REGULATORY APPLICATIONS OF THE RELATIONSHIPS BETWEEN NATURAL GAS USAGE AND WEATHER

James A. Gray
Dennis L. Patterson
Michael S. Proctor
Henry E. Warren

Energy Analysis & Forecasting Group
Economic Analysis Department
Missouri Public Service Commission
Jefferson City, Missouri

LIST OF CONTENTS

1997 ...

FOREWORD

This document provides the basics required for analysis and forecasting of natural gas usage for the determination of revenues and the revision of rates. The focus is on the Local Distribution Company (LDC) and its customers. Analysis of gas usage is required in a rate case in order to properly estimate the volumes and revenues that would be recoved from current rates under conditions of normal weather. Normal volumes are also required to evaluate proposed rates. In the context of the LDC's procurement of gas supply, forecasting of gas usage is required to determine both the daily usage profile and the peak day requirement.

The Energy Analysis and Forecasting Group within the Department of Economic Analysis at the Missouri Public Service Commission performs this analysis and produced this document. Contributions to the document were made by the following:

Jim Gray, Regulatory Economist
Dennis Patterson, Regulatory Economist
Mike Proctor, Senior Economist and Department Manager
Henry Warren, Regulatory Economist

This group does analysis and forecasting on both natural gas and electric utilities as well as large water utilities. For additional information, please contact:

Economic Analysis Department
Missouri Public Service Commission
Suite 530 Truman Building
Post Office Box 360
Jefferson City, Missouri 65102
Telephone: (573) 751-7275
Fax: (573) 751-0429
e-mail hwarren@mail.state.mo.us

Editors comment: Mr H.E. Warren (USA) is a member of the Working Group on Energy-Meteorology and Rapporteur on Energy-Meteorology and Economics for the WMO Commission for Climatology. This paper has been submitted as an example of a specific application of climate information to the energy sector.

I. BASIC CONCEPTS AND MEASUREMENTS

A. <u>The Measure of Usage</u>

Total usage can be measured either as volume (e.g. in cubic feet) of natural gas or therms[1] of natural gas, depending on the approved practice for a specific utility. Natural gas is metered on a cumulative basis. At the time the meter was installed, it most likely was set at a zero reading. From that point in time until the present, the meter has been measuring the total natural gas usage at a building or residence. The meters are read once a month by a utility meter reader. By taking the difference between this month's reading and last month's reading, the customers usage for the current month can be calculated. By taking the difference between this month's reading date and last month's reading date, the number of days over which the customer has taken service can also be calculated.

Meters are read on weekdays, excluding holidays. A "typical" month has four weeks plus an additional two or three days. Assuming that at least one of these two or three days is a weekday, there are normally at least twenty one (21) weekdays in each calendar month. Therefore, utilities will schedule to read all of their meters over a 21 day period (billing cycle), and groups of meters (approximately the total divided by 21 in size) are assigned to each of the 21 billing cycles. While the normal length (number of days between meter readings) of each billing cycle is thirty (30) days, the actual length of each billing cycle will vary from month to month.

The metered usage is aggregated (grouped) in two ways. First, each customer is assigned to a given rate class based on the eligibility requirements set out on the utility's tariffs. Second, each customer is assigned to a given billing cycle based on the utility's determination of the most efficient grouping of customers to assign to each billing cycle. Total usage and total days are measured for each rate class and billing cycle (see Figure I.1).

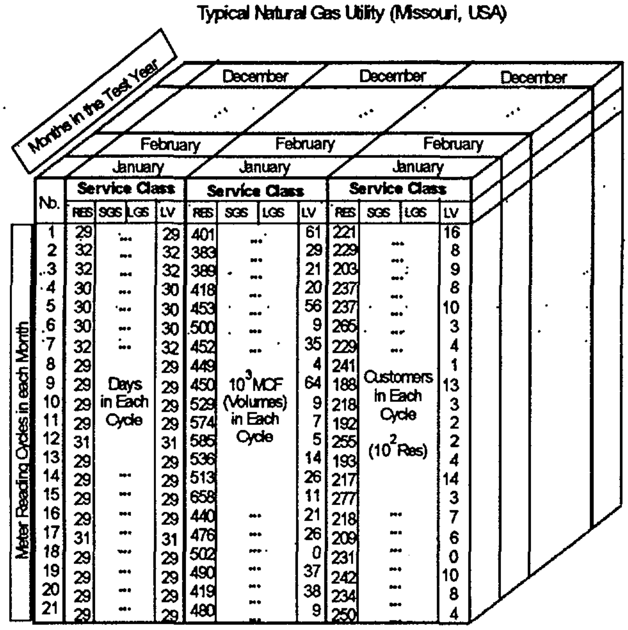
---

[1]Therm = Statutory unit of calorific value in gas supply.

FIGURE I.1

FIGURE I.1

BILLING CYCLE DATA BASE FOR A TEST YEAR
Typical Natural Gas Utility (Missouri, USA)



Service Classes

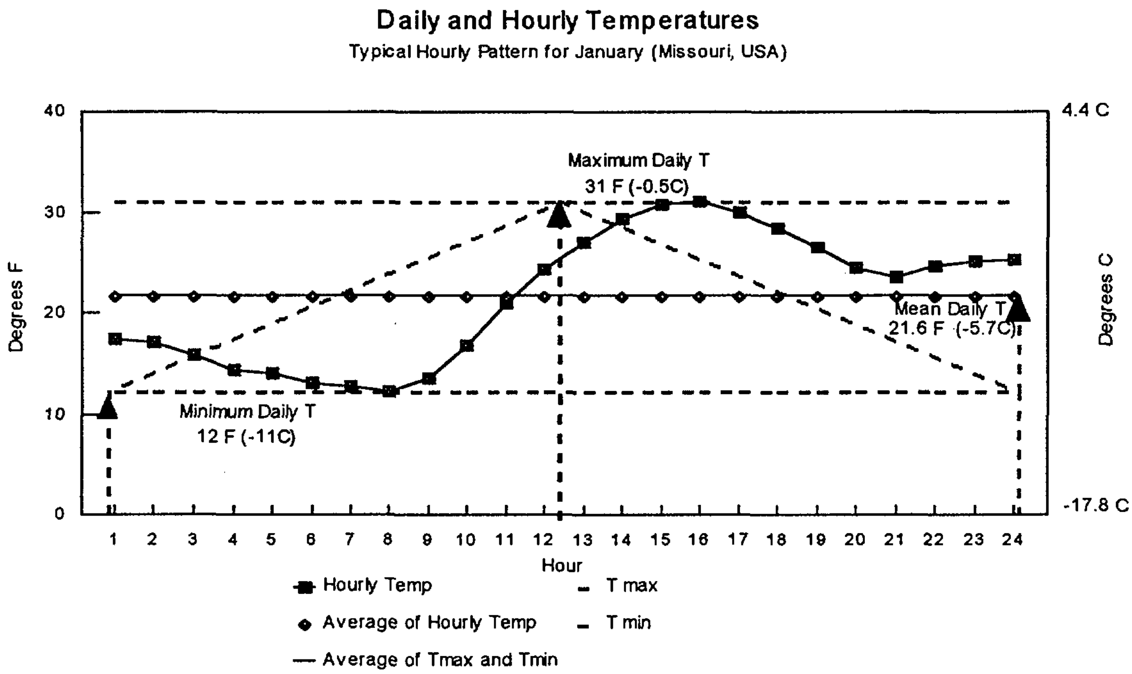
RES – Residential
SGS – Small General Service (< 1,000 MCF in all months)
LGS – Large General Service (>1,000 MCF in any month)
LV – Large Volume (>6,000 MCF in a year and Peak Day>25 MCF)

It is important that usage from one billing cycle not be simply added to the usage of another billing cycle. The two usage levels do not represent the same measures. Because each billing cycle can have a different number of days, the non-weather sensitive component of usage will be different. Because each billing cycle will have a different number of heating degree days, the weather sensitive component of usage will also be different. Thus, the basic measures of usage are: ***(1) Therms or volume (cubic feet) of usage; and (2) Number of days***; both of which are aggregated by: ***(1) Rate class; and (2) Billing cycle.***

B. <u>The Measure of Weather</u>

For natural gas usage, the most commonly used measure for weather is heating degree days (HDD). The national weather system collects daily information on temperatures, determines the daily maximum and minimum temperatures and calculates the daily mean temperature (DMT) as the average of the daily maximum and minimum temperatures. Conceptually, this can be pictured as a triangle on top of a rectangle (see Figure I.2). The height of the rectangle (the base) is the daily low temperature. The height of the triangle added to the height of the rectangle is the daily high temperature. The length of the rectangle is the 24 hours of the day. The daily mean temperature is half way from the top of the rectangle to the top of the triangle. In this conceptual simplification, the actual temperature is below the daily mean temperature for 12 hours and above the daily mean temperature for the other 12 hours.

FIGURE 1.2



If heating requirements respond linearly to temperature, then using daily mean temperature as the basic measure of weather will be a reasonably good predictor of daily natural gas usage. However, when it gets warm enough, there will be no heating requirements. Either the thermostat will keep the furnace from running or the user will shut off the furnace. A generally accepted criterion for this ceiling is 65° Fahrenheit. The difference between this ceiling and the daily mean temperature is defined as the heating degree days observed for a single day; i.e.,

$$HDD = \begin{cases} 65^o - DMT, & if\ DMT \leq 65^o \\ and & \\ 0, & if\ DMT > 65^o \end{cases}$$

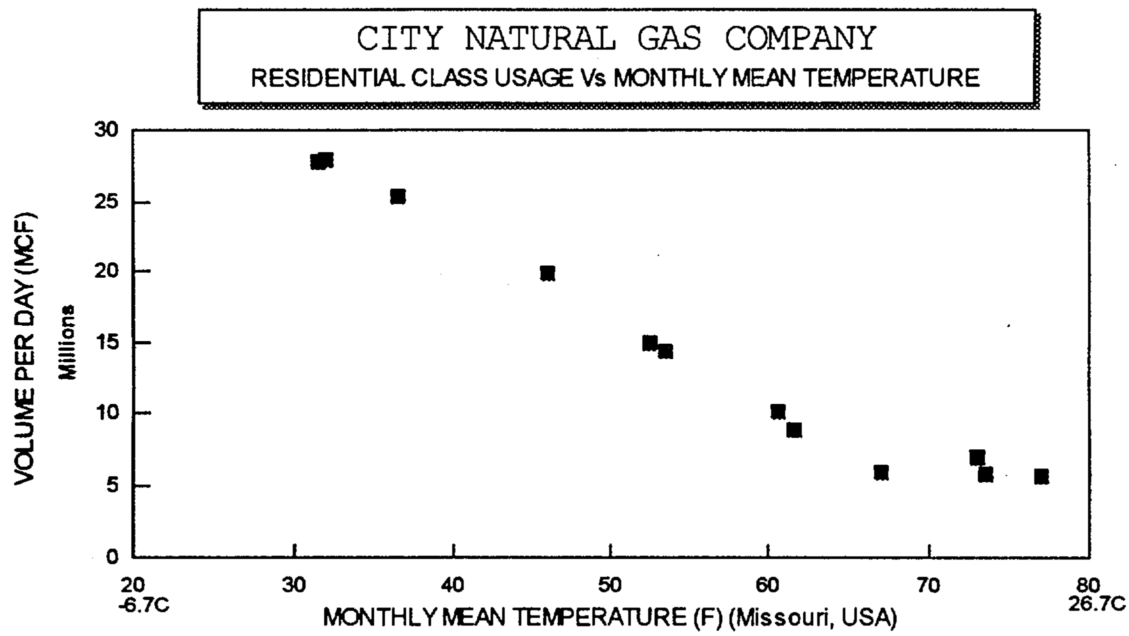
In theory, the heating requirements for one day having 10 HDDs or two days each having 5 HDDs will be the same.

It may be true that for a certain region of the country, using a different ceiling than 65°F provides a better predictor of usage; or different classes of service (e.g., residential versus commercial) may have different ceilings. But a well established approach has been to include a ceiling along with the use of daily mean temperature as the relevant measure of weather for explaining daily variations in natural gas usage.

C. The Relationship Between Usage and Weather

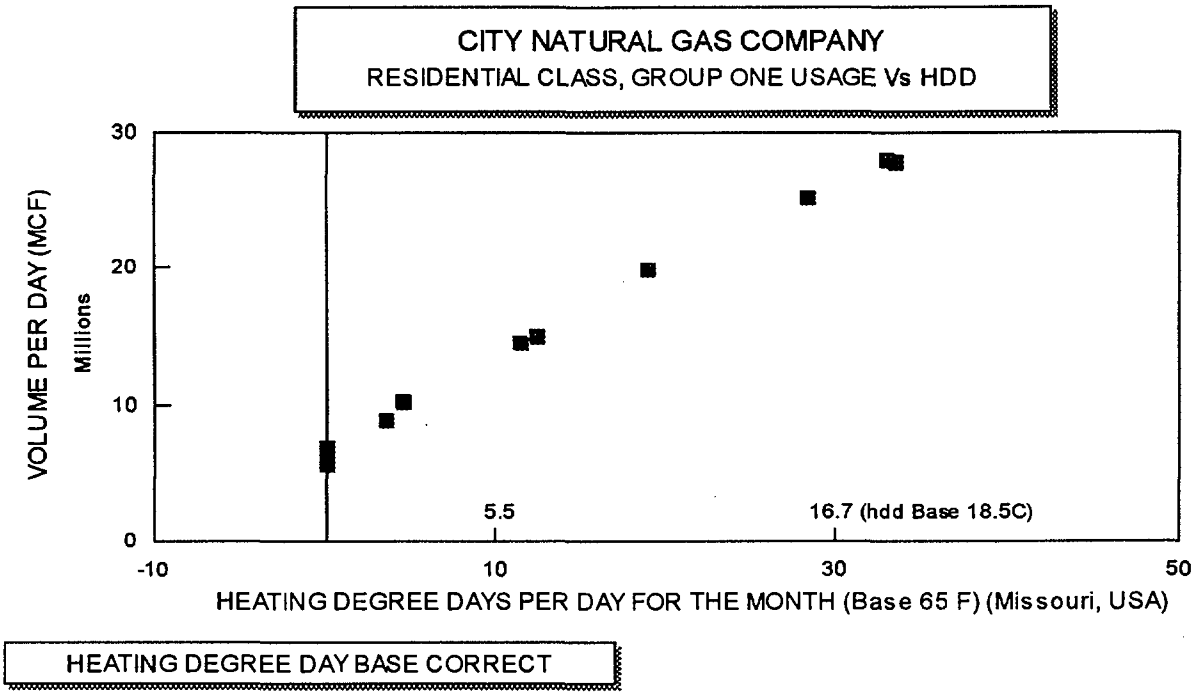
For a given billing cycle, the per day average usage can be calculated as the total billing cycle usage divided by the number of days in the billing cycle. Similarly, the per day average of daily mean temperature can be calculated as the sum of the daily mean temperatures for the billing cycle divided by the number of days in the billing cycle. If this is done for each month of a year and average usage per day is plotted against average mean temperature per day, the results will look similar to those shown on Figure I.3. Notice that there is a downward slope over the mean temperature range from 30°F to 65°F. Past 65°F, the average usage per day does not decrease because the primary use for gas within this range of temperature is for non-weather sensitive use.

FIGURE I.3



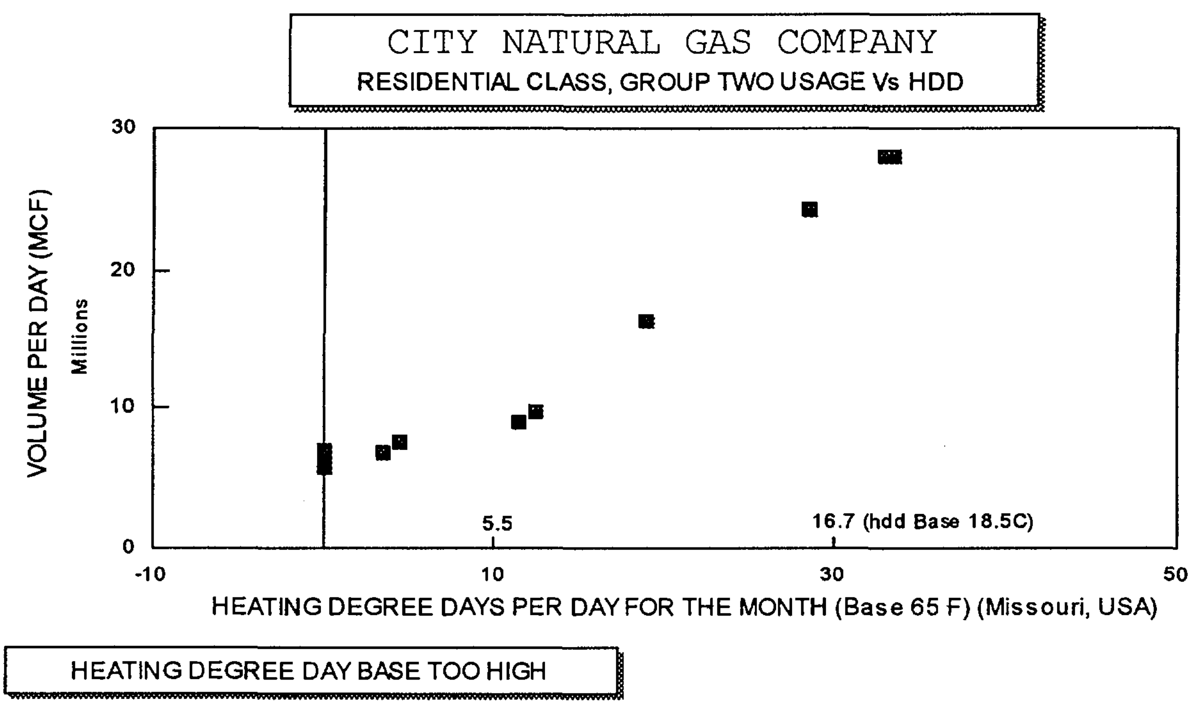
Instead of using an average of daily mean temperatures over a bill cycle, it is more appropriate to account for any days in which heating might have occurred by averaging over daily HDDs. This is done by dividing the HDDs in a bill cycle by the corresponding number of days. The relationship between usage and HDDs is shown in Figure I.4. For the months of June, July, August and September, the graph shows zero (or almost zero) heating degree days. The usage in these months will measure the non-weather sensitive component. The remaining months indicate increasing usage with increasing HDDs. The slope of the line going through these points is a measure of the incremental response of usage to variations in the weather.

FIGURE I.4



It should be pointed out that if the 65°F ceiling for calculating heating degree days is too high, then instead of a straight line connecting all of the points having positive heating degree days, the months with low levels of HDDs will appear to be on a line having a lower slope (left end of the plot) than the months with higher levels of HDDs, as illustrated in Figure I.5.

FIGURE I.5



D. The Weather Normalization of Natural Gas Usage

The objective of weather normalization is to provide a measure of the natural gas usage that would have occurred under normal weather conditions. In its simplest form this means estimating the change in usage associated with a change in weather, where the change in weather is the difference between normal weather and actual weather. Consider the following linear model of natural gas usage as a function of weather:

$$y = a + bx$$

where y = usage, and x = weather.

The intercept measures the **non-weather sensitive** component of the usage:

$$a = \textit{usage at zero weather}$$

The slope is a simple measure of the **change in usage** associated with a unit change in weather:

$$b = \frac{\Delta \; usage}{\Delta \; weather}$$

The slope times the weather variable measures the **weather sensitive** component of the usage:

$$bx = \textit{usage associated with "x" units of weather}$$

The error component measures that portion of the usage that is **not explained** by the linear model:

$$e = \textit{"random" (unexplained) usage}$$

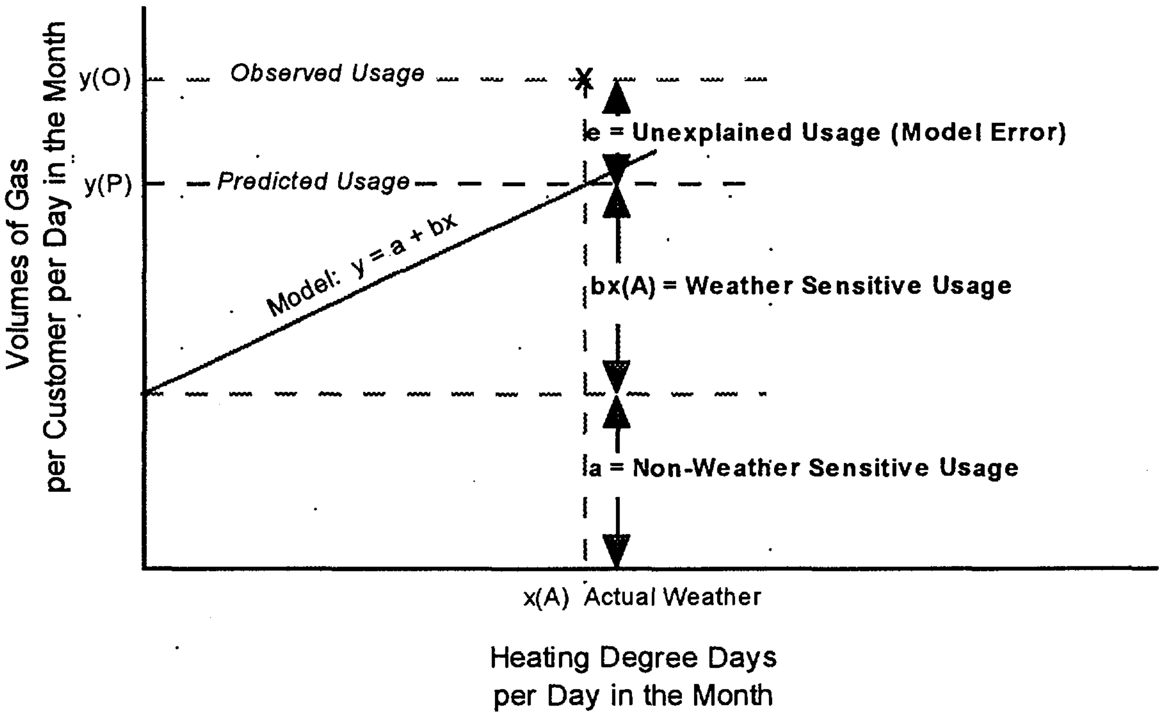
This linear model gives a good basic representation of the various components involved in the measurement (estimation) process. These three components are illustrated in Figure 6.

The importance of the three distinct components of usage is that only one of the three components is to be normalized. Specifically, the objective is to find the usage level which would have occurred at normal weather assuming all other factors have stayed the same. These other factors are those which caused the non-weather sensitive and random components of usage. The basic weather normalization model of usage does not attempt to normalize either of these two components of usage.

In Figure I.6, the measure for actual weather is specified as x(A) and the observed usage is specified as y(O).

FIGURE I.6

MODELING MONTHLY NATURAL GAS USAGE
Weather Sensitive Customer Class
Three Components of Gas Usage



Therefore it makes sense to think of this pair [x(A),y(O)] as the Actual, Observed values. In the linear model, the predicted value for usage at actual weather is given by

$$y(P) = a + bx(A) ,$$

and it makes sense to think of this pair [x(A),y(P)] as the *Actual, Predicted* values. If we introduce a normal level for weather, say x(N), then the predicted value for usage at normal weather would be given by

$$y(P)' = a + bx(N) ,$$

and this pair [x(N),y(P)'] would properly be called the *Normal, Predicted* values for usage. The normal, predicted values for weather and usage contain two normalized components of usage. First, the weather sensitive component of usage has been changed from bx(A) to bx(N), giving a net change (weather adjustment) of

$$b[x(N) - x(A)].$$

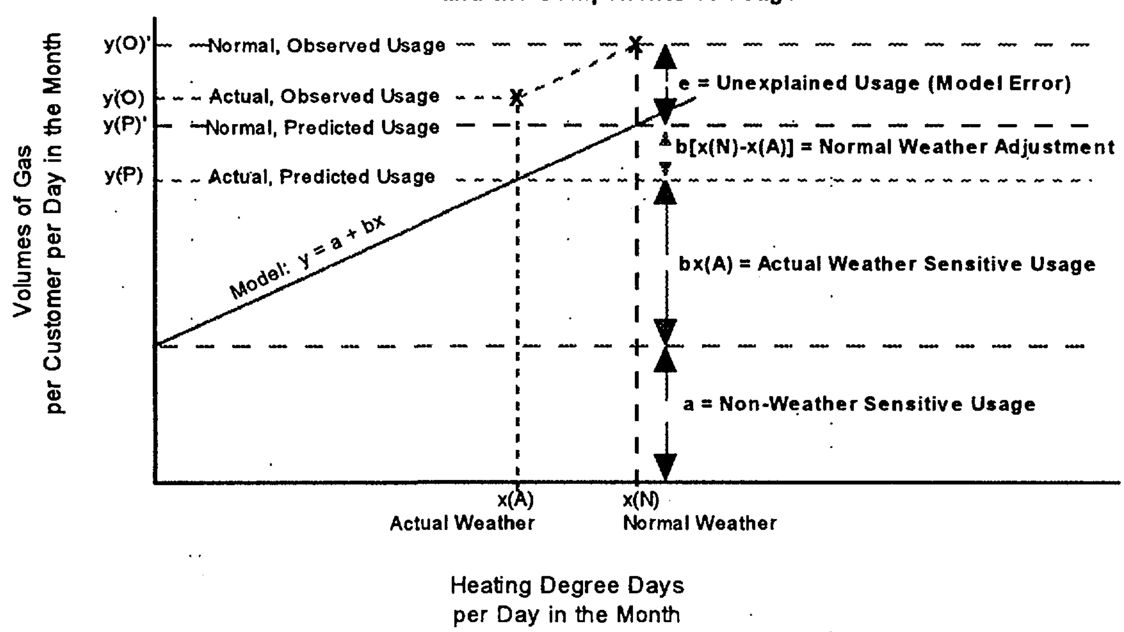
Second, the normal, predicted value for usage also assumes that the random component of usage is zero, and therefore implicitly normalizes that component of usage. If the objective is to normalize only the weather component, then the random component of usage should not be changed from its observed level; i.e., the error term should be added back. The following equations show this algebraically.

$$y(O)' = \begin{cases} = y(O) + b[x(N) - x(A)] \\ = a + bx(A) + e + bx(N) - bx(N) \\ = a + bx(N) + e \\ = y(P)' + e \end{cases}$$

The observed value for usage at normal weather is given by either adding the weather adjustment to the actual, observed level of usage, or by adding the error component to the normal, predicted level of usage. This pair [x(N), y(O)'] would properly be called the Normal, Observed values. Graphically, the weather adjustment to usage is shown in Figure I.7.

FIGURE I.7

ADJUSTING MOTHLY NATURAL GAS USAGE TO NORMAL WEATHER
Using the Model for the Weather Sensitive Customer Class
and the Components of Usage



The weather adjustment is the change in the predicted value of usage when the weather is changed from its actual to its normal level. In Figure I.7, this is shown by a movement along the straight line. The normal, observed level for usage can be thought of as drawing a parallelogram along the straight line. The height of the parallelogram is the error term. In Figure I.7, that height is measured above the straight line at the ordinate for actual weather, and to complete the parallelogram, that same height is added above the straight line at the ordinate for normal weather. The weather normalized value for observed usage can be calculated by either:

(1) Adding the weather adjustment for usage to the actual, observed level of usage; or

(2) Adding the error component to the normal, predicted level of usage.

In a careful application of weather normalization, the analyst should calculate the weather adjustment both ways and make sure the two calculations give the same answer. If they do not, then the analyst has made a mistake and should check both calculations.

II. THE CHOICE AND ANALYSIS OF WEATHER

A. <u>Tabulation of Temperature Data</u>

A. 1. Sources of temperature data.

Official temperature data are available for cooperative and first-order stations from the official NOAA publications "Climatological Data" (CD; c.f., figure II.1), "Local Climatological Data" (LCD), "Tape Deck TD-3200: Cooperative Summary of the Day" (TD-3200), and from digital data sets extracted from these by academic facilities such as the Atmospheric Sciences department at the University of Missouri.

FIGURE II.1

CLIMATOLOGICAL DATA

MISSOURI

MAY 1995

VOLUME 099 NUMBER 05

ISSN 0364-6068

TEMPERATURE AND PRECIPITATION EXTREMES

HIGHEST TEMPERATURE	91	MAY 2[illegible]	5 STATIONS
LOWEST TEMPERATURE	25	MAY 2	HAMILTON 2 W
GREATEST TOTAL PRECIPITATION	16.53		STOVER
LEAST TOTAL PRECIPITATION	1.66		PORTAGEVILLE
GREATEST 1 DAY PRECIPITATION	6.41	MAY [illegible]	VERSAILLES

"I certify that this is an official publication of the National Oceanic and Atmospheric Administration (NOAA). It is compiled using information from weather observing sites supervised by NOAA/National Weather Service and received at the National Climatic Data Center (NCDC), Asheville, North Carolina 28801."

DIRECTOR
NATIONAL CLIMATIC DATA CENTER

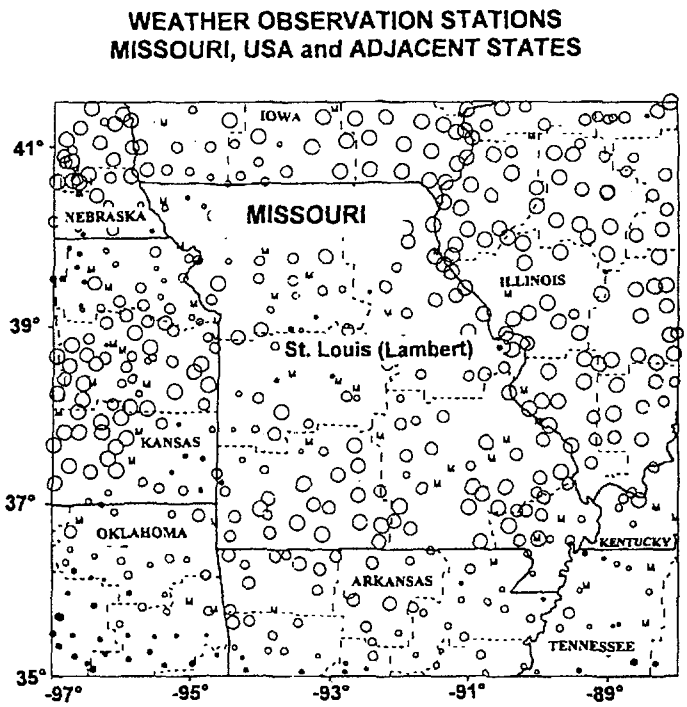
noaa — National Oceanic and Atmospheric Administration — National Environmental Satellite, Data and Information Service — National Climatic Data Center, Asheville, North Carolina

Very recent data (not yet deemed official) are available as digitized records from the National Weather Service (NWS) Climate Dial-up Computer service, by prearrangement from the NOAA Internet home page, and from photostatic copies of preliminary reports. The preliminary reports for either type of station may be requested from NOAA or directly from the observers.

A. 2. Choosing a weather station.

The choice of candidate stations depends upon the location with respect to the service area, the availability and quality of the data, and only then upon the type of station to be considered. Figure II.2 shows a map of the weather station locations for Missouri.

FIGURE II.2



(1) Location. A single-city service area may have shape that approaches that of an ordinary polygonal shape, such as a rectangle or circle, with no significant vacant areas within the perimeter. The candidate weather station for this shape should be within or very near to the perimeter. An example candidate station might be located at the city downtown airport, university, or federal building.

Another shape of service area resembles a crescent or doughnut, such as one comprising the suburban area which surrounds a city. A candidate weather station might be a regional or international airport also located in the suburbs, even though it may be closer to one side of the service area than to the center of mass. Weather from a protected, urban location would be less appropriate for this shape and type of service area. A utility may have more than one service area, with customers in two or more cities. A separate weather station will be required to normalize sales for each of the separate service areas.

(2) Availability and quality. For a weather station to qualify as a candidate, complete and accurate daily maximum and minimum temperature data must be made available for that station, not only for a test billing period but for a period long enough to determine the normal levels of the variables selected for modeling gas usage. The test billing period should be at least one full billing year, while the normals period should be the three decades used by NOAA to calculate the current published normals. Occasional (and inevitable) missing temperature observations will occur over these time spans. Isolated missing values are easily filled by inserting time interpolations, but any series of missing values are more properly filled by inserting unbiased average-of-three-neighbors values, calculated from observations at nearby stations. Thus, at least three alternate stations for each candidate must be chosen as well in order to maintain the quality of the data. In practice, all available daily temperature observations are tabulated for the candidate station and its alternates before the final tabulation for the candidate station.

(3) Type of station. Cooperative stations may shift locations every few years, and are usually manned by volunteers who record maximum temperature and minimum temperature for the previous 24 hours once each day at a scheduled hour (e.g., 7 AM). In contrast, a first-order station under NOAA administration is likely to have been located for many years at an airport near a population center, and is manned around the clock by professional observers who record the maximum and minimum temperatures for the 24-hour day which ends at midnight. If a cooperative station is better located with respect to the service area than an alternative first-order station, and if reasonably complete data are available for both, then the cooperative station is the preferred candidate. Data from first-order stations tend to be more complete, but are not inherently of better quality than that from cooperative stations. Weather normalization can be done satisfactorily with respect to weather from either class of station.

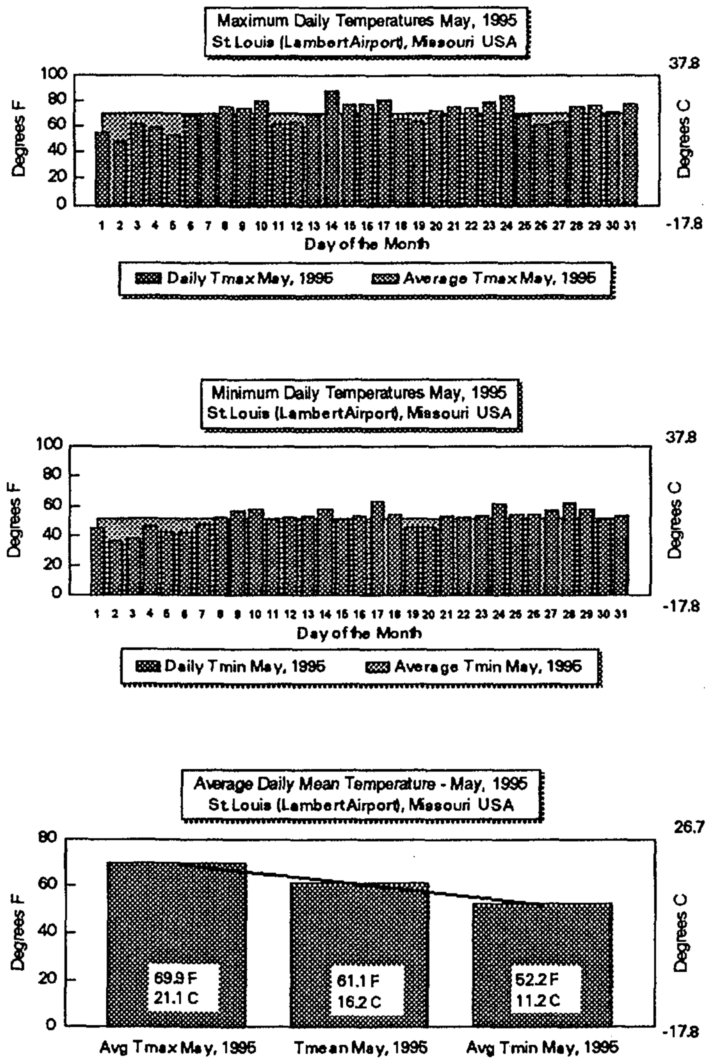
B. Determination of Normal Temperatures

B. 1. Thirty-year normals

By international agreement, three decades is the accepted time span for the calculation of normal weather. This length of time is held to be long enough to compensate for shorter-term cycles that may be present in the data, while not being so long that historical conditions which are no longer relevant may influence the calculations of normal weather. For temperatures, normals are calculated as the thirty-year average mean temperature for each month of the year. This calculation involves determination of the average maximum and minimum temperatures, each rounded to the nearest tenth of a degree. The average mean temperature is determined as the midpoint between the average maximum and the average minimum temperature, rounded to the nearest tenth of a degree. This process of averaging is illustrated in figure II.3.

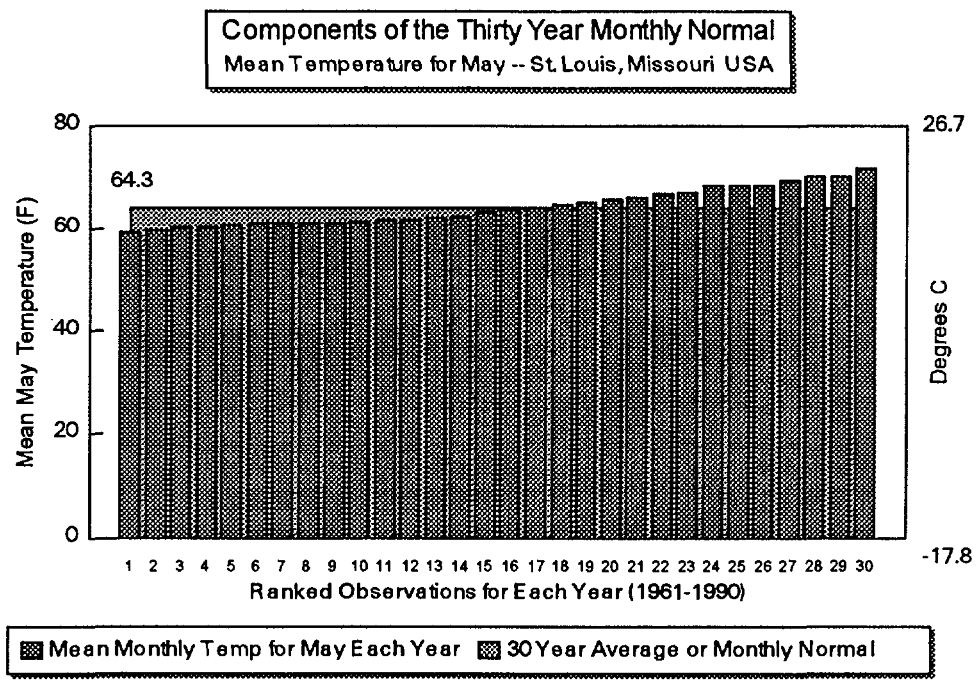
FIGURE II.3

COMPONENTS OF DAILY MEAN TEMPERATURE FOR A MONTH



Because the calculations of normals requires a substantial effort and commitment of NOAA's resources, they are done only once every ten years. The published normals remain the same for those ten years. NOAA performs adjustments and calculates new thirty-year normals at the end of each decade. The current edition is published as "Monthly Station Normals of Temperature, Precipitation and Degree Days MISSOURI 1961-90." While these monthly normals don't provide the daily values needed for utility billing cycle weather normalization, they do provide monthly benchmarks for the calculation of daily normals. Figure II.4 illustrates the 30 year average which NOAA applies to the average mean temperature for each month.

FIGURE II.4



B. 2. Exposure change adjustments

When NOAA weather normals are calculated for first-order stations and selected cooperative stations, special precautions are taken to insure that all the years of data in the calculations are consistent with the latest instrument installation. First, missing data are replaced with average-of-three neighbors calculations. If necessary, adjustments to historical data are then made for exposure changes such as change in observation practice, change in instrument type, and change in instrument location. For first-order and cooperative stations, the original daily weather data are published only once, shortly after the end of the month in which it occurs. When the daily data are published, monthly average temperatures and total precipitation are published with them. Before new normals are published at the end of each decade, the thirty-year series of monthly data is examined for inconsistencies. For example, the record might contain the information that the thermometer had been moved on a certain date. In order to calculate the effects of the move, the series of monthly averages before the move and after the move would be compared to the corresponding series at surrounding stations where no change occurred. Such effects are included by NOAA as adjustments in the thirty-year series of monthly average temperatures from which the normals are calculated.

B. 3. Monthly exposure change adjustments and daily temperatures

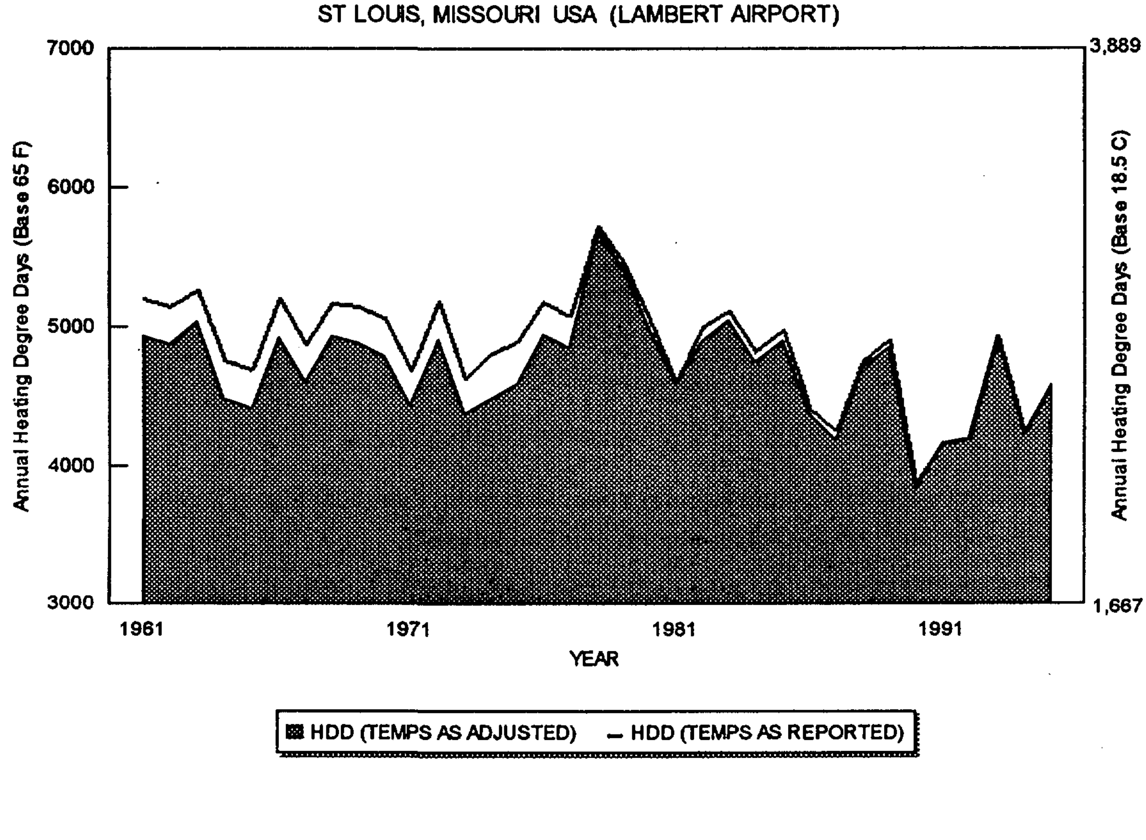
When thirty years of daily temperature data are used to calculate daily normals for a weather normalization, the series of monthly NOAA temperature adjustments must be incorporated to avoid bias in the results. First, it is necessary to consult the NOAA publication, "Tape Deck TD-9641: 1961-1990 Sequential Temperature and Precipitation" (TD-9641). TD-9641 contains a thirty-year series of adjusted monthly values for average daily maximum temperature, average daily minimum, and average mean daily temperature.

Then, daily 1961-1990 temperature data are tabulated from the sources listed above. Finally, adjustments are calculated which cause the daily temperature observations to add up to the NOAA adjusted monthly values over the thirty-year normals period.

Using the original daily observations, monthly average maximum temperature, average minimum temperature, and average mean daily temperature are calculated from observations for each month in the thirty years. The magnitudes of NOAA's adjustments may then be calculated for each month in the thirty years as the adjusted monthly average from the NOAA sequentials minus the directly calculated average from the daily observations. Finally, the monthly adjustments are added to each day's observation during the month to calculate adjusted daily temperatures. The standards for accuracy are that the monthly averages from the adjusted daily temperature data be equal to the monthly averages in NOAA's sequentials, and that the averages over the thirty years equal NOAA's published monthly normals. When the month-by-month and thirty-year averages by month are verified, daily normals may be calculated from the thirty-year time series of adjusted daily values. Figure II.5 shows a comparison of daily mean temperatures as reported to the adjusted daily mean temperatures that conform to NOAA's adjustments to monthly values.

FIGURE II.5

ADJUSTMENTS TO TIME SERIES OF HDD FOR INSTRUMENT EXPOSURE CHANGES
ST LOUIS, MISSOURI USA (LAMBERT AIRPORT)

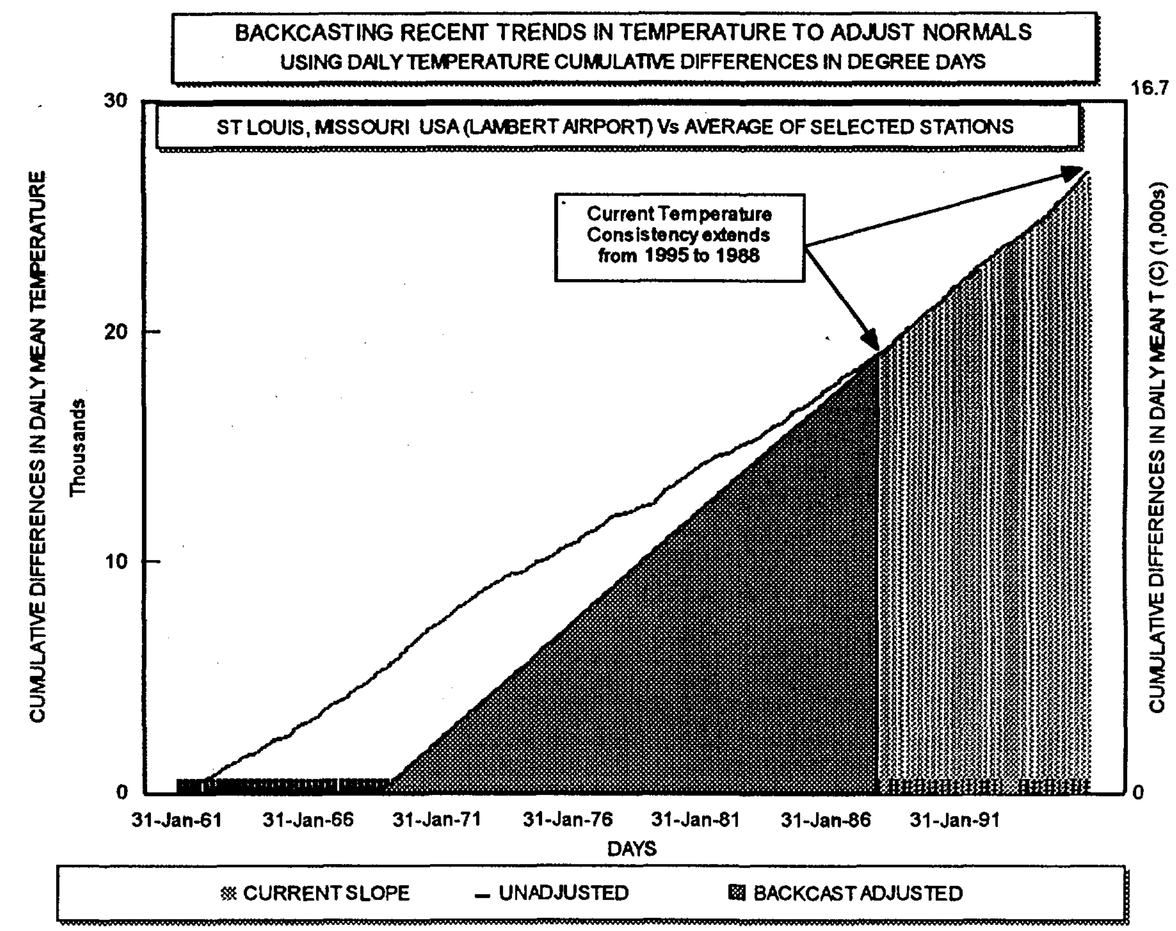


B. 4. Exposure change adjustments after normals have been published

If current temperature readings are no longer consistent with past readings, then the users of the data should determine the existence and magnitude of the disturbance, and apply appropriate adjustments either to the history before the disturbance or to readings after it occurs. Such disturbances may be caused by changes in instrumentation, by new construction in the vicinity of the instruments, by relocation of the instruments, or may simply be unexplained. A tool for finding and correcting disturbances in temperature data series is called double mass analysis (Dutcher and Hubbard, 1994). The analyst will prepare the series of daily mean temperatures for the candidate station for comparison with one or more concurrent temperature series known to be consistent.

Double mass temperature analysis is performed by accumulating the differences between the candidate and comparison temperature series, and then observing the slope of the accumulation. The slope will be constant if the candidate series is also consistent. If the candidate series experiences a disturbance, the slope will change at the point in time where it occurs, thus allowing the disturbance to be isolated. Figure II.6 illustrates this difference for monthly heating degree days for a station where the historical series through 1990 had been adjusted by NOAA. This series was compared to the average of selected Missouri and Illinois weather stations which were regarded as having sufficient and consistent data.

FIGURE II.6



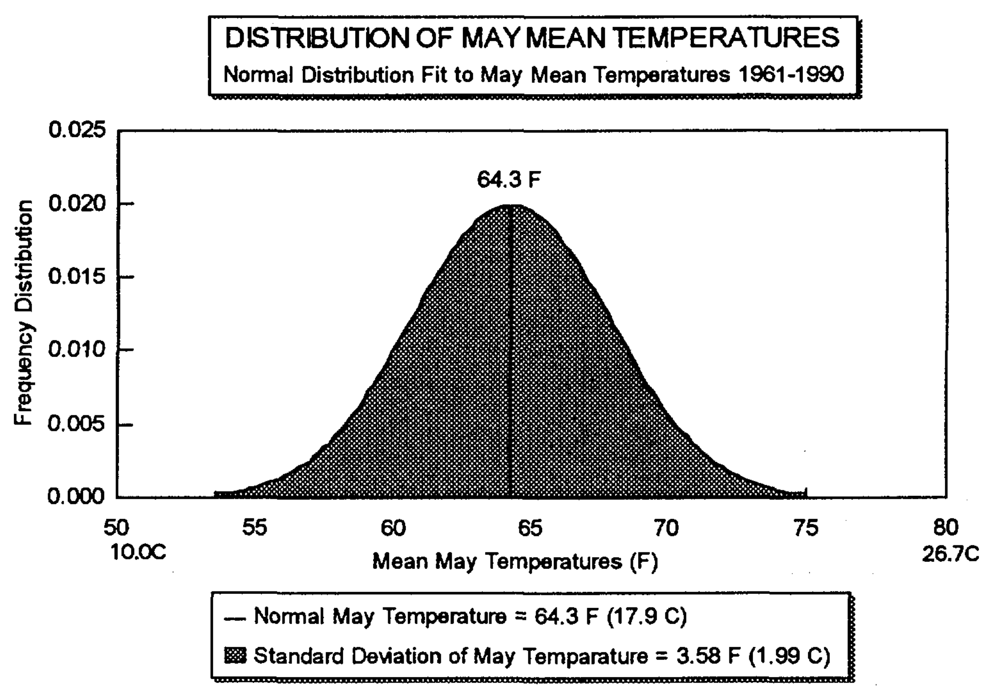
The difference between the slope before the disturbance and the slope afterward has the magnitude of the correction required to make the disturbed series consistent once more. The correction may be applied to all observations before the disturbance. In this case, normals must be recalculated if they are to be consistent with observations subsequent to the disturbance. Alternatively, the negative of the correction may be applied after the disturbance to make the subsequent observation consistent with those preceding the disturbance. In the latter case, if the disturbance occurs during the normals period, the normals must also be recalculated.

B. 5. Monthly heating degree day normals from adjusted daily temperatures

First, the entire temperature series must be consistent. Then, the adjusted temperature variables are calculated and verified for each day in the thirty-year normals period. Finally, heating degree day variables that are functions of daily temperature may be calculated for each day of the thirty years, summed by month, and then averaged for each of the twelve months to arrive at directly calculated monthly normals.

NOAA "Monthly Station Normals" are published for heating degree days (HDD) and cooling degree days (CDD) with base temperatures of 65°F. If desired, the thirty years of daily values may once again be adjusted (fine-tuned) to correspond exactly with NOAA published values for each of the twelve calendar months. A trivial difference will almost always exist because NOAA monthly HDD normals are statistical estimates based on the distribution of average temperatures for each calendar month in question, as compared to averages calculated directly from daily HDD. Figure II.7 illustrates the type of distribution used by NOAA in the calculation of normal heating degree days.

FIGURE II.7

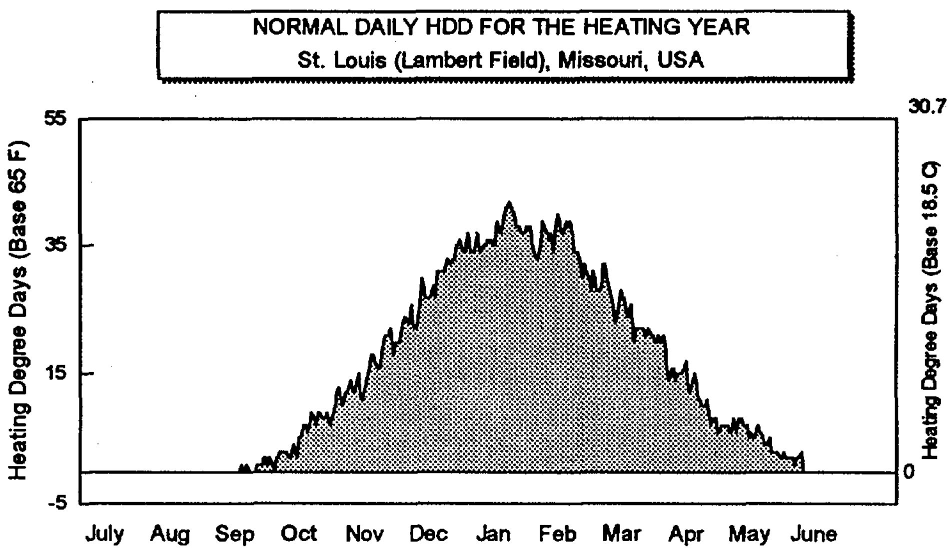


In order to get any HDDs for months having days that require heating as well as cooling (transition months), the estimated distribution around the mean allows NOAA to calculate HDDs for days with mean temperatures below 65 degrees.

B. 6. Daily HDD normals from daily HDD

Daily normal HDD are calculated from the 30 historical years of adjusted daily HDD. First, the 30 June observation is removed from leap years, since HDD do not usually occur in that season of the year. Then, leap year days from February 29th through June 29th are given new dates of March 1st through June 30th. In this way, the month and day of all 365 leap year values correspond to the month and day of values in the non-leap years, and all the leap year HDD are counted. Normal daily HDD may then be calculated over the 30 historical years, for each day, as simple averages. Figure II.8 shows the daily HDDs over an entire year, calculated as the simple average of daily values.

FIGURE II.8



For verification, these daily normal HDD are summed for each month and checked against the previously calculated monthly normals. The annual totals must be the same, as should the totals for the month of January and the months of July through December. However, small but compensating differences will exist between monthly totals for the spring months due to the way leap years were addressed. Daily values for these months may be adjusted to bring the monthly values in line. Daily values may also be adjusted in such a way that all values are whole degree days, by carrying positive or negative fractional remainders forward to the next non-zero value.

C. Billing Cycle Weather

C. 1. Billing cycle actual weather

Place and time information are found in the gas company's billing data, providing the means of matching actual weather information to usage. The billing information usually includes customer class, service area, billing month, cycle number and meter read date, volumes, and number of customers. Meter reading dates are usually assigned on a 21-cycle rotation, which corresponds to the number of working days in most months. For example, the billing month for cycle 1 of the January billing month may extend from the previous 29 November initial meter read date till the current 2 January read date, while the billing year for cycle 21 may extend from the previous 29 December of the first year to the current 31 January meter read date. Figure II.9 shows the reading cycle dates for 21 billing cycles that occurred during a recent test year.

FIGURE II.9

ACTUAL METER READING CYCLE DATES IN THE MONTHS

CYCLE	MONTHS IN A TEST YEAR (Oct - Sep)												
No.	-1	1	2	3	4	5	6	7	8	9	10	11	12
1	11-Sep	09-Oct	06-Nov	09-Dec	09-Jan	10-Feb	11-Mar	09-Apr	08-May	09-Jun	09-Jul	06-Aug	08-Sep
2	12-Sep	10-Oct	07-Nov	10-Dec	10-Jan	11-Feb	12-Mar	10-Apr	11-May	10-Jun	10-Jul	07-Aug	09-Sep
3	13-Sep	11-Oct	08-Nov	11-Dec	13-Jan	12-Feb	13-Mar	13-Apr	12-May	11-Jun	13-Jul	10-Aug	10-Sep
4	16-Sep	14-Oct	12-Nov	12-Dec	14-Jan	13-Feb	16-Mar	14-Apr	13-May	12-Jun	14-Jul	11-Aug	11-Sep
5	17-Sep	15-Oct	13-Nov	13-Dec	15-Jan	14-Feb	17-Mar	15-Apr	14-May	15-Jun	15-Jul	12-Aug	14-Sep
6	18-Sep	16-Oct	14-Nov	16-Dec	16-Jan	18-Feb	18-Mar	16-Apr	15-May	16-Jun	16-Jul	13-Aug	15-Sep
7	19-Sep	17-Oct	15-Nov	17-Dec	17-Jan	19-Feb	19-Mar	20-Apr	18-May	17-Jun	17-Jul	14-Aug	16-Sep
8	20-Sep	18-Oct	18-Nov	18-Dec	20-Jan	20-Feb	20-Mar	21-Apr	19-May	18-Jun	20-Jul	17-Aug	17-Sep
9	23-Sep	21-Oct	19-Nov	19-Dec	21-Jan	21-Feb	23-Mar	22-Apr	20-May	19-Jun	21-Jul	18-Aug	18-Sep
10	24-Sep	22-Oct	20-Nov	20-Dec	22-Jan	24-Feb	24-Mar	23-Apr	21-May	22-Jun	22-Jul	19-Aug	21-Sep
11	25-Sep	23-Oct	21-Nov	23-Dec	23-Jan	25-Feb	25-Mar	24-Apr	22-May	23-Jun	23-Jul	20-Aug	22-Sep
12	26-Sep	24-Oct	22-Nov	26-Dec	24-Jan	26-Feb	26-Mar	27-Apr	26-May	24-Jun	24-Jul	21-Aug	23-Sep
13	27-Sep	25-Oct	25-Nov	27-Dec	27-Jan	27-Feb	27-Mar	28-Apr	27-May	25-Jun	27-Jul	24-Aug	24-Sep
14	30-Aug	30-Sep	28-Oct	26-Nov	30-Dec	30-Jan	28-Feb	30-Mar	29-Apr	28-May	29-Jun	28-Jul	26-Aug
15	03-Sep	01-Oct	29-Oct	27-Nov	31-Dec	31-Jan	02-Mar	31-Mar	30-Apr	29-May	30-Jun	29-Jul	27-Aug
16	04-Sep	02-Oct	30-Oct	02-Dec	02-Jan	03-Feb	03-Mar	01-Apr	01-May	01-Jun	01-Jul	30-Jul	31-Aug
17	05-Sep	03-Oct	31-Oct	03-Dec	03-Jan	04-Feb	04-Mar	02-Apr	04-May	02-Jun	02-Jul	31-Jul	01-Sep
18	06-Sep	04-Oct	01-Nov	04-Dec	06-Jan	05-Feb	06-Mar	06-Apr	05-May	04-Jun	06-Jul	03-Aug	02-Sep
19	09-Sep	07-Oct	04-Nov	05-Dec	07-Jan	06-Feb	09-Mar	07-Apr	06-May	05-Jun	07-Jul	04-Aug	03-Sep
20	10-Sep	08-Oct	05-Nov	06-Dec	08-Jan	07-Feb	10-Mar	08-Apr	07-May	08-Jun	08-Jul	05-Aug	04-Sep
21	20-Sep	18-Oct	20-Nov	17-Dec	20-Jan	20-Feb	20-Mar	20-Apr	20-May	19-Jun	20-Jul	20-Aug	18-Sep

In order to match actual weather to usage and customers by both place (service area) and time (days in the billing cycle), it is necessary to aggregate actual degree day values for each day of each billing cycle for each month of the test year (c.f., figure II.10).

FIGURE II.10

ACTUAL HEATING DEGREE DAYS FOR THE METER READING CYCLES IN THE MONTHS

(HDD Base 65 F)

CYCLE	MONTHS IN A TEST YEAR (Oct - Sep)												CYCLE
No.	1	2	3	4	5	6	7	8	9	10	11	12	Year
1	158	419	887	874	950	586	576	260	94	5	0	12	4,818
2	161	462	862	889	946	586	544	261	93	5	0	12	4,818
3	161	518	826	949	900	581	555	220	93	5	0	16	4,822
4	180	628	706	980	895	615	487	220	93	5	0	20	4,827
5	199	622	720	1,001	879	594	476	220	93	5	0	20	4,827
6	207	624	802	952	932	521	460	223	84	5	2	18	4,827
7	187	638	815	957	930	518	465	191	84	5	4	17	4,807
8	178	673	811	1,021	845	521	471	163	84	5	11	10	4,791
9	193	639	837	1,003	844	571	420	143	84	5	11	10	4,756
10	189	655	854	988	879	534	414	132	89	0	11	19	4,761
11	182	674	922	929	882	517	413	118	89	0	11	25	4,760
12	176	698	987	878	872	506	447	104	52	0	11	34	4,763
13	179	817	889	937	807	511	432	109	35	0	11	42	4,767
14	95	183	840	949	928	739	563	369	122	20	0	11	4,817
15	95	194	853	952	923	744	555	350	130	12	0	11	4,817
16	95	223	952	880	917	697	580	322	135	7	0	12	4,818
17	95	249	971	863	908	682	607	295	132	7	0	12	4,818
18	105	268	991	899	854	661	649	251	124	7	0	12	4,818
19	158	351	885	888	865	653	633	260	108	7	0	12	4,818
20	158	384	878	883	877	659	604	262	99	5	0	12	4,818
21	178	704	741	1,060	845	521	443	191	84	5	11	10	4,791
AVG	158	506	858	939	889	596	514	222	95	6	4	16	4,803

C. 2. Billing cycle normal weather

It is also necessary to have normal degree day values by place and time. Normals must therefore be calculated for a station corresponding to the service area and for each day of the "normal" billing cycle year, where each of the 21 cycles would be exposed to exactly 365 days of "normal" weather. Figures II.11 and II.12 show billing cycle read dates and HDDs for the same test year used in figures II.9 and II.10.

FIGURE II.11

NORMAL METER READING CYCLE DATES IN THE MONTHS
(All Cycles Adjusted to 365 Day Year)

CYCLE	MONTHS IN A TEST YEAR (Oct - Sep)												
No.	-1	1	2	3	4	5	6	7	8	9	10	11	12
1	09-Sep	09-Oct	06-Nov	09-Dec	09-Jan	10-Feb	12-Mar	10-Apr	09-May	10-Jun	10-Jul	07-Aug	09-Sep
2	10-Sep	10-Oct	07-Nov	10-Dec	10-Jan	11-Feb	13-Mar	11-Apr	12-May	11-Jun	11-Jul	08-Aug	10-Sep
3	11-Sep	11-Oct	08-Nov	11-Dec	13-Jan	12-Feb	14-Mar	14-Apr	13-May	12-Jun	14-Jul	11-Aug	11-Sep
4	13-Sep	14-Oct	12-Nov	12-Dec	14-Jan	13-Feb	17-Mar	15-Apr	14-May	13-Jun	15-Jul	12-Aug	13-Sep
5	15-Sep	15-Oct	13-Nov	13-Dec	15-Jan	14-Feb	18-Mar	16-Apr	15-May	16-Jun	16-Jul	13-Aug	15-Sep
6	16-Sep	16-Oct	14-Nov	16-Dec	16-Jan	18-Feb	19-Mar	17-Apr	16-May	17-Jun	17-Jul	14-Aug	16-Sep
7	17-Sep	17-Oct	15-Nov	17-Dec	17-Jan	19-Feb	20-Mar	21-Apr	19-May	18-Jun	18-Jul	15-Aug	17-Sep
8	18-Sep	18-Oct	18-Nov	18-Dec	20-Jan	20-Feb	21-Mar	22-Apr	20-May	19-Jun	21-Jul	18-Aug	18-Sep
9	20-Sep	21-Oct	19-Nov	19-Dec	21-Jan	21-Feb	24-Mar	23-Apr	21-May	20-Jun	22-Jul	19-Aug	20-Sep
10	22-Sep	22-Oct	20-Nov	20-Dec	22-Jan	24-Feb	25-Mar	24-Apr	22-May	23-Jun	23-Jul	20-Aug	22-Sep
11	23-Sep	23-Oct	21-Nov	23-Dec	23-Jan	25-Feb	26-Mar	25-Apr	23-May	24-Jun	24-Jul	21-Aug	23-Sep
12	24-Sep	24-Oct	22-Nov	26-Dec	24-Jan	26-Feb	27-Mar	28-Apr	27-May	25-Jun	25-Jul	22-Aug	24-Sep
13	25-Sep	25-Oct	25-Nov	27-Dec	27-Jan	27-Feb	28-Mar	29-Apr	28-May	26-Jun	28-Jul	25-Aug	25-Sep
14	30-Aug	30-Sep	28-Oct	26-Nov	30-Dec	30-Jan	28-Feb	31-Mar	30-Apr	29-May	30-Jun	29-Jul	30-Aug
15	31-Aug	01-Oct	29-Oct	27-Nov	31-Dec	31-Jan	03-Mar	01-Apr	01-May	30-May	01-Jul	30-Jul	31-Aug
16	01-Sep	02-Oct	30-Oct	02-Dec	02-Jan	03-Feb	04-Mar	02-Apr	02-May	02-Jun	02-Jul	31-Jul	01-Sep
17	02-Sep	03-Oct	31-Oct	03-Dec	03-Jan	04-Feb	05-Mar	03-Apr	05-May	03-Jun	03-Jul	01-Aug	02-Sep
18	04-Sep	04-Oct	01-Nov	04-Dec	06-Jan	05-Feb	07-Mar	07-Apr	06-May	05-Jun	07-Jul	04-Aug	04-Sep
19	06-Sep	07-Oct	04-Nov	05-Dec	07-Jan	06-Feb	10-Mar	08-Apr	07-May	06-Jun	08-Jul	05-Aug	06-Sep
20	07-Sep	08-Oct	05-Nov	06-Dec	08-Jan	07-Feb	11-Mar	09-Apr	08-May	09-Jun	09-Jul	06-Aug	07-Sep
21	19-Sep	18-Oct	20-Nov	17-Dec	20-Jan	20-Feb	21-Mar	21-Apr	21-May	20-Jun	21-Jul	21-Aug	19-Sep

FIGURE II.12

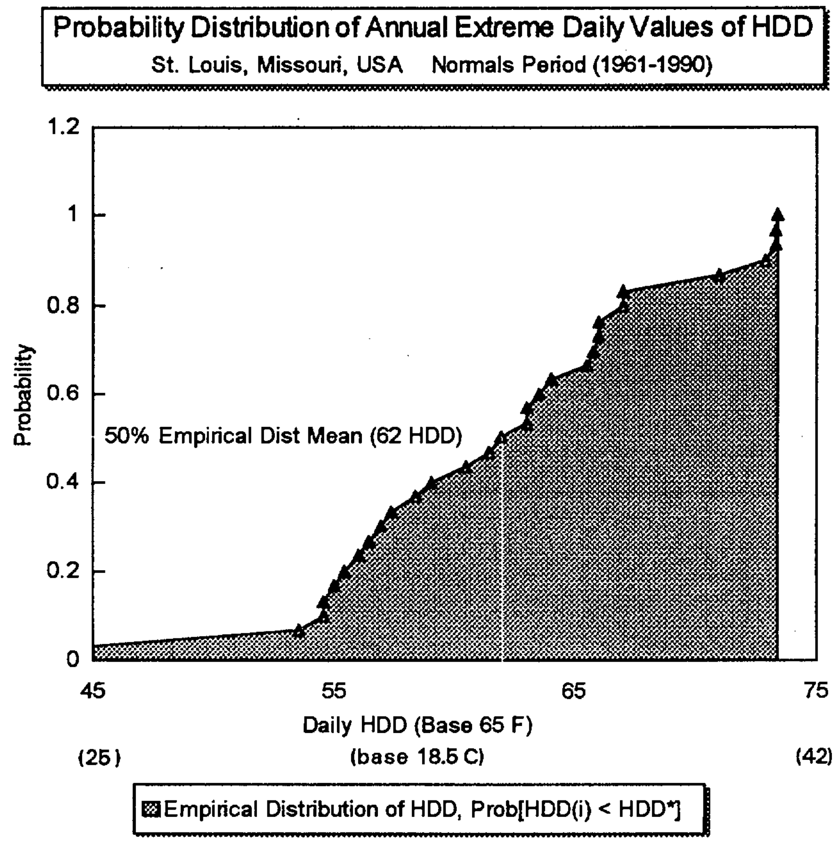
NORMAL HEATING DEGREE DAYS FOR THE METER READING CYCLES IN THE MONTHS
(HDD Base 65 F)

CYCLE	MONTHS IN A TEST YEAR (Oct - Sep)												CYCLE
No.	1	2	3	4	5	6	7	8	9	10	11	12	Year
1	88	302	811	1,159	1,237	861	520	215	72	1	0	7	5,273
2	95	310	829	1,168	1,230	847	509	220	56	1	0	8	5,273
3	102	319	847	1,257	1,141	836	517	194	51	1	1	7	5,273
4	120	378	802	1,262	1,135	866	466	185	48	1	2	8	5,273
5	123	393	816	1,267	1,127	855	454	180	45	0	2	11	5,273
6	129	407	903	1,198	1,215	749	441	175	42	0	2	12	5,273
7	136	417	921	1,200	1,207	740	448	155	34	0	2	13	5,273
8	144	467	897	1,280	1,121	732	433	151	31	0	2	15	5,273
9	175	457	909	1,283	1,112	764	379	147	28	0	2	17	5,273
10	179	472	923	1,281	1,171	690	366	144	24	0	2	21	5,273
11	188	486	1,008	1,207	1,166	679	352	141	21	0	2	23	5,273
12	197	497	1,100	1,127	1,161	667	354	128	14	0	2	26	5,273
13	207	558	1,062	1,206	1,073	655	345	124	12	0	3	28	5,273
14	40	228	546	1,149	1,213	984	676	304	119	10	0	4	5,273
15	44	235	561	1,161	1,212	1,020	617	297	113	8	0	5	5,273
16	48	241	698	1,092	1,251	929	607	290	109	3	0	5	5,273
17	53	245	719	1,101	1,252	917	595	292	92	2	0	5	5,273
18	57	252	737	1,187	1,173	935	600	237	87	2	0	6	5,273
19	75	279	719	1,201	1,171	969	537	230	83	2	0	7	5,273
20	81	291	731	1,214	1,168	951	529	221	79	1	0	7	5,273
21	143	513	815	1,316	1,121	732	426	161	28	0	2	16	5,273
AVG	115	369	826	1,206	1,174	828	484	200	57	2	1	12	5,273

D. Peak Day Weather

To calculate HDDs for the average coldest day, the highest HDD value from each of the 30 years of adjusted historical data is selected and the average of these 30 values is calculated. This average is the normal annual peak day HDD. For the St. Louis Lambert airport, 30 values and their mean are shown in figure II.13.

FIGURE II.13

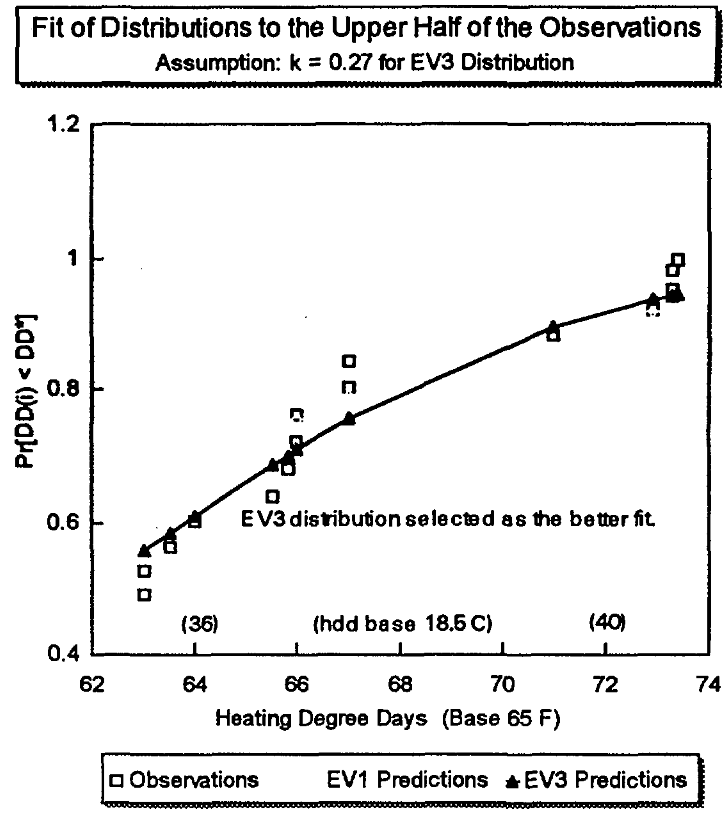


Peak day usage is an important concept for designing the gas supply system. In this context, reliability of gas supplies requires an analysis of the peak day weather that goes beyond the simple concept of the average of the coldest days over the past thirty years. NOAA has developed an extreme value analysis for calculating expected HDD maximum values which would be expected to recur over intervals such as 5, 10, or 20 years. In this analysis, a highly non-linear probability distribution is fit to the HDDs for the 30 years of coldest days. Based on this fitted probability distribution, an estimate of the probability that a specific extreme value for HDD will be exceeded can be made.

Based on experience with fitting extreme weather, NOAA's software prespecifies some of the parameters for each of the three different types of probability distribution functions, and uses the data to get the best fit for the remaining parameters. The resulting three candidate distributions are then compared and the one with the best fit is chosen as the appropriate probability distribution.

Figure II.13 shows two of the distributions compared to the actual data. It is important to note that it is only the values of the data above the mean (extreme values) that are of interest. Therefore the graph in Figure II.14 only shows the comparison for observations above the mean.

FIGURE II.14

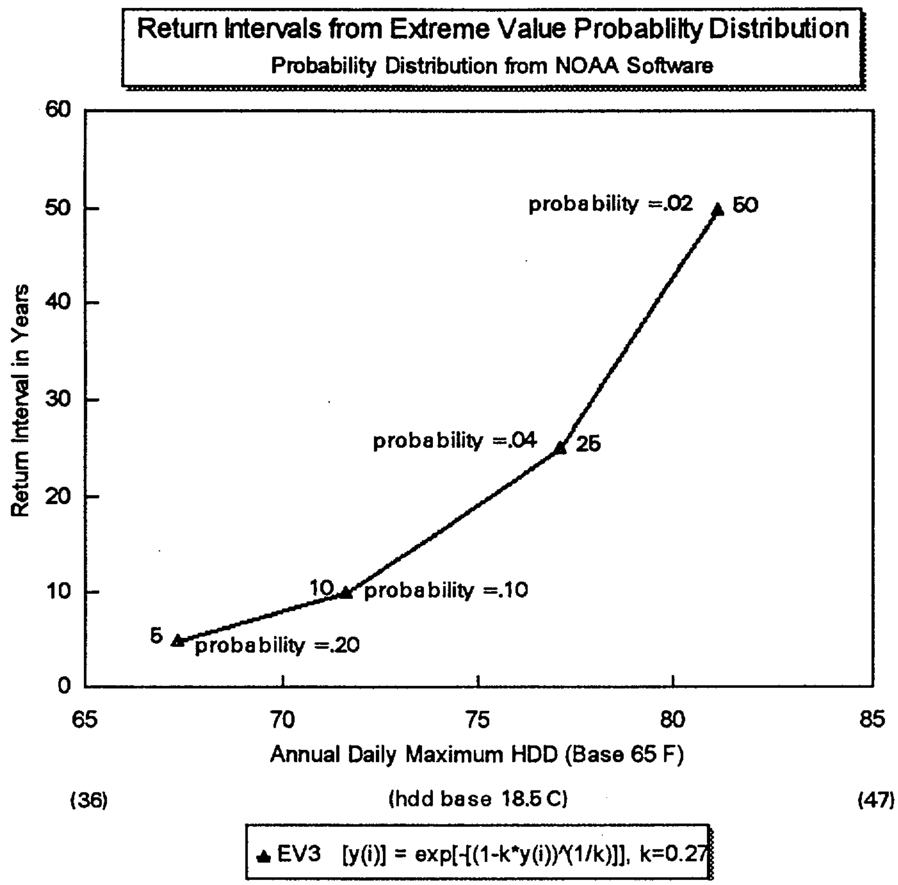


Once the probability function is chosen, the user can specify any extreme value for HDDs and the software will give the probability that the specified level will be exceeded. If the probability is say 5%, then the occurrence of having this extreme value exceeded will be once in twenty years; i.e.,

$$\frac{1}{20} = 0.05$$

In this example, the 20 years is called the "return interval." Figure II.15 shows the relationship of extreme values for HDD and the return intervals for the estimated distribution using NOAA's software (ANYEXR; c.f., attachment A).

FIGURE II.15



Return Intervals from Extreme Value Probablity Distribution
Probability Distribution from NOAA Software
Return Interval in Years
60
50
40
30
20
10
0
probability =.02
50
probability =.04
25
10
probability =.10
5
probability =.20
65
70
75
80
85
Annual Daily Maximum HDD (Base 65 F)
(36)
(hdd base 18.5 C)
(47)
EV3 [y(i)] = exp[-[(1-k*y(i))^(1/k)]], k=0.27

III. THE ANALYSIS OF USAGE

A. Organization of Usage Data

It would be ideal to have daily measures of both gas usage and weather. However, gas usage is only readily available from the monthly meter readings taken by the utility company. The meter readings are staggered throughout a month, with some billing cycle reading dates going from the middle of one month to the middle of the next month. Calendar month weather measures would be inappropriate for those billing cycles; however daily measures can be added over the known dates (days) between meter readings for each billing cycle. Figure III.1 shows the billing cycle read dates, days, number of customers and their usage for the billing months of April and May. Dividing the usage by the number of days between meter readings puts the usage on a per day basis.

FIGURE III.1

MONTHLY USAGE INFORMATION FROM METER READINGS
FOR MONTHLY BILLING

RESIDENTIAL CLASS, DIVISION B
CITY NATURAL GAS COMPANY, MISSOURI, USA

BILLING MONTH IN TEST YEAR	READ CYCLE NUMBER	STARTING READ DATE (A)	ENDING READ DATE (B)	DAYS (C=B-A)	No. of BILLS (CUST) (D)	USAGE VOLUMES (MCF) (E)	USAGE Per DAY (MCF) (F=E/C)
April	1	01-Mar	30-Mar	29	2,589	23,511	811
April	2	02-Mar	31-Mar	29	2,650	24,543	846
April	3	03-Mar	03-Apr	31	3,165	29,825	962
April	4	06-Mar	04-Apr	29	2,863	21,258	733
April	5	07-Mar	05-Apr	29	3,180	23,294	803
April	6	08-Mar	06-Apr	29	3,075	21,589	744
April	7	09-Mar	07-Apr	29	2,914	18,490	638
April	8	10-Mar	10-Apr	31	2,858	18,281	590
April	9	13-Mar	11-Apr	29	3,267	18,175	627
April	10	14-Mar	12-Apr	29	3,080	19,410	669
April	11	15-Mar	13-Apr	29	3,202	19,605	676
April	12	16-Mar	17-Apr	32	2,888	17,300	541
April	13	17-Mar	18-Apr	32	2,940	17,220	538
April	14	20-Mar	19-Apr	30	3,088	17,576	586
April	15	21-Mar	20-Apr	30	3,409	19,824	661
April	16	22-Mar	21-Apr	30	3,282	21,480	716
April	17	23-Mar	24-Apr	32	2,839	17,426	545
April	18	24-Mar	25-Apr	32	2,746	17,790	556
April	19	27-Mar	26-Apr	30	2,607	17,530	584
April	20	28-Mar	27-Apr	30	2,689	20,162	672
April	21	29-Mar	28-Apr	30	2,532	16,565	552
TOTALS for APRIL					61,863		14,050
May	1	30-Mar	01-May	32	2,564	16,749	523
May	2	31-Mar	02-May	32	2,618	18,259	571
May	3	03-Apr	03-May	30	3,142	20,084	669
May	4	04-Apr	04-May	30	2,830	16,737	558
May	5	05-Apr	05-May	30	3,176	19,337	645
May	6	06-Apr	08-May	32	3,052	20,142	629
May	7	07-Apr	09-May	32	2,900	19,060	596
May	8	10-Apr	10-May	30	2,844	19,134	638
May	9	11-Apr	11-May	30	3,252	20,542	685
May	10	12-Apr	12-May	30	3,080	20,066	669
May	11	13-Apr	15-May	32	3,210	19,527	610
May	12	17-Apr	16-May	29	2,890	16,694	576
May	13	18-Apr	17-May	29	2,948	15,533	536
May	14	19-Apr	18-May	29	3,080	15,466	533
May	15	20-Apr	19-May	29	3,406	17,654	609
May	16	21-Apr	22-May	31	3,299	18,121	585
May	17	24-Apr	23-May	29	2,827	12,827	442
May	18	25-Apr	24-May	29	2,724	11,715	404
May	19	26-Apr	25-May	29	2,586	11,656	402
May	20	27-Apr	26-May	29	2,654	13,470	464
May	21	28-Apr	30-May	32	2,503	11,531	360
TOTALS for MAY					61,585		11,704

The daily weather is also accumulated for each day between meter readings. Figure III.2 shows the total heating degree days (HDD) for each billing cycle in the billing months of April and May. The HDD are divided by the number of days between meter readings to put the weather on a per day basis (HDD/D). For purposes of customer weighting, the HDD/D are then multiplied by the number of customers in each billing cycle.

FIGURE III.2

HEATING DEGREE DAYS CONSISTENT WITH READ CYCLES AND MONTHLY BILLING CYCLES

RESIDENTIAL CLASS, DIVISION B
CITY NATURAL GAS COMPANY, MISSOURI, USA

BILLING MONTH IN TEST YEAR	READ CYCLE NUMBER	DAYS (A)	No. of BILLS (CUST) (B)	HDD IN THE CYCLE (C)	HDD Per DAY Per CUST (D=C/A)	HDDxCUST (E=BxD)
April	1	29	2,589	385.50	13.29	34,416
April	2	29	2,650	360.00	12.41	32,897
April	3	31	3,165	347.50	11.21	35,479
April	4	29	2,863	294.50	10.16	29,074
April	5	29	3,180	277.00	9.55	30,374
April	6	29	3,075	240.00	8.28	25,448
April	7	29	2,914	215.00	7.41	21,604
April	8	31	2,858	205.50	6.63	18,946
April	9	29	3,267	215.50	7.43	24,277
April	10	29	3,080	225.50	7.78	23,950
April	11	29	3,202	234.50	8.09	25,892
April	12	32	2,888	234.50	7.33	21,164
April	13	32	2,940	240.50	7.52	22,096
April	14	30	3,088	244.50	8.15	25,167
April	15	30	3,409	247.50	8.25	28,124
April	16	30	3,282	259.50	8.65	28,389
April	17	32	2,839	305.50	9.55	27,104
April	18	32	2,746	309.00	9.66	26,516
April	19	30	2,607	299.50	9.98	26,027
April	20	30	2,689	294.50	9.82	26,397
April	21	30	2,532	287.00	9.57	24,223
TOTAL for APRIL			61,863			557,563
May	1	32	2,564	299.50	9.36	23,997
May	2	32	2,618	297.50	9.30	24,339
May	3	30	3,142	283.50	9.45	29,692
May	4	30	2,830	270.00	9.00	25,470
May	5	30	3,176	269.50	8.98	28,531
May	6	32	3,052	269.00	8.41	25,656
May	7	32	2,900	269.00	8.41	24,378
May	8	30	2,844	273.50	9.12	25,928
May	9	30	3,252	261.50	8.72	28,347
May	10	30	3,080	247.50	8.25	25,410
May	11	32	3,210	235.50	7.36	23,624
May	12	29	2,890	231.00	7.97	23,020
May	13	29	2,948	222.00	7.66	22,567
May	14	29	3,080	215.50	7.43	22,888
May	15	29	3,406	217.00	7.48	25,486
May	16	31	3,299	206.50	6.66	21,976
May	17	29	2,827	155.50	5.36	15,159
May	18	29	2,724	145.00	5.00	13,620
May	19	29	2,586	145.00	5.00	12,930
May	20	29	2,654	129.50	4.47	11,851
May	21	32	2,503	122.00	3.81	9,543
TOTALS for MAY			61,585			464,412

In summary, customer rate class, aggregated daily weather measures, gas usage, number of customers, and number of days between meter readings are available for each billing cycle.

B. Aggregation of Usage and HDD Data

An average gas usage response to a customer weighted average weather measure is calculated by rate class. Since each billing cycle has different days between meter readings, it is important that HDD and gas usage be put on a per day basis. While this is a straightforward calculation for each billing cycle, in order to aggregate over a billing month, these per day measures of usage and HDD need to be averaged. Customer weights (% of customers) for each billing cycle are used to average all the billing cycles in each billing month. Since the usage is for all customers in each billing cycle, the usage per day is equivalent to the usage per customer per day multiplied by the number of customers (c.f., figure III.1). However, the HDD/D need to be multiplied by the number of customers in each billing cycle to put the weather on an equivalent basis (c.f., figure III.2). Then the usage per customer per day (U/C/D) and HDD/D which have been multiplied by customers in each billing cycle and summed over all billing cycles is divided by the total customers in each billing month to arrive at the customer weighted U/C/D and HDD/D for each of the twelve billing months. The customer weighted U/C/D by billing month is the dependent variable in the regressions while the customer weighted HDD/D is the independent variable. (See Figure III.3)

FIGURE III.3

AVERAGE USAGE AND HEATING DEGREE DAY INFORMATION
CONSISTENT WITH MONTHLY READ CYCLES
RESIDENTIAL CLASS, DIVISION B
CITY NATURAL GAS COMPANY, MISSOURI, USA

BILLING MONTH IN TEST YEAR	No. of BILLS (CUST) (Fig III.1) (A)	USAGE Per DAY (MCF) (Fig III.1) (B)	HDDxCUST (Fig III.2) (C)	USAGE Per CUST Per DAY (MCF) (D=B/A)	HDD Per CUST Per DAY (Base 65 F) (E=C/A)
Apr	61,863	14,050	557,563	0.23	9.01
May	61,585	11,704	464,412	0.19	7.54
Jun	61,294	5,663	58,547	0.09	0.96
Jul	61,137	4,077	0	0.07	0.00
Aug	60,862	3,590	0	0.06	0.00
Sep	60,962	3,910	33,485	0.06	0.55
Oct	61,527	7,231	245,000	0.12	3.98
Nov	62,218	14,718	696,513	0.24	11.19
Dec	62,473	25,084	1,350,543	0.40	21.62
Jan	62,766	36,042	1,938,574	0.57	30.89
Feb	62,987	37,857	1,935,638	0.60	30.73
Mar	63,128	26,793	1,354,574	0.42	21.46

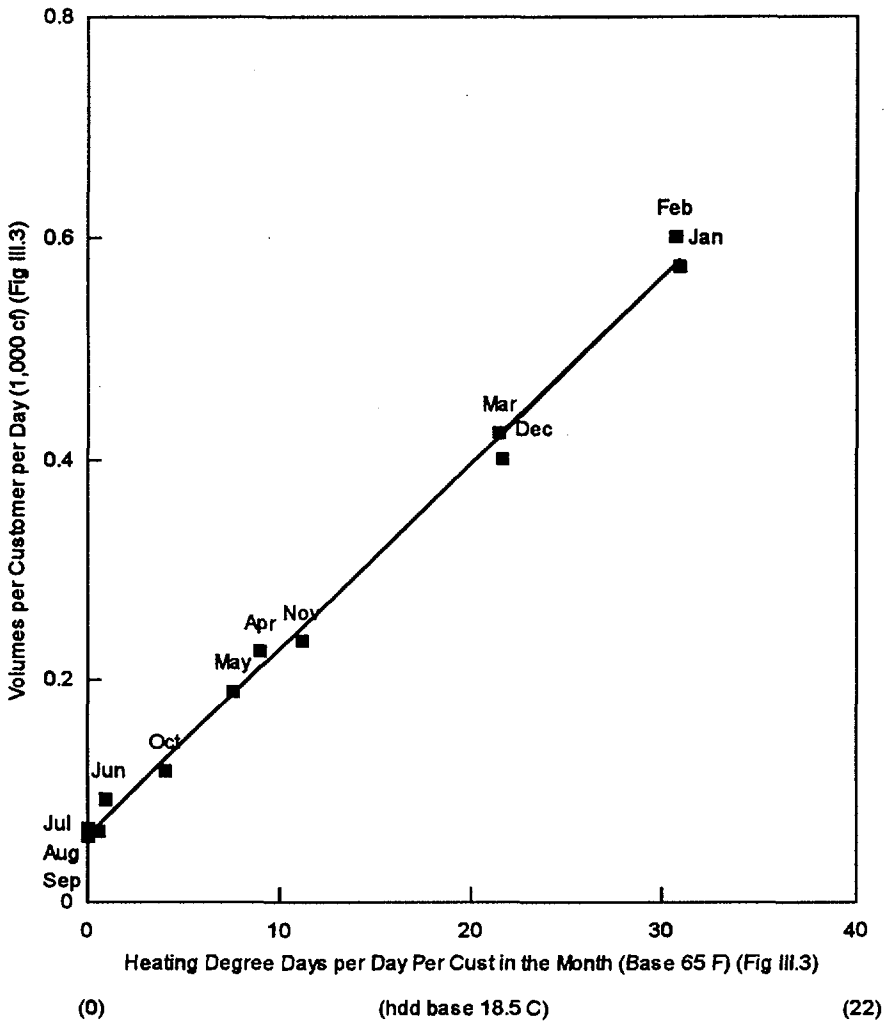
C. Regression Analysis on Usage

The regression model is a linear model as explained in section I. The linear model fits the data very closely. Figure III.4 shows the U/C/D billing month averages plotted against the corresponding HDD/D averages. The best fitting regression line is also shown going through the points.

The usage per customer per day for the months that do not have any heating degree days constitutes the base usage (shown in Figure III.4 as the “Constant” in the Regression Output). The higher the level for HDD/D, the higher is both the actual and predicted U/C/D. The slope of the line (shown in Figure III.4 as the "X Coefficient(s)" in the Regression Output) represents the change in U/C/D for a unit change in HDD/D.

FIGURE III.4

MONTHLY ACTUAL and ESTIMATED USAGE of NATURAL GAS
CONSISTENT with READ CYCLES in the TEST YEAR
RESIDENTIAL CLASS
CITY NATURAL GAS COMPANY, MISSOURI, USA



D. Weather and Days Adjustment to Usage

To adjust each billing cycle's actual usage to normal usage, the difference between normal and actual HDD are calculated for each billing cycle. These differences are multiplied by the estimated slope from the regression and by the number of customers in each billing cycle. The resulting adjustments to usage in each billing cycle are then summed over all the billing cycles for each billing month. (See Figure III.5).

FIGURE III.5

HDD DIFFERENCES to ADJUST the READ CYCLES
from the TEST YEAR to NORMAL

VOLUME (MCF) DIFFERENCES to ADJUST the
READ CYCLES COMPUTED USING the LINEAR MODEL
and NUMBER of BILLS in the CYCLE

RESIDENTIAL CLASS, DIVISION B
CITY NATURAL GAS COMPANY, MISSOURI, USA

	April		May	
CYCLE	Adjustment to Test Year		Adjustment to Test Year	
No.	HDD (65F)	MCF	HDD (65F)	MCF
1	2.16	5,599	-0.68	-1,748
2	2.49	6,601	-0.75	-1,961
3	2.60	8,230	-1.04	-3,252
4	2.94	8,408	-0.96	-2,715
5	3.11	9,901	-1.17	-3,715
6	3.60	11,075	-1.25	-3,801
7	3.84	11,182	-1.31	-3,807
8	3.95	11,280	-1.86	-5,289
9	3.34	10,914	-1.76	-5,719
10	2.95	9,097	-1.59	-4,899
11	2.67	8,542	-1.41	-4,511
12	2.85	8,239	-1.77	-5,107
13	2.01	5,913	-1.67	-4,912
14	1.74	5,379	-1.62	-5,002
15	1.47	5,020	-1.72	-5,847
16	1.09	3,563	-1.56	-5,136
17	0.29	836	-0.93	-2,641
18	-0.57	-1,571	-0.81	-2,201
19	-0.53	-1,382	-0.91	-2,350
20	-0.51	-1,380	-0.72	-1,898
21	-0.49	-1,236	-0.66	-1,643
Totals	41.01	124,209	-26.12	-78,154

Days adjustments are made for billing cycles that do not have 365 days. The HDD adjustment has already taken into account the correct number of days for the weather sensitive component of usage using the slope times the difference between normal and actual weather. However, there needs to be an additional adjustment made for the non-weather sensitive component of usage. This days adjustment is made by multiplying the intercept by the number of days different from 365 and by the number of customers. Days adjustments are assigned to either the first or last month of the test year and are calculated for each billing cycle. Figure III.6 shows these calculations by billing cycles for the first and last month of a given test year.

FIGURE III.6

ADJUSTMENTS to NORMALIZE
NATURAL GAS USAGE for DAYS

DAY DIFFERENCES to ADJUST the READ CYCLES
from the TEST YEAR to NORMAL (365)
FIRST and LAST MONTH ONLY ARE ADJUSTED

VOLUME (MCF) DIFFERENCES to ADJUST the
READ CYCLES COMPUTED USING the LINEAR MODEL
and NUMBER of BILLS in the CYCLE

RESIDENTIAL CLASS, DIVISION B
CITY NATURAL GAS COMPANY, MISSOURI, USA

	April		March	
CYCLE	Adjustment to Test Year		Adjustment to Test Year	
No.	DAYS	MCF	DAYS	MCF
1	0	0	0	0
2	0	0	0	0
3	-1	-193	-1	-193
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	-1	-175	-1	-178
9	0	0	0	0
10	0	0	0	0
11	0	0	0	0
12	0	0	0	0
13	-2	-359	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	-2	-336	0	0
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
Totals	-6	-1,063	-2	-371

E. Peak Day Estimates

Winter peak demands are estimated using the regression coefficients, the average number of winter (December through February) customers, and an average of thirty (1961 - 1990) annual peak heating degree days. Figure III.7 shows the average level for the maximum HDD for each month as well as for the year, where these averages have been calculated over the thirty year normal period. The maximum HDDs are multiplied by the HDD coefficient and added to the intercept to get the peak usage customer per day. This is then multiplied by the number of customers in each month to estimate the peak day usage for each month. Because the peak day for the year can occur in December, January, or February, the average number of customers over those three months are used to estimate normal peak day usage for the year.

FIGURE III.7

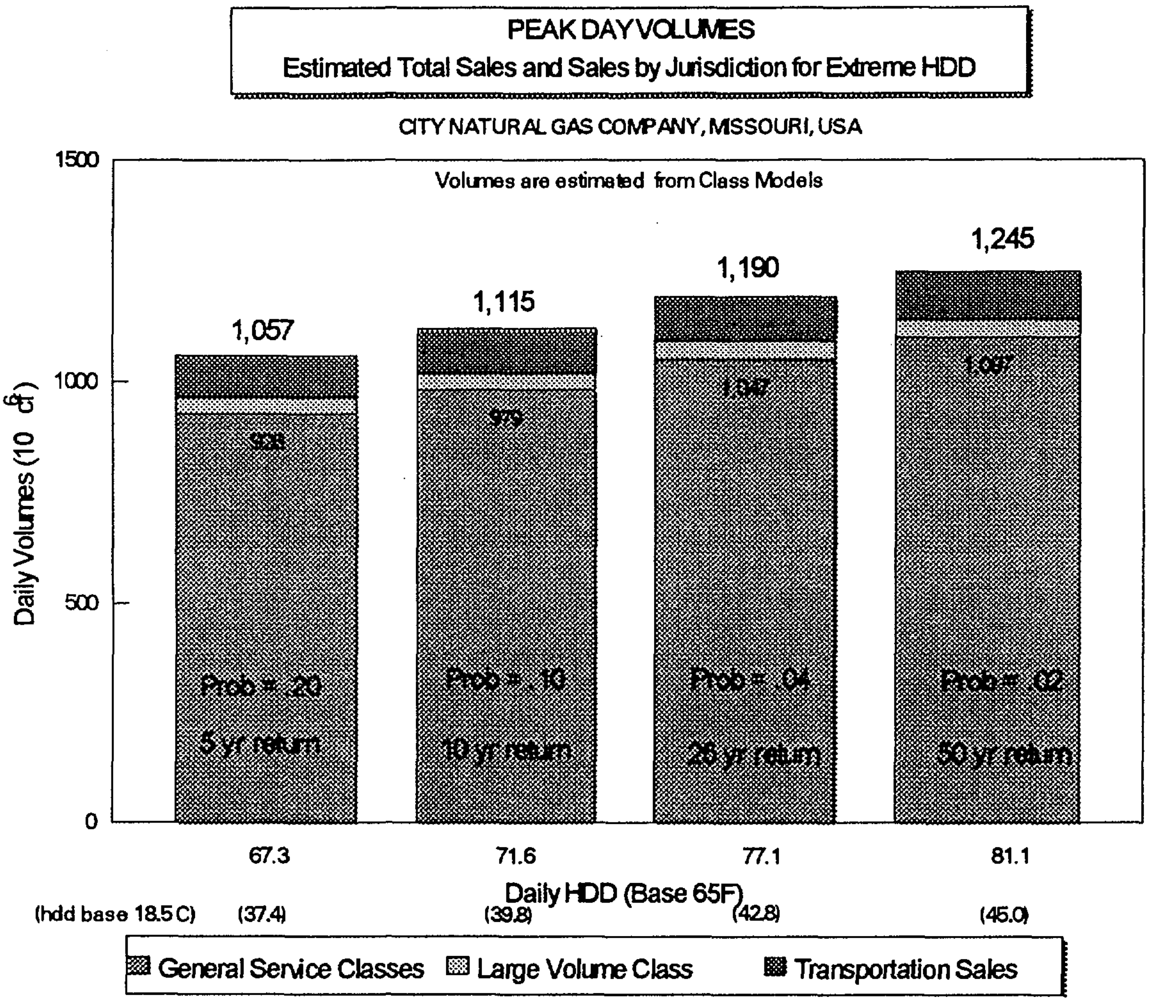
PEAK USAGE AND HEATING DEGREE DAY CONSISTENT WITH MONTHLY READ CYCLES

PEAK DAY VOLUMES COMPUTED USING LINEAR MODEL
RESIDENTIAL CLASS, DIVISION B
CITY NATURAL GAS COMPANY, MISSOURI, USA

BILLING MONTH IN TEST YEAR	No. of BILLS (CUST) (A)	AVG MAX DAY HDD (1961-90) (BASE 65 F) (B)	USAGE Per CUST Per DAY (From Model) (MCF) (C=a+bB)	USAGE Per DAY (MCF) (D=AxC)
Apr	61,863	24.22	0.47	28,995
May	61,585	12.04	0.26	16,240
Jun	61,294	2.32	0.10	6,141
Jul	61,137	0.00	0.06	3,736
Aug	60,862	0.69	0.07	4,422
Sep	60,962	10.70	0.24	14,698
Oct	61,527	21.53	0.42	26,053
Nov	62,218	36.22	0.67	41,729
Dec	62,473	50.88	0.92	57,319
Jan	62,766	55.48	0.99	62,445
Feb	62,987	49.67	0.90	56,502
Mar	63,128	37.80	0.70	44,022
Peak Day	62,742	59.85	1.07	67,030

While normal peak day usage gives a good measure of what to expect on the average, it ignores two important uncertainties that may need to be taken into account for certain applications. The first uncertainty is the probability of having extremely cold weather. Figure III.8 shows the estimated peak day demands for a system for various levels of extreme weather. Peak demands are shown for three major classes: (1) firm general service; (2) large volume; and (3) transportation. The firm general service customers are all estimated at the class level (residential, commercial, industrial) using the methods previously described. The large volume and transportation customer peak demands were estimated individually. Individual regressions were run for temperature sensitive customers. Maximum daily winter usage were estimated from monthly billing records for non-weather sensitive customers (e.g., divide monthly usage by the number of working days in the month). What is important to notice in Figure III.8 is that as the estimated peak demands increase, the probability that the estimates will be exceeded diminishes. Thus there is a direct link between the peak demand estimate and the level of reliability of any capacity (pipeline or gas supply) that has been planned to meet that level of peak demand.

FIGURE III.8



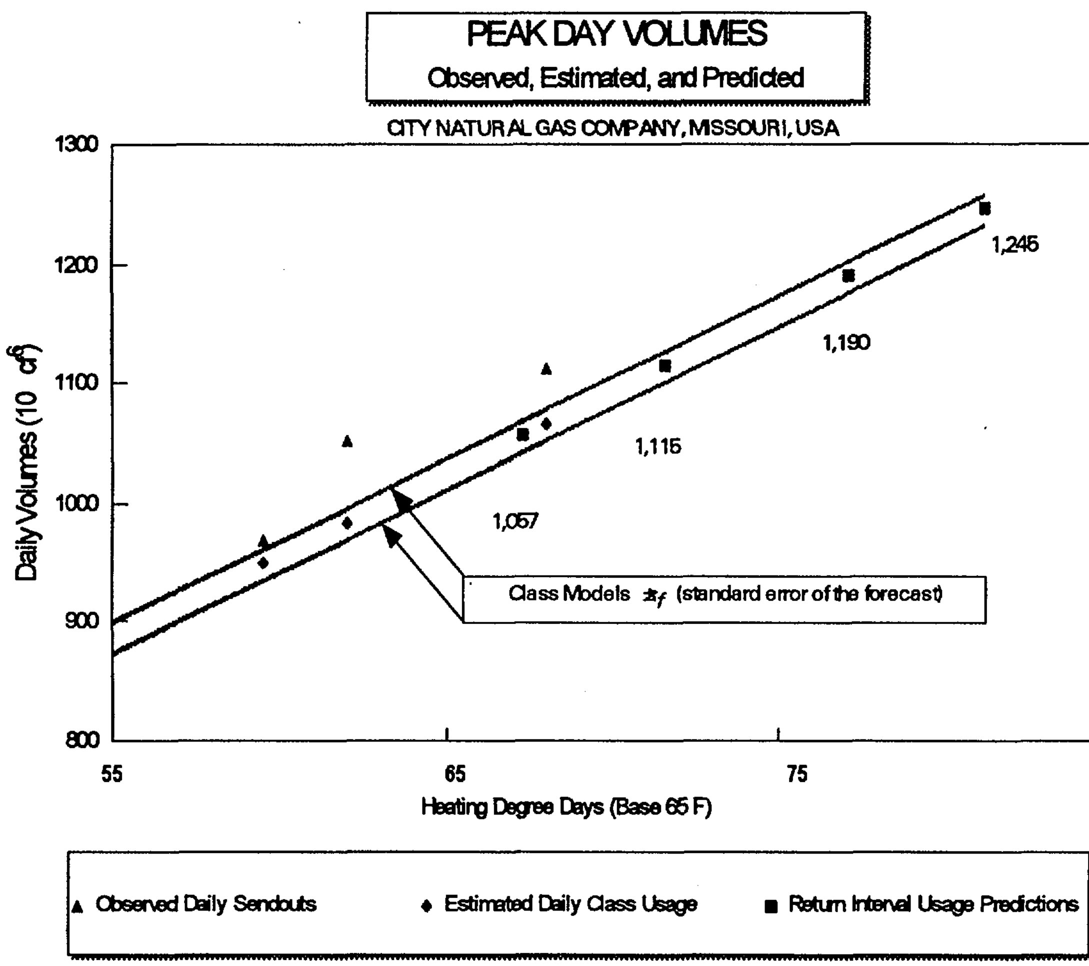
The second type of uncertainty occurs with the estimation process itself. While regressions on gas usage typically have very high correlations with heating degree days, they are not perfect predictors. Thus, the uncertainty of the regression model results also need to be taken into account. This uncertainty is measured through the regression errors, which are the difference between the observed and predicted levels of usage for each month. These errors are squared, summed over all twelve months, divided by ten (degrees of freedom = # observations - # coefficients) and finally, raised to the 1/2 power to give the standard error of the estimate of the dependent variable. Actually, this standard error is only applicable at the mean levels for the independent and dependent variables. If the estimate of the dependent variable moves away from the mean level, then there is an additional uncertainty introduced by the estimates of the regression coefficient from the sample being different from the regression coefficients that would be estimated from the entire population. The formula that accounts for both uncertainties is given by

$$s_f = s_e * [1 + (\frac{1}{n} + \frac{X_f - \overline{X}}{\sum (X_j - \overline{X})^2})]$$

where s_f is the forecasting error, s_e is the standard error, n is the number of observations, X_f is the level of the dependent variable at which the forecast is being made, and the X with a bar over it is the mean value from the observations on the dependent variable. Within the brackets, the term in the parenthesis takes into account the sample size and the difference between the value at which the forecast is being made and the value of the mean from the sample. As the forecast moves to the extremes for the dependent variable, the potential error from estimating the regression coefficients expands.

Figure III.9 combines both types of uncertianties into a single graph. The bands around the estimated usage account for the uncertianty in estimation. The higher levels for peak HDDs account for the extreme value uncertainty.

FIGURE III.9



REFERENCES

Dutcher, A. L. And K. G. Hubbard; *What's wrong with the data?;* **The Tripod** ; High Plains Climate Center. Department of Agriculture Meteorology. University of Nebraska-Lincoln; Fall, 1994; pp 1-4.

Lehman, Richard L.; *A New Software Product for Extreme Value Analysis: System Overview of ANYEXR*; Preprints, **Ninth Conference on Applied Climatology**, American Meteorological Society; Boston, MA; 1995, p.p. 73-76.

ATTACHMENT A

U. S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE

EXPERIMENTAL SOFTWARE PRODUCTS AVAILABLE FROM THE CLIMATE PREDICTION CENTER FOR TRIAL AND COMMENT
NOVEMBER 1995

In response to National Academy of Science committee recommendations for new methods to bridge the gap between climate information and user needs, the Climate Prediction Center(CPC) has developed a series of experimental software systems for personal computers. The present versions of these systems incorporate many features suggested by outside reviewers in the energy, hydrology, and agriculture areas, and are in the final stages of development. The systems are currently available in a limited basis from CPC for trial and comment. If there is sufficient interest, these systems may become available for a fee as separate entities, or as elements on the Climate Dial-Up Service.

a) The QPSO- [QPMO-] Systems (Acronym for Quick Projection of Seasonal [Monthly] Outlook implications in a region near any of 490 U.S. stations. The - is your 3-letter station code). These systems are designed to help you interpret and use the new long range seasonal [monthly] forecasts issued monthly by CPC. QPSO- requires user input of the local forecast values from any one of the thirteen long lead temperature and precipitation forecast maps, available on the Internet. QPMO- requires user input of the local forecast values from the current 0.5-month lead monthly Outlook, also on the Internet. By use of statistical data from disk files for your station and models for the forecast-conditional temperature and precipitation distributions, QPSO- [QPMO-] gives on-screen data on the probability of occurrence of a range of possible a) seasonal [monthly] average temperature outcomes and b) seasonal [monthly] total precipitation outcomes.

References:
R. L. Lehman, 1987, A Model for Decision Making Based on NWS Monthly Temperature Outlooks, *J. Clim. Appl. Meteor.* 26, 263-274.

D. S. Wilks and K. L. Eggleston, 1992, Estimating Monthly and Seasonal Precipitation Distributions Using the 30- and 90-Day Outlooks, *J. Clim.* 5, 252-259.

b) The DAYNORMS System (Acronym for Daily Normals). DAYNORMS uses six harmonics to convert an input sequence of 12 monthly means into 365 daily normals for the year. This product implements for personal computer the analytical work of Epstein, 1991. DAYNORMS lets you create a smooth series of daily normals that are 100% consistent with the input monthly means: the sum or mean of the output daily values for each month is exactly the input value for the month. DAYNORMS provides an internal check on the input/output consistency and allows you to compare on-screen graphs of the input and output data.

Reference: Edward S. Epstein, 1991, On Obtaining Daily Climatological Values from Monthly Means, *J. Climate* 4, 365.

c) The ANYEXR/ANYGAM Systems accept ANY single column ASCII data file of length 11 < L < 3000, and perform an on-screen graphical extreme value (EV) and gamma function analysis of the data. The input is the filespec of any such file of your choice. ANYEXR fits the data to two EV distributions, compares the results graphically, and uses a modified Chi-square statistic to measure the consistency of the data, weighted by the uncertainty implicit in its control bands, with the models. The models are rejected and further analysis halted if the statistical probability that the data are consistent with the best model is below 0.4. Otherwise, ANYEXR makes use of the model best fitting the data in the tail of interest to project the expected extreme values in the next 10 to 100+ years, and the mean wait times for the first occurrence of future extreme values you select. Sample files with 120 years of monthly precipitation data for Washington, DC (PRDC01=January, etc.) are provided for trial with ANYEXR. This product is designed to assist in your long range planning decisions involving any weather or climate variable. ANYEXR has been approved by the California Public Service Commission for use by gas distribution companies in defining return times for extreme events in 'peak day' studies.

ANYGAM provides statistics on the input data and fits the data to a standardized gamma distribution, estimating the single parameter by use of maximum likelihood analysis. Gamma model P(α,x) and dP/dx(α,x) results for a range of x values, including the data extremes, is provided for comparison with ANYEXR data.

Both systems have a referenced on-screen introduction to the theory and methods used, and may be tested with the external file GDEV041K, composed of 1000 random variates from a gamma distribution with mean and variance equal to 4.

Reference: R.L. Lehman, 1995, A New Software Product for Extreme Value Analysis: System Overview of ANYEXR, p.73-76 in *Preprints, Ninth Conference on Applied Climatology*, American Meteorological Society, Boston, MA.

d) The WESIM System (Acronym for Weather Simulator). WESIM simulates daily temperature during a year at any location that has a multi-year daily temperature record available. WESIM generates the simulated data by use of a model of a 1st-, 2nd-, or 3rd-order autoregressive process with parameters that change daily during the annual cycle. The models are driven by random daily 'white noise' shocks. As measured by comparisons of input and output statistics, WESIM appears to simulate daily temperature variables in a realistic way. Repeated runs of WESIM can quickly generate several hundred independent annual sets of daily data that may be examined for e.g. frequencies of runs of extreme daily values, thereby expanding the limited information available in the historical record at your station.

The WESIM system consists of two executable files WESIMZ and WESIM, and a specially formatted 3-letter (the station code, e.g. EWR for Newark, NJ) input file containing 45 years of daily temperature and precipitation data. Input files are available for 490 U.S. stations. WESIMZ accepts the input file and creates files of averaged monthly statistics that are used by WESIM. WESIM a) expands the 12 monthly values of each input statistic into 365 daily normal values, b) generates independent annual sets of simulated daily max, min, or average temperature data, and c) calculates statistics on the output data for comparison with statistics on the input data.

Reference: R.L. Lehman, 1994, Simulating Daily Weather During a Year at a Station: System Overview of the Interactive Software 'WESIM', *Tenth International Conference on Interactive Information Processing Systems, Preprints*, p.261-261, American Meteorological Society, Boston, MA.

The latest versions of these CPC software products are available for trial and comment via anonymous ftp in the pub/cac/econ directory of nic.fb4.noaa.gov; access and descriptive data are in the econ.doc file.
Further information on these and other products is available from Dr. Richard L. Lehman, Program Leader for Applications and Socioeconomic Assessment, National Meteorological Center, 5200 Auth Rd., Room 805, Washington DC, 20233. Telephone 301-763-4670, Fax 301-763-8680, E-mail rlehman@sgi43.wwb.noaa.gov

Internet Address

ftp://nic.fb4.noaa.gov
.pub
.cac
.econ