo Hcol_mo Hhot_mo 9.4;
     sum Havg_mo Hhot_mo Hcol_mo;
title1 'Avg, Cold and Hot Year Heating Degree-Days ';
run;


proc print;
     id month;
     var Cavgyr Ccolyr Chotyr Cavg_mo Ccol_mo Chot_mo tstat Cstd65;
     format Cavg_mo Ccol_mo Chot_mo 9.4;
     sum Cavg_mo Chot_mo Ccol_mo;
title1 'Avg, Cold and Hot Year Cooling Degree-Days ';
run;

*********  END:    Surpress execution of Stat. Analysis *************************/





quit ;
```

SoCalGas SAS Code #3: MinAvg-Freq(SoCalGas)ByMonth.sas

```
title1 "Calculate Min{Avg} (Minimum Average Daily Temp.) by Months for all data over a specified range of YEARS." ;

/***************************************************************************************

   File: S:\Weather\2018 Cgr\SoCalGas\MinAvg-Freq(SoCalGas)ByMonth.sas

 ***************************************************************************************/


options date number source notes ;
options mprint ;

/* %cour8l
   %cour8p; */

 options ls=160 ps=90 ;  *<To get PORTRAIT  and SAS-Monospace w/Roman 6pt. FONT >*;
*options ls=211 ps=69 ;  *<To get LANDSCAPE and SAS-Monospace w/Roman 6pt. FONT >*;

**options nomprint ;


     **<< Data set "DlySys_d.sas7bdat" was created by SAS program on file:              >>** ;
     **<<      "S:\Weather\2018 Cgr\SoCalGas\SysAvgVar(Daily).sas"                              >>** ;
     **<<                                                         >>** ;
libname in '/EDS_RB/Weather/2018 Cgr/SoCalGas';



   **<< Directory to save a copy of output data set! >>** ;
libname out '/EDS_RB/Weather/2018 Cgr/SoCalGas';


proc contents data=in.DlySys_d ;
run ;



%let startyr=1950;  ***<< Value of "Start Year" >>***;
%let lastyr=2015 ;  ***<< Value of "Last Year" >>***;

/********
%let tgtmonth=xx ;  ***<< Value of "Target Month", i.e., 1,2,3, ... 12 >>** ;
********/




proc format ;
   value mmmmmm    1='Jan'
                2='Feb'
                3='Mar'
                4='Apr'
                5='May'
                6='Jun'
                7='Jul'
                8='Aug'
                9='Sep'
               10='Oct'
               11='Nov'
               12='Dec'
               13='Min4Yr' ;
run ;
```

```
%macro FreqMon(name_mon,tgtmonth) ;

data combined ;
     set in.DlySys_d ;

year = year(date) ;
month= month(date) ;
day = day(date) ;
  **<< To "Select" Winter Months Only. >>**;
  **  if month in (1,2,3,11,12);

*** hdd = round(avg,1);  **<< Comment out ... Do NOT round data this time! >>** ;
if ((year >= &startyr) & (year <= &lastyr)) ;

%if (&tgtmonth >= 1 and &tgtmonth <= 12) %then
  %do ;
if (month = &tgtmonth) ;  **<< To select only a specific month! >>** ;
  %end ;
%else
  %do ;
month = 13 ;  **<< Set "month" variable to "13" and select ALL "months" of the YEAR! >>** ;
  %end ;
run;


proc sort data=combined;
 by year month day ;
run;

***proc contents data=combined ;
***run ;

proc means data=combined noprint;
  by year month ;
  var avg ;
  output out=&name_mon min=MinAvg;
  title1 "Minium{avg} (Minimum Avg. Daily Temp.) for &name_mon  by YEAR=&startyr to &lastyr";
run;

proc print data=&name_mon ;
run ;

%mend ;








%let startyr=1950;  ***<< Value of "Start Year" >>***;
%let lastyr=2017 ;  ***<< Value of "Last Year" >>***;


/************
%let name_mon='January ';  ***<< NAME of "Target Month" >>***;
%let tgtmonth= 1 ;    ***<< Value of "Target Month" >>***;
************/

%FreqMon(JAN,1) ;

data AllMonth ;
  set JAN ;
run ;
```

```
%FreqMon(FEB,2) ;

data AllMonth ;
  set AllMonth FEB ;
run ;


%FreqMon(MAR,3) ;

data AllMonth ;
  set AllMonth MAR ;
run ;


%FreqMon(APR,4) ;

data AllMonth ;
  set AllMonth APR ;
run ;


%FreqMon(MAY,5) ;

data AllMonth ;
  set AllMonth MAY ;
run ;


%FreqMon(JUN,6) ;

data AllMonth ;
  set AllMonth JUN ;
run ;


%FreqMon(JUL,7) ;

data AllMonth ;
  set AllMonth JUL ;
run ;


%FreqMon(AUG,8) ;

data AllMonth ;
  set AllMonth AUG ;
run ;


%FreqMon(SEP,9) ;

data AllMonth ;
  set AllMonth SEP ;
run ;


%FreqMon(OCT,10) ;

data AllMonth ;
  set AllMonth OCT ;
run ;


%FreqMon(NOV,11) ;
```

```
data AllMonth ;
  set AllMonth NOV ;
run ;


%FreqMon(DEC,12) ;

data AllMonth ;
  set AllMonth DEC ;
run ;


%FreqMon(ALL,13) ;

data AllMonth ;
  set AllMonth All ;
run ;



proc Tabulate data=AllMonth;
  class year month ;
  var MinAvg ;
  table year, MinAvg*f=6.2*(month)/rts=6 ;
  label MinAvg='Min{Avg} by Mo' ;
 *keylabel All ='Min4Yr';
  format month mmmmmm. ;
  title2 "Min{avg} by Months for all Years from YEAR=&startyr to &lastyr";
run;


proc sort data=AllMonth out=MinAvg_d;
  by month year MinAvg ;
run ;


proc print data=MinAvg_d ;
  by month ;
  pageby month ;
  var year MinAvg ;
run ;


data out.MinAvg_d ;  *<< Save a copy for later use!  >>** ;
  set MinAvg_d ;
 drop _freq_ _type_ ;
run ;

proc contents data=out.MinAvg_d ;
run ;


/************************
 **<< Export a copy as ".dbf" as well. >>** ;
 **<< Note: Must "delete" the prior version, othewise the save will not execute. >>** ;
proc dbload dbms=dbase data=MinAvg_d ;
path='S:\Weather\2016Tcap-Phase II\SoCalGas\MinAvg_d.dbf';
limit=0;
load;
run;
*************************/



quit ;
```

SoCalGas SAS Code #4: GEV4DlyTemp(NLReg2)ByMonthMACRO_Scg.sas

```
Title1 "Data Analysis for Maximum/Minimum Daily SysAvg Temperatures (Un-Rounded)." ;
Title2 "Fit GEV, Double-Exp and T-Dist Probability Models to Empirical CDF using NL-OLS Regression Methods." ;

/*****************************************************************************************************/
/*                                                                                                  */
/*                                                                                                  */
/*                                                                                                  */
/* FILE SAVED: "S:\Weather\2018 Cgr\SoCalGas\GEV4DlyTemp(NLReg2)ByMonthMACRO_Scg.sas"
            */
/*          Purpose: Annual Max of Negative of Min. Temp.                                           */
/*          for each of calendar months Jan-Dec, and the entire year(index=13).                     */
/*                                                                                                  */
/*          Fit GEV models (3-parameter and 2-parameter), plus a simple T-Distribution model.       */
/*                                                                                                  */
/*                                                                                                  */
/*****************************************************************************************************/



options mprint ;
/* %cour8p
   %cour8l */


 options ls=211 ps=69 ; **<<LANDSCAPE:  SAS-Monospace w/Roman 6pt. Font >>**;
*options ls=160 ps=90 ; **<<PORTRAIT:   SAS-Monospace w/Roman 6pt. Font >>**;

options date number notes ;



libname out1 '/EDS_RB/Weather/2018 Cgr/SoCalGas' ;  **<< Directory for daily weather variables as INPUT. >>**;

libname out2 '/EDS_RB/Weather/2018 Cgr/SoCalGas/MinTemp' ;  **<< Directory for estimation results OUPUT files. >>** ;


proc contents data=out1.DlySys_d ;
run ;

data seriesD ;
  set out1.DlySys_d ;
  year = year(date) ;
  month = month(date) ;
  posAvg = avg ;
  negAvg = -avg ;
run ;


proc means data=seriesD noprint nway ;
 class year month ;
  var posAvg negAvg ;
 output out=mostat
                 mean=posAvg  negAvg
                  max=MxPosAvg MxNegAvg
                  min=MnPosAvg MnNegAvg    ;
run;


proc sort data=mostat ;
   by year month ;
run ;
```

```
data mostat ;
  set mostat ;
  MxPRatio = MxPosAvg/ PosAvg  ;
  MnPRatio = MnPosAvg/ PosAvg  ;
  MxNRatio = MxNegAvg/ NegAvg  ;
  MnNRatio = MnNegAvg/ NegAvg  ;
run ;




/***************
***<< Print Summary Tables of Means/Minimums/Maximums of daily NEGATIVE-Temperatures (degrees-F). >>*** ;

proc transpose data=mostat out=AvTData prefix=AvT_ ;
  where (year < 2015) ;
  by year;
  id month ;
  var NegAvg  ;
run ;

data AvTData ;
  set AvTData ;

 if (mod(year,4)=0) then do ;
       AvT_13 = (AvT_1 + AvT_3 + AvT_5 + AvT_7 + AvT_8 + AvT_10 + AvT_12)*31
             + (AvT_4 + AvT_6 + AvT_9 + AvT_11)*30
             + (AvT_2)*29 ;
       AvT_13 = AvT_13 / 366 ;
       end ;
 else do ;
       AvT_13 = (AvT_1 + AvT_3 + AvT_5 + AvT_7 + AvT_8 + AvT_10 + AvT_12)*31
             + (AvT_4 + AvT_6 + AvT_9 + AvT_11)*30
             + (AvT_2)*28 ;
       AvT_13 = AvT_13 / 365 ;
       end ;

run ;

proc print data=AvTData ;
 id year ;
 var AvT_13 AvT_1-AvT_12 ;
title3 'Monthly Mean NEGATIVE Temperature (Deg-F) from 1950 thru 2015.';
run ;




proc transpose data=mostat out=MnTData prefix=Mn2016T_ ;
  where (year < 2015) ;
  by year;
  id month ;
  var MnNegAvg ;
run ;

data MnTData ;
  set MnTData ;
 MnT_13 = min(of MnT_1-MnT_12) ;
run ;

proc print data=MnTData ;
```

```
 id year ;
 var MnT_13 MnT_1-MnT_12 ;
title3 'Monthly MINIMUM NEGATIVE-Temperature (Deg-F) from 1950 thru 2015.';
run ;
****************/



proc transpose data=mostat out=MxTData prefix=MxT_ ;
   where (year < 2018) ;
   by year;
   id month ;
   var MxNegAvg ;
run ;

data MxTData ;
   set MxTData ;
 MxT_13 = max(of MxT_1-MxT_12) ;
run ;

proc print data=MxTData ;
 id year ;
 var MxT_13 MxT_1-MxT_12 ;
title3 'Monthly MAXIMUM NEGATIVE-Temperature (Deg-F) from 1950 thru 2017.';
run ;



/****************
***<< Descriptive Statistics: Maximums of daily NEGATIVE-Temperatures (Deg-F) for Year and each calendar month. >>*** ;


proc corr data=MxTData ;
  var MxTyr MxT_1 - MxT_12 ;
title3 'Correlation Matrix of Monthly Maximum NEGATIVE-Temperatures (Deg-F) within same year.';
run ;

proc arima data=MxTData ;
  identify var=MxT_13 ;
  identify var=MxT_1 ;
  identify var=MxT_2 ;
  identify var=MxT_3 ;
  identify var=MxT_4 ;
  identify var=MxT_5 ;
  identify var=MxT_6 ;
  identify var=MxT_7 ;
  identify var=MxT_8 ;
  identify var=MxT_9 ;
  identify var=MxT_10 ;
  identify var=MxT_11 ;
  identify var=MxT_12 ;
title3 "Auto-correlation analysis of each calendar month's Maximum NEGATIVE-Temperatures (Deg-F) within same year.";
run ;

proc univariate normal data=MxTData plot ;
  id year ;
  var MxT_13 MxT_1 - MxT_12 ;
title3 "Probability plots and tests for NORMALity by each calendar month's Maximun NEGATIVE-Temperatures (Deg-F) time series.";
```

```
run ;


proc means data=MxTData ;
  var MxT_1 - MxT_12  MxT_13 ;
run ;
****************/








***<< Statistical Estimation of GEV Models: Maximums of daily heating degrees for Year and each calendar month. >>*** ;

%macro RankIt(file=MxTData,var=MxT_13,rank=Rank,prob=PrMxT_13,Nobser=68,PltValue=0.375) ;
proc sort data=&file ;
  by &var ;
run ;

data &file ;
   set &file ;
   retain &rank 0     alpha &pltvalue ;

  &rank = &rank + 1 ;
  &prob =  (&rank - alpha) / (&Nobser +(1 - 2*alpha)) ;
run ;

proc print data=&file ;
  var &var &rank &prob alpha year ;
run ;
%mend RankIt ;




%macro GEVfit(file=MxTData,ofile=MxTNL1,outfit=fit1,outest=est1,depvar=PrMxT_13,var=MxT_13,typeGEV=1,
             KappaI=0.25,GammaI=-47.05,ThetaI=2.77,YrLo=1950,YrHi=2017) ;
proc sort data=&file ;
   by year ;
run ;



proc model data=&file converge=0.001
          maxit=500 dw ;  outmodel=&ofile ;
     range year = &YrLo to &YrHi ;  **<< Dropped any months of 2018 data. >>** ;


     y = (&var - Gamma) / Theta ;

    %if &typeGEV=1 %then %do ;  ***<< 3-parameter GEV Model. >>>*** ;
        &depvar = exp( -(1 - Kappa * (y))**(1/Kappa) ) ;
        %let typmod = 3-parameter GEV Model. ;
        %end ;

    %if &typeGEV=2 %then %do ;  **<< 2-parameter "Double Exponential" or  "Gumbel" Model. >>** ;
        &depvar = exp( -exp(-(y)) ) ;
        %let typmod = 2-parameter Double Exponential or Gumbel Model. ;
        %end ;

    %if (&typeGEV NE 1) AND (&typeGEV NE 2) %then  %do ; **<< 2-parameter "T-Dist" Model. >>** ;
        dft=(&YrHi - &YrLo) +1 -2 ;
        &depvar = probt(y,dft) ;
        %let typmod = 2-parameter T-Dist Model. ;
```

```
      %end ;


  %if &typeGEV = 1 %then %do ;
  parms
      Kappa  &KappaI
      Gamma  &GammaI
      Theta  &ThetaI ;
  %end ;

  %if (&typeGEV NE 1)  %then %do ;
  parms
      Gamma  &GammaI
      Theta  &ThetaI ;
  %end ;


  fit &depvar   /out=&outfit  outall
              outest=&outest corrb corrs outcov ;

title3 "Non-linear Estimation of &&typmod: for Maximum NEGATIVE Temperature (Deg-F)." ;
run ;
%mend GEVfit ;






%macro GEVbyMo(mm=_1) ;


*********<<< Analysis for "January"    (i.e., SUFIX "mm" = "_1"  >>>******************;
*********<<< Analysis for "February"   (i.e., SUFIX "mm" = "_2"  >>>******************;
*********<<< Analysis for "March"      (i.e., SUFIX "mm" = "_3"  >>>******************;
*********<<< Analysis for "April"      (i.e., SUFIX "mm" = "_4"  >>>******************;
*********<<< Analysis for "May"        (i.e., SUFIX "mm" = "_5"  >>>******************;
*********<<< Analysis for "June"       (i.e., SUFIX "mm" = "_6"  >>>******************;
*********<<< Analysis for "July"       (i.e., SUFIX "mm" = "_7"  >>>******************;
*********<<< Analysis for "August"     (i.e., SUFIX "mm" = "_8"  >>>******************;
*********<<< Analysis for "September"  (i.e., SUFIX "mm" = "_9"  >>>******************;
*********<<< Analysis for "October"    (i.e., SUFIX "mm" = "_10"  >>>******************;
*********<<< Analysis for "November"   (i.e., SUFIX "mm" = "_11"  >>>******************;
*********<<< Analysis for "December"   (i.e., SUFIX "mm" = "_12"  >>>******************;
*********<<< Analysis for "ALL Months"   (i.e., SUFIX "mm" = "_13"  >>>******************;

%RankIt(file=MxTData,var=MxT&mm,rank=Rank&mm,prob=PrMxT&mm,Nobser=68,PltValue=0.375) ;





%GEVfit(file=MxTData,ofile=MxTNL2,outfit=fit2,outest=est2,depvar=PrMxT&mm,var=MxT&mm,typeGEV=2,
      KappaI=0.25,GammaI=&&gamma&mm,ThetaI=&&theta&mm,YrLo=1950,YrHi=2017) ;


proc print data=fit2 ;
run ;




proc transpose data=fit2 out=pred2 prefix=probP_ ;
   where (_type_ = "PREDICT" ) ;
   by year;
   var prmxt&mm ;
run ;
```

```
data comb2 ;
   merge MxTData pred2 ;
  by year ;
   ProbP2 = ProbP_1 ;
  keep year MxT&mm PrMxT&mm ProbP2 ;
run ;


proc print data=comb2 ;
run ;


proc plot data=comb2 ;
  plot prmxt&mm*MxT&mm='*'
       probP2*MxT&mm='-' / overlay ;
run ;


proc print data=est2 ;
run ;



data out2.est2&mm ;   ***<<< Save a copy of the "Double Exponential Model" estimation results!  >>>*** ;
  set est2 ;
run ;





%GEVfit(file=MxTData,ofile=MxTNL0,outfit=fit0,outest=est0,depvar=PrMxT&mm,var=MxT&mm,typeGEV=0,
        KappaI=0.25,GammaI=&&gamma&mm,ThetaI=&&theta&mm,YrLo=1950,YrHi=2017) ;


proc print data=fit0 ;
run ;




proc transpose data=fit0 out=pred0 prefix=probP_ ;
   where (_type_ = "PREDICT" ) ;
   by year;
   var prmxt&mm ;
run ;

data comb0 ;
   merge MxTData pred0 ;
  by year ;
   ProbP0 = ProbP_1 ;
  keep year MxT&mm PrMxT&mm ProbP0 ;
run ;


proc print data=comb0 ;
run ;


proc plot data=comb0 ;
  plot prmxt&mm*MxT&mm='*'
       probP0*MxT&mm='-' / overlay ;
run ;


proc print data=est0 ;
```

```
run ;


data out2.est0&mm ;   ***<<< Save a copy of the 2-parameter "T-Distribution" Model estimation results!  >>>*** ;
  set est0 ;
run ;




%GEVfit(file=MxTData,ofile=MxTNL1,outfit=fit1,outest=est1,depvar=PrMxT&mm,var=MxT&mm,typeGEV=1,
        KappaI=0.25,GammaI=&&gamma&mm,ThetaI=&&theta&mm,YrLo=1950,YrHi=2017) ;


proc print data=fit1 ;
run ;



proc transpose data=fit1 out=pred1 prefix=probP_ ;
   where (_type_ = "PREDICT" ) ;
   by year;
   var prmxt&mm ;
run ;

data comb1 ;
   merge MxTData pred1 ;
  by year ;
   ProbP1 = ProbP_1 ;
  keep year MxT&mm PrMxT&mm ProbP1 ;
run ;


proc print data=comb1 ;
run ;


proc plot data=comb1 ;
  plot prmxt&mm*MxT&mm='*'
       ProbP1*MxT&mm='-' / overlay ;
run ;



proc print data=est1 ;
run ;



data out2.est1&mm ;   ***<<< Save a copy of the "G.E.V. Model" estimation results!  >>>*** ;
  set est1 ;
run ;

%mend GEVbyMo ;








/****************
```

```
****************/

proc means data=MxTData ;
  var MxT_1 - MxT_12  MxT_13;
 output out=VarStat
        mean=mean1-mean12 mean13
        std=stdev1-stdev12 stdev13;
title3 "Calc. Means and Standard Deviantions to use as Starting Values in Non-Linear Estimations." ;
run ;


proc print data=VarStat ;
run ;

data _null_ ;
   set VarStat ;

  call symput('gamma_13',mean13) ;
  call symput('theta_13',stdev13) ;

  call symput('gamma_12',mean12) ;
  call symput('theta_12',stdev12) ;

  call symput('gamma_11',mean11) ;
  call symput('theta_11',stdev11) ;

  call symput('gamma_10',mean10) ;
  call symput('theta_10',stdev10) ;

  call symput('gamma_9',mean9) ;
  call symput('theta_9',stdev9) ;

  call symput('gamma_8',mean8) ;
  call symput('theta_8',stdev8) ;

  call symput('gamma_7',mean7) ;
  call symput('theta_7',stdev7) ;

  call symput('gamma_6',mean6) ;
  call symput('theta_6',stdev6) ;

  call symput('gamma_5',mean5) ;
  call symput('theta_5',stdev5) ;

  call symput('gamma_4',mean4) ;
  call symput('theta_4',stdev4) ;

  call symput('gamma_3',mean3) ;
  call symput('theta_3',stdev3) ;

  call symput('gamma_2',mean2) ;
  call symput('theta_2',stdev2) ;

  call symput('gamma_1',mean1) ;
  call symput('theta_1',stdev1) ;

run ;


%GEVbyMo(mm=_13) ;   **<< Annual(Entire Year of Data.) >>** ;





%GEVbyMo(mm=_1) ;   **<< Jan Data. >>** ;
```

```
%GEVbyMo(mm=_2) ;   **<< Feb Data. >>** ;
%GEVbyMo(mm=_3) ;   **<< Mar Data. >>** ;


%GEVbyMo(mm=_4) ;   **<< Apr Data. >>** ;


%GEVbyMo(mm=_5) ;   **<< May Data. >>** ;
%GEVbyMo(mm=_6) ;   **<< Jun Data. >>** ;
%GEVbyMo(mm=_7) ;   **<< Jul Data. >>** ;


%GEVbyMo(mm=_8) ;   **<< Aug Data. >>** ;


%GEVbyMo(mm=_9) ;   **<< Sep Data. >>** ;
%GEVbyMo(mm=_10) ;   **<< Oct Data. >>** ;
%GEVbyMo(mm=_11) ;   **<< Nov Data. >>** ;
%GEVbyMo(mm=_12) ;   **<< Dec Data. >>** ;





quit ;
```

San Diego Gas & Electric Residential End-Use Model

I. Residential End-Use Model Description

Introduction:

San Diego Gas & Electric (SDG&E) used the End Use Forecaster model to generate annual gas demand forecasts for the residential market. The software's market segmentation and end-use modeling framework analyzes the impacts of competitive strategies (gas vs. electricity) and market scenarios on gas demand and market shares. The model separates the residential market into five building types (B-level).

These groups are identified by the premise code classification found in the company billing files. The four residential groups are:

• Single-Family(SF);
• Multi-Family (MF);
• Master Metered (MM); and
• Sub-Metered (SM).

The residential model identifies eight end-uses (N-level) that are the primary drivers of natural gas demand:

• Space heating;
• Water heating;
• Cooking;
• Drying;
• Pool heating;
• Spa heating;
• Fireplace; and
• Barbeque.

The model assumes two fuel choices (F-level) for end-uses:

• Natural gas; and
• Electricity.

The model assumes up to four efficiency levels (E-level) for the various end-uses. In general, the efficiency levels are:

• Stock;
• Standard;
• High efficiency; and
• Premium efficiency.

See Figure 1 for a classification of the number of efficiency levels for each end-use by customer segment type.

A set of post-model adjustments were applied to the model's annual demand forecast. The first adjustment calibrates to the recorded 2017 weather-adjusted demand. Next, the annual forecast was parceled out to a series of monthly forecasts by a process which involves two steps. These two steps consist of (1) using the fitted equation for customer demand to generate a forecast of use per customer that varies with the number of calendar days and heating degree days in a given month and (2) calculating a series of weights based on the customer's predicted monthly usage share in total annual consumption. The shares obtained from the latter step were then applied to annual totals to derive the stream of monthly forecasts which are conditional on the particular weather design specification for the entire year. An adjustment to the forecast offsets the throughput by the energy efficiency savings. Annual conservation benefits associated with AMI have been estimated by SDG&E to represent 1% of the core gas throughput.

Figures 2-5 illustrate the monthly forecasts for each weather scenario.

Data Sources:

The information used to perform the modeling and to generate the forecast includes historical 2017 consumption and customer counts; meter counts, growth, and decay; use per customer by vintage and unit energy consumption (UEC) values; fuel costs and price elasticity; equipment capital costs and availability; building and equipment lives and decay. The historical 2017 data is in Figure 6.

Meter Counts, Growth and Decay:

Regression equations were developed for each of the 4 building types. The meter count forecast is a company-specific forecast based on actual meter counts within the SDG&E service territory. Data on meter decay rates were obtained from the Energy Information Administration (EIA). See Figure 7 for the meter forecast.

Use Per Customer by Vintage and UEC:

Use per customer and Unit Energy Consumption (UEC) data were based on company marketing data and the California Measurement Advisory Council. See Figure 8 for the appliance UEC's.

Fuel Costs and Price Elasticity:

Average and marginal gas prices ($/therm) were calculated from forecasts of the residential rate components. Residential rates have two consumption tiers. We used the simple average of the second tiers' projected monthly prices for each forecast year as the marginal rate. The marginal rate was used for each housing segment type.

For a given housing segment type, the average gas commodity rate was calculated using a pair of weights for the two consumption tiers applied to the simple average of each tier's monthly rate. The average commodity rate in each forecast year was developed using the same consumption tier weights, but with the forecasts of rates for each residential rate tier. The average gas price each year was then calculated by including the non-volumetric customer charges with the year's average gas commodity price. Figure 9 illustrates the gas price forecasts.

Electric Price Data:

Both average prices (cents/kWh) and marginal prices (cents/kWh) were developed as electricity price inputs. Forecasts for the SDG&E residential customer class were developed based on the California Energy Commission's December 2017 updated forecast rates for California energy demand (forecast for the SDG&E planning area, under "Mid-Case" demand for electricity) for the SDG&E service area through our forecast time horizon.

To impute average electricity prices to each residential housing type, we simply calculated the ratio of the housing type's average gas price to the overall residential gas price for each housing type, then multiplied by the overall average electricity price.

The marginal prices for each residential housing type were calculated by multiplying each year's respective average price by a ratio. These ratios were 1.513 for the SF and MF housing types, 1.034 for the MM housing type and 1.125 for the SM housing type. These various ratios were the same as those used to construct the marginal electricity prices for the SoCalGas residential end-use model.

Equipment Capital Costs and Availability:

Data on equipment capital costs and availability were from EIA, the Residential Appliance Saturation Survey (RASS), Energy Star (EPA & DOE), and SDG&E company data. See Figures 11 and 12 for gas and electric appliance equipment
cost.

Building and Equipment Lives and Decay:

Building decay rates are based on the building shell lifetimes, where the lifetime is defined as the length of time it takes for either a demolition or a major renovation to occur. For single-family residential buildings, an exponential rate of decay of 0.3% per year was assumed. See Figure 13 for the building decay rates.
Data on equipment lives and decay rates are based on EIA, RASS, Energy Star, and SDG&E company data. See Figure 14 for the average lifetimes of gas appliances.

Saturations, Fuel and Efficiency Shares:

Saturation values, fuel shares, and efficiency shares were extracted from SDG&E company data files and RASS survey results. Please see Figures 15-18 for saturations, fuel, and efficiency shares.

AMI:
Mass deployment of AMI gas modules began in 2009. The conservation benefits estimated by SDG&E represent approximately 1% of core gas throughput.

II. Residential End-Use Model Data

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 1: Number of Efficiency Levels by End Use by Customer Segment

	Space Heating		Water Heating		Cooking		Drying		Pool		Spa		Fireplace		BBQ	
	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric	Gas	Electric
Single Family	4	4	4	4	2	2	2	2	1	1	1	1	1	1	1	1
Multi-Family	4	4	4	4	2	2	2	2	0	0	0	0	0	0	1	1
Master Meter	4	4	4	4	2	2	2	2	0	0	0	0	0	0	1	1
Sub-Meter	4	4	4	4	2	2	2	2	0	0	0	0	0	0	1	1

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 2: Average Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	4,373	3,873	3,510	2,879	2,131	1,619	1,536	1,528	1,490	1,793	2,834	4,488	32,052
2018	4,390	3,888	3,523	2,889	2,139	1,625	1,541	1,533	1,496	1,800	2,844	4,505	32,173
2019	4,357	3,858	3,496	2,867	2,122	1,612	1,530	1,522	1,484	1,787	2,823	4,470	31,929
2020	4,328	3,833	3,473	2,849	2,109	1,602	1,520	1,512	1,475	1,775	2,804	4,441	31,721
2021	4,284	3,794	3,438	2,819	2,087	1,585	1,504	1,496	1,460	1,757	2,775	4,396	31,394
2022	4,210	3,728	3,379	2,771	2,051	1,558	1,478	1,471	1,435	1,727	2,728	4,320	30,856

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 3: Cold Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	4,989	4,401	3,938	3,181	2,261	1,649	1,538	1,528	1,493	1,851	3,127	5,128	35,085
2018	5,009	4,420	3,955	3,195	2,271	1,656	1,544	1,534	1,499	1,859	3,140	5,149	35,230
2019	4,973	4,387	3,926	3,171	2,254	1,644	1,533	1,523	1,488	1,845	3,117	5,111	34,973
2020	4,942	4,361	3,902	3,152	2,240	1,634	1,523	1,514	1,479	1,834	3,098	5,080	34,759
2021	4,894	4,318	3,863	3,121	2,218	1,618	1,508	1,499	1,465	1,816	3,067	5,030	34,418
2022	4,812	4,246	3,799	3,069	2,181	1,591	1,483	1,474	1,440	1,786	3,016	4,946	33,845

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 4: Hot Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	3,758	3,345	3,081	2,576	2,000	1,588	1,533	1,527	1,487	1,736	2,541	3,848	29,019
2018	3,771	3,356	3,091	2,584	2,007	1,593	1,539	1,532	1,492	1,741	2,549	3,861	29,116
2019	3,741	3,329	3,067	2,564	1,991	1,581	1,526	1,520	1,480	1,728	2,529	3,830	28,884
2020	3,714	3,306	3,045	2,546	1,977	1,569	1,516	1,509	1,470	1,715	2,511	3,803	28,682
2021	3,674	3,270	3,012	2,518	1,955	1,552	1,499	1,493	1,454	1,697	2,484	3,762	28,370
2022	3,609	3,212	2,959	2,473	1,920	1,525	1,473	1,467	1,428	1,667	2,440	3,695	27,867

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 5: Base Temperature Year Demand Forecast (MDth)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2017	1,525	1,427	1,525	1,476	1,525	1,476	1,525	1,525	1,476	1,525	1,476	1,525	18,005
2018	1,526	1,427	1,526	1,476	1,526	1,476	1,526	1,526	1,476	1,526	1,476	1,526	18,013
2019	1,510	1,412	1,510	1,461	1,510	1,461	1,510	1,510	1,461	1,510	1,461	1,510	17,826
2020	1,494	1,398	1,494	1,446	1,494	1,446	1,494	1,494	1,446	1,494	1,446	1,494	17,645
2021	1,473	1,378	1,473	1,425	1,473	1,425	1,473	1,473	1,425	1,473	1,425	1,473	17,386
2022	1,441	1,348	1,441	1,394	1,441	1,394	1,441	1,441	1,394	1,441	1,394	1,441	17,011

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 6: 2017 Historical Data

Segment	2017 Therm Sales	2017 Meter Count	2017 Meter Count: Pre-1979 Customers	2017 Meter Count: 1979-2016 Customers	2017 Meter Count: 2017 "New" Customers	Avg Annual Consumption: Pre-1979 Customers	Avg Annual Consumption: 1979-2016 Customers	Avg Annual Consumption: 2017 "New" Customers	Price Elasticity
Single Family	228,117,239	652,014	550,259	98,853	2,902	338	417	362	-0.1053
Multi-Family	51,152,270	186,136	139,980	43,129	3,027	262	318	267	-0.07145
Master Meter	32,019,986	11,521	11,121	392	8	2,634	6,820	7,004	-0.0688
Sub-Meter	9,231,502	464	459	5	0	19,975	12,679	0	-0.1053

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 7: Meter Count Forecast

Year	Single Family	Multi-Family	Master Meter	Sub-Meter
2017	652,014	186,136	11,521	464
2018	656,655	187,461	11,521	464
2019	661,104	188,731	11,521	464
2020	665,746	190,056	11,521	464
2021	670,798	191,499	11,521	464
2022	676,004	192,985	11,521	464

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 8: Appliance Unit Energy Consumption (Gas in Therms, Electric in Kwh)

		Single Family		Multi-Family		Master Meter		Sub-Meter	
End-Use	**Vintage**	**Gas**	**Electric**	**Gas**	**Electric**	**Gas**	**Electric**	**Gas**	**Electric**
Space Heating	Stock	250	4,110	140	730	110	730	230	1,340
	Standard	230	3,730	130	670	100	670	210	1,210
	High	220	3,450	120	620	100	620	200	1,120
	Premium	200	3,170	110	570	90	570	180	1,030
Water Heating	Stock	150	2,440	150	2,440	120	2,440	150	2,010
	Standard	140	2,220	140	2,220	110	2,220	140	1,830
	High	130	2,110	130	2,110	100	2,110	130	1,740
	Premium	130	2,050	130	2,050	100	2,050	130	1,690
Cooking	Stock	28	574	26	465	26	465	27	514
	Standard	24	488	22	395	22	395	23	437
Drying	Stock	39	1,442	33	1,442	33	1,442	35	873
	Standard	37	1,370	31	1,370	31	1,370	33	830
Pool	Stock	110	3,431	-	-	-	-	-	-
Spa	Stock	100	290	-	-	-	-	-	-
Fireplace	Stock	17	0	-	-	-	-	-	-
BBQ	Stock	16	0	13	0	14	0	16	0

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 9: Average and Marginal Gas Prices ($/therm)

Year	Res Price Deflator	R SF Average Price	R SF Marginal Price	R MF Average Price	R MF Marginal Price	R MM Average Price	R MM Marginal Price	R SM Average Price	R SM Marginal Price
2017	97.7	1.3956	1.4185	1.3739	1.4185	1.3758	1.4185	1.3960	1.4185
2018	100.0	1.2994	1.3209	1.2789	1.3209	1.2807	1.3209	1.2997	1.3209
2019	102.4	1.3755	1.3973	1.3547	1.3973	1.3565	1.3973	1.3758	1.3973
2020	105.1	1.4267	1.4494	1.4052	1.4494	1.4071	1.4494	1.4271	1.4494
2021	107.8	1.5321	1.5564	1.5090	1.5564	1.5110	1.5564	1.5325	1.5564
2022	110.5	1.7497	1.7774	1.7234	1.7774	1.7256	1.7774	1.7502	1.7774

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 10: Average and Marginal Electricity Prices (Cents/KWh)

Year	R SF Average Price	R SF Marginal Price	R MF Average Price	R MF Marginal Price	R MM Average Price	R MM Marginal Price	R SM Average Price	R SM Marginal Price
2017	25.13	38.02	24.74	37.43	24.77	25.61	25.14	28.29
2018	25.75	38.97	25.35	38.36	25.38	26.25	25.76	28.99
2019	26.71	40.42	26.31	39.81	26.34	27.24	26.72	30.07
2020	27.67	41.87	27.25	41.24	27.29	28.22	27.68	31.15
2021	29.11	44.05	28.67	43.39	28.71	29.69	29.12	32.77
2022	29.20	44.19	28.76	43.53	28.80	29.78	29.21	32.87

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 11: Gas Appliance Equipment Cost (Nominal $)

End-Use	Customer Class	Stock Efficiency	Standard Efficiency	High Efficiency	Premium Efficiency
Space Heating	Single Family	4,000	4,600	4,800	5,000
	Multi-Family	1,600	1,840	1,920	1,980
	Master Meter	1,000	1,150	1,200	1,250
	Sub-metered	1,600	1,840	1,920	1,980
Water Heating	Single Family	550	650	700	750
	Multi-Family	330	390	420	450
	Master Meter	330	390	420	450
	Sub-metered	330	390	420	450
Cooking	Single Family	500	1,400	-	-
	Multi-Family	250	1,400	-	-
	Master Meter	250	1,400	-	-
	Sub-metered	250	1,400	-	-
Drying	Single Family	328	482	-	-
	Multi-Family	328	482	-	-
	Master Meter	328	482	-	-
	Sub-metered	328	482	-	-
Pool	Single Family	1,200	-	-	-
Spa	Single Family	2,000	-	-	-
Fireplace	Single Family	150	-	-	-
Barbecue	Single Family	1,000	-	-	-
	Multi-Family	600	-	-	-
	Master Meter	600	-	-	-
	Sub-metered	600	-	-	-

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 12: Electric Appliance Equipment Cost (Nominal $)

End-Use	Customer Class	Stock Efficiency	Standard Efficiency	High Efficiency	Premium Efficiency
Space Heating	Single Family	4,100	-	-	-
	Multi-Family	1,640	-	-	-
	Master Meter	1,025	-	-	-
	Sub-metered	1,640	-	-	-
Water Heating	Single Family	550	650	700	750
	Multi-Family	330	390	420	450
	Master Meter	330	390	420	450
	Sub-metered	330	390	420	450
Cooking	Single Family	500	1,400	-	-
	Multi-Family	250	1,400	-	-
	Master Meter	250	1,400	-	-
	Sub-metered	250	1,400	-	-
Drying	Single Family	328	482	-	-
	Multi-Family	328	482	-	-
	Master Meter	328	482	-	-
	Sub-metered	328	482	-	-
Pool	Single Family	1,200	-	-	-
Spa	Single Family	2,000	-	-	-
Fireplace	Single Family	150	-	-	-
Barbecue	Single Family	1,000	-	-	-
	Multi-Family	600	-	-	-
	Master Meter	600	-	-	-
	Sub-metered	600	-	-	-

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 13: Building Lives and Decay Rate

Building Type	Building Decay Rate
Single Family	0.003
Multi-Family	0.006
Master Meter	0.008
Sub-metered	0.008

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 14: Gas Appliance Age (Years)

		Single Family		Multi-Family		Master Meter		Sub-metered	
End-Use	**Vintage**	**Average**	**Max**	**Average**	**Max**	**Average**	**Max**	**Average**	**Max**
Space Heating	Pre-1979	16	16	23	23	20	20	16	16
	1979-2004	15	16	16	23	15	20	15	16
	2005-Current	5	16	4	23	4	20	5	16
Water Heating	Pre-1979	9	9	12	12	10	10	9	9
	1979-2004	9	9	10	12	10	10	9	9
	2005-Current	5	9	4	12	4	10	5	9
Cooking	Pre-1979	9	9	9	9	9	9	9	9
	1979-2004	9	9	9	9	8	9	9	9
	2005-Current	4	9	4	9	4	9	4	9
Drying	Pre-1979	7	7	6	7	7	7	7	7
	1979-2004	6	7	7	7	7	7	6	7
	2005-Current	4	7	3	7	3	7	4	7
Pool	Pre-1979	13	13	-	-	-	-	-	-
	1979-2004	9	13	-	-	-	-	-	-
	2005-Current	3	13	-	-	-	-	-	-
Spa	Pre-1979	11	11	-	-	-	-	-	-
	1979-2004	8	11	-	-	-	-	-	-
	2005-Current	3	11	-	-	-	-	-	-
Fireplace	Pre-1979	15	15	-	-	-	-	-	-
	1979-2004	15	15	-	-	-	-	-	-
	2005-Current	15	15	-	-	-	-	-	-
Barbecue	Pre-1979	7	7	5	7	5	6	7	7
	1979-2004	6	7	7	7	6	6	6	7
	2005-Current	4	7	3	7	4	6	4	7
Other	Pre-1979	15	15	15	15	15	15	15	15
	1979-2004	15	15	15	15	15	15	15	15
	2005-Current	15	15	15	15	15	15	15	15

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 15: End-Use Saturations

End-Use	Vintage	Single Family	Multi-Family	Master Meter	Sub-metered
Space Heating	Pre-1979	1.00000	1.00000	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	0.00000
Water Heating	Pre-1979	1.00000	1.00000	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	0.00000
Cooking	Pre-1979	1.00000	0.99633	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	0.00000
Drying	Pre-1979	0.85795	0.20040	0.47158	0.47158
	1979-2004	0.89516	0.42764	0.57182	0.57182
	2005-Current	0.92508	0.74161	0.74768	0.00000
Pool	Pre-1979	0.15644	-	-	-
	1979-2004	0.17913	-	-	-
	2005-Current	0.16916	-	-	-
Spa	Pre-1979	0.12651	-	-	-
	1979-2004	0.21695	-	-	-
	2005-Current	0.19134	-	-	-
Fireplace	Pre-1979	0.22973	-	-	-
	1979-2004	0.27252	-	-	-
	2005-Current	0.26269	-	-	-
Barbecue	Pre-1979	0.13716	0.04723	0.07424	0.07424
	1979-2004	0.25180	0.06165	0.10179	0.10179
	2005-Current	0.31442	0.07818	0.16198	0.00000
Other	Pre-1979	1.00000	1.00000	1.00000	1.00000
	1979-2004	1.00000	1.00000	1.00000	1.00000
	2005-Current	1.00000	1.00000	1.00000	NA

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 16: Gas Fuel Shares (average)

End-Use	Single Family	Multi-Family	Master Meter	Sub-metered
Space Heating	0.98200	0.91179	0.92461	0.92461
Water Heating	0.97630	0.89871	0.92997	0.92997
Cooking	0.83890	0.82622	0.81058	0.81058
Drying	0.80258	0.59654	0.70306	0.70306
Pool	0.49003	-	-	-
Spa	0.60804	-	-	-
Fireplace	0.56361	-	-	-
Barbecue	0.95008	0.85803	0.89234	0.89234
Other	1.00000	1.00000	1.00000	1.00000

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 17: Gas Efficiency Shares

End-Use	Customer Class	Stock Existing	Stock New	Standard Existing	Standard New	High Existing	High New	Premium Existing	Premium New
Space Heating	Single Family	0.06	0.06	0.78	0.78	0.14	0.14	0.02	0.02
	Multi-Family	0.41	0.41	0.46	0.46	0.01	0.01	0.04	0.04
	Master Meter	0.17	0.17	0.69	0.69	0.11	0.11	0.03	0.03
	Sub-metered	0.06	0.06	0.78	0.78	0.14	0.14	0.02	0.02
Water Heating	Single Family	0.00	0.00	0.64	0.64	0.34	0.34	0.02	0.02
	Multi-Family	0.00	0.00	0.61	0.61	0.37	0.37	0.02	0.02
	Master Meter	0.00	0.00	0.59	0.59	0.39	0.39	0.02	0.02
	Sub-metered	0.00	0.00	0.64	0.64	0.34	0.34	0.02	0.02
Cooking	Single Family	0.17	0.17	0.83	0.83	-	-	-	-
	Multi-Family	0.18	0.18	0.82	0.82	-	-	-	-
	Master Meter	0.17	0.17	0.83	0.83	-	-	-	-
	Sub-metered	0.17	0.17	0.83	0.83	-	-	-	-
Drying	Single Family	0.07	0.07	0.93	0.93	-	-	-	-
	Multi-Family	0.06	0.06	0.94	0.94	-	-	-	-
	Master Meter	0.06	0.06	0.94	0.94	-	-	-	-
	Sub-metered	0.07	0.07	0.93	0.93	-	-	-	-
Pool	Single Family	1.00	1.00	-	-	-	-	-	-
Spa	Single Family	1.00	1.00	-	-	-	-	-	-
Fireplace	Single Family	1.00	1.00	-	-	-	-	-	-
Barbecue	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-
Other	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-

San Diego Gas & Electric
2020 Triennial Cost Allocation Proceeding
Figure 18: Electric Efficiency Shares

End-Use	Customer Class	Stock Existing	Stock New	Standard Existing	Standard New	High Existing	High New	Premium Existing	Premium New
Space Heating	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-
Water Heating	Single Family	0.10	0.10	0.68	0.68	0.21	0.21	0.01	0.01
	Multi-Family	0.13	0.13	0.76	0.76	0.10	0.10	0.01	0.01
	Master Meter	0.13	0.13	0.76	0.76	0.10	0.10	0.01	0.01
	Sub-metered	0.10	0.10	0.68	0.68	0.21	0.21	0.01	0.01
Cooking	Single Family	0.90	0.90	0.10	0.10	-	-	-	-
	Multi-Family	0.95	0.95	0.05	0.05	-	-	-	-
	Master Meter	0.95	0.95	0.05	0.05	-	-	-	-
	Sub-metered	0.95	0.95	0.05	0.05	-	-	-	-
Drying	Single Family	0.75	0.75	0.25	0.25	-	-	-	-
	Multi-Family	0.75	0.75	0.25	0.25	-	-	-	-
	Master Meter	0.75	0.75	0.25	0.25	-	-	-	-
	Sub-metered	0.75	0.75	0.25	0.25	-	-	-	-
Pool	Single Family	1.00	1.00	-	-	-	-	-	-
Spa	Single Family	1.00	1.00	-	-	-	-	-	-
Fireplace	Single Family	1.00	1.00	-	-	-	-	-	-
Barbecue	Single Family	1.00	1.00	-	-	-	-	-	-
	Multi-Family	1.00	1.00	-	-	-	-	-	-
	Master Meter	1.00	1.00	-	-	-	-	-	-
	Sub-metered	1.00	1.00	-	-	-	-	-	-

Weather for SDG&E: Heating Degree Days – Average and Cold Year Designs; and Winter Peak Day Design Temperatures

I. Overview

San Diego Gas and Electric Company's service area for natural gas extends from southern Orange County throughout San Diego County to the Mexican border. To quantify the overall temperature experienced within this region, SDG&E aggregates daily temperature recordings from three U.S. Weather Bureau weather stations into one system average heating degree-day ("HDD") figure. The table below lists weather station locations along with its associated temperature zone(s).

Table 1

Representative Weather Stations with Temperature Zones

Station Location	**Weight**	**Temperature Zone**
1. El Cajon[1]	1/3	Coastal and Inland
2. San Diego's Lindberg Field	(2/3) x (#Coastal/(#Coastal + #Inland))	Coastal
3. Miramar Naval Air Station	(2/3) x (#Inland/(#Coastal + #Inland))	Inland

SDG&E uses 65° Fahrenheit to calculate the number of HDDs. One heating degree-day is accumulated for each degree that the daily average is *below* 65° Fahrenheit. To arrive at the system average HDDs figure for its entire service area, SDG&E weights the HDD figure for each zone using the weights[2] shown in Table 1. These weights are used in calculating the data shown from January 1998 to December 2017.

Daily maximum and minimum temperatures, for each individual weather station in the table above, are compiled from National Weather Service data. The web-site:

http://www.wrh.noaa.gov/sgx/obs/rtp/rtpmap.php?wfo=sgx

---

[1] The location of the station for El Cajon is at the boundary of the Coastal and Inland zones. Correspondingly, both the Coastal and Inland zones are considered represented in the data for the El Cajon station.

[2] As of December 2017, there were 498,013 gas customers associated with the Coastal temperature zone and 420,085 gas customers associated with the Inland temperature zone. The following URL shows a map of the SDG&E service area and temperature zones: http://www.sdge.com/tm2/pdf/ELEC_MAPS_Maps_-_Elec.pdf ; less than 0.04% of SDG&E's gas customers were in the mountain and desert zones.

provides easy access to temperature data for San Diego and parts of surrounding counties. For each station, the average temperature is computed as the (maximum + minimum)/2 and this value is used to compute the heating degrees (i.e., the *daily* HDD) for each station as well. System average values of HDD are then computed using the weights for each respective station. Annual and monthly HDDs for the entire SDG&E service area from 1998 to 2017 are listed in Table 2, below.

Table 2

Calendar Month Heating Degree-Days (Jan. 1998 through Dec. 2017)

	Month												Total
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	"Cal-Year"
1998	252	256	205	195	94	22	1	0	5	31	172	338	**1571**
1999	276	266	279	223	115	51	4	0	4	4	146	243	**1610**
2000	247	216	224	95	28	3	0	0	0	50	237	227	**1327**
2001	351	298	199	198	30	5	0	0	0	9	127	325	**1543**
2002	315	225	247	158	91	13	0	0	2	54	81	294	**1479**
2003	141	201	179	184	95	32	0	0	0	7	157	275	**1270**
2004	273	269	98	65	14	4	1	0	0	52	200	265	**1240**
2005	243	197	159	118	33	5	0	0	4	38	95	231	**1121**
2006	275	204	305	144	33	0	0	0	1	35	88	287	**1372**
2007	365	225	155	139	64	20	0	0	4	28	112	340	**1451**
2008	331	278	187	131	89	16	0	0	0	13	59	287	**1391**
2009	177	247	201	141	30	11	0	0	0	40	124	291	**1262**
2010	240	212	195	178	88	24	10	1	2	31	181	238	**1401**
2011	220	277	196	96	75	20	0	0	0	25	172	340	**1421**
2012	232	239	230	129	37	13	0	0	0	16	102	268	**1266**
2013	323	269	150	104	23	6	0	0	0	40	104	241	**1261**
2014	158	140	80	78	20	1	0	0	0	0	44	170	**691**
2015	161	87	58	44	46	0	0	0	0	0	104	259	**759**
2016	237	82	95	41	27	0	0	0	0	0	67	196	**746**
2017	242	156	82	27	42	4	0	0	0	1	38	149	**743**
20-Yr-Avg (Jan1998-Dec2017)													
Avg.	252.9	217.2	176.2	124.5	53.7	12.5	0.8	0.1	1.1	23.7	120.5	263.1	**1246.3**
St.Dev.	63.4	60.9	67.4	55.8	31.7	13.0	2.4	0.2	1.6	18.7	53.5	53.5	**288.9**
Min.	141.3	81.9	57.9	27.5	13.9	0.0	0.0	0.0	0.0	0.0	38.3	148.7	**691.0**
Max.	364.6	298.2	304.8	222.7	115.5	51.1	10.3	1.1	4.7	53.5	237.0	339.9	**1610.1**

II. Calculations to Define Our Average-Temperature Year

The simple average of the 20-year period (January 1998 through December 2017) was used to represent the Average Year total and the individual monthly values for HDD. In this TCAP, the standard deviation has been calculated using an approach that compensates for the annual HDD values for the years 2014-2017 in SDG&E's service territory being dramatically lower than in any preceding year going back to 1972. A regression with a time trend and a dummy variable for the years 2014-2017 has been used to estimate a shift in the level of annual HDD that occurred beginning in 2014. A dummy variable takes the value one for some observations to indicate the presence of an effect or membership in a group and zero for the remaining observations. Estimating the effect of the dummy variable gives an estimate of that effect or the impact of membership in that group. A dummy variable is used here to estimate the average effect on annual HDD of a given year having membership in the group of years 2014-2017. The dataset is SDG&E system-wide annual HDD for the years 1998-2017. The regression equation is:

$$HDD_t = \alpha + \beta * t + \beta_{2014-201} * D_{2014-2017} + \varepsilon$$

where $D_{2014-201}$ is a dummy variable for the years 2014-2017 and $\beta_{2014-2017}$ is the corresponding dummy coefficient. This regression equation estimates average HDD over the period 1998-2017 controlling for time trends in HDD and the warm weather regime of years 2014-2017. It's important to note that p-value for the estimate of $\beta_{2014-2017}$ is 0.0022% indicating an extremely low probability that membership in the group of years 2014-2017 had no effect on annual HDDs. Please see Table 3 below for the full regression output.

Table 3

Dummy Regression for Calculation of Heating Degree-Day Standard Deviation

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.929575794
R Square	0.864111157
Adjusted R Squa	0.848124234
Standard Error	112.6014212
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1370637.082	685318.5409	54.05112499	4.2861E-08
Residual	17	215544.361	12679.08006		
Total	19	1586181.443			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	28203.42414	12157.88847	2.319763355	0.033055018
Regime Dummy	-505.775872	87.39169403	-5.787459296	2.18633E-05
Trend	-13.37781348	6.062256732	-2.206738196	0.041372305

The dummy variable's estimated effect, $\beta_{2014-201}$, is subtracted from the actual annual HDD data for years 2014-2017 to adjust the data to remove the level shift. The standard deviation has been calculated using this adjusted dataset. This standard deviation has been used to design the two Cold Years based on a "1-in-10" and "1-in-35" chance, ***c***, that the respective annual "Cold Year" $\boldsymbol{hdd_c}$ value would be exceeded. A probability model for the annual HDD is based on a t-Distribution with N-1 degrees of freedom, where, N is the number of years of HDD data we use, μ is the average of the last 20 years of HDD, and S_{20} is the average of the standard deviations of the 20 most recent 20 year periods:

$U = (HDD_y - \mu)/S_{20}$, has a t-Distribution with N-1 degrees of freedom.

III. Calculating the Cold-Temperature Year Weather Designs

<u>Cold Year HDD Weather Designs</u>

For SDG&E, cold-temperature-year HDD weather designs are developed with a 1-in-35 year chance of occurrence. In terms of probabilities this can be expressed as the following for a "1-in-35" cold-year HDD value in equation 1 and a "1-in-10" cold-year HDD value in equation 2, with Annual HDD as the random variable:

(1) Prob { Annual HDD > "1-in-35" Cold-Yr HDD } = 1/35 = 0.0286

(2) Prob { Annual HDD > "1-in-10" Cold-Yr HDD } = 1/10 = 0.1000

An area of 0.0286 under one tail of the T-Distribution translates to 2.025 standard deviations *above* an average-year based on a t-statistic with 19 degrees of freedom. Using the standard deviation calculated as described earlier, 132.7 HDD, these equations yield values of about 1,515 HDD for a "1-in-35" cold year and 1,422 as the number of HDDs for a "1-in-10" cold year (an area of 0.1000 under one tail of the T-Distribution translates to 1.328 standard deviations *above* an average-year based on a t-statistic with 19 degrees of freedom). For example, the "1-in-35" cold-year HDD is calculated as follows:

(3) Cold-year HDD = 1,515 which equals approximately

1,246 average-year HDDs + 2.025 * 132.7

Table 4 shows monthly HDD figures for "1-in-35" cold year, "1-in-10" cold year and, average year temperature designs. The monthly average-temperature-year HDDs are calculated from weighted monthly HDDs from 1998 to 2017, as shown as the bottom of Table 2, above. For example, the average-year December value of 263.1 HDD equals the simple average of the 20 December HDD figures from 1998 to 2017 SDG&E calculates the cold--temperature-year monthly HDD values using the same shape of the average-year HDDs. For example, since 21.1 percent of average-temperature-year HDDs occurred in December, the estimated number of HDDs during December for a cold-year is equal to 1,515 HDDs multiplied by 21.1 percent, or 319.8 HDDs.

Table 4

Calendar Month Heating Degree-Day Designs

	Cold		Average	Hot	
	1-in-35 Design	1-in-10 Design		1-in-10 Design	1-in-35 Design
January	307.5	288.6	252.9	217.2	198.3
February	264.1	247.9	217.2	186.5	170.3
March	214.2	201.0	176.2	151.3	138.1
April	151.3	142.0	124.5	106.9	97.6
May	65.2	61.2	53.6	46.1	42.1
June	15.2	14.3	12.5	10.8	9.8
July	0.9	0.9	0.8	0.7	0.6
August	0.1	0.1	0.1	0.1	0.1
September	1.4	1.3	1.1	1.0	0.9
October	28.8	27.0	23.7	20.3	18.6
November	146.5	137.5	120.5	103.5	94.5
December	319.8	300.2	263.0	225.9	206.2
	1515	1422	1246	1070	977

IV. Calculating the Peak-Day Design Temperature

SDG&E's Peak-Day design temperature of 42.8 degrees Fahrenheit, denoted "Deg-F," is determined from a statistical analysis of observed annual minimum daily system average temperatures constructed from daily temperature recordings from the three U.S. Weather Bureau weather stations discussed above. Since we have a time series of daily data by year, the following notation will be used for the remainder of this discussion:

(1) $AVG_{y,d}$ = system average value of Temperature for calendar year "y" and day "d".

The calendar year, y, can range from 1972 through 2017, while the day, d, can range from 1 to 365, for non leap years, or from 1 to 366 for leap years. The "upper" value for the day, d, thus depends on the calendar year, y, and will be denoted by n(y)=365, or 366, respectively, when y is a non-leap year or a leap year.

For each calendar year, we calculate the following statistic from our series of daily system average temperatures defined in equation (1) above:

(2) $$MinAVG_y = \min_{d=1}^{n(y)} \{ AVG_{y,d} \}, \text{ for } y=1972, 1973, \ldots, 2017.$$

(The notation used in equation 2 means "For a particular year, y, list all the daily values of system average temperature for that year, then pick the smallest one.")

The resulting minimum annual temperatures are shown in Table 5, below. Note that most of the minimum temperatures occur in the months of December or January; however, for some calendar years the minimums occurred in other months (the observed minimum for 1991 was in March, and for 2004 it was in November).

The statistical methods we use to analyze this data employ software developed to fit three generic probability models: the Generalized Extreme Value (GEV) model, the Double-Exponential or GUMBEL (EV1) model and a 2-Parameter Students' T-Distribution (T-Dist) model. [The GEV and EV1 models have the same mathematical specification as those implemented in a DOS-based executable-only computer code that was developed by Richard L. Lehman and described in a paper published in the Proceedings of the Eighth Conference on Applied Climatology, January 17-22, 1993, Anaheim, California, pp. 270-273, by the American Meteorological Society, Boston, MA., with the title "Two Software Products for Extreme Value Analysis: System Overviews of ANYEX and DDEX." At the time he wrote the paper, Dr. Lehman was with the Climate Analysis Center, National Weather Service/NOAA in Washington, D.C., zip code 20233.] The Statistical Analysis Software (SAS) procedure for nonlinear statistical model estimation (PROC MODEL, from SAS V6.12) was used to do the calculations. Further, the calculation procedures were implemented to fit the probability models to observed *maximums* of data, like heating degrees. By recognizing that:

$$-\text{MinAVG}_y = -\min_{d=1}^{n(y)}\{\text{AVG}_{y,d}\} = \max_{d=1}^{n(y)}\{-\text{AVG}_{y,d}\}, \text{ for } y=1972, \ldots, 2017;$$

this same software, when applied to the *negative* of the minimum temperature data, yields appropriate probability model estimation results.

The calculations done to fit any one of the three probability models chooses the parameter values that provide the "best fit" of the parametric probability model's calculated cumulative distribution function (CDF) to the empirical cumulative distribution function (ECDF). Note that the ECDF is constructed based on the variable "-MinAVG_y" (which is a *maximum* over a set of *negative* temperatures) with values of the variable MinAVG_y that are the same as shown in Table 5, below.

In Table 5, the data for -MinAVG_y are shown after they have been sorted from "lowest" to "highest" value. The ascending *ordinal* value is shown in the column labeled "RANK" and the empirical cumulative distribution function is calculated and shown in the next column. The formula used to calculate this function is:

$$\text{ECDF} = (\text{RANK} - \alpha)/[\text{MaxRANK} + (1 - 2\,\alpha)],$$

where the parameter "α" (shown as *alpha* in Table 6) is a "small" positive value (usually less than ½) that is used to bound the ECDF away from 0 and 1.

Of the three probability models considered (GEV, EV1, and T_Dist) the results obtained for the T_Dist model were selected since the fit to the ECDF was better than that of either the GEV model or the EV1 model. (Although convergence to stable parameter estimates is occasionally a problem with fitting a GEV model to the ECDF, the T_Dist model had no problems with convergence of the iterative procedure to estimate parameters.)

The T_Dist model used here is a three-parameter probability model where the varable $z = (-\text{MinAVG}_y - \gamma) / \theta$, for each year, y, is presumed to follow a T_Dist with location parameter, γ, and scale parameter, θ, and a third parameter, v, that represents the number of degrees of freedom. For a given number of years of data, N, then v=N-2.

The following mathematical expression specifies the T_Dist model we fit to the data for "-MinAVG_y " shown in Table 5, below.

(3) $\text{ECDF}(-\text{MinAVG}_y) = \text{Prob}\{-T < -\text{MinAVG}_y\} = \text{T\_Dist}\{z;\, \gamma,\, \theta,\, v=N-2\}$,

where "T_Dist{ . }" is the cumulative probability distribution function for Student's T-Distribution[3], and

[3] A common mathematical expression for Student's T-Distribution is provided at http://en.wikipedia.org/wiki/Student%27s_t-distribution; with a probability density function

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$

(4) $z = (-MinAVG_y - \gamma) / \theta$, for each year, y, and

the parameters "γ" and "θ" are estimated for this model for given degrees of freedom v=N-2. The estimated values for γ and θ are shown in Table 6 along with the fitted values of the model CDF (the column: "Fitted" Model CDF).

Now, to calculate a *peak-day design temperature*, $TPDD_\delta$, with a specified likelihood, δ, that a value less than $TPDD_\delta$ would be observed, we use the equation below:

(5) $\delta = Prob\ \{ T \leq TPDD_\delta \}$, which is equivalent to

(6) $\delta = Prob\ \{ [(-T - \gamma) / \theta] \geq [(-TPDD_\delta - \gamma) / \theta] \}, = Prob\ \{ [(-T - \gamma) / \theta] \geq [z_\delta] \}$,

where $z_\delta = [(-TPDD_\delta - \gamma) / \theta]$. In terms of our probability model,

(7) $\delta = 1 – T\_Dist\{ z_\delta; \gamma, \theta, v=N-2\}$,

which yields the following equation for z_δ,

(7') $z_\delta = \{ TINV\_Dist\{ (1-\delta); \gamma, \theta, v=N-2\}$, where "TINV_Dist{ . }" is the inverse function of the T_Dist{ . } function[4]. The implied equation for $TPDD_\delta$ is:

(8) $TPDD_\delta = - [\gamma + (z_\delta)(\theta)]$.

To calculate the minimum daily (system average) temperature to define our extreme weather event, we specify that this COLDEST-Day be one where the temperature would be lower with a "1-in-35" likelihood. This criterion translates into two equations to be solved based on equations (7) and (8) above:

(9) solve for "z_δ" from equation (7') above with (1-δ) = (1 - 1/35) = 1 - 0.0286,

(10) solve for "$TPDD_\delta$" from $TPDD_\delta = - [\gamma + (z_\delta)(\theta)]$.

The value of $z_\delta = 1.959$ and $TPDD_\delta = - [\gamma + (z_\delta)(\theta)] = 42.8$ degrees Fahrenheit, with values for "v=N-2"; along with "γ" and "θ" in Table 6, below.

SDG&E's "1-in-10" peak-day design temperature of 44.5 degrees Fahrenheit, is calculated in a methodologically similar way as for the 42.8 degree "1-in-35" peak day temperature. The criteria specified in equation (9) above for a "1-in-35" likelihood would be replaced by a "1-in-10" likelihood.

(9') solve for "z_δ" from equation (7') above with (1-δ) = (1 - 1/10) = 1 - 0.1000,

which yields a "z_δ" value of $z_\delta = 1.303$ and, $TPDD_\delta = - [\gamma + (z_\delta)(\theta)] = 44.5$ with values for "v=N-2"; along with "γ" and "θ" in Table 6, below.

A plot of the cumulative distribution function for $MinAVG_y$ based on "v=N-2", the fitted model parameters, "γ" and "θ" with values in Table 6, below, is shown in Figure 1.

---

such that $T\_Dist\{z; \gamma, \theta, \nu=N-2\}=\int f(t)\, dt$, from t=-∞ to t=z. Also, the notation Γ(.) is known in mathematics as the GAMMA function; see http://www.wikipedia.org/wiki/Gamma_function for a description. Also, see *Statistical Theory*, 3rd Ed., B.W. Lindgren, MacMillian Pub. Inc, 1976, pp. 336-337.

[4] Computer software packages such as SAS and EXCEL have implemented statistical and mathematical functions to readily calculate values for T_Dist{ . } and TINV_Dist{ . } as defined above.

Table 5

YEAR	MINAVG	Month(MinAvg)
1972	46.7814	Dec
1973	46.1965	Jan
1974	44.2263	Dec
1975	44.1965	Jan
1976	45.0864	Jan
1977	50.6682	Mar
1978	42.7248	Dec
1979	45.1698	Jan
1980	53.8081	Jan
1981	49.8647	Jan
1982	48.8364	Dec
1983	51.5031	Jan
1984	48.4748	Dec
1985	46.1132	Dec
1986	50.1132	Feb
1987	41.5031	Dec
1988	45.4465	Dec
1989	45.1698	Jan
1990	43.7798	Feb
1991	48.7798	Mar
1992	47.1698	Dec
1993	46.7798	Jan
1994	48.0566	Nov
1995	51.1698	Dec
1996	48.7798	Feb
1997	49.0849	Dec
1998	46.7798	Dec
1999	48.8081	Jan
2000	50.3616	Jan
2001	47.6950	Jan
2002	45.7515	Jan
2003	49.0566	Dec
2004	47.7515	Nov
2005	47.8081	Jan
2006	48.3616	Dec
2007	43.3616	Jan
2008	48.7233	Dec
2009	48.4182	Feb
2010	48.1981	Dec
2011	49.0849	Feb
2012	48.1415	Dec
2013	44.1415	Jan
2014	47.7798	Dec
2015	48.2698	Jan
2016	50.3099	Feb
2017	51.2760	Jan

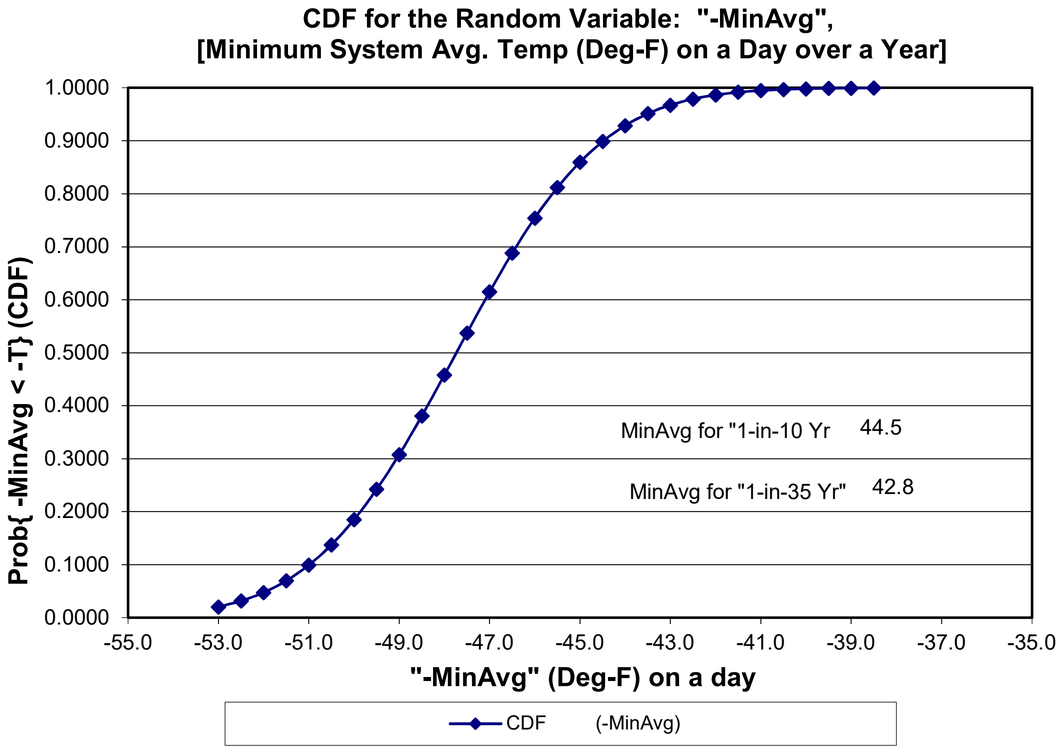
Table 6

alpha= 0.375

YEAR	Days/Yr	-MinAvg	Month (-MinAvg)	Rank	Emprical CDF	Fitted Model CDF
1980	366	-53.8081	Jan	1	0.0213	-2.089
1983	365	-51.5031	Jan	2	0.0426	-1.761
2017	365	-51.2760	Jan	3	0.0638	-1.553
1995	365	-51.1698	Dec	4	0.0851	-1.394
1977	366	-50.6682	Mar	5	0.1064	-1.264
2000	365	-50.3616	Jan	6	0.1277	-1.153
2016	365	-50.3099	Feb	7	0.1489	-1.053
1986	365	-50.1132	Feb	8	0.1702	-0.964
1981	365	-49.8647	Jan	9	0.1915	-0.881
1997	365	-49.0849	Dec	10	0.2128	-0.804
2011	365	-49.0849	Feb	11	0.2340	-0.732
2003	365	-49.0566	Dec	12	0.2553	-0.663
1982	365	-48.8364	Dec	13	0.2766	-0.598
1999	366	-48.8081	Jan	14	0.2979	-0.534
1991	366	-48.7798	Mar	15	0.3191	-0.473
1996	366	-48.7798	Feb	16	0.3404	-0.414
2008	365	-48.7233	Dec	17	0.3617	-0.356
1984	365	-48.4748	Dec	18	0.3830	-0.300
2009	365	-48.4182	Feb	19	0.4043	-0.244
2006	366	-48.3616	Dec	20	0.4255	-0.189
2015	365	-48.2698	Jan	21	0.4468	-0.135
2010	365	-48.1981	Dec	22	0.4681	-0.081
2012	365	-48.1415	Dec	23	0.4894	-0.027
1994	365	-48.0566	Nov	24	0.5106	0.027
2005	366	-47.8081	Jan	25	0.5319	0.081
2014	365	-47.7798	Dec	26	0.5532	0.135
2004	366	-47.7515	Nov	27	0.5745	0.189
2001	366	-47.6950	Jan	28	0.5957	0.244
1992	365	-47.1698	Dec	29	0.6170	0.300
1972	365	-46.7814	Dec	30	0.6383	0.356
1993	365	-46.7798	Jan	31	0.6596	0.414
1998	365	-46.7798	Dec	32	0.6809	0.473
1973	365	-46.1965	Jan	33	0.7021	0.534
1985	365	-46.1132	Dec	34	0.7234	0.598
2002	366	-45.7515	Jan	35	0.7447	0.663
1988	366	-45.4465	Dec	36	0.7660	0.732
1979	365	-45.1698	Jan	37	0.7872	0.804
1989	365	-45.1698	Jan	38	0.8085	0.881
1976	365	-45.0864	Jan	39	0.8298	0.964
1974	365	-44.2263	Dec	40	0.8511	1.053
1975	365	-44.1965	Jan	41	0.8723	1.153
2013	365	-44.1415	Jan	42	0.8936	1.264
1990	365	-43.7798	Feb	43	0.9149	1.394
2007	365	-43.3616	Jan	44	0.9362	1.553
1978	365	-42.7248	Dec	45	0.9574	1.761
1987	365	-41.5031	Dec	46	0.9787	2.089

"Gamma" (Fitted) =	**-47.74**
"Theta" (Fitted) =	**2.50**
Deg. Freedom=	**44**

Figure 1


CDF for the Random Variable: "-MinAvg",
[Minimum System Avg. Temp (Deg-F) on a Day over a Year]
Prob{ -MinAvg < -T} (CDF)
1.0000
0.9000
0.8000
0.7000
0.6000
0.5000
0.4000
0.3000
0.2000
0.1000
0.0000
-55.0
-53.0
-51.0
-49.0
-47.0
-45.0
-43.0
-41.0
-39.0
-37.0
-35.0
MinAvg for "1-in-10 Yr 44.5
MinAvg for "1-in-35 Yr" 42.8
"-MinAvg" (Deg-F) on a day
CDF (-MinAvg)

V. Estimating the Uncertainty in the Peak-Day Design Temperature

The calculated peak-day design temperatures in section IV above also have a statistical uncertainty associated with them. The estimated measures of uncertainty recommended for our use are calculated from the fitted model for the probability distribution and are believed to be reasonable, although rough, approximations.

The basic approach used the estimated parameters for the probability distribution (see the results provided in Table 6, above) to calculate the fitted temperatures as a function of the empirical CDF listed in Table 6. These fitted temperatures are then "compared" with the observed temperatures by calculating the difference = "observed" – "fitted" values. The full set of differences are then separated into the lower third (L), the middle third (M) and the upper third (U) of the distribution. Finally, calculate values of the root-mean-square error (RMSE) of the differences in each third of the distribution, along with the entire set of differences overall. The data in Table 6, below, show the temperature data and the resulting RMSE values.

The formula below is used to calculate the RMSE for a specified set of "N" data differences:

$$\text{RMSE} = \text{SQRT}\left\{ \left(\sum_{i=1, \ldots, N} e[\,i\,]^{2}\right) / \,(N-2)\right\},$$

where e[i] = *observed* less *fitted* value of temperature, T[i]. The number of estimated paramters (3 for the GEV model, 2 for the T-Dist and EV1 models) is subtracted from the respective number of data differences, N, in the denominator of the RMSE expression.

Since both the "1-in-35" and "1-in-10" peak-day temperature values are in the lower third quantile of the fitted distribution, the calculated standard error for these estimates is 0.44 Deg-F.

Table 7

Quantile: (Lower, Middle, Upper 3rd's)	Observed $T_{[i]}$ Temp. Ranked	Fitted Value of $T_{[i]}$	Residual $e_{[i]}$: Obs'd. less Fitted Value of $T_{[i]}$	Square of $e_{[i]}$:
U	53.8081	53.4596	0.3486	0.121493
U	51.5031	52.3780	-0.8749	0.765450
U	51.2760	51.7764	-0.5004	0.250395
U	51.1698	51.3409	-0.1711	0.029274
U	50.6682	50.9917	-0.3234	0.104618
U	50.3616	50.6957	-0.3341	0.111611
U	50.3099	50.4359	-0.1261	0.015894
U	50.1132	50.2023	-0.0891	0.007948
U	49.8647	49.9885	-0.1237	0.015310
U	49.0849	49.7900	-0.7051	0.497128
U	49.0849	49.6037	-0.5188	0.269166
U	49.0566	49.4273	-0.3708	0.137459
U	48.8364	49.2591	-0.4227	0.178665
U	48.8081	49.0976	-0.2895	0.083790
U	48.7798	48.9417	-0.1618	0.026183
M	48.7798	48.7903	-0.0105	0.000110
M	48.7233	48.6428	0.0804	0.006466
M	48.4748	48.4985	-0.0237	0.000562
M	48.4182	48.3567	0.0615	0.003784
M	48.3616	48.2169	0.1447	0.020943
M	48.2698	48.0786	0.1912	0.036559
M	48.1981	47.9414	0.2566	0.065860
M	48.1415	47.8048	0.3366	0.113320
M	48.0566	47.6685	0.3881	0.150645
M	47.8081	47.5319	0.2763	0.076321
M	47.7798	47.3947	0.3852	0.148366
M	47.7515	47.2564	0.4952	0.245188
M	47.6950	47.1166	0.5784	0.334511
M	47.1698	46.9748	0.1950	0.038012
M	46.7814	46.8304	-0.0491	0.002407
M	46.7798	46.6830	0.0969	0.009386
L	46.7798	46.5316	0.2482	0.061603
L	46.1965	46.3757	-0.1792	0.032108
L	46.1132	46.2142	-0.1010	0.010202
L	45.7515	46.0460	-0.2944	0.086674
L	45.4465	45.8696	-0.4231	0.179008
L	45.1698	45.6833	-0.5136	0.263764
L	45.1698	45.4848	-0.3151	0.099275
L	45.0864	45.2710	-0.1845	0.034056
L	44.2263	45.0374	-0.8110	0.657736
L	44.1965	44.7776	-0.5811	0.337649
L	44.1415	44.4816	-0.3402	0.115712
L	43.7798	44.1324	-0.3526	0.124325
L	43.3616	43.6969	-0.3353	0.112423
L	42.7248	43.0953	-0.3705	0.137270
L	41.5031	42.0137	-0.5106	0.260737
			Overall RMSE ($e_{[i]}$):	**0.38** °F
			Upper 3rd RMSE ($e_{[i]}$):	**0.45** °F
			Middle 3rd RMSE ($e_{[i]}$):	**0.30** °F
			Lower 3rd RMSE ($e_{[i]}$):	**0.44** °F

VI. The Relationship between Annual Likelihoods for Peak-Day Temperatures and "Expected Return Time"

The event whose probability distribution we've modeled is the likelihood that the minimum daily temperature over a calendar year is less than a specified value. And, in particular, we've used this probability model to infer the value of a temperature, our *peak-day design temperature* ($TPDD_\delta$), that corresponds to a pre-defined likelihood, δ, that the observed minimum temperature is less than or equal to this design temperature.

(1) δ = Prob{ Minimum Daily Temperature over the Year < $TPDD_\delta$ }.

For some applications, it is useful to think of how this specified likelihood (or "risk level" δ) relates to the expected number of years until this Peak-Day event would first occur. This expected number of years is what is meant by the *return period*. The results stated below are found in the book: ***Statistics of Extremes***, E.J. Gumbel, Columbia University Press, 1958, on pages 21-25.

(2) E[#Yrs for Peak-Day Event to Occur] = 1 / δ,

1 / Prob{ Minimum Daily Temperature over the Year < $TPDD_\delta$ }.

For our peak-day design temperature (42.8°F) associated with a 1-in-35 annual likelihood, the return period is 35 years (δ=1/35). For the 44.5°F peak-day design temperature, the return period is 10 years (δ=1/10). Occasionally, a less precise terminology is used. For example, the 42.8°F peak-day design temperature may be referred to as a "1-in-35 year cold day"; and the 44.5°F peak-day design temperature may be referred to as a "1-in-10 year cold day."

The probability model for the *return period*, as a random variable, is a geometric (discrete) distribution with positive integer values for the *return period*. The parameter δ = Prob{ Minimum Daily Temperature over the Year < $TPDD_\delta$ }.

(3) Prob{ *return period* = r } = $(1 – \delta)^{(r-1)}\,\delta$, for r = 1, 2, 3, . . .

The expected value of the *return period* is already given in (2) above; the variance of the *return period* is:

(4) Var[*return period*] = $(E[\text{ return period }])^2 \times (1- (1 / E[\text{ return period }]))$,

(4') Var[*return period*] = $(E[\text{ return period }]) \times (E[\text{ return period }] - 1)$.

Equations (4) and (4') indicate that the standard deviation (square root of the variance) of the *return period* is nearly equal to its expected value. Thus, there is substantial variability about the expected value—a *return period* is not very precise.

SDG&E SAS Code #1: SysAvTmp(SCG-Method) CustCntWgts-v2.sas

```
title1 "Calculation of 'System Average' Temperature for SDGandEs Gas System Load";
title2 "(Weighting of three stations: El-Cajon, San Diego (Lingberg Fld) and Miramar, Naval Air Station)" ;


****************************************************************************************************;
****                                                             ****;
****   [System-wide daily temperature series: ]                                                      ****;
****                                                             ****;
****   SECOND Alternative Weighting Method (Alt2)                          ****;
****                                                             ****;
****                                                             ****;
****************************************************************************************************;



options date number source notes ;

options mprint ;
/* %cour8l
   %cour8p */

options nomprint;

**<<LANDSCAPE: Set Line and Page Size for "SAS-Monospace and Roman 6pt.>>**;
options ls=211 ps=69;

**<<PORTRAIT: Set Line and Page Size for "SAS-Monospace and Roman 6pt.>>**;
**options ls=160 ps=90;

options date number source notes ;




****<< Directory for "SDWeathr" worksheet file:                              >>****;
****<< "S:\SDGandE-Econometric-Models\Weather-Data\Cycle\Sdge-Cycle-Methodology\"        >>****;




**<< Weather data file now maintained by SDGandE's Electric Demand Forecasting group.        >> **;
**<<   URL for SDG&E's data: "http://strategicanalysisd.sempra.com/Weather/Default.aspx"       >> **;


libname lb2   '/EDS_RB/Weather/2018 Cgr/SDGE' ; ***<< Directories for 2018 CGR  weather data analsysis. >>***;
libname LbDly '/EDS_RB/Weather/2018 Cgr/SDGE' ;

proc import out = SDWeathr
        datafile = "/EDS_RB/Weather/2018 Cgr/SDGE/SanDiegoWeather-20180202-MODIFIED.xls"
        dbms = xls replace;
                sheet = "Temp";
                getnames = yes;
run;



data DailySD ;
        set SDWeathr;
```

```
  *infile "/data/home/scgdf/2016_CGR_Weather/SDGE/sdweathr.sas7bdat" ;
  *** informat Date date12.   S42706Mx   S42706Mn   S47740Mx   S47740Mn   S93107Mx   S93107Mn    6.2 ;
  ***   << Format above is for "WthrbkInfo.xls" or "Wthrbkinfo-Rev.xls"  DATE variable!  >>  *** ;

  *informat Date mmddyy10.   S42706Mx   S42706Mn   S47740Mx   S47740Mn   S93107Mx   S93107Mn    6. ;

   ***<<< Note: For S# for El Cajon is 42706=CoopId# from NWS Weather Station Id# list, and              >>>*** ;
   ***<<<       the S# for San Diego, Lindberg Field is 47740=CoopId# from NWS Weather Station Id# list.  >>>*** ;
   ***<<<    But, the S# for Mirimar, Naval Air Station is 93107=WBanID# from NWS Weather Station Id# list. >>** ;

*   input Date      S42706Mx   S42706Mn   S47740Mx   S47740Mn   S93107Mx   S93107Mn  ;
*   format Date mmddyy8. ;

  if (date >= mdy(01,01,1972)
     AND date <= mdy(01,31,2018)) ; **<< Select only dates through January 31st, 2018  >>**;

  month = month(date) ;
  year = year(date) ;

  label S42706Mx  ='El Cajon - Max';
  label S42706Mn  ='El Cajon - Min';
  label S93107Mx  ='Miramar NAS - Max';
  label S93107Mn  ='Miramar NAS - Min';
  label S47740Mx  ='S.D. (Lindberg Fld.) - Max';
  label S47740Mn  ='S.D. (Lindberg Fld.) - Min';

  label avg = "Syst-Avg. Avg" ;
  label max = "Syst-Avg. Max" ;
  label min = "Syst-Avg. Min" ;
  label hdd = "Syst-Avg. Hdd" ;
  label cdd = "Syst-Avg. Cdd" ;


    year  = year(date);
    month = month(date);
    ARRAY A_Mx(I) S42706Mx   S47740Mx    S93107Mx  ;
    ARRAY A_Mn(I) S42706Mn   S47740Mn    S93107Mn  ;
    ARRAY A_Av(I) S42706Av   S47740Av    S93107Av  ;
    ARRAY A_Hd(I) S42706Hd   S47740Hd    S93107Hd  ;
    ARRAY A_Cd(I) S42706Cd   S47740Cd    S93107Cd  ;

    DO OVER A_Mx;
       A_Av = (A_Mx+A_Mn)/2;
      IF A_Av-INT(A_Av)=.5 THEN IF MOD(INT(A_Av),2)=0
      THEN A_Av=INT(A_Av);
      ELSE A_Av=INT(A_Av)+1;
      a_hd = max(0,65-a_av);
      a_cd = max(0,a_av-65);
    END;

   ***<<<  The equations below were used initially, but they are not exactly the correct method we have  >>>*** ;
   ***<<<  used with the SoCalGas temperature data anslysis work.                                 >>>*** ;
   ***Avg = 0.35*(S42706Mx   + S42706Mn  ) + 0.30*(S47740Mx   + S47740Mn  )
            + 0.35*(S93107Mx   +  S93107Mn  ) ;
   ***Avg = Avg/2 ;    **<< Do not round! >>** ;
   ***Hdd = max(0,65-Avg) ;  **<< Do not round! >>** ;
   ***Cdd = max(0,Avg-65) ;  **<< Do not round! >>** ;


/*****************************************
   ***<<< The equation below for "Avg" is the one we used perviously! >>>*** ;

   Avg = 0.35*( S42706Av ) + 0.30*( S47740Av )
            + 0.35*( S93107Av ) ;
   Max = 0.35*( S42706Mx ) + 0.30*( S47740Mx )
            + 0.35*(S93107Mx ) ;
   Min = 0.35*( S42706Mn ) + 0.30*( S47740Mn )
```

```
             + 0.35*( S93107Mn ) ;
******************************************/



/*****************************************
******************************************/
 ***<<<                                                                   >>>*** ;
 ***<<< The equation below for "Avg" is the one we NOW use.                           >>>*** ;
 ***<<<                                                                   >>>*** ;
 ***<<<     1). GKK indicated recently (per his e-mail on 7/31/2007) at 12:09 PM,       >>>*** ;
 ***<<<         that he had changed the station weights to the same (1/3)              >>>*** ;
 ***<<<         for each station.                                           >>>*** ;
 ***<<<         (Apparently, he did this "some time ago"!!!)                          >>>*** ;
 ***<<<     2). BMW believes this SHOULD HAVE been done a LOT sooner.                   >>>*** ;
 ***<<<         The prior weighting always seemed to be arbitrary and could            >>>*** ;
 ***<<<         not be documented as being based on customer/meter count              >>>*** ;
 ***<<<         data.                                                       >>>*** ;
 ***<<<     3). BMW believes this equal weighting for each station will be             >>>*** ;
 ***<<<         more in line with the six-zone appraoch we use for SoCalGas.           >>>*** ;
 ***<<<         With "system integration" on the horizon, we would then               >>>*** ;
 ***<<<         treat SDG&E as simply a "7th" temperature zone, with a seventh         >>>*** ;
 ***<<<         zone-weight based on the number of SDG&E active residential            >>>*** ;
 ***<<<         meter/customer count that is around  804,545 for June 2007            >>>*** ;
 ***<<<         from Al Burye "R2Page" gas data.                                  >>>*** ;
 ***<<<                                                                   >>>*** ;


 **<< Equal wgts for each station. >>** ;
 ** Avg =    (( S42706Av ) + ( S47740Av ) +  ( S93107Av )) / 3 ;
 ** Max =    (( S42706Mx ) + ( S47740Mx ) +  ( S93107Mx )) / 3 ;
 ** Min =    (( S42706Mn ) + ( S47740Mn ) +  ( S93107Mn )) / 3 ;

 ** do over a_hd ;  ***<<< Round Degree-Days for each station. >>>*** ;
 **    a_hd = round(a_hd) ;
 **    a_cd = round(a_cd) ;
 ** end ;

     **<< Equal wgts for each station. >>** ;
 ** Hdd =    (( S42706Hd ) + ( S47740Hd ) +  ( S93107Hd )) / 3 ;    **<< Do not round! >>** ;
 ** Cdd =    (( S42706Cd ) + ( S47740Cd ) +  ( S93107Cd )) / 3 ;    **<< Do not round! >>** ;


/**** Per Downie Beckett's e-mail of February 2, 2015  ***************

(We can assume all the SDGE's customers with gas services also have electric services.
Here is the count for gas services by climate zone.
Representative for December-2014, billing month.

ClimateZone      Gas Service Points        Electric Service Points (including the gas customers)

Coastal       469,087                          806,105
Mountain          12                           17,644
Desert           313                           3,584
Inland        398,037                          589,279
*****************************************************************/


 **<< SECOND Alternative Weighting Method (Alt2)             >>** ;
 **<< Weighted using Coastal & Inland counts ONLY.           >>** ;
 **<<                                               >>** ;
 **<< Lindberg Field    ---->  Coastal (SdTz1)               >>** ;
 **<< El Cajon (Santee) ---->  Inland  (SdTz2)               >>** ;
 **<< Mirimar           ---->  Coastal + Inland              >>** ;

 **<<  Based on analysis of this and method "Alt1", it makes the better         >>** ;
 **<<  transition from the simple "1/3" weighting for each location to one      >>** ;
```

```
**<<  based on current SDGE gas-customer count data. (BMW, 4/8/2011.}          >>** ;
**<<  (See work done at "S:\Weather\2010Cgr\SDGandE-Alt1Wgt\"                  >>** ;
**<<   and "S:\Weather\2010Cgr\SDGandE-Alt2Wgt\")                             >>** ;
**<<                                                         >>** ;
**<<  Formula:                                                   >>** ;
**<<     Sys-Avg = (1/3) * (Mirimar)                                     >>** ;
**<<            [+ (1/3) * (Inland_Cnt)  * (El-Cajon/Santee)                >>** ;
**<<             + (1/3) * (Coastal_Cnt) * (Lindberg Field)]                >>** ;
**<<                 / [Inland_Cnt + Coastal_Cnt]                          >>** ;
**<<                                                         >>** ;

 Coastal = 498013 ;    **<< Per Feb-2018 data from Mehdi Danesh                                    >>** ;
 Inland  = 420085 ;


  label S42706Mx  ='El Cajon - Max';
  label S42706Mn  ='El Cajon - Min';
  label S93107Mx  ='Miramar NAS - Max'; **<< It turns out that Miramar NAS is      >>** ;
  label S93107Mn  ='Miramar NAS - Min'; **<<  on the border of SdTz1 and SdTz2.    >>** ;
  label S47740Mx  ='S.D. (Lindberg Fld.) - Max';
  label S47740Mn  ='S.D. (Lindberg Fld.) - Min';

 SdTz1Mx = (1*S93107Mx + 2*S47740Mx) / (1+2) ;  **<< Coastal (SdTz1):  Miramar NAS & Lindberg Fld. >>** ;
 SdTz2Mx = (1*S93107Mx + 2*S42706Mx) / (1+2) ;  **<< Inland  (SdTz2):  Miramar NAS & El Cajon(Santee) >>** ;

 SdTz1Mn = (1*S93107Mn + 2*S47740Mn) / (1+2) ;  **<< Coastal (SdTz1):  Miramar NAS & Lindberg Fld. >>** ;
 SdTz2Mn = (1*S93107Mn + 2*S42706Mn) / (1+2) ;  **<< Inland  (SdTz2):  Miramar NAS & El Cajon(Santee) >>** ;

 SdTz1Av = (1*S93107Av + 2*S47740Av) / (1+2) ;  **<< Coastal (SdTz1):  Miramar NAS & Lindberg Fld. >>** ;
 SdTz2Av = (1*S93107Av + 2*S42706Av) / (1+2) ;  **<< Inland  (SdTz2):  Miramar NAS & El Cajon(Santee) >>** ;

 Avg =    ((Coastal) * SdTz1Av + (Inland) * SdTz2Av) / (Coastal + Inland) ;   **<< Do not round! >>** ;
 Max =    ((Coastal) * SdTz1Mx + (Inland) * SdTz2Mx) / (Coastal + Inland) ;   **<< Do not round! >>** ;
 Min =    ((Coastal) * SdTz1Mn + (Inland) * SdTz2Mn) / (Coastal + Inland) ;   **<< Do not round! >>** ;

 do over a_hd ;  ***<<< Round Degree-Days for each station. >>>*** ;
   a_hd = round(a_hd) ;
   a_cd = round(a_cd) ;
 end ;

 SdTz1Hd = (1*S93107Hd + 2*S47740Hd) / (1+2) ;  **<< Coastal (SdTz1):  Miramar NAS & Lindberg Fld. >>** ;
 SdTz2Hd = (1*S93107Hd + 2*S42706Hd) / (1+2) ;  **<< Inland  (SdTz2):  Miramar NAS & El Cajon(Santee) >>** ;

 SdTz1Cd = (1*S93107Cd + 2*S47740Cd) / (1+2) ;  **<< Coastal (SdTz1):  Miramar NAS & Lindberg Fld. >>** ;
 SdTz2Cd = (1*S93107Cd + 2*S42706Cd) / (1+2) ;  **<< Inland  (SdTz2):  Miramar NAS & El Cajon(Santee) >>** ;

 Hdd =    ((Coastal) * SdTz1Hd + (Inland) * SdTz2Hd) / (Coastal + Inland) ;    **<< Do not round! >>** ;
 Cdd =    ((Coastal) * SdTz1Cd + (Inland) * SdTz2Cd) / (Coastal + Inland) ;    **<< Do not round! >>** ;


 drop Coastal Inland ;

run ;




proc sort data=DailySD ;
  by year month ;
run ;

data LbDly.DailySD ;  **<< Save a copy of ALL the Daily Temperature Variables by Station and Sys-Avg for later use. >>** ;
  set DailySD ;
run ;


data LbDly.SavgSDGE ;  **<< Save  JUST Daily System Avgerage Temperature Variables for later use. >>** ;
```

```
   set DailySD ;
  format date date9. ;
  keep date  avg max min hdd cdd ;
run ;



proc means data=DailySD nway noprint;
     class year month;
     var hdd
         S42706Hd   S47740Hd    S93107Hd
         cdd
         S42706Cd   S47740Cd    S93107Cd  ;
     output out=MonSDGE sum=;
title3 "(Heating Degree-Days by Calendar Month)" ;
run ;



data MonSDGE ;
     set MonSDGE ;   *** Used to be named "MonSD.sd2" on 12/07/2004. *** ;

   date = mdy(month,1,year) ;
*** NOTE: Previously, (12/07/2004), exported as ".dbf" file a copy of "MonthSD.sd2" under the name "MonthSCG.dbf" ... *** ;
*** Subsequent UPDATES using the SoCalGas (SCG) system-avg-calc method are to "MonSDGE.dbf" ... *** ;
run;



/********************
proc dbload dbms=dbase data=MonSDGE ;   **<< Export a ".dbf" version of "MonSDGE.sd2" for later use.   >>** ;
                                       **<< Note:  Need to DELETE the previously-created ".dbf" file! >>** ;
 ** path='S:\Weather\2016Tcap-Phase II\SDGandE-Alt2Wgt\MonSDGE.dbf';
  path='C:\Weather\2016Tcap-Phase II\SDGandE-Alt2Wgt\MonSDGE.dbf';
  limit=0;
  load;
run;
********************/



proc sort data=MonSDGE  out=monthly(keep=date year month hdd cdd);
     by year month;
run;

* For saving non-rounded monthly numbers;
data lb2.MonSDGE;
              set MonSDGE;
run;



/***********
***********/
data lb2.mn72_17 ;  **<< Save a copy of the System Average Monthly Hdd/Cdd data for later use. >>** ;
   set monthly ;
  keep date year month hdd cdd ;
run ;




data ByStnDD ;
   set MonSDGE ;

  dateYYMM = mdy(month,1,year) ;
```

```
  if (dateYYMM < mdy(02,01,2018)) ;  **<< Only Pre February-2018 observations! >>** ;

  hdd = round(hdd) ;  **<< Round NOW for Monthly Reporting! >>** ;
  cdd = round(cdd) ;  **<< Round NOW for Monthly Reporting! >>** ;


   label S42706Hd  ='El Cajon - Hdd';
   label S42706Cd  ='El Cajon - Cdd';
   label S93107Hd  ='Miramar NAS - Hdd';
   label S93107Cd  ='Miramar NAS - Cdd';
   label S47740Hd  ='S.D. (Lindberg Fld.) - Hdd';
   label S47740Cd  ='S.D. (Lindberg Fld.) - Cdd';

   label hdd = "Syst-Avg. Hdd" ;
   label cdd = "Syst-Avg. Cdd" ;

run ;




proc sort data=ByStnDD ;
   by year month ;
run ;

data lb2.ByStnDD ;  **<< Save a copy of the MONTHLY Hdd/Cdd by Station for later use. >>** ;
   set ByStnDD ;
run ;




proc print data=ByStnDD  uniform split="/" ;
 id year;
 by year;
 sumby year;
     var dateYYMM hdd
         S42706Hd   S47740Hd    S93107Hd ;
     sum hdd
         S42706Hd   S47740Hd    S93107Hd ;
     format  dateYYMM worddate3.  hdd
         S42706Hd   S47740Hd    S93107Hd   6. ;

  label s42706Hd = "El Cajon   /              / ----------"
        s93107Hd = "Miramar    / Naval Air St./ ----------"
        s47740Hd = "San Diego  / Lindberg Fld./ ----------"

        hdd      = "System Avg./ for SDGandE  / ----------" ;


    ***<<< Note: For S# for El Cajon is 42706=CoopId# from NWS Weather Station Id# list, and               >>>*** ;
    ***<<<       the S# for San Diego, Lindberg Field is 47740=CoopId# from NWS Weather Station Id# list.  >>>*** ;
    ***<<<    But, the S# for Mirimar, Naval Air Station is 93107=WBanID# from NWS Weather Station Id# list. >>** ;

title3 "(Heating Degree-Days by Calendar Month)" ;

run ;



proc means data=ByStnDD noprint ;
   class year ;
```

```
    var hdd
        S42706Hd   S47740Hd    S93107Hd ;
    output out=HdSmByYr sum=;
run ;

/**
data HdSmByYr ;
  set HdSmByYr ;
  if (year=.) then delete ;
run ;
**/

proc print data=HdSmByYr  uniform split="/" ;
  where year ne . ;
    var year hdd
        S42706Hd   S47740Hd    S93107Hd ;
    format   hdd
        S42706Hd   S47740Hd    S93107Hd   6. ;

  label s42706Hd = "El Cajon   /              / ----------"
        s93107Hd = "Miramar    / Naval Air St./ ----------"
        s47740Hd = "San Diego  / Lindberg Fld./ ----------"

        hdd      = "System Avg./ for SDGandE  / ----------" ;


    ***<<< Note: For S# for El Cajon is 42706=CoopId# from NWS Weather Station Id# list, and              >>>*** ;
    ***<<<       the S# for San Diego, Lindberg Field is 47740=CoopId# from NWS Weather Station Id# list.  >>>*** ;
    ***<<<    But, the S# for Mirimar, Naval Air Station is 93107=WBanID# from NWS Weather Station Id# list. >>** ;

title3 "(Heating Degree-Days by Calendar Month)" ;

run ;
```

```
proc print data=ByStnDD  uniform split="/" ;
 id year;
 by year;
 sumby year;
    var dateYYMM Cdd
        S42706Cd   S47740Cd    S93107Cd ;
    sum Cdd
        S42706Cd   S47740Cd    S93107Cd ;
    format  dateYYMM worddate3.  Cdd
        S42706Cd   S47740Cd    S93107Cd   6. ;

  label s42706Cd = "El Cajon   /              / ----------"
        s93107Cd = "Miramar    / Naval Air St./ ----------"
        s47740Cd = "San Diego  / Lindberg Fld./ ----------"

        Cdd      = "System Avg./ for SDGandE  / ----------" ;


    ***<<< Note: For S# for El Cajon is 42706=CoopId# from NWS Weather Station Id# list, and              >>>*** ;
    ***<<<       the S# for San Diego, Lindberg Field is 47740=CoopId# from NWS Weather Station Id# list.  >>>*** ;
    ***<<<    But, the S# for Mirimar, Naval Air Station is 93107=WBanID# from NWS Weather Station Id# list. >>** ;

title3 "(Cooling Degree-Days by Calendar Month)" ;
```

```
run ;



proc means data=ByStnDD noprint ;
   class year ;
     var Cdd
         S42706Cd   S47740Cd    S93107Cd ;
     output out=CdSmByYr sum=;
run ;

/**
data CdSmByYr ;
  set CdSmByYr ;
  if (year=.) then delete ;
run ;
**/

proc print data=CdSmByYr  uniform split="/" ;
  where year ne . ;
     var year Cdd
         S42706Cd   S47740Cd    S93107Cd ;
     format   Cdd
         S42706Cd   S47740Cd    S93107Cd   6. ;

  label s42706Cd = "El Cajon   /              / ----------"
        s93107Cd = "Miramar    / Naval Air St./ ----------"
        s47740Cd = "San Diego  / Lindberg Fld./ ----------"

        Cdd      = "System Avg./ for SDGandE  / ----------" ;


    ***<<< Note: For S# for El Cajon is 42706=CoopId# from NWS Weather Station Id# list, and               >>>*** ;
    ***<<<       the S# for San Diego, Lindberg Field is 47740=CoopId# from NWS Weather Station Id# list.  >>>*** ;
    ***<<<    But, the S# for Mirimar, Naval Air Station is 93107=WBanID# from NWS Weather Station Id# list. >>** ;

title3 "(Cooling Degree-Days by Calendar Month)" ;

run ;









***<< Print Summary Tables of Hdd and Cdd by month  with Annual Totals >>*** ;

proc transpose data=monthly out=hddSmry prefix=mon ;
   by year;
   id month ;
   var hdd ;
run ;

data hddSmry ;
   set hddSmry ;
 hddTot = sum(of mon1-mon12) ;
run ;

proc print data=hddSmry ;
 id year ;
 var hddTot mon1-mon12 ;
title1 'Monthly Heating Degree-Days from 1972 thru 2017(Month-to-Date)';
title2 "   " ;
```

```
run ;



proc transpose data=monthly out=cddSmry prefix=mon ;
   by year;
   id month ;
   var cdd ;
run ;

data cddSmry ;
   set cddSmry ;
 cddTot = sum(of mon1-mon12) ;
run ;

proc print data=cddSmry ;
 id year ;
 var cddTot mon1-mon12 ;
title1 'Monthly Cooling Degree-Days from 1972 thru 2017(Month-to-Date)';
run ;








quit ;
```

SDG&E SAS Code #2: <u>MinAvg-Freq(SDGandE)ByMonth.sas</u>

```
title1 "Calculate Min{Avg} (Minimum Average Daily Temp.) by Months for all data over a specified range of YEARS." ;

/*****************************************************************************************************

    File: S:\Weather\2016 CGR\SDGE123!@#qwe1
\MinAvg-Freq(SDGandE)ByMonth.sas

 ****************************************************************************************************/


options date number source notes ;
options mprint ;

/* %cour8l
   %cour8p; */

 options ls=160 ps=90 ;  *<To get PORTRAIT  and SAS-Monospace w/Roman 6pt. FONT >*;
*options ls=211 ps=69 ;  *<To get LANDSCAPE and SAS-Monospace w/Roman 6pt. FONT >*;

**options nomprint ;


      **<< Data set "SysAvgSD.sd2" was created by SAS program on file:                                   >>** ;
      **<<     "S:\Weather\2016Tcap-Phase II\SDGandE-Alt2Wgt\SysAvTmp(SCG-Method)_CustCntWgts-v2.sas"       >>** ;
      **<<                                                                                      >>** ;
libname in '/EDS_RB/Weather/2018 Cgr/SDGE';



   **<< Directory to save a copy of output data set! >>** ;
libname out '/EDS_RB/Weather/2018 Cgr/SDGE';


proc contents data=in.SAvgSDGE ;
run ;


       *** << Note: Need to change here and below!  Don't forget! >> *** ;
%let startyr=1972;  ***<< Value of "Start Year" >>***;
%let lastyr=2017 ;  ***<< Value of "Last Year" >>***;

/********
%let tgtmonth=xx ;  ***<< Value of "Target Month", i.e., 1,2,3, ... 12 >>** ;
********/




proc format ;
   value mmmmmm    1='Jan'
                 2='Feb'
                 3='Mar'
                 4='Apr'
                 5='May'
                 6='Jun'
                 7='Jul'
                 8='Aug'
                 9='Sep'
                10='Oct'
                11='Nov'
                12='Dec'
                13='Min4Yr' ;
run ;
```

```
%macro FreqMon(name_mon,tgtmonth) ;

data combined ;
    set in.SAvgSDGE ;

year = year(date) ;
month= month(date) ;
day = day(date) ;
  **<< To "Select" Winter Months Only. >>**;
  **  if month in (1,2,3,11,12);

*** hdd = round(avg,1);  **<< Comment out ... Do NOT round data this time! >>** ;
if ((year >= &startyr) & (year <= &lastyr)) ;

%if (&tgtmonth >= 1 and &tgtmonth <= 12) %then
  %do ;
if (month = &tgtmonth) ;  **<< To select only a specific month! >>** ;
  %end ;
%else
  %do ;
month = 13 ;  **<< Set "month" variable to "13" and select ALL "months" of the YEAR! >>** ;
  %end ;
run;


proc sort data=combined;
 by year month day ;
run;

***proc contents data=combined ;
***run ;

proc means data=combined noprint;
  by year month ;
  var avg ;
  output out=&name_mon min=MinAvg;
  title1 "Minium{avg} (Minimum Avg. Daily Temp.) for &name_mon  by YEAR=&startyr to &lastyr";
run;

proc print data=&name_mon ;
run ;

%mend ;









%let startyr=1972;  ***<< Value of "Start Year" >>***;
%let lastyr=2017 ;  ***<< Value of "Last Year" >>***;


/************
%let name_mon='January ';  ***<< NAME of "Target Month" >>***;
%let tgtmonth= 1 ;    ***<< Value of "Target Month" >>***;
************/

%FreqMon(JAN,1) ;

data AllMonth ;
  set JAN ;
```

```
run ;


%FreqMon(FEB,2) ;

data AllMonth ;
  set AllMonth FEB ;
run ;


%FreqMon(MAR,3) ;

data AllMonth ;
  set AllMonth MAR ;
run ;


%FreqMon(APR,4) ;

data AllMonth ;
  set AllMonth APR ;
run ;


%FreqMon(MAY,5) ;

data AllMonth ;
  set AllMonth MAY ;
run ;


%FreqMon(JUN,6) ;

data AllMonth ;
  set AllMonth JUN ;
run ;


%FreqMon(JUL,7) ;

data AllMonth ;
  set AllMonth JUL ;
run ;


%FreqMon(AUG,8) ;

data AllMonth ;
  set AllMonth AUG ;
run ;


%FreqMon(SEP,9) ;

data AllMonth ;
  set AllMonth SEP ;
run ;


%FreqMon(OCT,10) ;

data AllMonth ;
  set AllMonth OCT ;
run ;


%FreqMon(NOV,11) ;
```

```
data AllMonth ;
  set AllMonth NOV ;
run ;


%FreqMon(DEC,12) ;

data AllMonth ;
  set AllMonth DEC ;
run ;


%FreqMon(ALL,13) ;

data AllMonth ;
  set AllMonth All ;
run ;



proc Tabulate data=AllMonth;
  class year month ;
  var MinAvg ;
  table year, MinAvg*f=6.2*(month)/rts=6 ;
  label MinAvg='Min{Avg} by Mo' ;
 *keylabel All ='Min4Yr';
  format month mmmmmm. ;
  title2 "Min{avg} by Months for all Years from YEAR=&startyr to &lastyr";
run;


proc sort data=AllMonth out=MinAvg_d;
  by month year MinAvg ;
run ;


proc print data=MinAvg_d ;
  by month ;
  pageby month ;
  var year MinAvg ;
run ;


data out.MinAvg_d ;  *<< Save a copy for later use!  >>** ;
  set MinAvg_d ;
 drop _freq_ _type_ ;
run ;

proc contents data=out.MinAvg_d ;
run ;


/************************
 **<< Export a copy as ".dbf" as well. >>** ;
 **<< Note: Must "delete" the prior version, othewise the save will not execute. >>** ;
proc dbload dbms=dbase data=MinAvg_d ;
path='/data/home/scgdf/2016_CGR_Weather/SDGE/MinAvg_d.dbf';
limit=0;
load;
run;
*************************/



quit ;
```

SDG&E SAS Code #3: GEV4DlyTemp(NLReg2)ByMonthMACRO_Sdge.sas

```
Title1 "Data Analysis for Maximum/Minimum Daily SysAvg Temperatures (Un-Rounded)." ;
Title2 "Fit GEV Probability Model to Empirical CDF using NL-OLS Regression Methods." ;

/**************************************************************************************************************/
/*                                                                                    */
/*                                                                                    */
/*                                                                                    */
/* FILE SAVED: "S:\Weather\2016 Cgr\SDGE\GEV4DlyTemp(NLReg2)ByMonthMACRO_Sdge.sas"                      */
/*          for Annual Max of Negative of Min. Temp.                                  */
/*          for each of calendar months Jan-Dec, and the entier year(index=13).                 */
/*                                                                                    */
/*          Fit GEV models (3-parameter and 2-parameter), plus a simple T-Distribution model.         */
/*                                                                                    */
/*                                                                                    */
/**********************************************************************************************&**************/



options mprint ;
/* %cour8p
   %cour8l */


 options ls=211 ps=69 ; **<<LANDSCAPE:  SAS-Monospace w/Roman 6pt. Font >>**;
*options ls=160 ps=90 ; **<<PORTRAIT:   SAS-Monospace w/Roman 6pt. Font >>**;

options date number notes ;



libname out1 '/EDS_RB/Weather/2018 Cgr/SDGE' ;  **<< Directory for daily weather variables as INPUT. >>**;

libname out2 '/EDS_RB/Weather/2018 Cgr/SDGE/MinTemp' ;  **<< Directory for estimation results OUPUT files. >>** ;



proc contents data=out1.SAvgSDGE ;
run ;

data seriesD ;
  set out1.SAvgSDGE ;
  year = year(date) ;
  month = month(date) ;
  posAvg = avg ;
  negAvg = -avg ;
run ;


proc means data=seriesD noprint nway ;
 class year month ;
  var posAvg negAvg ;
 output out=mostat
                mean=posAvg  negAvg
                 max=MxPosAvg MxNegAvg
                 min=MnPosAvg MnNegAvg    ;
run;


proc sort data=mostat ;
   by year month ;
run ;
```

```
data mostat ;
  set mostat ;
  MxPRatio = MxPosAvg/ PosAvg  ;
  MnPRatio = MnPosAvg/ PosAvg  ;
  MxNRatio = MxNegAvg/ NegAvg  ;
  MnNRatio = MnNegAvg/ NegAvg  ;
run ;






/****************
***<< Print Summary Tables of Means/Minimums/Maximums of daily NEGATIVE-Temperatures (degrees-F). >>*** ;

proc transpose data=mostat out=AvTData prefix=AvT_ ;
  where (year < 2016) ;
  by year;
  id month ;
  var NegAvg  ;
run ;

data AvTData ;
  set AvTData ;

 if (mod(year,4)=0) then do ;
       AvT_13 = (AvT_1 + AvT_3 + AvT_5 + AvT_7 + AvT_8 + AvT_10 + AvT_12)*31
             + (AvT_4 + AvT_6 + AvT_9 + AvT_11)*30
             + (AvT_2)*29 ;
       AvT_13 = AvT_13 / 366 ;
       end ;
 else do ;
       AvT_13 = (AvT_1 + AvT_3 + AvT_5 + AvT_7 + AvT_8 + AvT_10 + AvT_12)*31
             + (AvT_4 + AvT_6 + AvT_9 + AvT_11)*30
             + (AvT_2)*28 ;
       AvT_13 = AvT_13 / 365 ;
       end ;

run ;

proc print data=AvTData ;
 id year ;
 var AvT_13 AvT_1-AvT_12 ;
title3 'Monthly Mean NEGATIVE Temperature (Deg-F) from 1972 thru 2015.';
run ;





proc transpose data=mostat out=MnTData prefix=MnT_ ;
  where (year < 2016) ;
  by year;
  id month ;
  var MnNegAvg ;
run ;

data MnTData ;
  set MnTData ;
 MnT_13 = min(of MnT_1-MnT_12) ;
run ;

proc print data=MnTData ;
```

```
 id year ;
 var MnT_13 MnT_1-MnT_12 ;
title3 'Monthly MINIMUM NEGATIVE-Temperature (Deg-F) from 1972 thru 2015.';
run ;
****************/




proc transpose data=mostat out=MxTData prefix=MxT_ ;
   where (year < 2018) ;
   by year;
   id month ;
   var MxNegAvg ;
run ;

data MxTData ;
   set MxTData ;
 MxT_13 = max(of MxT_1-MxT_12) ;
run ;

proc print data=MxTData ;
 id year ;
 var MxT_13 MxT_1-MxT_12 ;
title3 'Monthly MAXIMUM NEGATIVE-Temperature (Deg-F) from 1972 thru 2017.';
run ;




/****************
***<< Descriptive Statistics: Maximums of daily NEGATIVE-Temperatures (Deg-F) for Year and each calendar month. >>*** ;


proc corr data=MxTData ;
  var MxTyr MxT_1 - MxT_12 ;
title3 'Correlation Matrix of Monthly Maximum NEGATIVE-Temperatures (Deg-F) within same year.';
run ;

proc arima data=MxTData ;
  identify var=MxT_13 ;
  identify var=MxT_1 ;
  identify var=MxT_2 ;
  identify var=MxT_3 ;
  identify var=MxT_4 ;
  identify var=MxT_5 ;
  identify var=MxT_6 ;
  identify var=MxT_7 ;
  identify var=MxT_8 ;
  identify var=MxT_9 ;
  identify var=MxT_10 ;
  identify var=MxT_11 ;
  identify var=MxT_12 ;
title3 "Auto-correlation analysis of each calendar month's Maximum NEGATIVE-Temperatures (Deg-F) within same year.";
run ;

proc univariate normal data=MxTData plot ;
  id year ;
  var MxT_13 MxT_1 - MxT_12 ;
title3 "Probability plots and tests for NORMALity by each calendar month's Maximun NEGATIVE-Temperatures (Deg-F) time series.";
```

```
run ;


proc means data=MxTData ;
  var MxT_1 - MxT_12  MxT_13 ;
run ;
****************/









***<< Statistical Estimation of GEV Models: Maximums of daily heating degrees for Year and each calendar month. >>*** ;

%macro RankIt(file=MxTData,var=MxT_13,rank=Rank,prob=PrMxT_13,Nobser=46,PltValue=0.375) ;
proc sort data=&file ;
  by &var ;
run ;

data &file ;
   set &file ;
   retain &rank 0     alpha &pltvalue ;

  &rank = &rank + 1 ;
  &prob =  (&rank - alpha) / (&Nobser +(1 - 2*alpha)) ;
run ;

proc print data=&file ;
  var &var &rank &prob alpha year ;
run ;
%mend RankIt ;




%macro GEVfit(file=MxTData,ofile=MxTNL1,outfit=fit1,outest=est1,depvar=PrMxT_13,var=MxT_13,typeGEV=1,
             KappaI=0.25,GammaI=-47.05,ThetaI=2.77,YrLo=1972,YrHi=2017) ;
proc sort data=&file ;
   by year ;
run ;



proc model data=&file converge=0.001
           maxit=500 dw ;  outmodel=&ofile ;
     range year = &YrLo to &YrHi ;  **<< Dropped Jan-Feb 2016 data. >>** ;


     y = (&var - Gamma) / Theta ;

    %if &typeGEV=1 %then %do ;  ***<< 3-parameter GEV Model. >>>*** ;
        &depvar = exp( -(1 - Kappa * (y))**(1/Kappa) ) ;
        %let typmod = 3-parameter GEV Model. ;
        %end ;

    %if &typeGEV=2 %then %do ;  **<< 2-parameter "Double Exponential" or  "Gumbel" Model. >>** ;
        &depvar = exp( -exp(-(y)) ) ;
        %let typmod = 2-parameter Double Exponential or Gumbel Model. ;
        %end ;

    %if (&typeGEV NE 1) AND (&typeGEV NE 2) %then  %do ; **<< 2-parameter "T-Dist" Model. >>** ;
        dft=(&YrHi - &YrLo) +1 -2 ;
        &depvar = probt(y,dft) ;
        %let typmod = 2-parameter T-Dist Model. ;
```

```
      %end ;


  %if &typeGEV = 1 %then %do ;
  parms
       Kappa  &KappaI
       Gamma  &GammaI
       Theta  &ThetaI ;
  %end ;

  %if (&typeGEV NE 1)  %then %do ;
  parms
       Gamma  &GammaI
       Theta  &ThetaI ;
  %end ;


  fit &depvar   /out=&outfit  outall
               outest=&outest corrb corrs outcov ;

title3 "Non-linear Estimation of &&typmod: for Maximum NEGATIVE Temperature (Deg-F)." ;
run ;
%mend GEVfit ;






%macro GEVbyMo(mm=_1) ;


*********<<< Analysis for "January"    (i.e., SUFIX "mm" = "_1"  >>>******************;
*********<<< Analysis for "February"   (i.e., SUFIX "mm" = "_2"  >>>******************;
*********<<< Analysis for "March"      (i.e., SUFIX "mm" = "_3"  >>>******************;
*********<<< Analysis for "April"      (i.e., SUFIX "mm" = "_4"  >>>******************;
*********<<< Analysis for "May"        (i.e., SUFIX "mm" = "_5"  >>>******************;
*********<<< Analysis for "June"       (i.e., SUFIX "mm" = "_6"  >>>******************;
*********<<< Analysis for "July"       (i.e., SUFIX "mm" = "_7"  >>>******************;
*********<<< Analysis for "August"     (i.e., SUFIX "mm" = "_8"  >>>******************;
*********<<< Analysis for "September"  (i.e., SUFIX "mm" = "_9"  >>>******************;
*********<<< Analysis for "October"    (i.e., SUFIX "mm" = "_10"  >>>******************;
*********<<< Analysis for "November"   (i.e., SUFIX "mm" = "_11"  >>>******************;
*********<<< Analysis for "December"   (i.e., SUFIX "mm" = "_12"  >>>******************;
*********<<< Analysis for "ALL Months"   (i.e., SUFIX "mm" = "_13"  >>>******************;

%RankIt(file=MxTData,var=MxT&mm,rank=Rank&mm,prob=PrMxT&mm,Nobser=46,PltValue=0.375) ;





%GEVfit(file=MxTData,ofile=MxTNL2,outfit=fit2,outest=est2,depvar=PrMxT&mm,var=MxT&mm,typeGEV=2,
      KappaI=0.25,GammaI=&&gamma&mm,ThetaI=&&theta&mm,YrLo=1972,YrHi=2017) ;


proc print data=fit2 ;
run ;




proc transpose data=fit2 out=pred2 prefix=probP_ ;
   where (_type_ = "PREDICT" ) ;
   by year;
   var prmxt&mm ;
run ;
```

```
data comb2 ;
   merge MxTData pred2 ;
  by year ;
   ProbP2 = ProbP_1 ;
  keep year MxT&mm PrMxT&mm ProbP2 ;
run ;


proc print data=comb2 ;
run ;


proc plot data=comb2 ;
  plot prmxt&mm*MxT&mm='*'
       probP2*MxT&mm='-' / overlay ;
run ;


proc print data=est2 ;
run ;



data out2.est2&mm ;    ***<<< Save a copy of the "Double Exponential Model" estimation results!  >>>*** ;
  set est2 ;
run ;





%GEVfit(file=MxTData,ofile=MxTNL0,outfit=fit0,outest=est0,depvar=PrMxT&mm,var=MxT&mm,typeGEV=0,
        KappaI=0.25,GammaI=&&gamma&mm,ThetaI=&&theta&mm,YrLo=1972,YrHi=2017) ;


proc print data=fit0 ;
run ;




proc transpose data=fit0 out=pred0 prefix=probP_ ;
   where (_type_ = "PREDICT" ) ;
   by year;
   var prmxt&mm ;
run ;

data comb0 ;
   merge MxTData pred0 ;
  by year ;
   ProbP0 = ProbP_1 ;
  keep year MxT&mm PrMxT&mm ProbP0 ;
run ;


proc print data=comb0 ;
run ;


proc plot data=comb0 ;
  plot prmxt&mm*MxT&mm='*'
       probP0*MxT&mm='-' / overlay ;
run ;


proc print data=est0 ;
```

```
run ;



data out2.est0&mm ;   ***<<< Save a copy of the 2-parameter "T-Distribution" Model estimation results!  >>>*** ;
  set est0 ;
run ;




%GEVfit(file=MxTData,ofile=MxTNL1,outfit=fit1,outest=est1,depvar=PrMxT&mm,var=MxT&mm,typeGEV=1,
        KappaI=0.25,GammaI=&&gamma&mm,ThetaI=&&theta&mm,YrLo=1972,YrHi=2017) ;


proc print data=fit1 ;
run ;



proc transpose data=fit1 out=pred1 prefix=probP_ ;
   where (_type_ = "PREDICT" ) ;
   by year;
   var prmxt&mm ;
run ;

data comb1 ;
   merge MxTData pred1 ;
  by year ;
   ProbP1 = ProbP_1 ;
  keep year MxT&mm PrMxT&mm ProbP1 ;
run ;


proc print data=comb1 ;
run ;


proc plot data=comb1 ;
  plot prmxt&mm*MxT&mm='*'
       ProbP1*MxT&mm='-' / overlay ;
run ;



proc print data=est1 ;
run ;



data out2.est1&mm ;   ***<<< Save a copy of the "G.E.V. Model" estimation results!  >>>*** ;
  set est1 ;
run ;

%mend GEVbyMo ;









/****************
```

```
****************/

proc means data=MxTData ;
  var MxT_1 - MxT_12  MxT_13;
 output out=VarStat
        mean=mean1-mean12 mean13
        std=stdev1-stdev12 stdev13;
title3 "Calc. Means and Standard Deviantions to use as Starting Values in Non-Linear Estimations." ;
run ;


proc print data=VarStat ;
run ;

data _null_ ;
   set VarStat ;

  call symput('gamma_13',mean13) ;
  call symput('theta_13',stdev13) ;

  call symput('gamma_12',mean12) ;
  call symput('theta_12',stdev12) ;

  call symput('gamma_11',mean11) ;
  call symput('theta_11',stdev11) ;

  call symput('gamma_10',mean10) ;
  call symput('theta_10',stdev10) ;

  call symput('gamma_9',mean9) ;
  call symput('theta_9',stdev9) ;

  call symput('gamma_8',mean8) ;
  call symput('theta_8',stdev8) ;

  call symput('gamma_7',mean7) ;
  call symput('theta_7',stdev7) ;

  call symput('gamma_6',mean6) ;
  call symput('theta_6',stdev6) ;

  call symput('gamma_5',mean5) ;
  call symput('theta_5',stdev5) ;

  call symput('gamma_4',mean4) ;
  call symput('theta_4',stdev4) ;

  call symput('gamma_3',mean3) ;
  call symput('theta_3',stdev3) ;

  call symput('gamma_2',mean2) ;
  call symput('theta_2',stdev2) ;

  call symput('gamma_1',mean1) ;
  call symput('theta_1',stdev1) ;

run ;


%GEVbyMo(mm=_13) ;   **<< Annual(Entire Year of Data.) >>** ;





%GEVbyMo(mm=_1) ;   **<< Jan Data. >>** ;
```

```
%GEVbyMo(mm=_2) ;   **<< Feb Data. >>** ;
%GEVbyMo(mm=_3) ;   **<< Mar Data. >>** ;


%GEVbyMo(mm=_4) ;   **<< Apr Data. >>** ;


%GEVbyMo(mm=_5) ;   **<< May Data. >>** ;
%GEVbyMo(mm=_6) ;   **<< Jun Data. >>** ;
%GEVbyMo(mm=_7) ;   **<< Jul Data. >>** ;


%GEVbyMo(mm=_8) ;   **<< Aug Data. >>** ;


%GEVbyMo(mm=_9) ;   **<< Sep Data. >>** ;
%GEVbyMo(mm=_10) ;   **<< Oct Data. >>** ;
%GEVbyMo(mm=_11) ;   **<< Nov Data. >>** ;
%GEVbyMo(mm=_12) ;   **<< Dec Data. >>** ;





quit ;
```