1977

Load Forecasting for an Electrical Utility

Thomas Edmund Washburn
University of Central Florida

LOAD FORECASTING FOR AN ELECTRICAL UTILITY

BY

THOMAS EDMUND WASHBURN
B.S., Georgia Institute of Technology, 1972

Research Report

Submitted in partial fulfillment of the requirements for the degree of Master of Science: Electrical Energy Systems in the Graduate Studies Program of the College of Engineering of Florida Technological University

Orlando, Florida
1977

ACKNOWLEDGEMENT

I wish to express my appreciation to the members of my supervisory committee: Dr. Herbert C. Towle and Dr. Fred O. Simons for their time and assistance during the preparation of this research paper. Also, I am indebted to Dr. Benjamin W. Lin, my committee chairman, for his guidance and his time spent helping me throughout the development and preparation of this research paper.

Finally, I wish to express my thanks and appreciation to Sandra Anderson and my wife, Rachelle, for their large amounts of time spent typing and proofreading my paper.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

Introduction

The purpose of this paper is to explain load forecasting for an electrical utility. As in any business there is a desire to know the future demand of that business's product. For an electrical utility this product is electricity, and the instantaneous need for electricity is known as peak load, peak demand or simply peak. The daily, monthly, or annual need for electricity is called load or energy. These two measurements of usage are related by a number called a load factor. If the load factor is known one usage can be calculated from the other measurement of usage. Energy usage is used for fuel management, maintenance scheduling and budgeting. While peak usage must be known for planning the amount of future capacity needed to meet this instantaneous peak.

Forecasting techniques vary as the projection time period into the future changes, because the variables causing the variance change. This paper presents different methods of forecasting, and some of the usages of these projections.

Then a long-term forecast is developed using actual data for an electrical utility. This shows one methodology of projecting future loads. Multiple regression programs are used to aid in the calculations. After the regression equations are obtained statistical significance is tested. Then elasticity of the dependent variable relative to the independent variables is examined. Finally, there is a subjective analysis of the developed mathematical model.

Chapter II

Load Forecasting in General

The need for a peak demand or energy forecast for an electric utility can fall into three major areas - short-term, a few days to several months; intermediate-term, a year to five or ten years; and long-term, any longer period. Each different time period is sensitive to different data and requires varying degrees of accuracy.

All electric loads are comprised of a base load and a weather-sensitive load. In most techniques these two parts of the load are assumed to be statistically independent and normally distributed, which means that the sum of the mean values of these two components is the mean value of the total peak demand. Also, the sum of the variance of these two parts is the variance of the total load. Hence, there are three factors which contribute to the variance of the load forecast - (1) random variations of the base load about a trend curve, (2) uncertainty of the coefficients of the equation of the trend curve, and (3) random variations of the weather variable or uncertainty of the future weather.

In the short-term, daily and monthly peak demand predictions are necessary for the system operator to decide which units can be off for maintenance, whether he can sell power to another system, if he will need to use his peaking units, etc. Therefore, it follows that the short-term forecast is mostly for the basic operation of the power system, and the major reasons for variance are the random variations and future uncertainty of weather. The common forecasting method for the short-term is to obtain historical weather data that causes load variations - temperature, humidity, rainfall, etc. Then the weather-sensitive load must be removed from the historical loads. This can be done by examining monthly all daily peaks related to the weather data. Now a trend line can be fit to the base peak load and by correlating weather data with the weather-sensitive portion of the load, a complete model can be developed. By inputing normal, extreme or expected (may be different than normal) weather data in the model, it can yield a range of peaks based on weather uncertainty. Extreme data will give the upper limit for expected peaks.

It used to be conventional just to use historical demand data and a least-squares fit for a trend of the data. Then this curve would be projected into the future with corresponding probability limits. However, as more electric air-conditioning, heating and other appliances became common, this method became no longer acceptable for the short-term. It is still used for some long-term projections, under the assumption that weather normalizes itself over a period of time. However, for the short-range forecast, this curve fit method can still be used for the base load trend line.

Another technique is to correlate the base load with economic or system characteristic variables in a multiple regression and project the base load. In fact, for the short-term projection, this can be done weekly or monthly instead of annually, if the accuracy warrants the additional effort. If this is only to be done annually, then the weather-sensitive load must be done in the same format. Then percentages or factors need to be developed to relate this work to a shorter time period, that is - daily, weekly, monthly. In any case, the accurate forecast of the peak of today, tomorrow, next week, and/or next month is the final result. The energy is also a desired result, and in the short-term is a by-product of the peak by using load factors. In the longer range, forecasted peaks are still important for future plant expansion, but due to inaccuracy of forecasting the peak, energy is often projected and the peak is the by-product answer using load factors.

Now for the intermediate-term forecast a complete detailed analysis of weather correlated with load is usually too time consuming to be worthwhile. This does not mean ~~to~~ neglect the effect of weather on historical loads, but that more attention should be focused on what produces peak and energy growth. As the time range of the forecast becomes longer, more effort should be spent on explaining the random variations of the base load and the uncertainty of the coeficients of the equation. This mid-term projection has a different importance today than about fifteen years ago because of the longer lead times for generation and transmission plant additions. When it took only a few years to install a steam turbine generator, there was not the need for a long-term forecast there is today; the intermediate-term forecast of today was the long-term at that time. At that time most forecasters used a least-squares trend line for a projection. However, today with a lead time of ten years or longer, the plant needs to be under construction sooner in order to meet the load. Where mid-range forecasting is useful is to check the need for more generation. If it is determined that the load is growing faster than expected, then purchasing power, building gas turbines, or other means of meeting the load, must be built or obtained, because base load capacity can not be built in time. Likewise, if the growth is slower, a unit may be delayed or part of it sold to other utilities.

For planning of future generation and transmission systems a long-range forecast of twenty years or more is needed. In transmission planning an horizon year, about twenty years in the future, will be examined and a transmission network built for that year, and then the intermediate years will be looked at for timing or possible other lines needed. On the other hand, generation planning involves planning the next unit at least ten years ahead. Also, it is important to plan the next couple units in order to determine the best fuel type, site, construction timing, etc. A low forecast could result in insufficient capacity on line, meaning expensive purchased power must be bought or costly gas turbine generation added. While a high projection of load would cause more capacity in operation giving a loss of economy.

Intermediate-term and long-term forecasting techniques are often about the same. More time will be spent explaining variations of the weather variable in mid-range than in long-range projecting. But the major difference between intermediate and long-range forecasting is that in the long-term there can be a gradual change in system load factor (relationship between annual peak load and annual energy) that may not influence the mid-term projections. The energy load factor method is often used for long-range forecasting because annual energy can be projected with less uncertainty than annual peak demand. This paper will expand on this technique and perform an actual forecast using this method. By multiplying the energy by the load factor and dividing by the number of hours in a year (8760 hours), the annual energy can be converted to an annual peak demand. It is difficult with this method to get estimates of the variance of the peak demand forecast. While extrapolation of the peak demands using a trend curve will yield an estimate of the confidence limits, but even if the weather-sensitive load is removed, the trend line will explain less of the variance than the energy-load-factor technique. Also, the effort needed to examine the weather-sensitive portion of the peak may be too extensive for a long-term forecast, because more than just a multiple regression with weather variables is necessary for an accurate representation of weather-sensitive peak. However, with an energy forecast in the long-term only a temperature related variable is needed, because it explains the greatest amount of the weather variance. Usually cooling and/or heating degree days result in a good energy weather model. The rest of the regression variables should be economical or technical in nature. This can include technical breakthroughs such as the electric car, home solar energy, etc. But these variables can affect the load factor and energy more than they may affect the peak demand. Economic variables can include most anything that affects energy consumption - price of electricity, inflation, income, population, business activity index, etc.

Research on the effect of any non-weather variables to be included in the model should be done in order to prevent too many variables in the model. Appliance surveys and information about new housing construction as to size, appliances installed, and type of heating and air conditioning equipment installed can be useful. For that matter, the more information about customer usage, number of customers and economic conditions that can be obtained, the more confidence there is in a peak demand or energy forecast. However, once again too many different variables can lead to cross-correlation and possible confusion for the forecaster or the person interpreting the model or results. It is important in the regression analysis that close attention be paid to the F-value, standard error of regression coefficient, multiple correlation, standard error of estimate, and other pertinent results of the multiple regression, in order not to include too many variables.

In answer to the question of uncertainty in the future value of the variables, there are two possible ways to get a handle on this uncertainty. For one, each variable can be given a probability distribution, and then using game theory, a probabilistic simulation can be done. But once again these distributions are only as good as a person's estimate. If a computer program for this simulation is not available or this method is not acceptable, some pertinent "what if" questions can be asked. This second technique is done by changing the future values of an independent variable to simulate a possible event occurring, that is - large increases in the price of electricity, lower inflation, another oil embargo, etc. But one pitfall to this method is too many scenarios may result from changing variable values, which can be confusing. Only the most relevant questions should be addressed and investigated.

Different areas of the country will have different variables in their model. In fact, a large electric utility may divide its service area into regions with similar weather, economics and customer classes. Then in the final analysis, sum the results to obtain a system-wide answer. (The utility that this paper will work with is not large enough to warrant such a detailed analysis, but the principle for each analysis is the same.)

Before the oil embargo in October, 1973, there was no measurable price increase of electricity. But this price increase over the past three years, along with the economic recession, have reduced the per customer use of electricity. This energy crisis gives a dramatic demonstration of how price and personal conservation can affect electric demand and energy. Until this time, there was no event that showed the negative effect the economy could have on the load. The only thing that could have been done before was to conjecture what the price elasticity of electricity was. This problem of guessing what effect an event or development will have on the load can only be solved by a person's best estimate. For example, the effect of the electric car on energy usage or peak demand can not be determined by historical multiple regression. A detailed analysis of the best data available will yield the additional energy and peak that will be demanded by the electric car. Even more important is the effect on the load factor. As one can see, there is still a large amount of forecasting that can not be done simply by using mathematical tools. When using a multiple regression, all the independent variables must be projected into the future in order to yield the dependent variable. This may be done by the forecaster or this information may be obtained from other expert sources.

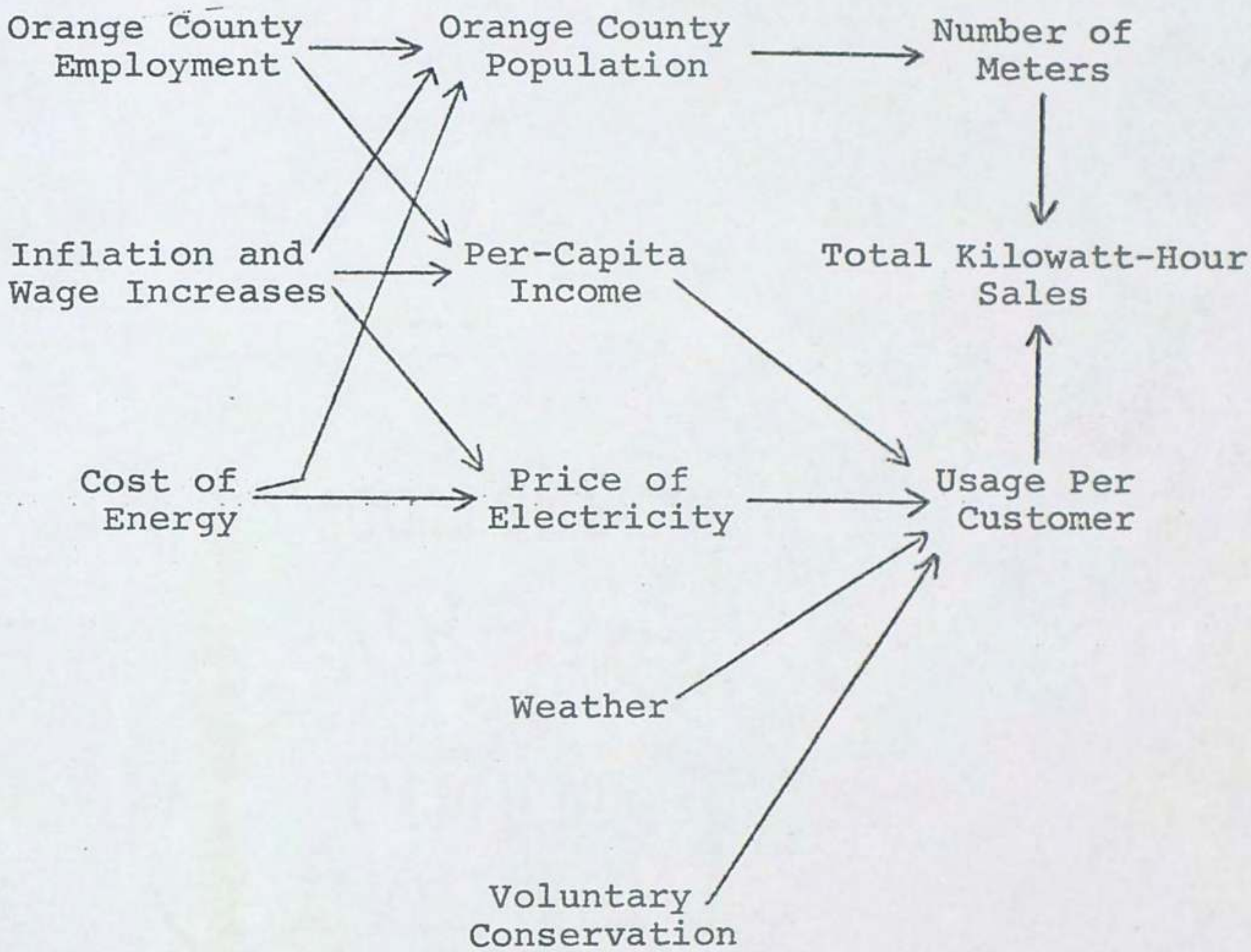
Chapter III

Actual Long-Term Load Forecast

The forecast presented in this paper is a long-term projection using actual historical data of the Orlando Utilities Commission. Any results developed or conclusions drawn are those of this forecaster alone, and should not be misconstrued as those of any other person or of the Orlando Utilities Commission. An attempt to show two different forecasts is made - one based on all customers together, and the other from projecting residential and commercial customers separately, then adding the results to arrive at the total. This second procedure is valid because residential usage does not affect commercial usage, being independent of each other they can be added together to obtain the total. Both forecasts are done by projecting separately annual average usage kilowatt-hour per customer and the average number of active meters (one meter per customer), then multiplying the results of each to yield the total energy for the year. Looking at only total usage does not give the true picture of the changes taking place. That is the average usage can vary independent of the number of meters. Likewise, the number of meters is independent of the average usage. This is why the total usage can be calculated using these two independent variables. In order to get the peak demand a load factor is applied. Data is not available for a load factor analysis, so it is assumed that the expected load factor is the average of the past nine years.

First, an important decision must be made as to which variables to include in the multiple regression mode. Too many variables cause too much effort spent on data collection and cross correlation, while not enough variables yield an inaccurate model. In exhibit 1, it is shown what is considered to be the important causes of changes in the number of meters and usage per customer. From these causes, the conclusions are drawn that lead to the development of the data in Tables 1 and 2. In the regression analysis, the forecaster relates the number of active services to the Orange County population. For the usage per customer, the independent variables are: degree days (cooling, heating and total), a Handy-Whitman Index, the Orange County per-capita income, a conservation index, and the cost of electricity. The Handy-Whitman Index is an average of the plant index and the labor index for the Southeast. These indices are pub-

Exhibit 1



Arrows Indicate Influence

TABLE 1 REGRESSION DATA

YEAR	HANDY-WHITMAN PLANT INDEX	HANDY-WHITMAN LABOR INDEX	WEIGHT 50-50 INDEX
1965	167	192	179.5
1966	173	197	185.0
1967	179	206	192.5
1968	184	216	200.0
1969	199	245	222.0
1970	213	270	241.5
1971	230	302	266.0
1972	247	350	298.5
1973	260	361	310.5
1974	313	389	351.0
1975	367	432	399.5

YEAR	DEGREE DAYS COOLING	DEGREE DAYS HEATING	DEGREE DAYS TOTAL
1965	3417	481	3898
1966	3119	655	3774
1967	3394	410	3804
1968	3081	1009	4090
1969	3339	840	4179
1970	3625	760	4385
1971	3891	427	4318
1972	3870	364	4234
1973	3720	556	4276
1974	3474	399	3873
1975	3601	443	4044

YEAR	ORANGE COUNTY POP. (1000)	ORANGE COUNTY PER-CAPITA INCOME	CONSERVE INDEX
1965	308.9	2480	1
1966	315.1	2586	1
1967	319.0	2877	1
1968	327.4	3135	1
1969	341.1	3476	1
1970	344.3	3704	1
1971	363.1	4219	1
1972	385.0	4824	1
1973	408.4	5421	0.83
1974	424.0	5664	0
1975	424.6	5774	0

TABLE 2 REGRESSION DATA

RESIDENTIAL

YEAR	KWH/CUST.	ACTIVE METERS	$/MWH
1965	7548	45043	22.56
1966	8170	45897	22.02
1967	8572	47133	21.66
1968	9606	48509	20.40
1969	10477	50013	20.03
1970	11366	51732	20.21
1971	11650	53819	20.94
1972	12040	56133	21.48
1973	12868	59603	21.58
1974	11456	62587	28.21
1975	11382	64315	35.97

COMMERCIAL

YEAR	KWH/CUST.	ACTIVE METERS	$/MWH
1965	59727	7099	20.04
1966	64589	7091	19.75
1967	71661	7075	19.17
1968	76806	7200	18.33
1969	83772	7401	18.27
1970	89089	7543	18.75
1971	93864	7692	19.54
1972	105230	7954	19.64
1973	113729	8318	19.96
1974	107871	8677	27.07
1975	110051	8905	34.12

ALL CUSTOMERS

YEAR	KWH/CUST.	ACTIVE METERS	$/MWH
1965	14680	52042	21.21
1966	15721	52988	20.88
1967	16806	54208	20.40
1968	18291	55709	19.39
1969	19925	57414	19.20
1970	21257	59275	19.58
1971	21931	61511	20.39
1972	23609	64087	20.64
1973	25220	67921	20.86
1974	23195	71264	27.91
1975	23382	73220	35.32

lished semi-annually with 1949 being the base (100). The cooling degree days, heating degree days and the sum of the two are figured using $65^{o}F$ as the base. In order to see whether the oil embargo of 1973 and the following conservation were significant, relative to usage per customer, a dummy variable was used. This variable has a value of one per month before the embargo then divided by twelve to put it on an annual basis. The cost of electricity is the annual average in dollars per megawatt-hour. Usage per customer is measured in annual kilowatt-yours per customer.

For the total number of customers regressed against the Orange County population, using a regression analysis program, the resulting equation is:

$$M_a = 167.7\ P + 488.5 \qquad (1)$$

where M_a is the total number of active meters for all customers and P is the Orange County population. This equation has a multiple correlation of 0.994, which means 99% of the variance of M is explained by P. Also, the F-value (810.0), the standard error of estimate (816.9), and the standard error of the regression coefficient imply that the equation is in the 99% confidence area. In examining the residuals it is found that all unit normal deviates are between the limits (-2,2). This approximates a normal distribution, N (0,1), which is the assumption for residuals of a least-squares regression model. A plot of the residuals versus the dependent variable (Graph 1) shows one possible point of concern the 1975 residual. However, the 1975 population estimate is about the same as 1974 which is unusual. Being only an estimate and not a census value, this high residual value, though it is not an outliner, may be caused by a bad measurement of the independent variable.

Now for the number of residential meters and the number of commercial meters, the equations are:

$$M_r = 152.6\ P - 1796.5 \qquad (2)$$

and

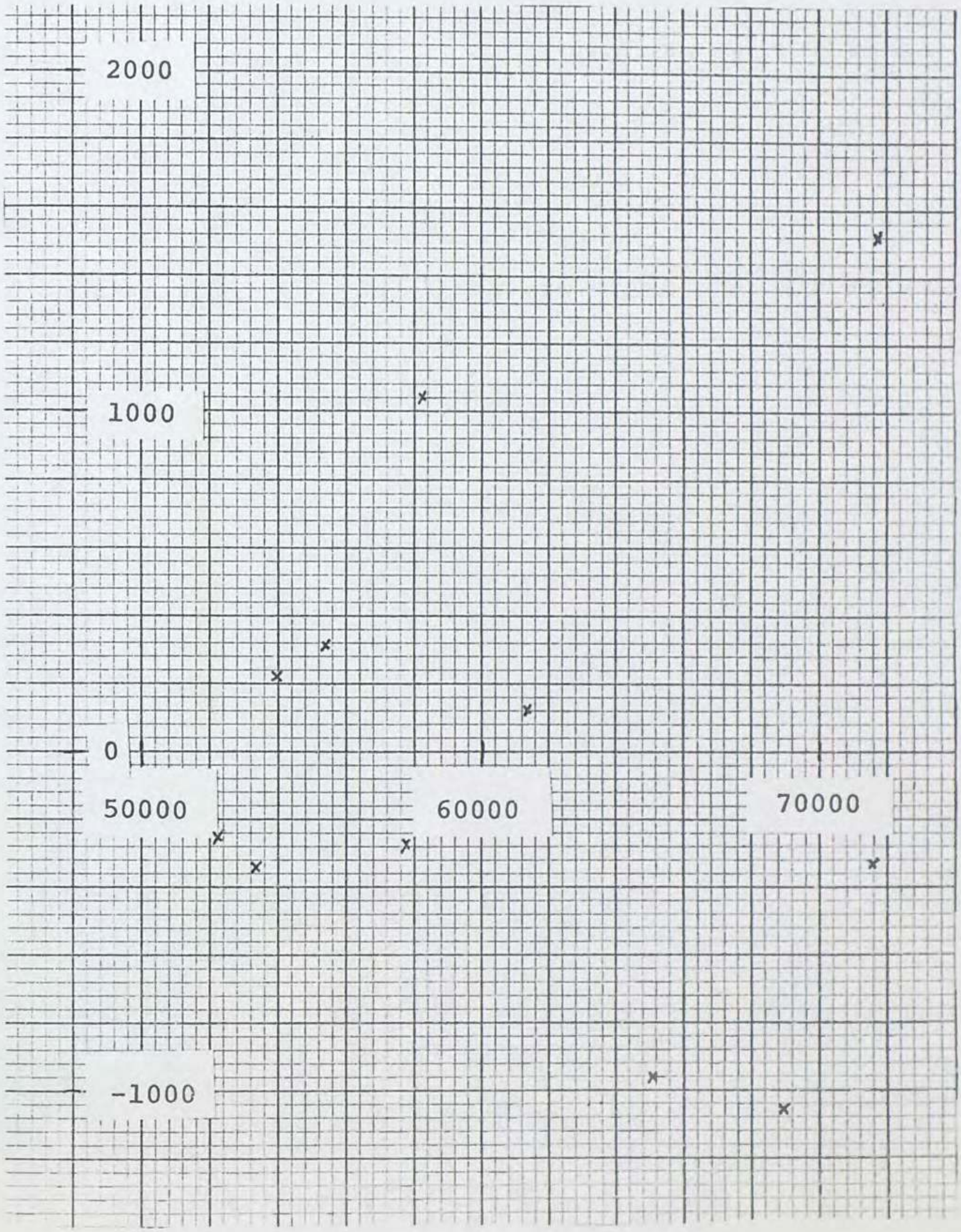
$$M_c = 14.8\ P + 2390.0 \qquad (3)$$

where M_r is the number of residential active meters and M_c is the number of commercial active meters. Likewise, with multiple correlation of 0.995 and 0.986, and F-values of 828.6 and 323.1, these equations are in the 99% confidence limits. Once again the unit normal deviates for both equations approximate a normal distribution. The plots of the residuals versus the dependent variable (Graphs 2 and 3) show that same high 1975 residual point, though not an outliner, to have the most deviation. But as before this could be caused by an error in estimating the population.

Since these were regressions of only one variable there is not a problem of which variables to include. But

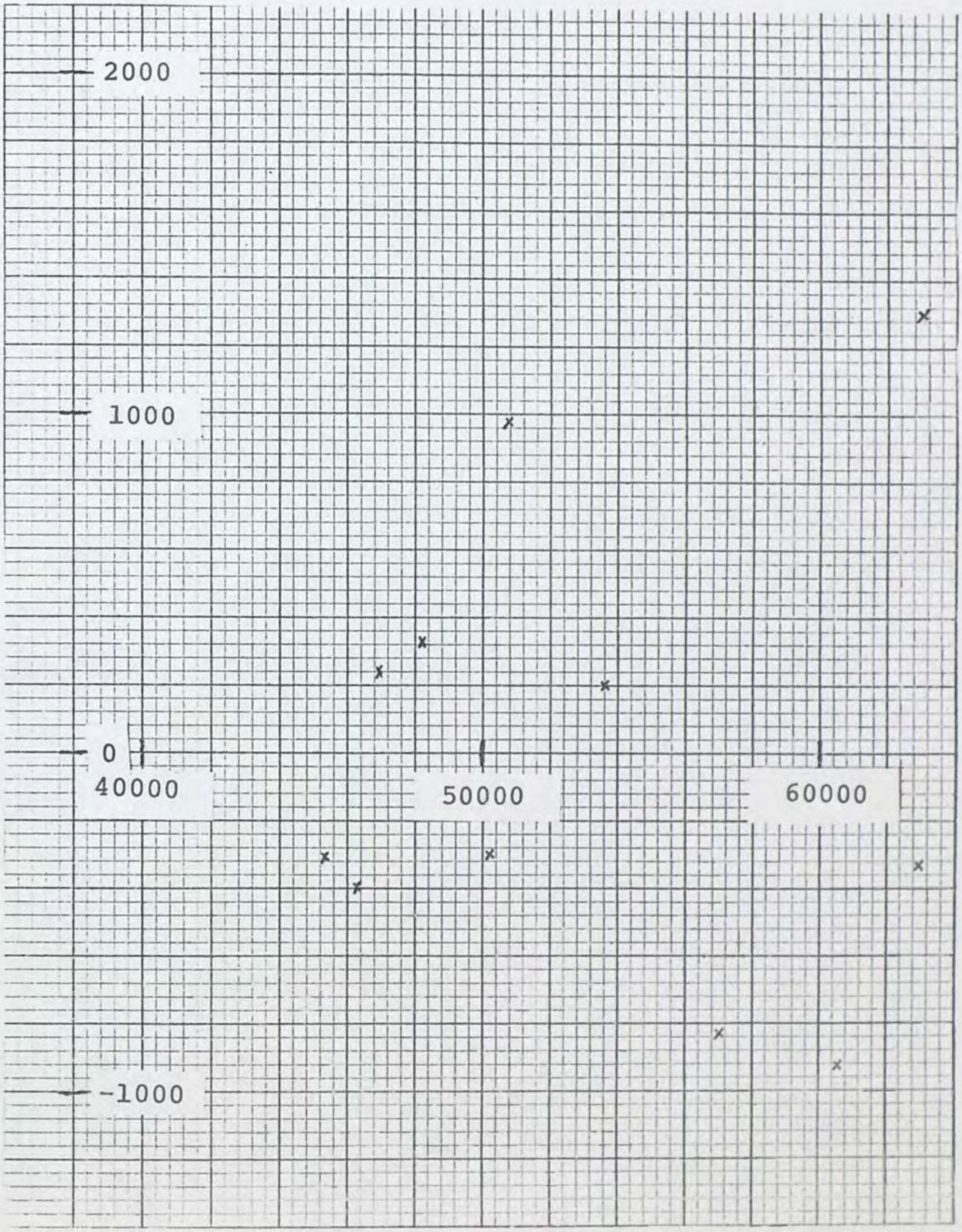
GRAPH 1

ALL METERS RESIDUAL -
DEPENDENT VARIABLE PLOT



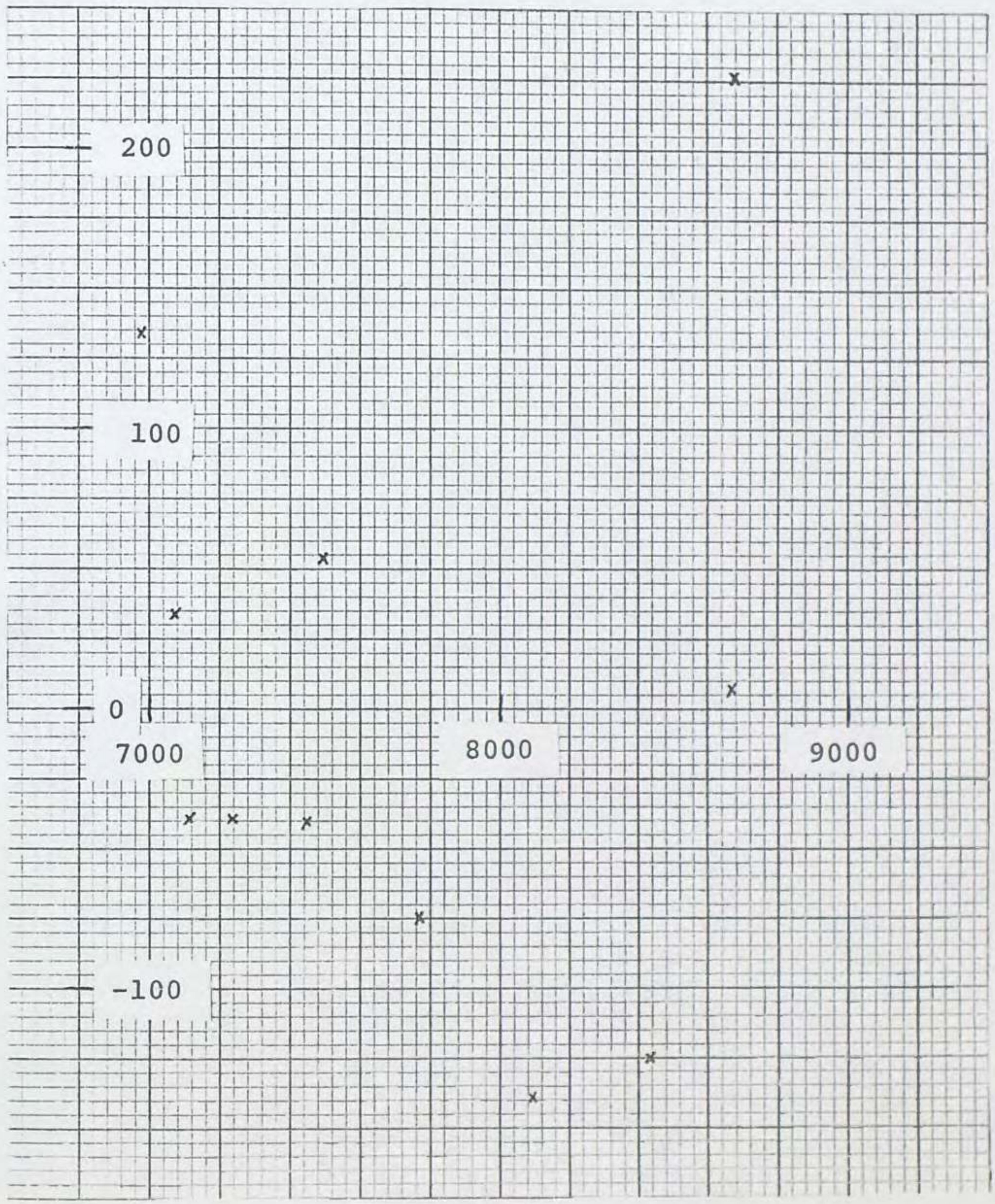
GRAPH 2

RESIDENTIAL METERS RESIDUAL -
DEPENDENT VARIABLE PLOT



GRAPH 3

COMMERCIAL METERS RESIDUAL -
DEPENDENT VARIABLE PLOT



for the usage per customer a program is needed to aid in deciding the pertinent variables. Hence, for this analysis, a step-wise multiple regression program is used, which is only one of several possible methods. The best method is the opinion of the statistician. This program enters the independent variable in order of reducing the sum of the squares (variance) until reaching the constant to limit the entering variable (input from the user).

For all three step-wise regressions a limiting value of 0.001 is used, in order to see the effects of several variables and then omit those which do not improve the F-value, standard error of estimate and standard error of regression coefficient. Looking at the results of the residential usage per customer regression, per-capita income, total degree days and cost of electricity explain 99% of the variance of the usage variable. Also, each variable improves the F-value and standard errors. However, the next two independent variables entering into the analysis do not improve the F-value or standard errors, so they are excluded from the final equation. This results in the following equation:

$$U_r = 1.283\ I + 2.772\ D - 102.4\ C_r - 3615.4 \qquad (4)$$

where U_r is the residential usage per customer, I is per-capita income, D is the total number of degree days, and

C_r is the cost of electricity for residential customers.

In a long-term model to be used in forecasting, weather variables, of course, can not be predicted. But for this equation the total degree days are normalizing or taking out the weather-sensitive part of usage for the historical data. In the forecast the normal number of degree days will be used for each year. Graph 4 shows the actual residential usage per customer, this usage adjusted for the number of degree days and the estimated usage from the model. This deweatherizing effect can be seen in the graph (tabular data in Table 3).

Next, for the commercial usage per customer, the per-capita income and the conservation index explain 99% of the variance of the usage variable. While heating degree days do yield lower standard error of estimate and standard error of regression coefficient, its presence in the analysis also gives a lower F-value, which is not favorable. By examining the runs of the residuals of the equation without the heating degree days variable in the model an unusual pattern of positive and negative residuals is noticeable (Table 4). But including the heating degree day variable helps to solve this problem by eliminating the runs.

Thus, the following is the final equation:

$$U_c = 18.625\ I + 13308.8\ CI + 6.338\ H - 345.2 \qquad (5)$$

GRAPH 4

RESIDENTIAL USAGE PER CUSTOMER

LEGEND: . Actual Usage
x Estimated Usage
\+ Actual Usage Adjusted for Weather

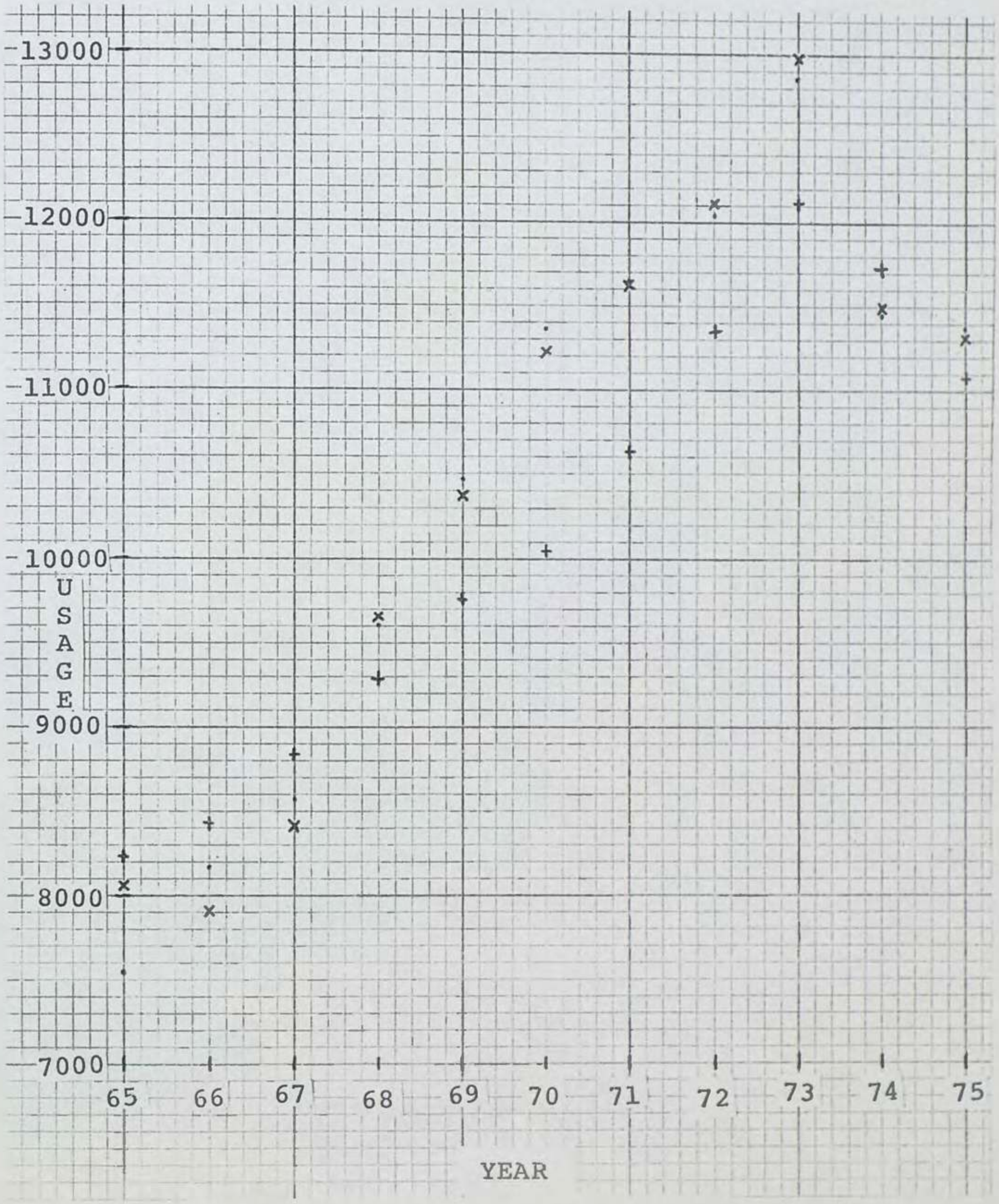


TABLE 3

RESIDUALS

Residential Usage Model

Year	Y-Value	Y-Estimate	Residual	Adjusted To Normal Weather
1965	7,548	8,062	-514	8,231
1966	8,170	7,909	261	8,422
1967	8,572	8,402	170	8,832
1978	9,606	9,655	- 49	9,292
1979	10,477	10,377	100	9,768
1970	11,366	11,223	143	10,042
1971	11,650	11,623	27	10,628
1972	12,040	12,111	- 71	11,349
1973	12,868	12,983	-115	12,104
1974	11,456	11,499	- 43	11,737
1975	11,382	11,319	63	11,084

All Customers Usage Model

Year	Y-Value	Y-Estimate	Residual	Adjusted To Normal Weather
1965	14,680	15,495	-815	15,723
1966	15,721	15,386	335	16,077
1967	16,806	16,402	404	16,980
1968	18,291	18,362	- 71	17,873
1969	19,925	19,695	230	18,874
1970	21,257	21,054	203	19,463
1971	21,931	22,143	-212	20,803
1972	23,609	23,513	96	22,486
1973	25,220	25,335	-115	24,151
1974	23,195	23,418	-223	23,739
1975	23,382	23,209	173	22,892

TABLE 4

RESIDUALS FOR
COMMERCIAL USAGE MODEL
WITHOUT HEATING DEGREE DAY VARIABLE

Year	Y-Value	Y-Estimate	Residual
1965	59,727	63,443	-3,716
1966	64,589	65,379	- 790
1967	71,661	70,695	966
1968	76,806	75,408	1,398
1969	83,772	81,637	2,135
1970	89,089	85,802	3,287
1971	93,864	95,209	-1,345
1972	105,230	106,261	-1,031
1973	113,729	114,830	1,101
1974	107,871	107,865	6
1975	110,051	109,874	177

where U_c is the commercial usage per customer, I is per-capita income, CI is the conservation index, and H is the number of heating degree days.

Like in the residential usage model, the heating degree days in this model normalizes the usage. Probably the reason that total degree days do not have as much effect as just heating degree days for the commercial usage model is that commercial establishments before the oil embargo seem to have their air conditioning on all the time. Remember how cold the stores used to be during the summer?

Finally looking at all customers as one group, per-capita income, total degree days and cost of electricity explain 99% of the variance of the usage variable. Once again, the Handy-Whitman Index and heating degree day variables improve the sum of the squares reduced, but these variables do not improve the F-value or standard errors. The exclusion of these two variables gives this equation:

$$U_a = 2.847\ I + 3.734\ D - 156.6\ C_a - 2798.8 \qquad (6)$$

where U_a is the usage per customer of all customers, and C_a is the cost of electricity for all customers. Once again the total degree days in this model normalizes the usage (Table 3).

Another fact to keep in mind, in determining the variables to be included in the final equation, is the reliability of the historical data. For example, in equation (4) for residential usage, the per-capita income and degree days explain about 96% of the variance. Therefore, if the historical cost of electricity had been suspect in terms of its validity, it could have been left out of the equation, because it does not explain a major portion of the variance. However, the historical cost of electricity is reliable; hence it is included.

Assuming there is little cross-correlation, the regression coefficient of a variable will show the effect of a change in that variable on the usage (dependent variable). Once again, in equation (4), if the cost of electricity goes up one dollar per megawatt-hour then usage per residential customer will go down 102.4 kilowatt-hours. This type of analysis can show the sensitivity of the usage to changes in each of the independent variables. Of course any cross-correlation between independent variables will impact the regression coefficient of each, giving distortion in the sensitivity study.

Now it is possible to calculate the elasticity of usage with respect to each independent variable. The elasticity for a linear model is different at every point, but it is common to calculate the elasticity using the mean values of the variables, which yields:

$$E_I = \frac{\partial U}{\partial I}\frac{\overline{I}}{\overline{U}} = 1.283\ \frac{4014.5}{10466.8} = 0.4921$$

$$E_C = \frac{\partial U}{\partial C}\frac{\overline{C}}{\overline{U}} = -102.4\ \frac{23.187}{10466.8} = -0.2268$$

$$E_D = \frac{\partial U}{\partial D}\frac{\overline{D}}{\overline{U}} = 2.772\ \frac{4079.5}{10466.8} = 1.0804$$

Actually the elasticity of usage with respect to price ranges from -0.17 to -0.33. Likewise, the other two elasticities have ranges, but the mid point gives the necessary information. The price elasticity is inelastic, which means electricity usage will go down little with a large rise in price with all other variables constant. Since the income elasticity is positive but less than one (about 0.5), it is low. This means that as per-capita income rises the cost of electricity becomes a smaller portion of the income.

Once the equations are developed, future data for the independent variables is researched. In this case, Orange County population and per-capita income are from the University of Florida statistics and the cost of electricity is Orlando Utilities Commission's best estimate (Table 5). The University of Florida used 1967 dollars as a base for their future per-capita income, then these must be inflated to current dollars. Inflation

TABLE 5

FUTURE DATA

Year	Orange County Pop. (1000)	Orange County Per-Capita Income 1967 $	Per-Capita Income Current $
1976	432.6	3572	6532
1977	440.6	3796	7323
1978	448.7	4021	8145
1979	463.0	4245	8986
1980	477.2	4470	9841
1985	537.9	5132	13098
1990	588.5	5794	17142

Year	All Customers $/MWH	Residential $/MWH	Commercial $/MWH
1976	39.12	39.97	38.12
1977	42.89	43.74	41.89
1978	43.65	44.50	42.65
1979	44.75	45.60	43.75
1980	46.18	47.03	45.18
1985	53.24	54.09	52.24
1990	59.50	60.35	58.50

Year	Degree Days Heating (Normal)	Degree Days Total (Normal)	Load Factor
1976	733	3959	0.5397
1977	733	3959	0.5397
1978	733	3959	0.5397
1979	733	3959	0.5397
1980	733	3959	0.5397
1985	733	3959	0.5397
1990	733	3959	0.5397

factors of 6%, 5.5%, 5%, 4.5%, and 4% are used for the years 1976 through 1980, then for 1981 through 1990 3% is assumed. The number of degree days (normal is assumed for the future) from the United States Weather Service is used.

Now with these best estimates, the usage and number of meters are projected. Using equations (1) and (6), and the data from Table 5 for all customers, the average usage, number of meters, total usage, and the sales peak are developed (Table 6). Likewise, for commercial and residential, equations (3) and (5) and equations (2) and (4) yield similar results. Adding together total sales for these two classes gives the total usage for all customers. Finally, by applying the load factor to the total usage, the annual peaks are figured.

TABLE 6

FUTURE FORECAST

All Customers

Year	Active Meters	KWH/Cust.	Sales (MWH)	Annual Peak (MW)
1976	73,036	24,455	1,786,095	378
1977	74,377	26,116	1,942,430	411
1978	75,735	28,337	2,146,103	454
1979	78,134	30,559	2,387,697	505
1980	80,515	32,770	2,638,477	558
1985	90,694	40,936	3,712,650	785
1990	99,180	51,470	5,104,795	1,080

Residential

Year	Active Meters	KWH/Cust.	Sales (MWH)
1976	64,218	11,646	747,883
1977	65,439	12,275	803,264
1978	66,675	13,252	883,577
1979	68,857	14,219	979,078
1980	71,024	15,169	1,077,363
1985	80,287	18,625	1,495,345
1990	88,009	23,172	2,039,345

Commercial

Year	Active Meters	KWH/Cust.	Sales (MWH)
1976	8,792	125,959	1,107,432
1977	8,911	140,691	1,253,698
1978	9,031	156,001	1,408,845
1979	9,242	171,665	1,586,528
1980	9,453	187,589	1,773,279
1985	10,351	248,251	2,569,646
1990	11,100	323,570	3,591,627

Year	Sum of Residential and Commercial Sales (MWH)	Annual Peak (MW)
1976	1,855,315	392
1977	2,056,962	432
1978	2,292,422	485
1979	2,565,606	543
1980	2,850,642	603
1985	4,064,991	860
1990	5,630,972	1,191

Chapter IV

Summary and Conclusions

The method of projecting the whole and the method of forecasting the parts and then totaling, both yield similar results with the difference between the two widening through time, but still within ten percent of each other even in 1990. It seems that the more detailed work in projecting by parts would be more accurate.

Using only mathematical processes can cause a problem between historical data fitting smoothly into projected data. Notice the projected active meter counts for all three groups; they are lower in 1976 than 1975, but it is possible that the 1975 population is out of line. This is further supported by the fact that 1975 meter count residuals are high for all three equations. Even historical data is an estimation of what was or is there. Some historical counts are better than others, that is, population estimates are less reliable than costs of electricity. More active service growth occurred in 1975 than is indicated by the active meter versus population curve.

The 1976 actual data for active meters and usage is now available. For all customers the total usage is

1,820,000 megawatt-hours with 77017 active meters, giving an average use of 2363 kilowatt-hours per meter. For residential customers, these figures are 786,037 megawatt-hours, 67,873 active services, and 11,581 kilowatt-hours per service, and for commercial - 1,033,963 megawatt-hours, 9,144 customers, and 113,075 kilowatt-hours per customer. The total sales is between the two total usage figures calculated. However, the actual average usage is lower for all groups and the number of active meters is higher for each of the three groups. Not knowing any of the actual 1976 data for the independent variables, it is not possible to explain this difference.

Finally, a "what if" study to develop different scenarios is performed. For example, what if inflation goes to zero during the 1980's and the cost of electricity in 1990 is the same as 1980? This results in a residential per customer usage of 18,996 kilowatt-hours and a commercial usage of 237,406 kilowatt-hours, giving total sales of 4,307,026 megawatt-hours. This seems to say that as inflation increases per-capita income, it is raising the cost of electricity less in proportion to income, which means more money available to spend on electricity. Other such questions can be asked, and then answered by the model.

After all the regressions are done and the "what if" study completed, the actual forecast may be a combination of several different scenarios, depending on the judgment of the forecaster and management. Forecasting, especially long-term, is not only mathematical formulas and analysis, but also judgment decisions. Of course these subjective decisions are based on the results of the models and on variables that are not or can not be figured in the mathematical model.

Looking at this forecast, the number of active services is growing at an average compound growth rate of 2.0%. Unless some large industry comes into the area and changes the population projection, this growth subjectively seems reasonable. However, the average usage per customer is the big question. From 1965 until 1973 the average usage per residential customer grew at an average compound growth rate of 6.9%, and projected 1990 usage yields a 4.9% rate from 1975. Of course, during the late 1960's and early 1970's the percentage of homes with electric air conditioners and new electric appliances (trash compactors, dishwashers, etc.) grew tremendously. This caused the large per customer usage growth then. The question now is what will continue to cause this usage per customer to grow? This is when subjective reasoning is used. Appliance surveys can be taken to aid in the decision. But the future of the electric car or other new electric gadgetry is a question mark. It is possible that there will be little usage per customer increases in the future due to more efficient appliances and more customer awareness of efficient operation of electric appliances. This would be a break in the historical trend which means the historical model would have to be modified. Hence, the reader can see the reasoning used after all the statistical analysis is done. In a way, the creation of a model is only the beginning of the problem of projecting what the future holds.

BIBLIOGRAPHY

Asbury, C. E. "Weather Load Model for Electric Demand and Energy Forecasting." IEEE Power Engineering Society, Summer Meeting. Vancouver, Canada, July 1973, pp. 1-6.

Bureau of Economic and Business Research. College of Business Administration. Florida Statistical Abstract. Gainesville: University of Florida, 1974-6.

Clair, S. S., and Einwechter, W. S. "The Weather and Daily Forecasting of Philadelphia Electric System Loads." American Power Conference Proceedings 23 (1961): 853-7.

Corpening, S. L.; Reppen, N. D.; and Ringlee, R. J. "Experience with Weather Sensitive Load Models for Short and Long-Term Forecasting." IEEE Transactions on Power Apparatus and Systems 92 (November/December 1973): 1966-72.

Gupta, P. C., and Yamada, Keigo. "Adaptive Short-Term Forecasting of Hourly Loads Using Weather Information." IEEE Power Engineering Society, Winter Meeting. New York, February 1972, pp. 2085-94.

Heinemann, G. T.; Nordman, D. A.; and Plant, E. C. "The Relationship Between Summer Weather and Summer Loads - A Regression Analysis." IEEE Transactions on Power Apparatus and Systems 85 (November 1966): 1144-54.

Latham, James, Jr.; Nordman, D. A.; Plant, E. C.; and Voorhis, J. S. "Probability Approach to Electric Utility Load Forecasting." IEEE Transactions on Power Apparatus and Systems 87 (February 1968): 496-504.

Orlando Utilities Commission. "Financial Statements." 1965-1975.

Sergel, R. P., and Perey, Delia. "Load Forecasting." EPRI-EEI Power Growth Workshop. Chicago, 1976, pp. IV-76-83.

Stanton, K. N.; Gupta, P. C.; and El-Abiad, A. H. "Long Range Demand Forecasting for the Electric Utility Industry." Power Industry Computer Applications Conference Proceedings. Chicago, 1969, pp. 565-71.

Stanton, K. N., and Gupta, P. C. "Forecasting Annual or Seasonal Peak Demand in Electric Utility Systems." IEEE Transactions on Power Apparatus and Systems 89 (May/June 1970): 951-9.

United States Federal Power Commission. The 1970 National Power Survey. Washington, D. C.: Government Printing Office, 1970.

Whitman, Requardt and Associates. Handy-Whitman Index of Public Utility Construction Costs. 1965-1976.