# ESTIMATION OF AGGREGATE MILES OF EHV TRANSMISSION LINE NEEDS

L. C. Markel

Report #MIT-EL 73-011 August 1972

# ENERGY LABORATORY

The Energy Laboratory was established by the Massachusetts Institute of Technology as a Special Laboratory of the Institute for research on the complex societal and technological problems of the supply, demand and consumption of energy. Its full-time staff assists in focusing the diverse research at the Institute to permit undertaking of long term interdisciplinary projects of considerable magnitude. For any specific program, the relative roles of the Energy Laboratory, other special laboratories, academic departments and laboratories depend upon the technologies and issues involved. Because close coupling with the normal academic teaching and research activities of the Institute is an important feature of the Energy Laboratory, its principal activities are conducted on the Institute's Cambridge Campus.

This study was done in association with the Electric Power Systems Engineering Laboratory and the Department of Civil Engineering (Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics and the Civil Engineering Systems Laboratory).

## ABSTRACT

Regression analysis is used to develop models for the total number of miles of EHV line needed in a power system. These models are not meant to be used to design a transmission system but to examine existing EHV systems to see if general patterns or relationships exist on a system-wide basis.

The U.S. was divided into regions and data on regional load and generation characteristics was obtained for the years 1960, 1970, 1980, and 1990. Regressions were performed on this combination cross-sectional and time-series data to develop equations for circuit miles of EHV line and for gigawatt-miles (miles of line X power-carrying capability) of EHV line. t- and F-tests were used to determine the statistical significance of the model parameters.

The independent variables (system characteristics) found to be most significant in determining the miles of EHV line include the generating capacity of the system, the area (square miles) of the region, the percent of area that is metropolitan, the number of generating plants in the system, the percent of energy used for industrial purposes, the percent of generating capacity which is hydroelectric, and the average distance between plants and load centers.

The equations developed are multiplicative, of the form miles of line = $K \cdot \prod_i X_i^{a_i}$, where the $X_i$'s are system characteristics. "Expansion" models attempting to determine the additions to the existing grid in a 10 year period were also postulated. However, they are not as statistically significant as the "static" models.

Methods of using the models to investigate new plant siting strategies, such as power parks or offshore nuclear plants, are discussed. Analysis of the effect of trends in plant siting and construction, such as the growing scarcity of potential plant sites near load centers, is also mentioned. The limitations, uses, and possible extensions of this type of model are also described.

## ACKNOWLEDGEMENTS

I would like to express my appreciation to Professor Fred Schweppe who supervised this work, for his support and guidance. As well as by giving suggestions and criticisms, he was invaluable in helping to put things in their proper perspective, and it was a pleasure to work with him.

I'd also like to thank Dr. Martin Baughman for his advice with the statistical and econometric aspects of this research. Discussions with him were always a great help.

The work described in this report was partly supported by a National Science Foundation Grant (#GI-34936) as part of a research program at MIT entitled "Dynamics of Energy Systems".

The report essentially constitutes a master's thesis submitted to the Electrical Engineering Department. Marilyn Pierce and Karen Keefe spent many hours typing this thesis and good-naturedly met deadlines and made last minute changes.

# TABLE OF CONTENTS

# CHAPTER 1

## Introduction

With the demand for electricity growing at about 8% a year, the number of generating facilities and transmission lines to deliver this power to the users must also grow accordingly. As the number of plants and transmission lines increases, and as the environmental protection constraints on new construction become stricter, it is becoming harder to site new plants and lines.

In any attempt to predict, model or plan the growth of an electric power system, the characteristics of the associated bulk transmission system are an important factor. A transmission system is shaped by the social, economic, and political constraints of a region as well as by load carrying, system stability, etc. constraints. Urban areas, the load centers, are already overcrowded. Land is at a premium, with both the economic and social costs of clearing the right-of-way becoming larger.

New plants tend to be larger and farther from the load centers. Air quality standards make it difficult to locate a large fossil steam plant near a city. Public opinion is not amenable to nearby nuclear plants. Water requirements can often only be met, especially for pumped hydro plants, far from a city. This requires even more high voltage line, often in wilderness, forest or park-land areas. While a cleared right-of-way does have some multiple purpose uses, such as bridle paths or bicycle and hiking trails, it usually has significant adverse aesthetic and environmental effects. (A description of the costs and land requirements for EHV lines is found in Appendix I and Tables 1 and 3.)

Thus, an estimation of the additional line requirements for various 10 or 20 year expansion strategies could be useful in system planning. This study's purpose is to predict the miles of EHV transmission line needed to meet the growth of the electrical system of the U.S. One use for an explicit model of this type is to analyze the effects of different plant siting or generation mix strategies or the growing scarcity of new plant sites on the expansion of the transmission grid.

Transmission line siting is a complicated trade-off of engineering, economic, right-of-way, and environmental considerations. In a report to the FPC by the Advisory Committee on Reliability of Electric Bulk Power Supply, the following recommendations on the design of bulk transmission networks were made:[1]

1. Maintain proper relationship between size and capacity of all system elements, or between systems.
2. Plan for adequate margins of transmission reserve capability through EHV lines.
3. Avoid excessive concentration of transmission capacity on a given right-of-way.
4. Maintain adequate interconnections among systems.
5. Avoid concentration of critical circuits at substation switching facilities.
6. Use relay schemes of least complexity that provide the required protection with the least hazard in the event of faulty operation or testing.
7. Examine proper transmission network contingencies when investigating system performance.

These principles are not all quantifiable. They involve determining minimum levels of acceptance and involve many trade-offs, especially economic. The general models developed here are not meant to locate lines on an individual basis or to take the place of the system planner in planning grid construction. Rather, this study attempts to determine whether the present techniques used in planning the transmission grid on a line-by-line basis can be generalized or aggregated to give some system-wide conclusions, characteristics, or "rules of thumb". The purpose is to predict the number of transmission line miles, based on such more easily predicted and measured qualities as the generating capacity of the system, the physical area covered by the system, the number of major load centers, etc.

"EHV line" for this study is taken to be line at 230 KV and higher. The years examined are 1960, 1970, and projections for 1980 and 1990. 1955 to 1960 saw the start of the present rapid growth in EHV line, with developing technology leading to the opening of the first 500 KV line in 1964 and the first 765 KV line in 1968 (see figure 1). Research on higher voltages is underway. Although costs and land requirements of line construction increase with rated voltage, the power carrying capacity grows even faster, roughly as the square of the rated voltage. (See table 1 and Appendix I) Thus, the models developed here estimate two quantities: the circuit miles of EHV line (CM), and the power-carrying capability, or gigawatt-equivalent miles (EM), of a grid.

To develop a model of grid size, equations relating miles of line to various system characteristics were hypothesized. Using multiple linear regression, the parameters for the equations were estimated and t- and F-tests

(Appendix V has a description of these statistical tests) were used to measure the statistical significance of each parameter. Those parameters which were found to be statistically significant were incorporated in the final model. Those found not to have a high level significance (i.e., the null hypothesis, that the coefficient of that parameter equals zero, could not be rejected) were dropped from the model.

For each region of the country, data for the years 1960 and 1970, and Federal Power Commission projections for 1980 and 1990 were taken as observations for the regression. Data from all four years could not be used in each regression since it was not always available, especially for the year 1960.

Chapter 2 and Appendix II describe the data used for the models. Chapter 3 gives the models developed. (And Appendix IV contains some of the results of regressions on previous models leading up to the final ones.) It was found that equations with a fairly simple structure:

$$\text{Miles of line} = K \cdot \prod_i X_i^{a_i}$$

$X_i$ = independent variables

gave good results. Only a comparatively small number of independent variables were needed to estimate the miles of line to within 25% accuracy. Generating capacity, physical area of a region, average distance between plants and load centers, and per cent of metropolitan area were found to be the most significant parameters.

Also, some models concerned explicitly with system expansion were hypothesized and tested. The results, however, were a good deal less encouraging, and no acceptable expansion models were developed.

From the first type, or "static," models, some generalizations can be made regarding system growth in the light of current trends, such as the increasing rarity of sites, and for possible new siting strategies, such as power parks. Some of the results, and the limitations, of the models developed are discussed in Chapter 4.

# CHAPTER 2

## Data Used in Models

### A. Division of the U.S. into Regional Systems

The entire country is interconnected; but, for the purpose of this study, the U.S. can be viewed as a number of fairly independent electrical regions connected by tie lines. Possible divisions are by power pools, FPC power supply areas, reliability councils, census regions, etc.

The six FPC national power survey regions are used in this study rather than the 50 states, since dividing the U.S. into 50 states would result in areas too interdependent. North Dakota, for example, has an extensive 230 KV network, but it is used to ship power out of the state, to places such as Minneapolis. Similarly, large pumped hydro projects being planned in Maine or New Hampshire are to be used to meet the peak of the Boston area.

The FPC national power survey regions(figure 2) closely follow the boundary lines of the reliability councils (figure 3) of the country in most cases. The utilities comprising a reliability council cooperate in system planning and dispatch, so they are fairly good divisions of the country into electrical systems.

Of course, some of these regions are much more closely bound than others. Region II, the East Central region, follows the boundaries of ECAR, and it is a very closely interconnected region. Region IV, the South Central region, includes ERCOT, which is practically unconnected with the rest of the country, as well as the Southwest Power Pool. The West region (region VI) combines the dissimilar areas of the mountain states, a large, thinly settled region, with the thickly settled, electricity-intensive Pacific states.

The Northeast region (region I) contains three smaller power pools, New England Power Pool (NEPOOL), New York Power Pool, and P-J-M (Pennsylvania-New Jersey-Maryland). Because data for each of these smaller, well-defined regions was available, the Northeast was divided into New England, New York, and PJM for this study. Since in the regression models the Northeast then counts as 4 points (Northeast, New England, New York, and PJM), the models' coefficients and parameters are biased toward the northeast. However, the three northeast subregions were used mainly to help in model development, to decide which variables of the models were statistically significant. Also, the three areas, while all are thickly settled, have many differences in other system characteristics.

To develop a model, nine regions were used, the Northeast (NE), New England (ND), New York (NY), PJM (PJ), the East Central (EC), the Southeast (SE), the South Central (SC), the West Central (WC), and the West (W). These regions are shown in figure 2. Data for these regions was taken for the years 1960 and 1970, and FPC projections were used for values for 1980 and 1990. The projections are valid data points as long as they are self consistent. Knowing that, for instance, the East Central region will require 58 new plants and 11508 more miles of EHV line by 1990 to meet a peak demand of 148,000 MW is valid as data. If the peak in 1990 turns out to be 160,000 MW and construction delays result in only 50 new plants and of a different generation mix than was originally planned, then obviously the additional mileage needed might not be 11500 miles. As the system characteristics change, so do the transmission needs. The projections are accurate if, (if the projection describes a viable system) should all the other projections <u>be</u> fulfilled, the transmission grid size is as predicted.

B. Selecting Independent Variables for the Model

Regression is used to determine the relationship between one variable, the "dependent" one, and other variables it is correlated with, the "independent" variables. However, correlation between two numbers does not necessarily prove that a casual relationship exists between them. In choosing regional characteristics as independent variables, an effort was made to choose those which logically should influence EHV grid size. The number and location of a region's load centers is probably more of a determinant of the transmission network of that region than vice versa. The load characteristics can also be predicted fairly independently of the planned electric generation and distribution system. This independence, of course, is not always found. Thus, in predicting the number and location of future plants and the reserve requirements of a system, implicit assumptions and decisions about the transmission system must be made. Nothing, of course, works out perfectly, so the real world never matches exactly a mathematical model. Some interactions of the actual system can not be easily modelled.

Some obvious problems can be avoided, however, The circuit miles of line in a grid would no doubt prove to be highly correlated with the number of transmission towers in the system. Telling a system planner that if his system has "x" towers, it will require "y" miles of line is not helpful, since he is more apt to estimate the second number by first estimating the first. The causal effect is clearly reversed in saying "y" is a function of "x", but a correlation coefficient can not indicate that.

The less obvious interactions are the ones which in effect provide "spurious" correlations, since few people would attempt to try to predict the line mileage needed by the number of transmission towers planned to be built. In a model, a strong correlation may exist between the percent of generating capacity that is nuclear and the miles of transmission line. However, this might simply be a coincidence, since developing nuclear technology, stricter air quality standards, and fossil fuel shortages are presently leading to a larger nuclear portion of the generation mix. At the same time, there is a trend towards stronger transmission networks and greater interconnection. Those regions increasing interconnection the most may also be installing more nuclear plants. The two trends may be coincident but not related, and if for some reason no new nuclear plants were constructed, the grid size might not grow very differently from the present predictions. Again, a correlation coefficient does not imply causality.

Causality can be inferred only from knowledge of how the system is planned. Thus, fear of nuclear accidents may necessitate locating nuclear plants farther from cities than fossil installations, requiring more transmission line.

In this case, a correlation, and a causal relationship, between the portion of generation that is nuclear and the transmission requirements is to be expected. In actuality, a correlation in the model may represent a combination of causal dependence, coincidence, and a third, often unquantifiable, factor or interaction. There is no sure way to separate one from the other.

Appendix II lists the independent variables investigated as possible components of a transmission model. It also states the source of the data and reasons for investigating each variable.

Data for 1960 and 1970 and projections for 1980 and 1990 were compiled. Some data for 1965 is given in table 5 along with the 1960, 1970, 1980 and 1990 data, but 1965 data was not used in any of the regressions.

The dependent variables, the quantities to be estimated, were circuit miles of EHV line, (CM), and gigawatt-equivalent miles of EHV line, (EM). For each voltage class of line there is a scaling factor, proportional to the square of the rated voltage, which indicates the power-carrying capability of that line. (The load scale factors (L) used are given in table 1.) The derivation of EM is

$$EM = \sum_{i} L_i \cdot CM_i \quad i = 230\ KV,\ 345\ KV,\ 500\ KV,\ 765\ KV$$

The independent variables used to develop the models were:

| Symbol | Variable |
|---|---|
| A | area; the physical size (in square miles) of a region |
| G | the generating capacity of a system (in gigawatts) |
| P | the number of generating plants with a nameplate capacity of 400 MW or above |
| LC-500 and LC-1000 | the number of load centers with a peak demand of 500 MW or over and 1000 MW or over |
| I, R, and C | the percent of delivered power used for industrial, residential, and commercial uses |
| F, H, and N | the percent of generating capacity which is fossil fueled, hydroelectric, and nuclear |
| $F_{new}$, $H_{new}$, and $N_{new}$ | the percent of generating capacity built within the last 10 years which is fossil fueled, hydroelectric, and nuclear |
| METRO | the percent of land area included in metropolitan areas |
| APL | the average plant to load center distance between plants of 400 MW or over and load centers of 1000 MW or over. |

| | |
|---|---|
| TPL | the <u>total plant to load center distance</u> |
| | TPL = P · APL |
| NAPL and NTPL | APL and TPL for <u>new</u> plants, built within the last 10 years |
| OAPL and OTPL | APL and TPL for <u>old</u> plants, built over 10 10 years previously |
| AAPL | <u>adjusted average plant to load center</u> distance |

$$AAPL_{1980} = \frac{APL_{1970}}{OAPL_{1980}} x\ APL_{1980}$$

$$AAPL_{1990} = \frac{AAPL_{1980}}{OAPL_{1990}} x\ APL_{1990}$$

The purpose of this study was to predict EHV line needs on the basis of more readily predicted system characteristics. Therefore, the independent variables picked were objective and easily computable and quantifiable. Other system characteristics such as reliability also may have a major influence on the size of the EHV grid. It would be very useful to be able to include a reliability value in the model and thus determine explicitly the tradeoff between reliability and grid size. However, there is no agreed-upon measure of a transmission system's reliability, and little system-wide data on reliability is available. The number of customers per year who experience an interruption or degredation (as in voltage reduction) in service due to lack of adequate transmission facilities is simply not an available datum.

Some characteristics are unquantifiable; environmental constraints are an example of this. Air quality standards may result in more plants located far from the urban load centers. Court decisions on licensing may delay nuclear, pumped storage, or transmission line construction, resulting in an increase in gas turbines on sites near the existing grid. However, a variable such as "maximum allowable $SO_2$ emission" or "average lead time for new plant construction" would not be very practical. Aside from the difficulty in projecting federal standards or court decisions, the impact of these "unquantifiable" developments is not known. Strict air quality standards could lead to the use of low sulphur fuels, the development of better air scrubbers, the location of new plants further from cities, the encouragement of nuclear and pumped hydro technology or, more probably, a combination of all of these. Thus, the effect of stricter environmental statutes on plant size, location, type and number should be predicted, and the resulting effect on EHV grid size can be determined from changes in these parameters.

For ease in handling the data, data was divided by various factors of 10 so that all variables would be of a magnitude comparable to 10. Thus, for instance, percents were actually expressed as "10's of percents" for the regression (58.6% would become 5.86). Generating capacity was expressed as "10's of gigawatts"; miles of line was expressed as "thousands of miles of line", etc. These scaling factors are given in table 6.

# CHAPTER 3

## Model Development

The purpose of this study was to estimate the total number of circuit miles (CM) and gigawatt miles (EM) of EHV transmission line in a system. In general, miles of line, M, is given by:

$$M = f\ (X_1, X_2, \ldots X_i) \quad (1)$$

where the X's, the independent variables, are the system characteristics, G, A, F, I, METRO, APL, etc. listed in Chapter 2 of this paper. It was assumed that the general model was multiplicative, that is:

$$M = e^K \cdot \prod_i f_i\ (X_i)^* \quad (2)$$

Each factor of the equation may be a function of several variables, however:

$$f_i = f_i\ (X_i, X_1, X_2, \ldots)$$

The general model assumed was:

$$f_i = X_i^{\ a_o + \sum_j (a_j \cdot X_j)} \quad (3)$$

To develop a workable model, an equation was postulated, its parameters (K, b's, and a's) estimated by multiple linear regression, and the statistical significance of

---

* $e^K$ was used as a constant instead of simply "K" for ease of notation, since the logrithm of this equation was used in the regression.

those parameters determined in order to include only relevant variables in the final model.

Because linear regression was used, the general model of equations (2) and (3) could not be used immediately. The number of possible parameters (a's and b's) to be estimated was too large. Also, in addition to developing an accurate model, it was desired to include as few independent variables as possible. (Using 30 independent variables to estimate 36 values of M would be highly suspect.)

Thus, the first equations examined were much simpler than (2) or (3), while still within their general framework.

The IBM Scientific Subroutine Package[1] was used to perform the regression. The SSP programs also computed t- and F-statistics in order to test the significance of the estimated parameters. A description of the t- and F-tests and their meaning is given in Appendix V. These tests were used to "weed out" the unrelated or nonsignificant variables.

While t- and F-tests may reveal whether a variable is "significant" or not, a user of a regression model is also interested in "how far off" the model can be. Thus, in addition to using critical points in the t- and F-tables, the percentage of error (1-predicted value/actual value) for each region was computed.

The results of regressions done on some preliminary models are given in Appendix IV.

The first type of model examined is a straight multiplicative model of the form:

$$M = e^K \cdot \prod_i X_i^{a_i} \tag{4}$$

Logrithms were taken of both sides in order to perform a linear regression:

$$\ell n\ M = K + \sum_i a_i\ \ell n\ X_i \tag{5}$$

After eliminating non-significant variables (see Appendix IV), the resulting equations are:

$$C\ M = e^K A^{a_1} P^{a_2} I^{a_3} G^{a_4} METRO^{a_5} H^{a_6} AAPL^{a_7} \tag{6}$$

and

$$E\ M = e^K A^{a_1} G^{a_2} METRO^{a_3} AAPL^{a_4} \tag{7}$$

The coefficients, statistics, and predictions of these models are given in tables 3-1 and 3-2. The signs of the a's are as we would expect; M is directly proportional to A, I, G, METRO, H, and APL and inversely proportional to P:

1. The greater the physical area (A) a system serves, the more line is needed to "cover" it.

2. A larger percentage of industrial use (I) indicates a concentration of demand in specific areas. Higher capacity line is needed to bring power to industrial than to residential areas. Conversely, the coefficient for residential use, R, is negative. (See Appendix IV, Table IV-2.) (R and I are strongly correlated, so both could not be used in the same equation. Tables IV - 3 & 4)

3. The greater the generating capacity (G), the more power has to be transported and the greater the grid size needed.

4. Metropolitan or urban areas (METRO) have greater load concentrations than rural areas, so we would expect the same relationship as in "2" above.

5. Hydroelectric plants (H) are usually located relatively further from load centers than steam plants and thus require more line.

TABLE 3-1-A

$$CM = e^{K} A^{a_1} P^{a_2} I^{a_3} G^{a_4} METRO^{a_5} H^{a_6} AAPL^{a_7}$$

Degrees of freedom (Df)*: 7/17

Regions & Years used (R & Y): (NE, ND, NY, PJ, EC, SE, SC, W) 70, 80, 90

(WC) 70

F-Statistic/Significance (F)*: 67.51/.01

Average % Error (AE): 15.6%

| Coefficient | Value (V) | Standard Deviation (S) | t-value/α* (t) |
|---|---|---|---|
| K | -1.946 | | |
| $a_1$ | .583 | .113 | 5.16/.01 |
| $a_2$ | - .471 | .280 | 1.68/.2 |
| $a_3$ | 1.325 | .480 | 2.76/.05 |
| $a_4$ | .756 | .218 | 3.47/.02 |
| $a_5$ | .566 | .188 | 3.01/.05 |
| $a_6$ | .223 | .100 | 2.23/.1 |
| $a_7$ | .341 | .157 | 2.17/.1 |

* See Appendix V for definition of these terms.

TABLE 3-1-B

| Region | CM Actual Value | CM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 7,452 | 6,502 | 13 |
| NE 80 | 10,855 | 10,935 | 1 |
| NE 90 | 15,610 | 14,326 | 8 |
| ND 70 | 1,397 | 12,140 | 13 |
| ND 80 | 2,580 | 2,608 | 1 |
| ND 90 | 3,370 | 3,585 | 6 |
| NY 70 | 1.600 | 1,981 | 24 |
| NY 80 | 2,310 | 2,869 | 24 |
| NY 90 | 3,595 | 3,480 | 3 |
| PJ 70 | 4,455 | 3,832 | 14 |
| PJ 80 | 5,965 | 6,070 | 2 |
| PJ 90 | 8,645 | 7,673 | 11 |
| EC 70 | 6,160 | 6,881 | 12 |
| EC 80 | 13,730 | 11,661 | 15 |
| EC 90 | 17,030 | 18,148 | 7 |
| SE 70 | 7,160 | 10,417 | 46 |
| SE 80 | 21,770 | 16,360 | 25 |
| SE 90 | 30,340 | 25,746 | 15 |
| SC 70 | 5,700 | 6,699 | 18 |
| SC 80 | 12,780 | 14,311 | 12 |
| SC 90 | 18,270 | 24,674 | 35 |
| WC 70 | 8,770 | 6,149 | 30 |
| W 70 | 29,280 | 28,355 | 3 |
| W 80 | 42,910 | 37,455 | 13 |
| W 90 | 51,970 | 60,372 | 16 |

TABLE 3-2-A

$EM = e^{K} A^{a_1} G^{a_2} METRO^{a_3} AAPL^{a_4}$

Df: 4/20

R&Y: (NE, ND, NY, PJ, EC, SE, SC, W) 70, 80, 90; (WC) 70

F: 62.39/.01

AE: 23.2%

| | V | S | t |
|---|---|---|---|
| K | -1.384 | | |
| $a_1$ | .502 | .128 | 3.91/.02 |
| $a_2$ | .708 | .171 | 4.15/.02 |
| $a_3$ | .582 | .251 | 2.32/.1 |
| $a_4$ | .490 | .214 | 2.29/.1 |

TABLE 3-2-B

| Region | EM Actual Value | EM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 3,124 | 3,653 | 17 |
| NE 80 | 5,331 | 8,019 | 50 |
| NE 90 | 13,104 | 14,845 | 13 |
| ND 70 | 618 | 442 | 28 |
| ND 80 | 1,328 | 1,284 | 3 |
| ND 90 | 2,467 | 2,443 | 1 |
| NY 70 | 561 | 834 | 49 |
| NY 80 | 987 | 1,606 | 63 |
| NY 90 | 3,411 | 2,772 | 19 |
| PJ 70 | 1,945 | 1,713 | 12 |
| PJ 80 | 3,016 | 3,433 | 14 |
| PJ 90 | 7,226 | 6,261 | 13 |
| EC 70 | 5,550 | 3,325 | 40 |
| EC 80 | 13,818 | 6,592 | 52 |
| EC 90 | 17,058 | 1,170 | 31 |
| SE 70 | 2,529 | 4,234 | 67 |
| SE 80 | 9,385 | 9,033 | 4 |
| SE 90 | 16,154 | 16,755 | 4 |
| SC 70 | 3,778 | 4,067 | 8 |
| SC 80 | 8,260 | 10,166 | 23 |
| SC 90 | 15,011 | 19,317 | 29 |
| WC 70 | 3,232 | 3,062 | 5 |
| W 70 | 11,996 | 10,956 | 9 |
| W 80 | 22,608 | 18,656 | 17 |
| W 90 | 31,946 | 34,409 | 8 |

6. It is expected that the greater the average distance between generation and load centers (APL), the more line will be needed to transport power to the load.

7. The more plants there are (P) for a given G, the more dispersed the generation pattern is. Thus, plants are more likely to be located near load centers.

The independent variables used in the equations are not all uncorrelated, however. In general, the larger the physical area, the more generation which one can expect to find in that area. Likewise, the larger A is, the larger the plant to load center distance can be (APL). A larger percentage of metropolitan area (METRO) could indicate higher load concentrations, and thus more generating capacity. Thus, A, G, METRO, and APL are all intercorrelated.

In addition, many variables are correlated with time. The number of plants, the percent of nuclear capacity, the miles of EHV line, the generating capacity, and the number of load centers all increase with time. The percentage of hydroelectric and fossil fuel generation is declining with time. The data for equations (6) and (7) was for 9 regions of the country for the years 1970, 1980, and 1990; thus it was both cross-sectional and time series data. In an attempt to determine whether the same variables were significant on a purely cross-sectional basis, regressions using data for only a single year were done. (See Table IV-1-A.) Some results are given in Tables IV-9 to 12 and in Tables 3-3 to 3-6. The correlation between A, G, METRO, and APL, as mentioned before, is apparent here, as equations with several of these variables show different ones as significant in different years.

Because there were only 8 or 9 data points for each year, fewer variables were used. The results of the final year-by-year equation:

$$C\,M = e^{K} A^{a_1} G^{a_2} \tag{8}$$

$$E\,M = e^{K} A^{a_1} G^{a_2} \tag{9}$$

are given in Tables 3-3 and 3-4.

The results for 1960 were not very accurate and the parameters not very significant. This is probably because EHV line was just beginning to be a major power carrier in 1960. Before that, grids of 115 KV, 138 KV and 161 KV handled most of the bulk power transmission

TABLE 3-3-A

$CM = e^K A^{a_1} G^{a_2}$ (BY YEAR)

| | 1960 | | | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Df: | 2/6 | | | 2/6 | | | 2/6 | | | 2/6 | | |
| R&Y: | NE, ND, NY, PJ, EC, SE, SC, WC, W | | | Same | | | Same | | | Same | | |
| F: | 1.92/-- | | | 16.92/.01 | | | 40.50/.01 | | | 51.97/.01 | | |
| AE: | 145.7% | | | 25.2% | | | 19.6% | | | 16.7% | | |
| | V | S | t | V | S | t | V | S | t | V | S | t |
| K | - .863 | | | .361 | | | .243 | | | .001 | | |
| $a_1$ | .231 | .524 | .44/.7 | .418 | .197 | 2.12/.2 | .436 | .151 | 2.89/.2 | .355 | .140 | 2.53/.2 |
| $a_2$ | 1.306 | 1.099 | 1.19/.5 | .750 | .385 | 1.95/.2 | .844 | .292 | 2.89/.2 | .879 | .257 | 3.42/.1 |

TABLE 3-3-B

| Region | CM Actual Value | CM Fitted Value | % Error |
|---|---|---|---|
| NE 60 | 2,130 | 2,385 | 12 |
| ND 60 | 266 | 2,436 | 8 |
| NY 60 | 371 | 474 | 28 |
| PJ 60 | 1,393 | 618 | 56 |
| EC 60 | 2,221 | 2,117 | 5 |
| SE 60 | 1,318 | 2,437 | 85 |
| SC 60 | 143 | 1,544 | 979 |
| WC 60 | 3,417 | 1,293 | 62 |
| W 60 | 13,962 | 3,333 | 76 |
| NE 70 | 7,452 | 7,274 | 2 |
| ND 70 | 1,397 | 1,445 | 3 |
| NY 70 | 1,600 | 1,928 | 21 |
| PJ 70 | 4,455 | 2,594 | 42 |
| EC 70 | 6,160 | 6,912 | 12 |
| SE 70 | 7,160 | 9,773 | 37 |
| SC 70 | 5,700 | 9,384 | 65 |
| WC 70 | 8,770 | 8,581 | 2 |
| W 70 | 29,280 | 16,761 | 43 |
| NE 80 | 10,855 | 13,191 | 44 |
| ND 80 | 2,580 | 2,548 | 7 |
| NY 80 | 2,310 | 2,675 | 35 |
| PJ 80 | 5,965 | 4,437 | 22 |
| EC 80 | 13,730 | 12,416 | 49 |
| SE 80 | 21,770 | 19,554 | 23 |
| SC 80 | 12,780 | 20,401 | 42 |
| WC 80 | 14,780 | 15,449 | 2 |
| W 80 | 42,910 | 30,634 | 21 |
| NE 90 | 15,610 | 18,232 | 17 |
| ND 90 | 3,370 | 3,696 | 10 |
| NY 90 | 3,595 | 3,754 | 4 |
| PJ 90 | 8,645 | 6,619 | 23 |
| ES 90 | 17,030 | 16,921 | 1 |
| SE 90 | 30,340 | 27,019 | 11 |
| SC 90 | 18,270 | 28,121 | 54 |
| WC 90 | 22,060 | 19,887 | 10 |
| W 90 | 51,970 | 41,330 | 20 |

TABLE 3-4-A

$EM = e^{K} A^{a_1} G^{a_2}$

| | 1960 | | | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Df: | 2/6 | | | 2/6 | | | 2/6 | | | 2/6 | | |
| R&Y: | NE, ND, NY, PJ, EC, SE, SC, WC, W | | | Same | | | Same | | | Same | | |
| F: | 2.02/-- | | | 14.27/.01 | | | 24.65/.01 | | | 52.48/.01 | | |
| AE: | 152.4% | | | 33.9% | | | 27.1% | | | 18.8% | | |
| | V | S | t | V | S | t | V | S | t | V | S | t |
| K | -2.202 | | | -.577 | | | -.611 | | | .962 | | |
| $a_1$ | .126 | .532 | .24/.9 | .399 | .227 | 1.76/.3 | .436 | .208 | 2.10/.2 | .860 | .089 | 10.21/.01 |
| $a_2$ | 1.546 | 1.114 | 1.39/.3 | .878 | .443 | 1.98/.2 | .970 | .403 | 2.41/.2 | .158 | .045 | 3.51/.1 |

TABLE 3-4-B

| Region | EM Actual Value | EM Fitted Value | % Error |
|---|---|---|---|
| NE 60 | 542 | 786 | 45 |
| ND 60 | 67 | 61 | 8 |
| NY 60 | 117 | 141 | 20 |
| PJ 60 | 358 | 185 | 48 |
| EC 60 | 1,189 | 673 | 43 |
| SE 60 | 324 | 731 | 126 |
| SC 60 | 38 | 396 | 942 |
| WC 60 | 874 | 324 | 63 |
| W 60 | 3,599 | 886 | 75 |
| NE 70 | 3,124 | 3,556 | 14 |
| ND 70 | 618 | 587 | 5 |
| NY 70 | 561 | 845 | 51 |
| PJ 70 | 1,945 | 1,166 | 40 |
| EC 70 | 5,550 | 3,320 | 40 |
| SE 70 | 2,529 | 4,728 | 87 |
| SC 70 | 3,778 | 4,312 | 14 |
| WC 70 | 3,232 | 3,903 | 21 |
| W 70 | 11,996 | 7,966 | 34 |
| NE 80 | 5,331 | 7,651 | 22 |
| ND 80 | 1,328 | 1,234 | 1 |
| NY 80 | 987 | 1,330 | 16 |
| PJ 80 | 3,016 | 2,336 | 26 |
| EC 80 | 13,818 | 7,092 | 10 |
| SE 80 | 9,385 | 11,518 | 10 |
| SC 80 | 8,260 | 11,715 | 60 |
| WC 80 | 8,384 | 8,542 | 5 |
| W 80 | 22,608 | 17,843 | 29 |
| NE 90 | 13,104 | 14,639 | 12 |
| ND 90 | 2,467 | 3,106 | 26 |
| NY 90 | 3,411 | 3,285 | 4 |
| PJ 90 | 7,226 | 5,754 | 20 |
| EC 90 | 17,058 | 13,360 | 22 |
| SE 90 | 16,154 | 20,284 | 26 |
| SC 90 | 15,011 | 19,819 | 32 |
| WC 90 | 16,423 | 13,854 | 16 |
| W 90 | 31,946 | 27,044 | 15 |

and continued to do so through 1960, especially in the Southeast and South Central regions.

Another "year-by-year" regression with very significant parameters was

$$C\ M = e^{K}\ G^{a_1}\ APL^{a_2} \tag{10}$$

$$E\ M = e^{K}\ G^{a_1}\ APL^{a_2} \tag{11}$$

Equations (10) and (11) could tell a system planner more about how his decisions will affect the system, as APL is a variable dependent on plant siting policy while A is a constant. However, A and APL are obviously correlated with each other, as regressions using A, G, and APL gave much poorer results than equations (8) - (11). (Tables IV-9 and IV-10, for example) This makes sense, since in regions of larger area, the distance between plants and load centers can be larger.

Equations (8) and (9) were also used as a model for the cross sectional and time series data. While the parameters were significant, two variables were not sufficient to estimate the 25 "regions" used with accuracy comparable to equations (6) and (7). (See Tables VI-5 and IV-6.)

TABLE 3-5-A

$CM = e^{K} G^{a_1} APL^{a_2}$

| | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|
| Df: | 2/6 | | | 2/5 | | | 2/5 | | |
| R&Y: | NE, ND, NY, PJ, EC, SE, SC, WC, W | | | NE, ND, NY, PJ, EC, SE, SC, W | | | NE, ND, NY, PJ, EC, SE, SC, W | | |
| F: | 34.09/.01 | | | 22.10/.01 | | | 32.53/.01 | | |
| AE: | 21.0% | | | 23.1% | | | 20.6% | | |
| | V | S | t | V | S | t | V | S | t |
| K | -1.164 | | | -1.650 | | | -1.463 | | |
| $a_1$ | 1.065 | .202 | 5.27/.05 | 1.428 | .239 | 5.97/.05 | 1.401 | .177 | 7.93/.02 |
| $a_2$ | .976 | .265 | 3.68/.1 | .864 | .547 | 1.58/.3 | .319 | .293 | 1.09/.5 |

TABLE 3-5-B

| Region | CM Actual Value | CM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 7,452 | 7,032 | 6 |
| ND 70 | 1,397 | 1,552 | 11 |
| NY 70 | 1,600 | 1,417 | 11 |
| PJ 70 | 4,455 | 3,849 | 14 |
| EC 70 | 6,160 | 6,362 | 3 |
| SE 70 | 7,160 | 9,104 | 27 |
| SC 70 | 5,700 | 9,539 | 67 |
| WC 70 | 8,770 | 6,802 | 22 |
| W 70 | 29,280 | 21,420 | 27 |
| NE 80 | 10,855 | 16,622 | 53 |
| ND 80 | 2,580 | 2,886 | 12 |
| NY 80 | 2,310 | 1,965 | 15 |
| PJ 80 | 5,965 | 5,884 | 1 |
| EC 80 | 13,730 | 12,766 | 7 |
| SE 80 | 21,770 | 20,342 | 7 |
| SC 80 | 12,780 | 18,556 | 45 |
| W 80 | 42,910 | 23,672 | 45 |
| NE 90 | 15,610 | 22,267 | 43 |
| ND 90 | 3,370 | 3,776 | 12 |
| NY 90 | 3,595 | 3,289 | 8 |
| PJ 90 | 8,645 | 7,181 | 17 |
| EC 90 | 17,030 | 17,673 | 4 |
| SE 90 | 30,340 | 26,957 | 11 |
| SC 90 | 18,270 | 24,774 | 36 |
| W 90 | 51,970 | 34,222 | 34 |

TABLE 3-6-A

$EM = e^{K} G^{a_1} APL^{a_2}$

| | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|
| Df: | 2/6 | | | 2/5 | | | 2/5 | | |
| R&Y: | NE, ND, NY, PJ, EC, SE, SC, WC, W | | | NE, ND, NY, PJ, EC, SE, SC, W | | | NE, ND, NY, PJ, EC, SE, SC, W | | |
| F: | 27.38/.01 | | | 16.49/.01 | | | 33.14/.01 | | |
| AE: | 20.4% | | | 29.2% | | | 20.2% | | |
| | V | S | t | V | S | t | V | S | t |
| K | -2.095 | | | -2.536 | | | -1.209 | | |
| $a_1$ | 1.159 | .240 | 4.84/.05 | 1.552 | .299 | 5.19/.05 | 1.284 | .158 | 8.10/.02 |
| $a_2$ | .992 | .314 | 3.16/.1 | .893 | .684 | 1.30/.5 | .106 | .263 | .40/.9 |

TABLE 3-6-B

| Regions | EM Actual Value | EM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 3,124 | 3,358 | 8 |
| ND 70 | 618 | 640 | 4 |
| NY 70 | 561 | 608 | 8 |
| PJ 70 | 1,945 | 1,711 | 12 |
| EC 70 | 5,550 | 3,004 | 46 |
| SE 70 | 2,529 | 4,373 | 73 |
| SC 70 | 3,778 | 4,472 | 18 |
| WC 70 | 3,232 | 3,150 | 2 |
| W 70 | 11,996 | 10,474 | 13 |
| NE 80 | 5,331 | 9,581 | 80 |
| ND 80 | 1,328 | 1,409 | 6 |
| NY 80 | 987 | 962 | 3 |
| PJ 80 | 3,016 | 3,092 | 3 |
| EC 80 | 13,818 | 7,227 | 48 |
| SE 80 | 9,385 | 11,914 | 27 |
| SC 80 | 8,260 | 10,670 | 29 |
| W 80 | 22,608 | 13,839 | 39 |
| NE 90 | 13,104 | 16,660 | 27 |
| ND 90 | 2,467 | 2,914 | 18 |
| NY 90 | 3,411 | 3,113 | 9 |
| PJ 90 | 7,226 | 6,229 | 14 |
| EC 90 | 17,058 | 13,946 | 18 |
| SE 90 | 16,154 | 20,536 | 27 |
| SC 90 | 15,011 | 17,624 | 17 |
| W 90 | 31,946 | 22,163 | 31 |

Models of the type of equations (4) and (5) were used to narrow down the number of relevant variables, resulting in equations (6) and (7). With the number of independent variables to be examined decreased, a more general model, of the type of equations (2) and (3), could be developed. The single most significant variable was the generating capacity, G. This makes sense intuitively, as the transmission system is laid out to "serve" the generation system, to be able to deliver power. Appendix III consists of graphs of miles of EHV line vs. generating capacity for the 9 systems used in the study. A strong correlation between the two can be seen.

Thus, it would seem to make sense to talk of grid size depending mostly on G, (i.e., of the function $f_G$ (equation (3)) "dominating" equation (2)). The resulting equation is:

$$M = e^K \cdot \left( \prod_{\substack{i \\ i \neq G}} X_i^{a_i} \right) \cdot G^{b_o + \sum_i b_i X_i} \qquad (12)$$

or, taking logrithms to linearize for the regression:

$$\ln M = K + \sum_{i \neq G} a_i \ln X_i + b_o \ln G + \sum_i b_i (\ln G \cdot X_i) \qquad (13)$$

After eliminating the non-significant variables (Appendix IV, Table IV-1-B), the resulting equations were

$$C\,M = e^{K} A^{a_1} APL^{a_2} G^{a_3} G^{a_4 \cdot METRO} \qquad (14)$$

or $$C\,M = e^{K} A^{a_1} AAPL^{a_2} G^{a_3 \cdot I} G^{a_4 \cdot METRO} \qquad (15)$$

or $$C\,M = e^{K} A^{a_1} AAPL^{a_2} G^{a_3 \cdot METRO} \qquad (16)$$

and $$E\,M = e^{K} A^{a_1} AAPL^{a_2} G^{a_3 \cdot METRO} \qquad (17)$$

or $$E\,M = e^{K} A^{a_1} AAPL^{a_2} G^{a_3} G^{a_4 \cdot METRO} \qquad (18)$$

(The coefficients are given in Tables 3-7 to 3-11.) Equations (14) and (15) give comparable accuracy. Equation (16) is slightly less accurate, but it uses one less independent variable. Similarly, equation (17) is only slightly less accurate than (18), but it has 3 independent variables while (18) has 4. Once again, the coefficients of A, APL, AAPL, G, I and METRO are all positive as expected; each of these characteristics tends to increase the amount of transmission line needed in a system. The results of each model are summarized in Tables 3-7, through 3-11.

TABLE 3-7-A

$$CM = e^{K} A^{a_1} G^{a_2} APL^{a_3} G^{a_4 METRO}$$

Df: 4/20

F: 72.01/.01

AE: 20%

| | V | S | t |
|---|---|---|---|
| K | -.185 | | |
| $a_1$ | .490 | .129 | 3.81/.02 |
| $a_2$ | .424 | .248 | 1.71/.2 |
| $a_3$ | .510 | .195 | 2.62/.1 |
| $a_4$ | .105 | .046 | 2.30/.1 |

TABLE 3-7-B

| Region | CM Actual Value | CM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 7,452 | 7,410 | 1 |
| NE 80 | 10,855 | 12,709 | 17 |
| NE 90 | 15,610 | 21,286 | 36 |
| ND 70 | 1,397 | 1,414 | 1 |
| ND 80 | 2,580 | 2,368 | 8 |
| ND 90 | 3,370 | 3,673 | 9 |
| NY 70 | 1,600 | 1,526 | 5 |
| NY 80 | 2,310 | 2,363 | 2 |
| NY 90 | 3,595 | 4,026 | 12 |
| PJ 70 | 4,455 | 3,020 | 32 |
| PJ 80 | 5,965 | 5,176 | 13 |
| PJ 90 | 8,645 | 8,089 | 6 |
| EC 70 | 6,160 | 6,833 | 11 |
| EC 80 | 13,730 | 11,531 | 16 |
| EC 90 | 17,030 | 20,100 | 18 |
| SE 70 | 7,160 | 9,525 | 33 |
| SE 80 | 21,770 | 15,100 | 31 |
| SE 90 | 30,340 | 22,891 | 25 |
| SC 70 | 5,700 | 10,756 | 89 |
| SC 80 | 12,780 | 15,775 | 23 |
| SC 90 | 18,270 | 24,286 | 33 |
| WC 70 | 8,770 | 8,785 | 0 |
| W 70 | 29,280 | 24,202 | 17 |
| W 80 | 42,910 | 26,140 | 39 |
| W 90 | 51,970 | 44,977 | 13 |

TABLE 3-8-A

$$CM = e^K A^{a_1} AAPL^{a_2} G^{a_3 I} G^{a_4 METRO}$$

Df: 4/20

F: 70.81/.01

AE: 20%

| | V | S | t |
|---|---|---|---|
| K | .219 | | |
| $a_1$ | .639 | .095 | 6.73/.01 |
| $a_2$ | .301 | .143 | 2.10/.2 |
| $a_3$ | .0448 | .0283 | 1.59/.2 |
| $a_4$ | .125 | .036 | 3.52/.05 |

TABLE 3-8-B

| Region | CM Actual Value | CM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 7,452 | 6,465 | 13 |
| NE 80 | 10,855 | 11,321 | 4 |
| NE 90 | 15,610 | 18,673 | 20 |
| ND 70 | 1,397 | 1,560 | 12 |
| ND 80 | 2,580 | 2,553 | 1 |
| ND 90 | 3,370 | 3,361 | 0 |
| NY 70 | 1,600 | 1,526 | 5 |
| NY 80 | 2,310 | 2,439 | 6 |
| NY 90 | 3,595 | 3,973 | 11 |
| PJ 70 | 4,455 | 2,833 | 36 |
| PJ 80 | 5,965 | 5,177 | 13 |
| PJ 90 | 8,645 | 9,434 | 9 |
| EC 70 | 6,160 | 6,720 | 9 |
| EC 80 | 13,730 | 11,725 | 15 |
| EC 90 | 17,030 | 21,399 | 26 |
| SE 70 | 7,160 | 8,803 | 23 |
| SE 80 | 21,770 | 14,422 | 34 |
| SE 90 | 30,340 | 23,926 | 21 |
| SC 70 | 5,700 | 10,736 | 88 |
| SC 80 | 12,780 | 17,346 | 36 |
| SC 90 | 18,270 | 26,416 | 45 |
| WC 70 | 8,770 | 8,960 | 2 |
| W 70 | 29,280 | 22,193 | 24 |
| W 80 | 42,910 | 29,301 | 32 |
| W 90 | 51,970 | 44,257 | 15 |

TABLE 3-9-A

$$CM = e^{K} A^{a_1} AAPL^{a_2} G^{a_3 METRO}$$

| | |
|---|---|
| Df: | 3/21 |
| F: | 87.28/.01 |
| AE: | 22% |

| | V | S | t |
|---|---|---|---|
| K | .304 | | |
| $a_1$ | .756 | .063 | 12.08/.01 |
| $a_2$ | .319 | .148 | 2.16/.2 |
| $a_3$ | .172 | .021 | 8.24/.01 |

TABLE 3-9-B

| Region | CM Actual Value | CM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 7,452 | 6,556 | 12 |
| NE 80 | 10,855 | 11,796 | 9 |
| NE 90 | 15,610 | 20,832 | 34 |
| ND 70 | 1,397 | 1,608 | 15 |
| ND 80 | 2,580 | 2,437 | 5 |
| ND 90 | 3,370 | 3,080 | 9 |
| NY 70 | 1,600 | 1,513 | 5 |
| NY 80 | 2,310 | 2,550 | 10 |
| NY 90 | 3,595 | 4,555 | 27 |
| PJ 70 | 4,455 | 2,811 | 37 |
| PJ 80 | 5,965 | 5,207 | 13 |
| PJ 90 | 8,645 | 10,121 | 17 |
| EC 70 | 6,160 | 6,255 | 2 |
| EC 80 | 13,730 | 10,770 | 22 |
| EC 90 | 17,030 | 20,275 | 19 |
| SE 70 | 7,160 | 8,604 | 20 |
| SE 80 | 21,770 | 12,997 | 40 |
| SE 90 | 30,340 | 20,468 | 32 |
| SC 70 | 5,700 | 11,588 | 103 |
| SC 80 | 12,780 | 16,334 | 28 |
| SC 90 | 18,270 | 22,138 | 21 |
| WC 70 | 8,770 | 9,860 | 12 |
| W 70 | 29,280 | 26,422 | 10 |
| W 80 | 42,910 | 32,698 | 24 |
| W 90 | 51,970 | 43,025 | 17 |

TABLE 3-10-A

$$EM = e^{K} A^{a_1} AAPL^{a_2} G^{a_3 METRO}$$

Df: 3/21

F: 89.00/.01

AE: 24%

| | V | S | t |
|---|---|---|---|
| K | -.936 | | |
| $a_1$ | .728 | .070 | 10.44/.01 |
| $a_2$ | .546 | .165 | 3.31/.05 |
| $a_3$ | .242 | .023 | 10.43/.01 |

TABLE 3-10-B

| Region | EM Actual Value | EM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 3,124 | 3,316 | 6 |
| NE 80 | 5,331 | 7,971 | 50 |
| NE 90 | 13,104 | 18,162 | 39 |
| ND 70 | 618 | 654 | 6 |
| ND 80 | 1,328 | 1,275 | 4 |
| ND 90 | 2,467 | 1,816 | 26 |
| NY 70 | 561 | 621 | 11 |
| NY 80 | 987 | 1,355 | 37 |
| NY 90 | 3,411 | 3,113 | 9 |
| PJ 70 | 1,945 | 1,453 | 25 |
| PJ 80 | 3,016 | 3,560 | 18 |
| PJ 90 | 7,226 | 9,238 | 28 |
| EC 70 | 5,550 | 3,014 | 46 |
| EC 80 | 13,818 | 6,585 | 52 |
| EC 90 | 17,058 | 16,140 | 5 |
| SE 70 | 2,529 | 3,975 | 57 |
| SE 80 | 9,385 | 7,229 | 23 |
| SE 90 | 16,154 | 13,789 | 15 |
| SC 70 | 3,778 | 5,328 | 41 |
| SC 80 | 8,260 | 8,897 | 8 |
| SC 90 | 15,011 | 13,687 | 9 |
| WC 70 | 3,232 | 4,223 | 31 |
| W 70 | 11,996 | 13,920 | 16 |
| W 80 | 22,608 | 18,926 | 16 |
| W 90 | 31,946 | 27,943 | 12 |

TABLE 3-11-A

$$EM = e^{K} A^{a_1} AAPL^{a_2} G^{a_3} G^{a_4 METRO}$$

Df: 4/20

F: 72.44/.01

AE: 21%

| | V | S | t |
|---|---|---|---|
| K | -.999 | | |
| $a_1$ | .563 | .123 | 4.59/.02 |
| $a_2$ | .420 | .177 | 2.37/.1 |
| $a_3$ | .377 | .234 | 1.61/.2 |
| $a_4$ | .163 | .054 | 3.02/.05 |

TABLE 3-11-B

| Region | EM Actual Value | EM Fitted Value | % Error |
|---|---|---|---|
| NE 70 | 3,124 | 3,437 | 10 |
| NE 80 | 5,331 | 8,074 | 52 |
| NE 90 | 13,104 | 17,851 | 36 |
| ND 70 | 618 | 593 | 4 |
| ND 80 | 1,328 | 1,310 | 1 |
| ND 90 | 2,467 | 2,129 | 14 |
| NY 70 | 561 | 664 | 18 |
| NY 80 | 987 | 1,367 | 39 |
| NY 90 | 3,411 | 2,966 | 13 |
| PJ 70 | 1,945 | 1,369 | 30 |
| PJ 80 | 3,016 | 3,271 | 9 |
| PJ 90 | 7,226 | 7,944 | 10 |
| EC 70 | 5,550 | 3,105 | 44 |
| EC 80 | 13,818 | 6,718 | 51 |
| EC 90 | 17,058 | 15,443 | 9 |
| SE 70 | 2,529 | 4,160 | 65 |
| SE 80 | 9,385 | 8,265 | 12 |
| SE 90 | 16,154 | 16,415 | 2 |
| SC 70 | 3,778 | 4,742 | 26 |
| SC 80 | 8,260 | 9,381 | 14 |
| SC 90 | 15,011 | 16,398 | 9 |
| WC 70 | 3,232 | 3,856 | 19 |
| W 70 | 11,996 | 11,226 | 6 |
| W 80 | 22,608 | 17,311 | 23 |
| W 90 | 31,946 | 29,855 | 6 |

As before, year-by-year regressions were done on the equations

$$C\ M = e^{K}\ A^{a_1}\ G^{a_2 \cdot METRO} \tag{19}$$

and $$E\ M = e^{K}\ A^{a_1} \cdot G^{a_2 \cdot METRO} \tag{20}$$

the results of which are given in Tables 3-12 and 3-13. Equations (19) and (20) were used for the entire data set, too, with results similar to those for equations (8) and (9): Two independent variables, though significant, were not as accurate for estimation as equations (14) - (18).

The previous equations have been static ones; that is, they estimate the miles of lines in a system without any knowledge of that system at a prior time. It was desired to determine whether a model for system expansion could be developed. Such a model would use knowledge of the state of the system at time $t_o$ and the values of certain system characteristics (the independent variables) at time $t > t_o$ to predict the miles of line at time t.

TABLE 3-12-A

$CM = e^{K} G^{a_1} APL^{a_2}$

| | 1960 | | | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Df: | 2/6 | | | 2/6 | | | 2/6 | | | 2/6 | | |
| R&Y: | NE, ND, NY, PJ, EC, SE, SC, WC, W | | | Same | | | Same | | | Same | | |
| F: | 2.49/-- | | | 28.10/.01 | | | 52.48/.01 | | | 61.42/.01 | | |
| AE: | 131.0% | | | 22.7% | | | 18.8% | | | 15.9% | | |
| | V | S | t | V | S | t | V | S | t | V | S | t |
| K | -.998 | | | .489 | | | .962 | | | 1.268 | | |
| $a_1$ | .553 | .385 | 1.43/.3 | .751 | .104 | 7.20/.02 | .860 | .084 | 10.21/.01 | .851 | .077 | 11.08/.01 |
| $a_2$ | .839 | .556 | 1.51/.3 | .273 | .089 | 3.05/.1 | .158 | .045 | 3.51/.1 | .107 | .028 | 3.84/.1 |

TABLE 3-12-B

| Region | CM Actual Value | CM Fitted Value | % Error |
|---|---|---|---|
| NE 60 | 2,130 | 3,815 | 79 |
| ND 60 | 266 | 238 | 11 |
| NY 60 | 371 | 392 | 6 |
| PJ 60 | 1,393 | 692 | 50 |
| EC 60 | 2,221 | 2,053 | 7 |
| SE 60 | 1,318 | 1,793 | 36 |
| SC 60 | 143 | 1,364 | 854 |
| WC 60 | 3,417 | 1,265 | 63 |
| W 60 | 13,962 | 3,742 | 73 |
| NE 70 | 7,452 | 8,302 | 11 |
| ND 70 | 1,397 | 1,272 | 9 |
| NY 70 | 1,600 | 1,874 | 17 |
| PJ 70 | 4,455 | 2,990 | 33 |
| EC 70 | 6,160 | 7,273 | 18 |
| SE 70 | 7,160 | 8,386 | 17 |
| SC 70 | 5,700 | 8,792 | 54 |
| WC 70 | 8,770 | 7,930 | 10 |
| W 70 | 29,280 | 19,040 | 35 |
| NE 80 | 10,855 | 12,828 | 18 |
| ND 80 | 2,580 | 2,183 | 15 |
| NY 80 | 2,310 | 2,965 | 28 |
| PJ 80 | 5,965 | 5,059 | 15 |
| EC 80 | 13,730 | 13,160 | 4 |
| SE 80 | 21,770 | 16,102 | 26 |
| SC 80 | 12,780 | 18,435 | 44 |
| WC 80 | 14,780 | 15,341 | 4 |
| W 80 | 42,910 | 37,139 | 13 |
| NE 90 | 15,610 | 16,583 | 6 |
| ND 90 | 3,370 | 3,095 | 8 |
| NY 90 | 3,595 | 4,448 | 24 |
| PJ 90 | 8,645 | 7,208 | 17 |
| EC 90 | 17,030 | 18,658 | 10 |
| SE 90 | 30,340 | 22,307 | 26 |
| SC 90 | 18,270 | 25,612 | 40 |
| WC 90 | 22,060 | 20,626 | 6 |
| W 90 | 51,970 | 48,903 | 6 |

TABLE 3-13-A

$EM = e^{K} A^{a_1} G^{a_2 METRO}$

| | 1960 | | | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Df: | 2/6 | | | 2/6 | | | 2/6 | | | 2/6 | | |
| R&Y: | NE, ND, NY, PJ, EC, SE, SC, WC, W | | | Same | | | Same | | | Same | | |
| F: | 2.50/-- | | | 23.38/.01 | | | 31.29/.01 | | | 162.73/.01 | | |
| AE: | 139.1% | | | 28.6% | | | 22.1% | | | 8.9% | | |
| | V | S | t | V | S | t | V | S | t | V | S | t |
| K | -2.311 | | | -.418 | | | .199 | | | .947 | | |
| $a_1$ | .512 | .394 | 1.30/.5 | .787 | .121 | 6.48/.05 | .925 | .118 | 7.86/.02 | .776 | .043 | 17.91/.01 |
| $a_2$ | .932 | .569 | 1.69/.3 | .315 | .104 | 3.03/.1 | .185 | .063 | 2.94/.1 | .125 | .016 | 7.99/.02 |

TABLE 3-13-B

| Region | EM Actual Value | EM Fitted Value | % Error |
|---|---|---|---|
| NE 60 | 542 | 1,251 | 131 |
| ND 60 | 67 | 64 | 5 |
| NY 60 | 117 | 114 | 2 |
| PJ 60 | 358 | 208 | 42 |
| EC 60 | 1,189 | 622 | 48 |
| SE 60 | 324 | 505 | 56 |
| SC 60 | 38 | 355 | 834 |
| WC 60 | 874 | 328 | 62 |
| W 60 | 3,599 | 1,011 | 72 |
| NE 70 | 3,124 | 4,118 | 32 |
| ND 70 | 618 | 511 | 17 |
| NY 70 | 561 | 818 | 46 |
| PJ 70 | 1,945 | 1,372 | 29 |
| EC 70 | 5,550 | 3,507 | 37 |
| SE 70 | 2,529 | 3,948 | 56 |
| SC 70 | 3,778 | 4,009 | 6 |
| WC 70 | 3,232 | 3,575 | 11 |
| W 70 | 11,996 | 9,244 | 23 |
| NE 80 | 5,331 | 7,474 | 40 |
| ND 80 | 1,328 | 1,020 | 23 |
| NY 80 | 987 | 1,495 | 52 |
| PJ 80 | 3,016 | 2,732 | 9 |
| EC 80 | 13,818 | 7,641 | 55 |
| SE 80 | 9,385 | 9,234 | 2 |
| SC 80 | 8,260 | 10,393 | 26 |
| WC 80 | 8,384 | 8,420 | 0 |
| W 80 | 22,608 | 2,226 | 1 |
| NE 90 | 13,104 | 13,753 | 5 |
| ND 90 | 2,467 | 2,399 | 3 |
| NY 90 | 3,411 | 3,919 | 15 |
| PJ 90 | 7,226 | 6,485 | 10 |
| EC 90 | 17,058 | 15,463 | 9 |
| SE 90 | 16,154 | 16,791 | 4 |
| SC 90 | 15,011 | 17,706 | 18 |
| WC 90 | 16,423 | 13,909 | 15 |
| W 90 | 31,946 | 32,251 | 1 |

Taking the derivative of equation (4),

$$\frac{dM}{dt} = e^{K} \cdot \sum_i a_i \frac{\dot{X}_i}{X_i} \left(\prod_j X_j^{a_j}\right) \qquad (21)$$

By approximating $\dot{X}$ by $X_t - X_o$ and dividing by $M_o$, equation (19) becomes

$$\frac{M_t - M_o}{M_o} = \sum_i a_i \left( \frac{X_{it} - X_{io}}{X_{io}} \right) \qquad (22)$$

Similarly, the derivative of equation (12) was taken

$$\frac{dM}{dt} = e^{K} \left[ \sum_i a_i \frac{\dot{X}_i}{X_i} \left(\prod_j X_j^{a_j}\right)\right] \cdot G^{\sum_i b_i X_i} +$$

$$\left(\prod_i X_i^{a_i}\right) \left[ \left(\sum_i b_i X_i\right) G^{\sum_i b_i X_i} \cdot \frac{\dot{G}}{G} + (\ln G) \cdot G^{\sum_i b_i X_i} \cdot \left(\sum_i b_i \dot{X}_i\right)\right]$$

(23)

Again, approximating $\dot{X}$ by $X_t - X_o$ and dividing by equation (12) gives

$$\frac{M_t-M_o}{M_o} = \sum_i \left(a_i \frac{X_{it}-X_{io}}{X_{io}}\right) + \frac{G_t-G_o}{G_o}\left(\sum_i b_i X_{it}\right) + \ell n G \sum_i b_i (X_{it}-X_{io}) \qquad (24)$$

Regressions were performed on these equations (Table IV-I-C); however, no set of independent variables was found to be statistically significant.

The average % error in estimating the amount of line added to the system in the previous 10 years varied from 67% to 52%, with F-values from 0.1 to 1.5 for equations for EM. Some equations for CM did have significant F-values, such as

$$\frac{CM_t-CM_o}{CM_t} = K + a_1 \frac{AAPL_t-AAPL_o}{AAPL_o} + a_2 \frac{METRO_t-METRO_o}{METRO_o} + a_3 \frac{P_t-P_o}{P_o} + a_4 \frac{I_t-I_o}{I_o} \qquad (25)$$

with an F-statistic of 11.70. For equation 25, the regression results are given in Table 3-14. However, the average % of error was 38%, the same percentage as for other equations with low F-values. The % of error is about the same (11%) for total line in the system as for the "static" models used earlier (15%). No significant improvement was made in model accuracy by these "expansion" model equations.

TABLE 3-14-A

$$\frac{CM_t - CM_o}{CM_o} = K + a_1 \frac{AAPL_t - AAPL_o}{AAPL_o} + a_2 \frac{METRO_t - METRO_o}{METRO_o} + a_3 \frac{P_t - P_o}{P_o} + a_4 \frac{I_t - I_o}{I_o}$$

Df: 4/11
F: 11.70/.01
R&Y: (NE, ND, NY, PJ, EC, SE, SC, W) 80, 90

AE in Line Added in last 10 years: 38.0%

AE in Total Line: 10.5%

| | V | S | t |
|---|---|---|---|
| K | - .508 | | |
| $a_1$ | - .549 | .325 | 1.69/.2 |
| $a_2$ | 2.391 | 1.285 | 1.86/.2 |
| $a_3$ | 1.626 | .444 | 3.66/.05 |
| $a_4$ | -3.537 | 1.517 | 2.33/.1 |

TABLE 3-14-B

| Region | CM Actual Value (Line Added) | CM Fitted Value (Line Added) | % Error (Line Added) | CM Actual Value (Total Line) | CM Fitted Value (Total Line) | % Error (Total Line) |
|---|---|---|---|---|---|---|
| NE 80 | 3,403 | 3,647 | 7 | 10,855 | 11,099 | 2 |
| NE 90 | 4,755 | 2,984 | 37 | 15,610 | 13,839 | 11 |
| ND 80 | 1,183 | 1,204 | 2 | 2,580 | 2,601 | 1 |
| ND 90 | 790 | 955 | 21 | 3,370 | 3,535 | 5 |
| NY 80 | 710 | 612 | 10 | 2,310 | 2,242 | 3 |
| NY 90 | 1,285 | 932 | 27 | 3,595 | 3,242 | 10 |
| PJ 80 | 1,510 | 2,664 | 76 | 5,965 | 7,119 | 19 |
| PJ 90 | 2,680 | 1,330 | 50 | 8,645 | 7,295 | 16 |
| EC 80 | 7,570 | 8,059 | 7 | 13,730 | 14,219 | 4 |
| EC 90 | 3,300 | 6,382 | 93 | 17,030 | 20,112 | 18 |
| SE 80 | 14,610 | 11,132 | 24 | 21,770 | 18,292 | 16 |
| SE 90 | 8,570 | 10,910 | 27 | 30,340 | 32,680 | 8 |
| SC 80 | 7,080 | 7,590 | 7 | 12,780 | 13,290 | 4 |
| SC 90 | 5,490 | 7,007 | 28 | 18,270 | 19,787 | 8 |
| W 80 | 13,630 | 23,363 | 71 | 42,910 | 52,643 | 23 |
| W 90 | 9,060 | - 1,755 | 119 | 51,970 | 41,155 | 21 |

One reason for this may be the approximation of the derivative of the variables. A derivative was approximated by

$$\dot{X} \approx (X_t - X_o)/(t-t_o)$$

where (t-0) is 10 years; in a 10 year period, most systems approximately double their generating capacity and grid size. Thus, the incremental change a derivative describes is not what is being described. If the growth rate is exponential (as in curve b, figure 3-1) then the derivative

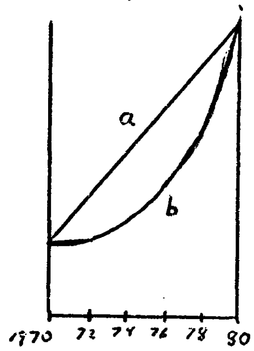



fig. 3-1

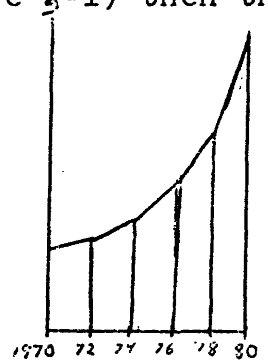



fig. 3-2

approximation for a 10 year $\Delta t$ (curve a, figure 3-1) may be quite wrong. Using a smaller time period, such as 2 years, will result in a much more realistic approximation (figure 3-2).

A workable system expansion model may result from regressing ($Miles_t$ - $Miles_o$) for two-year intervals and, to estimate the state of the system 10 years from the present, the "output" or estimate for the equation for $(M_t - M_o)$ could be used as input for the equation for $(M_{t+1} - M_t)$. This will, of course, necessitate more detailed predictions for growth of generating capacity, number of plants, plant sites, load classification, metropolitan area, etc. The result would be a more dynamic model of transmission expansion as compared to the relatively static models developed here. The results of these "expansion" models will be discussed in greater detail in Chapter 4.

# CHAPTER 4

## Limitations of Regression Models

Regression anslysis can be a useful tool, but the models developed here do have shortcomings and limitations. The most important limitation of a regression model has already been mentioned: it is no proof of causality. The correlation of two variables may be the result of causality, coincidence, or common correlation with a third variable. A regression coefficient has no "sign"; even in a causal relationship it does not indicate which variable is the dependent one. This must be inferred from theory and knowledge of the system being modeled.

Even "causal" correlations, however, may have many factors. "METRO" influences the size of the EHV grid because of its connection with land use and availability, load characteristics, or the "energy-intensiveness" of a region. These factors can not be separated out. The fact that the variable METRO affects the number of miles of EHV line in a system can provide insight into what determines the size of the EHV grid of a region, but it does not give a definitive causal relationship.

The data used in the regressions was both cross-sectional and time-series data. As was mentioned previously, regressions using data for only one year were done for the more important variables. It was seen that such variables as G, APL, and METRO were indeed "significant" and not just related to CM and EM through a common correlation with time. (In addition, regressions were performed using the <u>1970</u> value of METRO for all years, and METRO was still found to be significant.) AAPL is much more time-correlated than APL, but both have about the same significance and result in models with similar parameter values. However, in all the models, the parameters do reflect a certain amount of time correlation which can not be separated from their other casual relationship to miles of EHV line.

Factors not included in a model can drastically change the results. For example, EHV line was not used in the South Central region until later than most of the rest of the country; 115 KV and 161 KV line continued to handle the majority of bulk power transmission. Thus, predictions for the region "SC 60" are much higher than the actual EHV mileage. The effects of this "low" data point on the model are shown in tables 4-1, and 4-2, where the same regression equation, $CM = e^{K} A^{a_1} G^{a_2 \text{ METRO}}$, was used for data with and without the SC 60 region. The

TABLE 4-1

Model: $EM = e^{K} A^{a_1} G^{a_2 METRO}$

Regions & years used: a) (NE, ND, NY, PJ, EC, SE, SC, WC, W) 60,70,80,90
b) Same as (a) except for SC 60
c) (NE, ND, NY, PJ, EC, SE, SC, WC, W) 70,80,90

| | | | |
|---|---|---|---|
| Degrees of freedom*: | a) 2/33 | b) 2/32 | c) 2/24 |
| F-Statistic/Significance*: | a) 34.45/.01 | b) 59.05/.01 | c) 85.18/.01 |
| Average % of Error: | a) 189.6% | b) 67.3% | c) 31.5% |

| | Coefficient | Value | Standard Deviation | t-value/$\alpha$* |
|---|---|---|---|---|
| a) | K | -1.077 | | |
| | $a_1$ | .855 | .160 | 5.34/.05 |
| | $a_2$ | .399 | .060 | 6.71/.05 |
| b) | K | - .887 | | |
| | $a_1$ | .943 | .117 | 8.03/.02 |
| | $a_2$ | .359 | .044 | 8.18/.02 |
| c) | K | - .032 | | |
| | $a_1$ | .867 | .075 | 11.60/.01 |
| | $a_2$ | .235 | .029 | 8.19/.02 |

TABLE 4-2

Model: $CM = e^{K} A^{a_1} G^{a_2 METRO}$

Regions & years used: a) (NE, ND, NY, PJ, EC, SE, SC, WC, W) 60,70,80,90
b) Same as (a) except for SC 60
c) (NE, ND, NY, PJ, EC, SE, SC, WC, W) 70,80,90

| | | | |
|---|---|---|---|
| Degrees of freedom*: | a) 2/33 | b) 2/32 | c) 2/24 |
| F-Statistic/Significance:* | a) 32.31/.01 | b) 77.81/.01 | c) 116.97/.01 |
| Average % of Error: | a) 140.5% | b) 47.5% | c) 25.7% |

| | Coefficient | Value | Standard Deviation | t-value/$\alpha$* |
|---|---|---|---|---|
| a) | K | .040 | | |
| | $a_1$ | .826 | .133 | 6.20/.05 |
| | $a_2$ | .293 | .050 | 5.92/.05 |
| b) | K | .214 | | |
| | $a_1$ | .907 | .087 | 10.47/.01 |
| | $a_2$ | .256 | .032 | 7.91/.02 |
| c) | K | .830 | | |
| | $a_1$ | .835 | .058 | 14.52/.01 |
| | $a_2$ | .168 | .022 | 7.62/.02 |

* See Appendix V for definition of these terms

large errors when including SC 60 "weaken" the entire model and bias all other regions' predictions. This comes about because regression, using a least squares cost criterion, penalizes the rest of the data points a little in order to prevent the error due to incorrectly estimating SC 60 being too large. This large error comes about because no provision for the late introduction of EHV line, a technological factor, was incorporated in the model.

Mine-mouth coal plants are another important factor not included in the models. "Improvements in high voltage transmission generally tend to improve the competitive position of coal"[1]. This facilitates the development of inaccessible or mine-mouth reserves. The amount of coal-fueled generating capacity was not incorporated in the model due to lack of data. (There were no "% generating capacity using coal" predictions for 1980 and 1990 for most of the regions.) However, further research along these lines should attempt to include coal use.

As was mentioned, a correlation between two numbers may be a combination of causality, coincidence, and correlation with a third variable. Thus, the coefficients for the same variables in tables 3-1 and IV-13 are different, although the 2 regression equations,

$$CM = e^K A^{a_1} P^{a_2} I^{a_3} G^{a_4} METRO^{a_5} H^{a_6} AAPL^{a_7} \tag{26}$$

and

$$CM = e^K A^{a_1} P^{a_2} I^{a_3} G^{a_4} METRO^{a_5} H^{a_6} APL^{a_7} \tag{27}$$

differ only by using AAPL or APL. For example, the value of $a_2$, the coeffieient for P, the number of plants, is -.471 in equation (26) and -.586 in equation (27). These two values of $a_2$ reflect varying degrees of the three "factors" of a correlation, and using APL instead of AAPL in the equation changes the "proportions" of those factors. Thus, in comparing two systems or strategies, the -.471 and -.586 coefficients are both equally valid to use in determining the effect of changing the number of generation sites, (as in building a power park instead of 7 or 8 separate plants). An examination of the results of regressions show that one exact number can not be assigned to each variable; you can not say with certainty that, for instance, CM is proportional to $P^{-.471}$. However, the variance of the coefficients is such that you can say CM is proportional to $P^{-X}$, where X is between .45 and .60; in other words, you can get an approximate relationship or a range of values; (i.e., saying, for instance,"changing the number of plants from 40 to 50 will change the amount of EHV line needed by between 5 to 7%.")

The "expansion" type models did not provide any improvement over the more "static" models. This may be because the 10 year period used was too large. Possibly the information gained in knowing the line mileage 10 years previously was offset by the inaccurate derivative approximation. The models for EM (the best is given in table IV-20) did not have significant F-values. Even with the models for CM, the difference in accuracy (Average %Error) between equations with high and low F-values was small.

These models were dependent on the structure already postulated for the models:

$$M = K \cdot \prod_i X_i^{a_i} \tag{28}$$

and

$$M = K \cdot G^{\sum_j b_j X_j} \cdot \prod_i X_i^{a_i} \tag{29}$$

A derivative approximation of these equations was used for the expansion models. Cutting the 10 year $\Delta t$ into 2 year periods may improve the derivative approximation, but an entirely different type of expansion model, such as

$$\frac{M_t}{M_o} = K \cdot \prod_i \left( \frac{X_{ti}}{X_{oi}} \right)^{a_i} \qquad (30)$$

may give better results.

The models of equations 1 through 20 were multiplicative. Taking the logrithms of these equations gives equations linear in logs of the variables, and therefore suitable for multiple linear regression. A more complicated, non-linear model is possible, but another computational tool would have to be used to estimate its parameters. On the basis of the research so far, there is no other form which, _à priori,_ appears to be better than a multiplicative model. However, it is possible that a better form exists. For the expansion type models aimed at determining the additions to the existing grid, models not of the "derivative" form of equations 22 and 24, such as equation 30, should be investigated.

# CHAPTER 5

## Applications and Uses of Models Developed

This study has shown that predictions can be made of the EHV transmission grid size of a system from information about other system characteristics. Accuracy is almost always to within 50% and usually within 25% of the actual values. A relatively small number of system variables were found to be relevant. Changes in technology, laws, and siting strategies can be related to these variables (generating capacity, number of plants, average plant to load center distance, etc.). Together with predictions on the character of the load (% industrial) and the characteristics of the area served (% metropolitan area), the size of the transmission grid can be predicted.

As mentioned before, however, this study is not an attempt to decide where the transmission line goes, so its "absolute" numbers of EHV line miles are not accurate enough for a "hard" prediction to be made. The models can give estimates to within a factor of two, but it is as a "relative" comparison or ranking that they may be most useful. The models quantify some of the tradeoffs involved in system planning. Thus, the economies of scale of building larger plants (and, therefore, fewer total plants) can be compared with the cost of the extra transmission needed (equation (6) and table 3-1). To meet air quality standards, the difference in transmission between continuing the present generation siting strategy (using low sulfur fuel or air scrubbing equipment) and a new policy of locating the power plants far from the cities or using offshore nuclear plants (thus increasing APL) can be examined.

Specifically, a power park, in reducing the number of plant sites, is expected to increase the amount of transmission line needed. Using equation (6) and table 3-1, a system with, for example, 60 plants in 1980 may decide that its generation needs in 1990 can be met by building 20 new plants or 9 plants and a power park (with a capacity of about 11 plants.*) From table 3-1, comes the relationship:

$$\frac{CM_{strategy^1}}{CM_{strategy^2}} = \left( \frac{80}{70} \right)^{-.471} \approx .94 \qquad (31)$$

or 6% more line will be needed for the power park. In addition, the effect on the average plant to load center

*These numbers (one power park ≈ 11 separate plants, etc.) are used for illustrative purposes and are not necessarily indicative of the actual size of a power park.

distance (APL) must also be included. From table 3-1 comes:

$$\frac{CM_{strategy^1}}{CM_{strategy^2}} = \left(\frac{P_1}{P_2}\right)^{-.471} \left(\frac{AAPL_1}{AAPL_2}\right)^{.341} \qquad (32)$$

Chapter 4 brought out the fact that table IV-13 (equation (27)) gives -.586 as a value for the coefficient of P. Using equation (27) for the basic model, equation (32) becomes

$$\frac{CM_{strategy^1}}{CM_{strategy^2}} = \left(\frac{P_1}{P_2}\right)^{-.586} \left(\frac{APL_1}{APL_2}\right)^{.392} \qquad (33)$$

However, the difference between $(P_1/P_2)^{-.471}$ and $(P_1/P_2)^{-.586}$ is .94 and .925. Thus, the "extra" line needed for a power park is between 6 and 8%, the range given by using equations (26) and (27). In examining any alternative, "hard" predictions can not be made, but a range of possible effects (as in "about 6%" or "10 to 15%", etc.) can be determined.

The effect of the growing scarcity of acceptable power plant sites on P and APL can be estimated. Then the resulting effect on EHV line needs can be calculated using the preceding procedure.

In similar manner, the effect of building offshore nuclear plants on the plant to load center distances can be determined and the relative "cost" in transmission line estimated. However, the transmission line carrying power from an offshore plant may have relevant characteristics not included in this study. Because of special reliability problems, several lines may be laid. A variable, "% of capacity in offshore plants", may be what is lacking to describe the special characteristics of these plants, as "% hydroelectric capacity", (H), was used in this study to indicate the role of hydroelectric plants.

The results of this study are limited in that offshore plants and power parks were not in the data used. Thus, a relationship which may be valid for the tradeoff between 3 400 MW plants or 2 600 MW plants may not be valid for the tradeoff between 15 600 MW plants or a 6000 MW power park and 3 1000 MW plants. The 6% increase in line from the example of using the power park strategy should be taken with "a grain of salt". The 6 to 8% difference between using a power park or not is a

<u>hypothesis</u> resulting from this model. Extensive analysis of system alternatives, such as planning the system with and without power parks and comparing the two, will be needed to determine the validity of this application of the model.

# CHAPTER 6

## Conclusions

The models developed in this paper are summarized in Table 8. While the limitations of regression models must be kept in mind, the equations of Table 8 do represent relationships for estimating EHV transmission line needs using a relatively small number of system characteristics as independent variables.

In addition to determining a region's EHV line needs, these models can also be used to evaluate different generation construction and siting strategies or predict the effect of trends in technology, site availability, or environmental constraints. The effect of a particular "strategy" or "trend" on the independent variables of the models can be estimated, and the difference in transmission line needs computed from the changes in those variables. However, it must be kept in mind that there may be other relevant variables (% coal generation, number of offshore nuclear plants, etc.) which are not included in these models.

Two other aspects of transmission lines not covered in this research are reliability and cost. Reliability data was simply not adequate enough to assign a reliability "value" to a transmission system. If such a number could be developed, however, the information on the tradeoff between reliability and line mileage could be quite useful.

Costs of transmission line, both right-of-way and construction costs, depend quite a bit on the characteristics of the surrounding area. A study similar to this one attempting to estimate EHV line cost may be fruitful. The degree of settlement or urbanization of the area in which the line is built would be expected to be the most significant parameter of such a model.

In general, the models developed here show that different plant siting strategies don't have radically different effects on the EHV needs of a region. Things "beyond the control" of the system planner, such as load size and characteristics, metropolitan area, etc. seem to be more important. A more useful tool might be an improved expansion model, concerned more directly with the <u>additions</u> to the power system and the <u>changes</u> in load and regional characteristics.

Further research along this line could consist of examining other possibly significant variables, developing a model concerned explicitly with additions and expansion of an existing transmission grid, developing a model for transmission line costs, or exploring other forms of a transmission model than the multiplicative ones developed here and using other curve fitting techniques to compute the model parameters.

# APPENDIX I

## Description of EHV Transmission Line

EHV line, for this paper, is defined as transmission line with a rated nominal voltage of 230 KV or higher. The National Standards Institute, Inc., has established standard voltage levels for transmission systems. Standard nominal voltages include 230 KV, 345 KV, 500 KV, and 765 KV. Associated with each standard nominal voltage is a standard maximum voltage, the designed limit at which the system can operate.[1]

EHV construction rapidly increased starting about 1957, as 345 KV line was introduced, followed by 500 KV line in 1964 and 765 KV line in 1968. The amount of EHV line in the U.S. for the years 1940 to 1990 is given in table 2. Figure 1 illustrates the rise in 345, 500, and 765 KV line mileages. Research to develop higher rated voltages is currently underway.[2]

As the nominal voltage of a line increases, the power carrying capability of that line grows approximately as the square of the voltage.[3] Construction costs and right-of-way requirements also increase with voltage. These characteristics are summarized in tables 1 and 3. Construction and right-of-way costs vary greatly according to the existing uses of the land, as shown in table 4;

right-of-way costs are as much as $60,000 per mile in urban areas and as low as $2,000 per mile in desert areas. Construction costs are higher in urban areas, too, because of the larger number of towers required, increased clearance requirements, foundation problems, and aesthetic considerations.[4] Table 4 is meant to be illustrative, and no attempt to estimate costs of line is made in this paper. Representative costs are roughly proportional to rated voltage for lines 230 KV and above.[5]

The maximum plant size is constantly increasing, from 208 MW in 1950 to 1068 MW in 1969[6], and that growth is continuing. As generation is concentrated in one place, so transmission is being concentrated in large corridors.[7] One 765 KV line has the power capacity of 5 345 KV or 30 138 KV lines. The equivalent reactance of a 765 KV line is about 1/5 that of a 345 KV line, effectively cutting the "electrical" distance between the ends of the line by 1/5 and thus decreasing resistance losses. Also, while one 765 KV line requires about 200 feet of right of way, 5 345 KV lines require 750 feet, and 30 138 KV lines require 3000 feet.[8] Thus, with the increase in the amount of electric generation in the U.S., and with land available for power facilities becoming scarcer, EHV line is becoming more and more widely used.

Presently, EHV line is used in 3 cases:

1. Long distance energy transfer from remote generating sources to load centers.

2. Interconnection between areas previously isolated from each other to achieve economies in utilization of available generation resources.

3. Higher voltage overlay of an existing well-developed lower-voltage system so as to allow such an overlay to take over the major tasks of bulk power transfer between generating points and load centers within the system and to permit the continued integrated operation of the overall system in an economical and reliable manner.[9]

An example of the first use is the 735 KV line of Hydro Quebec to deliver power from the planned northern hydro-electric plants to load centers 380 miles away in Quebec City and Montreal. The second type of use, interconnection, is exemplified by the 850 mile $\pm$ 400 KV DC and 500 KV AC lines comprising the Pacific Northwest - Southwest intertie. AEP's 1200 circuit mile 765 KV grid overlays an extensive existing 345 KV grid in the east central region.

# APPENDIX II

## Data: Description and Sources

The variables chosen for the regression models are described in this section. The data used to estimate the model parameters are listed in table 5. Data was assembled for the years 1960, 1965, 1970, 1980 and 1990. However, data for each year was not used in all the regression models.

Two measures of EHV transmission line mileage were used as dependent variables. The first, circuit miles of EHV line (CM), is the total number of circuit miles of line at a nominal voltage of 230 KV and above. The second measure is that of the power carrying capability of the line. The capacity of a line increases as the square of the rated voltage. Thus, the number of miles of each voltage class of line was multiplied by a load factor. NEMA[1] load scale factors were used to determine the "gigawatt" equivalent miles" of EHV line (EM). The load factors for each voltage class are given in table 1. For example, one mile of 230 KV line quals 0.25 GW-equivalent miles, one mile of 500 KV line equals 1.2 GW-equivalent miles, etc. Data for the years 1970, 1980 and 1990 were obtained from the 1970 National Power Survey. Data for 1960 was obtained from the FPC Statistics of Privately Owned Electric Utilities in the U.S., Statistics of Publicly Owned Electric Utilities in the U.S., Electrical World's Directory of Electric Utilities, the 1964 FPC National Power Survey, and FPC "Principal Electric Facilities" maps.

The population of each region was taken from the 1960 and 1970 Censuses and the 1970 National Power Survey. The regional advisory committees for the most part assumed a 1.5% per year growth of population, and this was used to project future population for 1980 and 1990.

The physical area (A) of a system may have a large effect on grid size. In a larger area, generation sites and load centers are usually scattered more widely, and there are usually longer plant to load center distances.

The generating capacity (G) of a system obviously is an indication of the power needs of that region. Since the purpose of transmission line is to transfer and distribute the generated power, the generating capacity was expected to be a significant input to the model. Appendix III contains graphs of generating capacity vs. miles of line for the regions studied. A strong correlation is seen to exist between them.

Sources for data on generating capacity were the FPC Statistics of Privately Owned Electric Utilities in the United States - 1960 and Statistics of Publicly Owned Electric Utilities in the United States - 1960 for the year

1960. The generating capacity for 1965 and 1970 and projections for 1980 and 1990 were taken from the 1970 National Power Survey.

The number of generating plants and the number of load centers were investigated as possible model inputs. Because the study was concerned with EHV line, only large plants and load centers were considered, as small plants and loads might only require 115 KV or lower line. The variables used were the number of plants over 400 MW (P), the number of load centers over 500 MW peak load, (LC-500), and the number of load centers over 1000 MW peak load, (LC-1000). These sizes were used mainly to facilitate data collection, as the FPC National Power Survey maps included those classifications. Data for 1965, 1970, 1980 and 1990 were taken from the 1970 National Power Survey, from plant and load listings for regions when available and from regional maps when not. The number of load centers for the East Central region in 1960 was also given in the National Power Survey.

The West Central region did not include a plant list or map, but rather a list of "generation areas". These included zero, one, or several plants with capacities over 400 MW. The capacity for each generation area was broken down to peaking units, hydroelectric units, fossil units, and nuclear units. Where any of these subdivisions was in excess of 400 MW, a plant over the minimum size was assumed to exist. This approximation yielded fairly accurate results for 1970, the only year it could be checked, as the above method yielded 25 plants, and the 1971 FPC Principal Electric Facilities maps showed 28 plants in July, 1970.

The load classification was also investigated for possible significance. Of the yearly generation (in gigawatt hours) actually delivered to the customer (losses were not included) the percentage in each of three classification was computed. Those classifications were:

% Industrial ( I ), including industrial use and electric transportation;

% Residentual and Rural ( R ), including residential, rural, agricultural (irrigation) and street and highway lighting; and

% Commercial (C), consisting of commercial use.

All three variables were not used at the same time in a regression, since any one can be expressed as a function of the other two. This linear dependence would cause the data matrix to be singular. Since matrix inversion is used in multiple regression, a zero determinant would result. Thus, only the variables R and I were used.

Sources for the data were the 1970 National Power Survey for the years 1965, 1970, 1980 and 1990 and EEI's Statistical Year Book of the Electric Utility Industry for the year 1960.

Another characteristic of each region which was examined was the generation mix of the system. There are a number of justifications for this:

1. Hydroelectric and pumped storage plants tend to be located farther from major load centers than fossil plants simply because the necessary water resources and land required to develop them are seldom found near cities.

2. Because of public opinion, nuclear power plants are usually not located in metropolitan areas because of the chance of a radiation accident. Thus, they are often farther from the load centers than conventional plants.

3. Nuclear plants are often being planned in conjunction with pumped hydro-peaking plants. The pumped hydro plant is run by the nuclear plant at night. Thus, there must be a strong tie between the two. More EHV line is often required. Also, the nuclear plant may be located near the pumped storage plant, and thus farther from the load center.

Information on the generation mix is in the form of _% of capacity fossil fueled (F)_, _% of capacity hydroelectric (H)_, and _% of capacity nuclear (N)_. The makeup of generation additions (per cent F, H, and N) in the 1971-1980 and 1981-1990 periods was also calculated. Sources for the data were the FPC's 1970 _National Power Survey_ for the years 1970, 1980, and 1990 and EEI's _Statistical Year Book for 1960_ and _1965_. As with the load characteristic variables (I, R, and C), any variable can be determined if the other two are known. Because of this linear dependence, only 2 of these 3 variables could be used in any regression. Also, since N was zero for many cases, it was not used as an independent variable in models where its logrithm would be used, as its log would be infinite.

Urban areas provide unique problems in EHV line construction. In addition to the ever-present problem of minimizing environmental impact, the right-of-way and construction costs are highest in urban areas, with typical costs of $330,000/mile line near an urban center as opposed to $126,000/mile in a desert area.[2] Thus, the percent of land area occupied by urban settlements was investigated as a possible parameter for the model.

The variable used was % of area of a region that is included in the Census Bureau's Standard Metropolitan Statistical Areas with over 100,000 population. The definition of percent metropolitan area (METRO) is arbitrary, but was designed to be more easily and objectively measured and predicted. The census defines 243 Standard Metropolitan Statistical Areas (SMSA's) in the U.S. in 1970. Except in New England, a SMSA is a county or group of contiguous counties which contain at least one city of 50,000 or more or "twin" cities with a combined population of 50,000 or more. Contiguous counties are included if they are socially and economically integrated with the central city. In New England, SMSA's are towns and cities instead of counties. Urbanized areas are the thickly settled core of SMSA's. However, the land area of urban areas was not given in the census, so SMSA's were used. A further qualification imposed for this study in order to help "weed out" the "less urban" SMSA's was that a SMSA, to be included, must have a population of 100,000 or more. Source for the 1960 data was the 1960 Census of Population, Volume 1: Characteristics of the Population, Part 1: U.S.Summary, tables 34 and 36. 1970 data was taken from the 1970 Census of Population, Volume 1: Number of Inhabitants, Part 1: U.S. Summary, tables 35 and 36. Projections for 1980 and 1990 were made by linear extrapolation of the 1960 and 1970 values. Projections were also made by calculating the yearly growth rate from 1960 to 1970 and assuming that rate of growth would continue. The yearly growth rates calculated were:

| | |
|---|---|
| North East | 1.9% |
| New England | 1.9% |
| New York | 2.2% |
| P J M | 1.7% |
| East Central | 3.9% |
| South East | 3.8% |
| South Central | 4.1% |
| West Central | 2.2% |
| West | 1.6% |

Regressions were done using both the linear and exponential projections, and no significant difference was found in the resulting models. A comparison of two such regressions is given in table 7. For the other models presented in this paper, the linear projections for 1980 and 1990 were used.

Because transmission line's ultimate use is to bring the power from the plant to the load center, two variables dealing with plant to load center distance were used. The first is the <u>average plant to load center distance (APL),</u> and is the average distance between plants of over 400 MW capacity and the nearest load center with over 1000 MW peak load. The second variable, total plant to load center distance (TPL), is the sum of the distances between plants over 400 MW and load centers over 1000 MW.

Because the number of load centers of over 1000 MW is increasing, the yearly average plant to load center distance for both new and old plants is decreasing. However, it was desired to compare the APL and TPL of existing plants with those of new plants, to see if new plants were being located in relatively more or less remote areas than existing plants. Thus, for the year 1980, for instance, the APL and TPL for plants constructed before 1970 was calculated, as was the APL and TPL for plants between 1970 and 1980. The APL for "old" plants in 1980 (Old APL) was compared with the APL for 1970, and all 1980 values were scaled up accordingly to give adjusted APL (AAPL) and adjusted TPL (TAPL). Thus, AAPL and TAPL represent values scaled to 1970. The relationship used was:

$$AAPL_{1980} = \frac{APL_{1970}}{\text{Old } APL_{1980}} \times APL_{1980}$$

$$AAPL_{1990} = \frac{AAPL_{1980}}{\text{Old } APL_{1990}} \times APL_{1990}$$

The values of APL and TPL were obtained by measuring the distance, on maps in the 1970 National Power Survey, between plants and load centers. In many cases, this necessitated making a composite map showing both plants and load centers. The West Central region, as was mentioned previously, did not include a plant list, but rather a "generation area" list. Thus, actual plant sites could not be located to obtain the values of APL and TPL for the West Central region; the 1971 "Principal Electric Facilities" maps were used. Thus, the only values of APL and TPL for the West Central region is for the year 1971.

For the East Central region, some generation addition sites weren't specified exactly, but an area in which they would be located was given. In this case, the center of the proposed area was used as the plant location.

# APPENDIX III

## Graphs of Generating Capacity vs. EHV Grid Size

The graphs in this section are of circuit miles of EHV line (CM) and gigawatt-equivalent miles of EHV line (EM) vs. gigawatts of generating capacity (G).

Nine regions of the country are plotted, and in graphs 1 through 10 the symbol for each section is used for the plot. The regions and symbols used are:

1. Northeast
   - A. New England
   - B. New York
   - C. P J M

   (A–C: Subregions of the Northeast)
2. East Central
3. Southeast
4. South Central
5. West Central
6. West

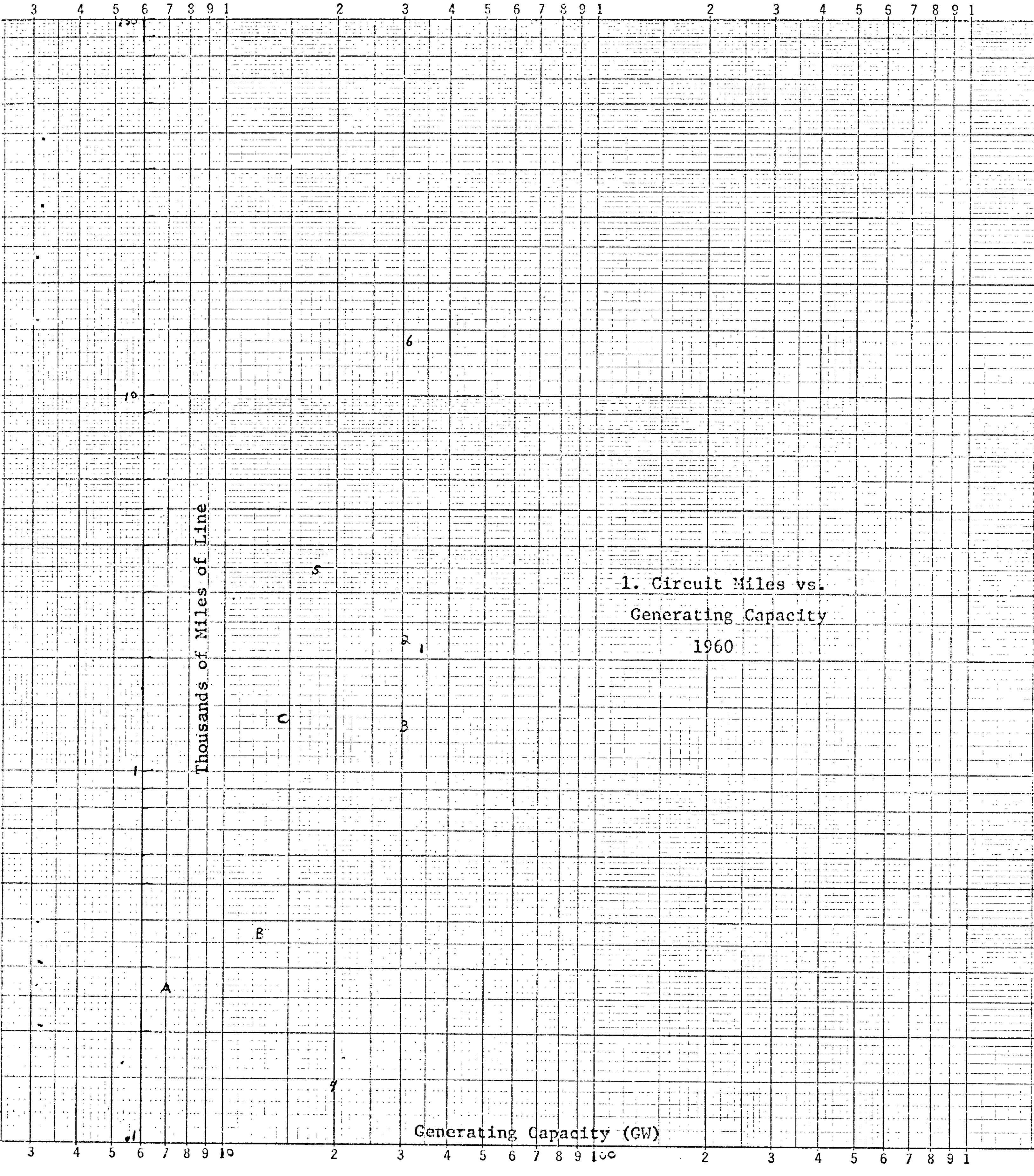
The graphs are log-log graphs with thousands of miles of line on the vertical axis and gigawatts on the horizontal axis. For each region, there are 4 points on graphs 9 through 19. These points are for the years 1960, 1970, 1980 and 1990, with CM, EM and GW all increasing with time. Graphs 11 through 19 use the following notation:

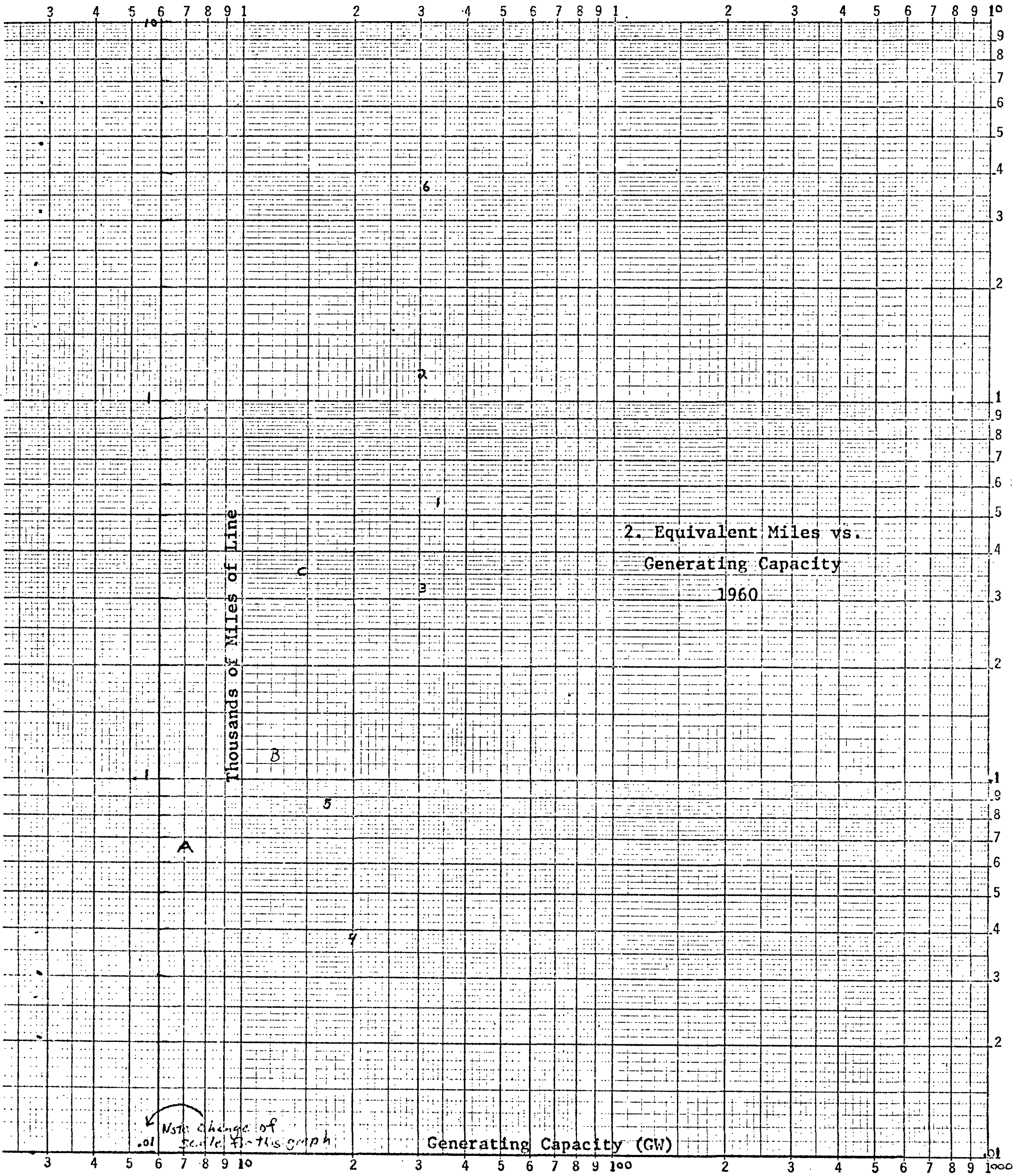
• = thousands of circuit miles of line (CM)

× = thousands of gigawatt miles of line (EM)

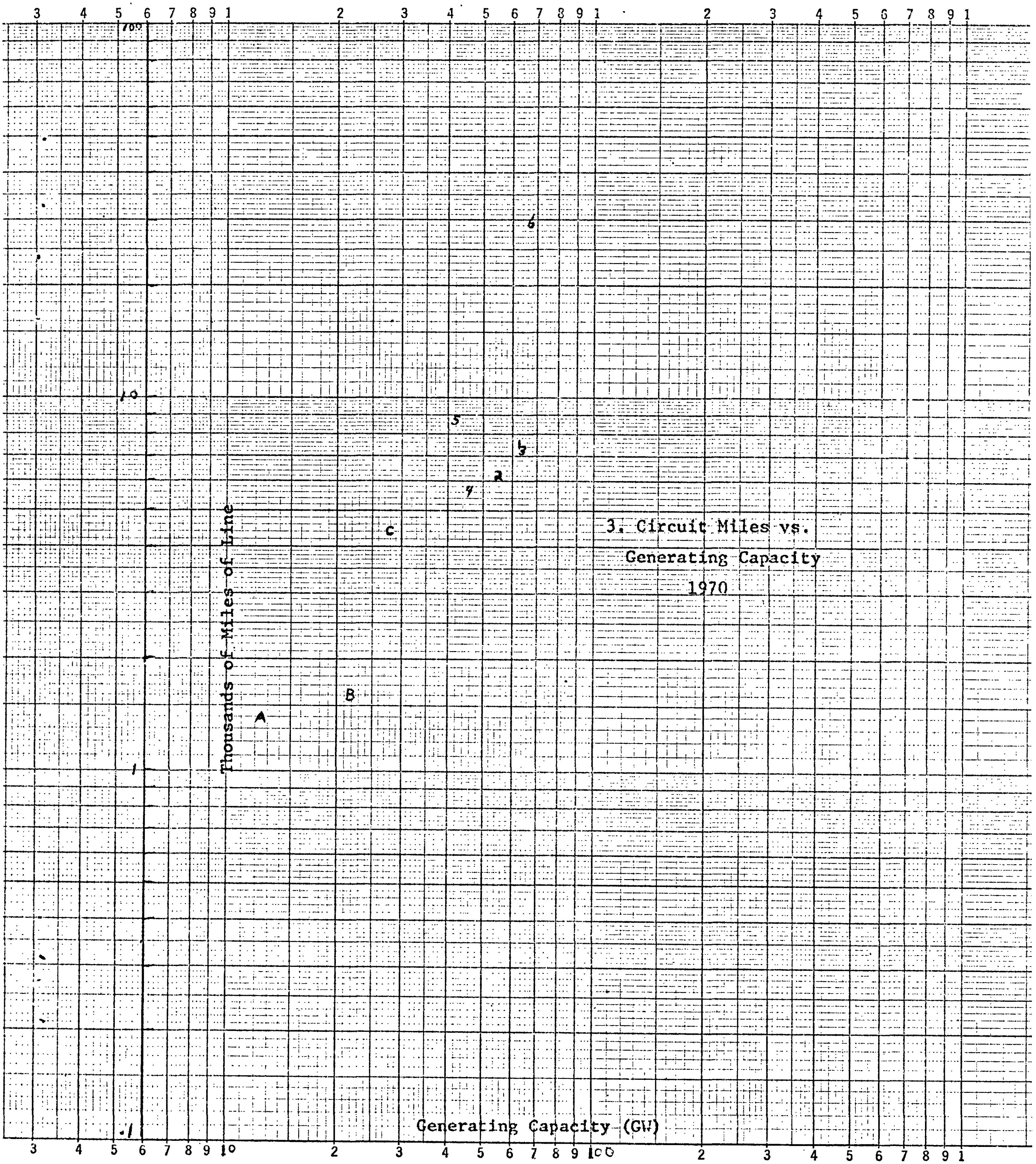
Data for this section was obtained from the _1970 National Power Survey_, _1960 Statistics of Publicly Owned Electric Utilities in the U.S._ and _1960 Statistics of Privately Owned Electric Utilities in the U.S._ and the _Directory of Electric Utilities_.

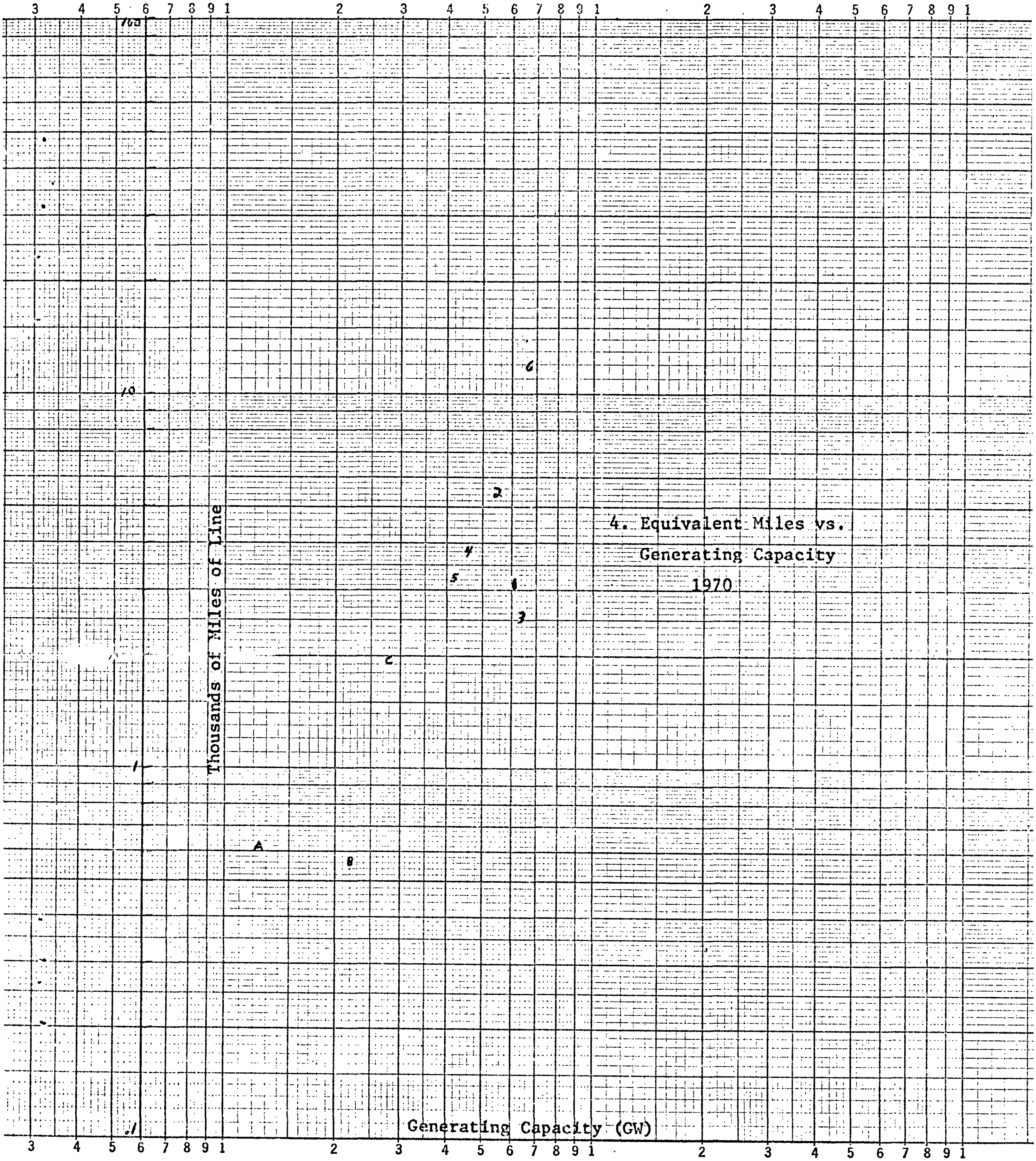
**Values for EM under 100 miles are not plotted in graphs.**


1. Circuit Miles vs.
Generating Capacity
1960
Thousands of Miles of Line
Generating Capacity (GW)
100
10
1
.1
6
5
2
1
C
3
B
A
4

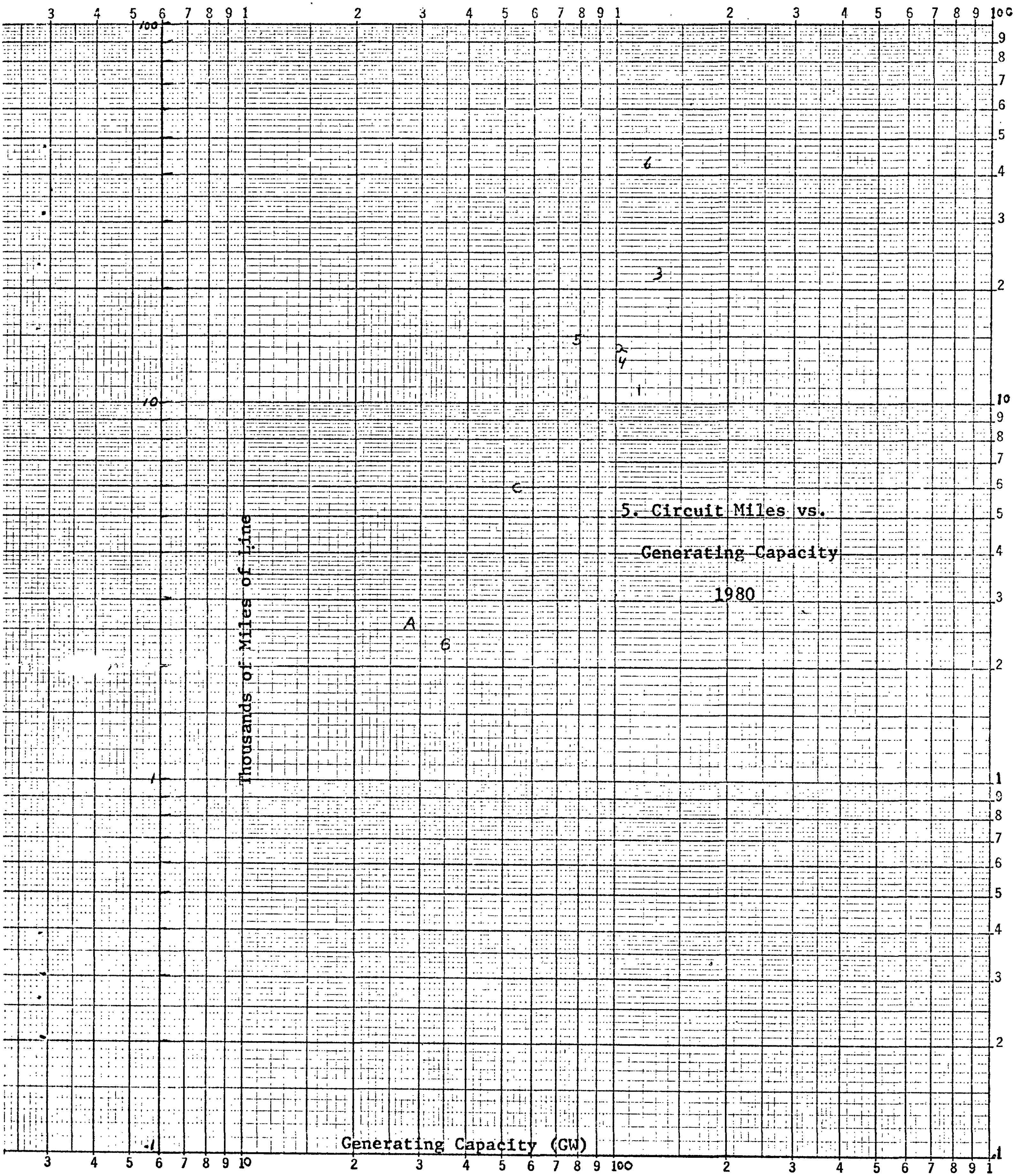


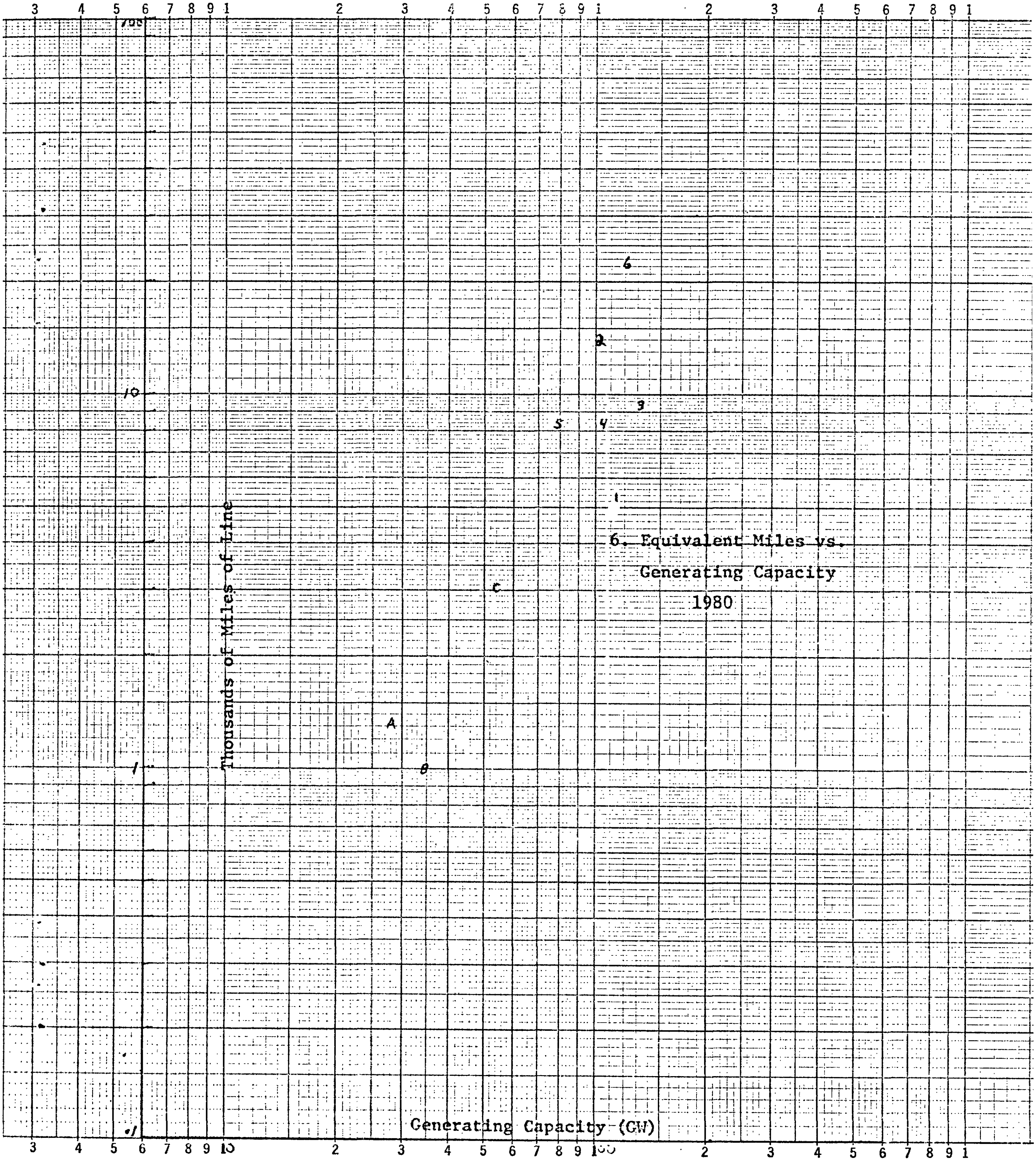
2. Equivalent Miles vs.
Generating Capacity
1960
Thousands of Miles of Line
Generating Capacity (GW)
Note change of scale for this graph
A
B
C


3. Circuit Miles vs.
Generating Capacity
1970
Thousands of Miles of Line
Generating Capacity (GW)

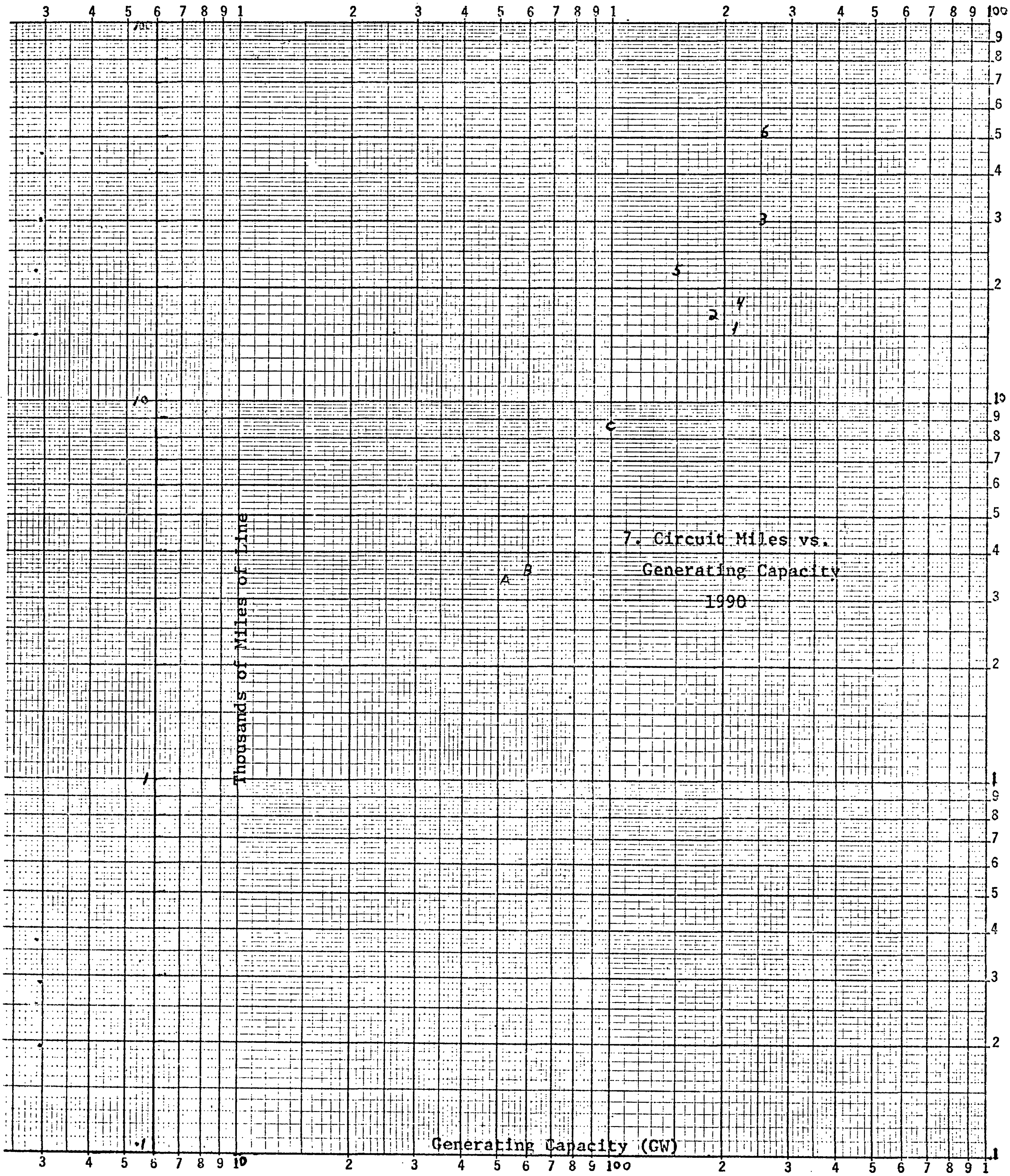


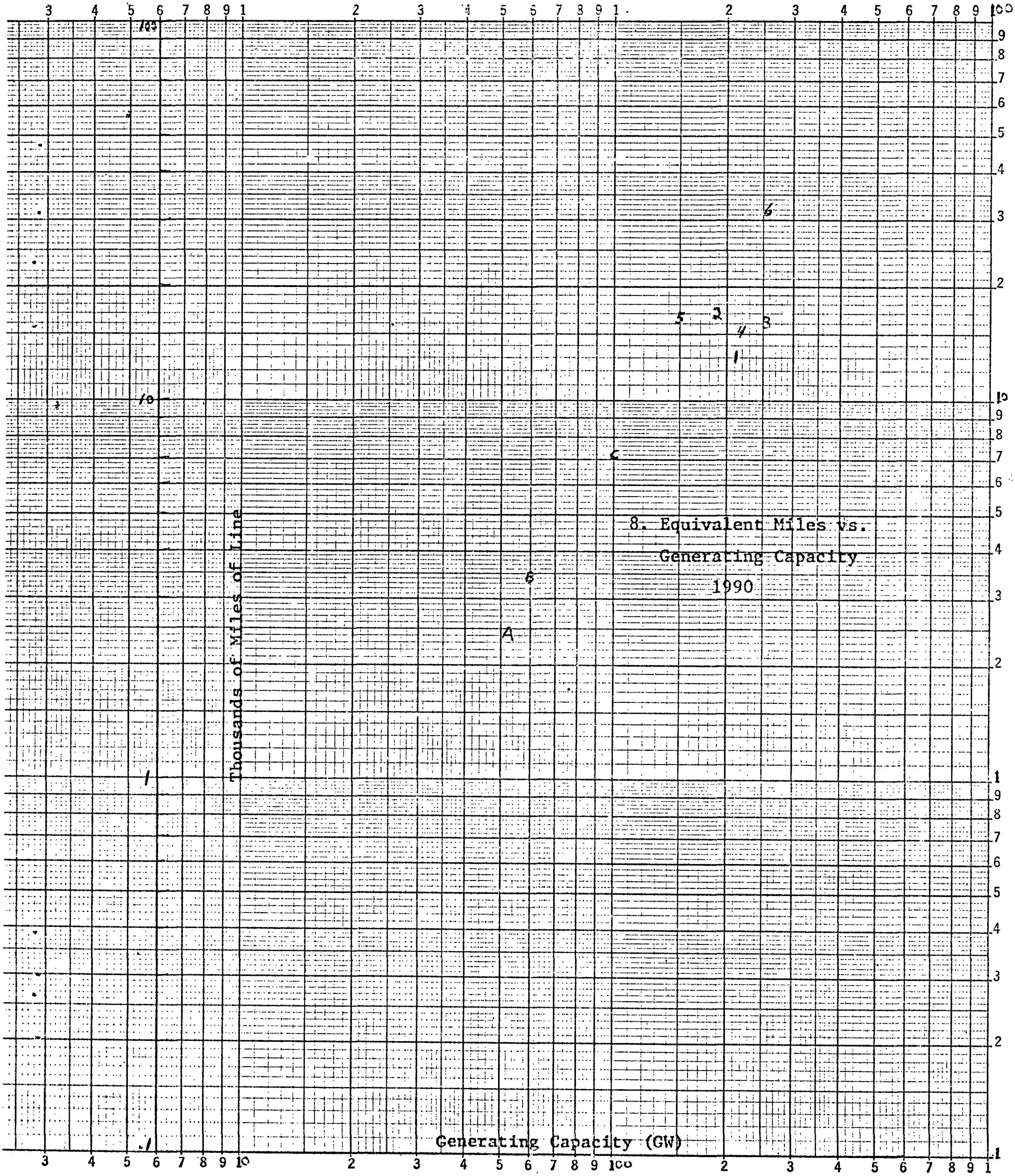
4. Equivalent Miles vs.
Generating Capacity
1970
Thousands of Miles of Line
Generating Capacity (GW)
100
10
1
.1
6
2
4
5
1
3
C
A
B

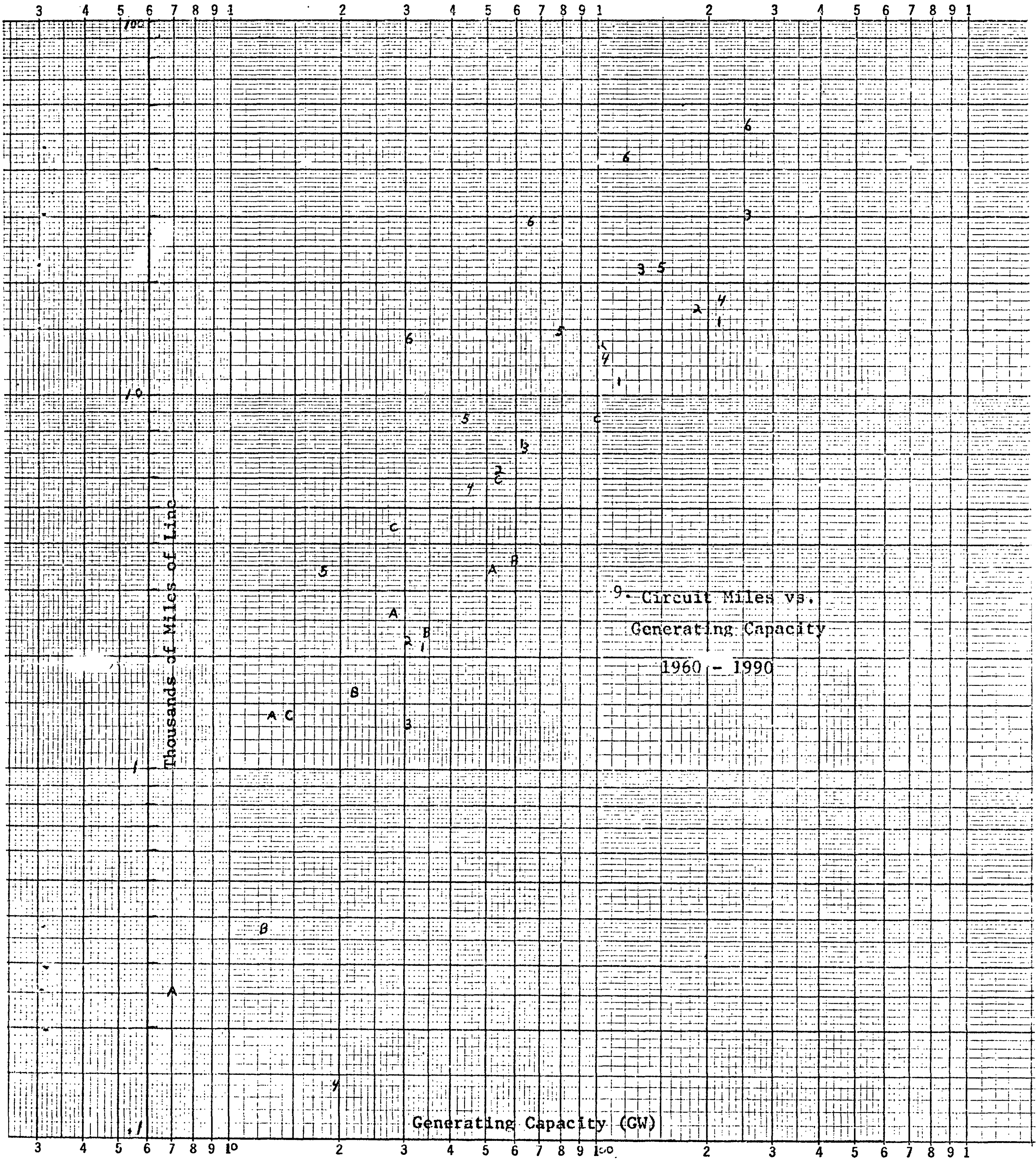

5. Circuit Miles vs.
Generating Capacity
1980
Thousands of Miles of Line
Generating Capacity (GW)

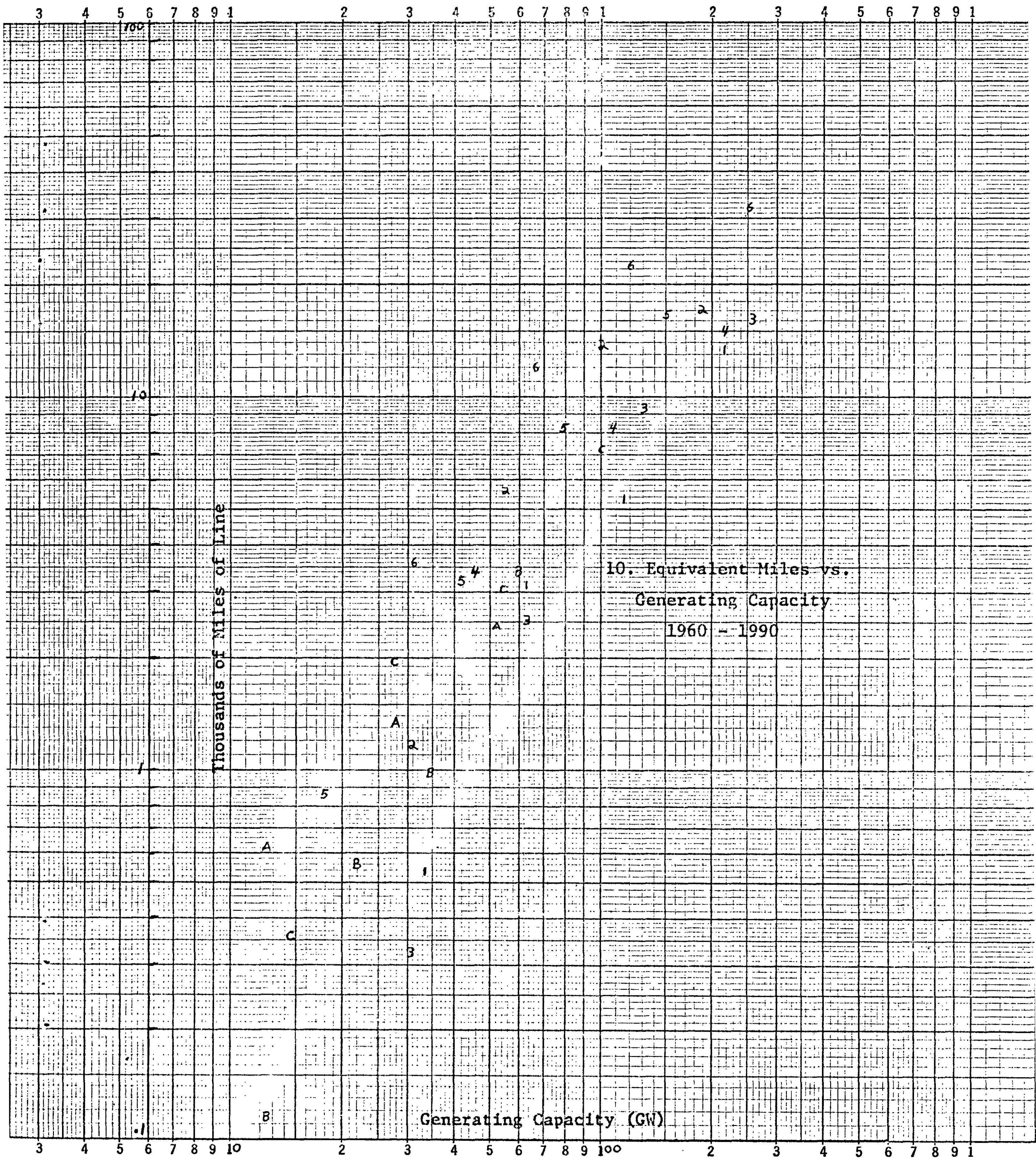



6. Equivalent Miles vs.
Generating Capacity
1980

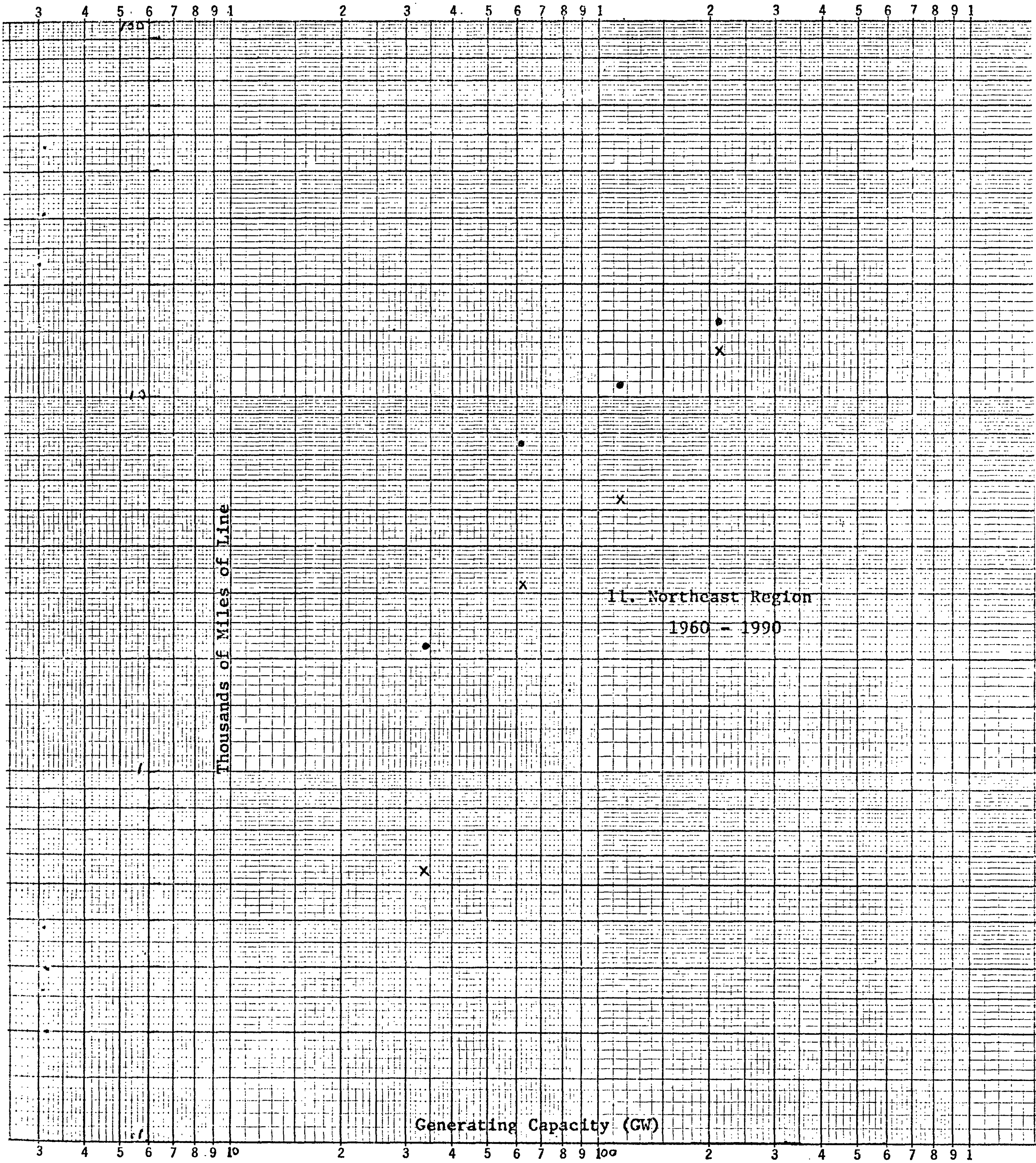

7. Circuit Miles vs.
Generating Capacity
1990
Thousands of Miles of Line
Generating Capacity (GW)
6
3
5
2
4
1
C
A
B


8. Equivalent Miles vs.
Generating Capacity
1990
Thousands of Miles of Line
Generating Capacity (GW)

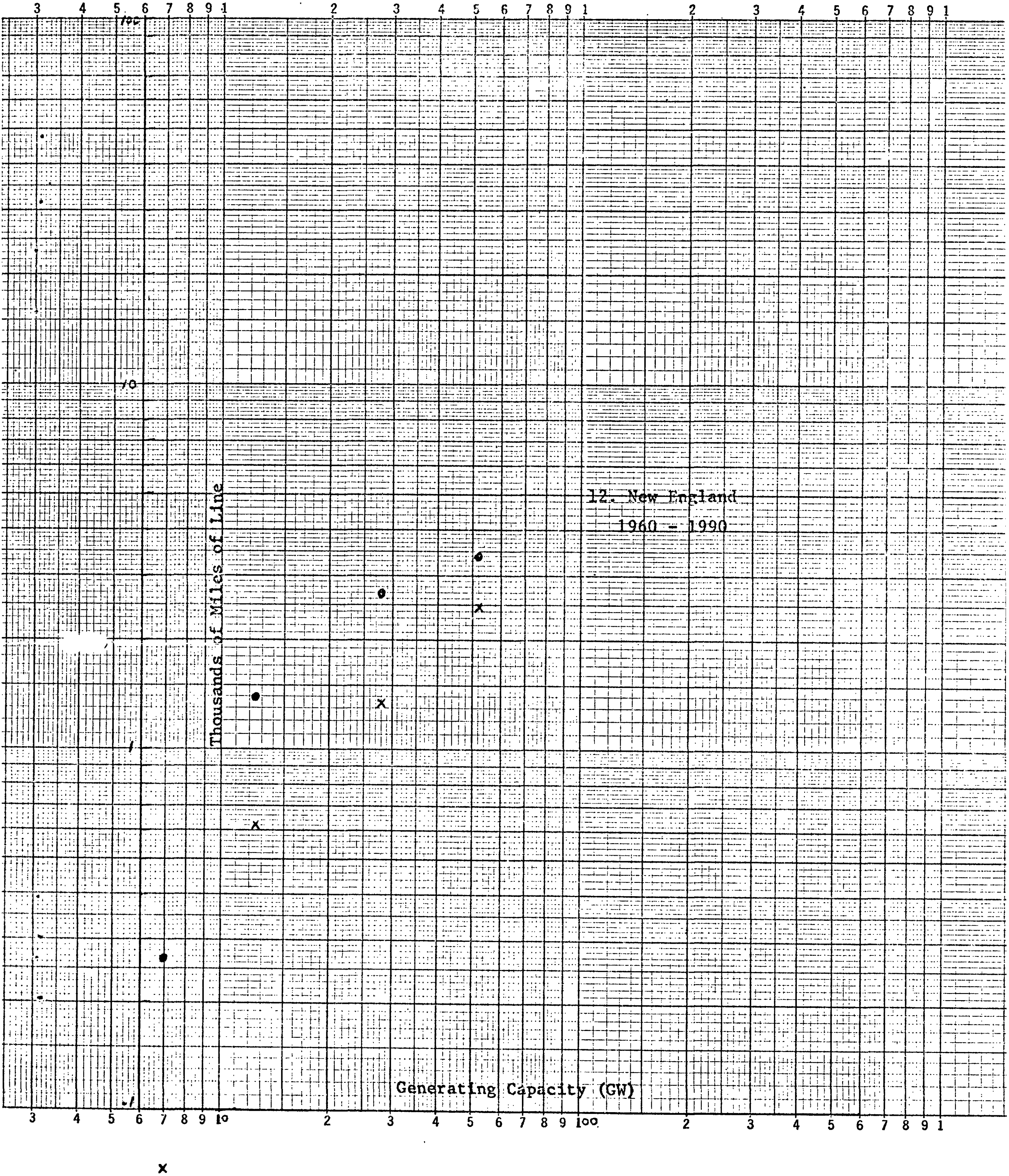

9. Circuit Miles vs.
Generating Capacity
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)
100
10
1
10
100

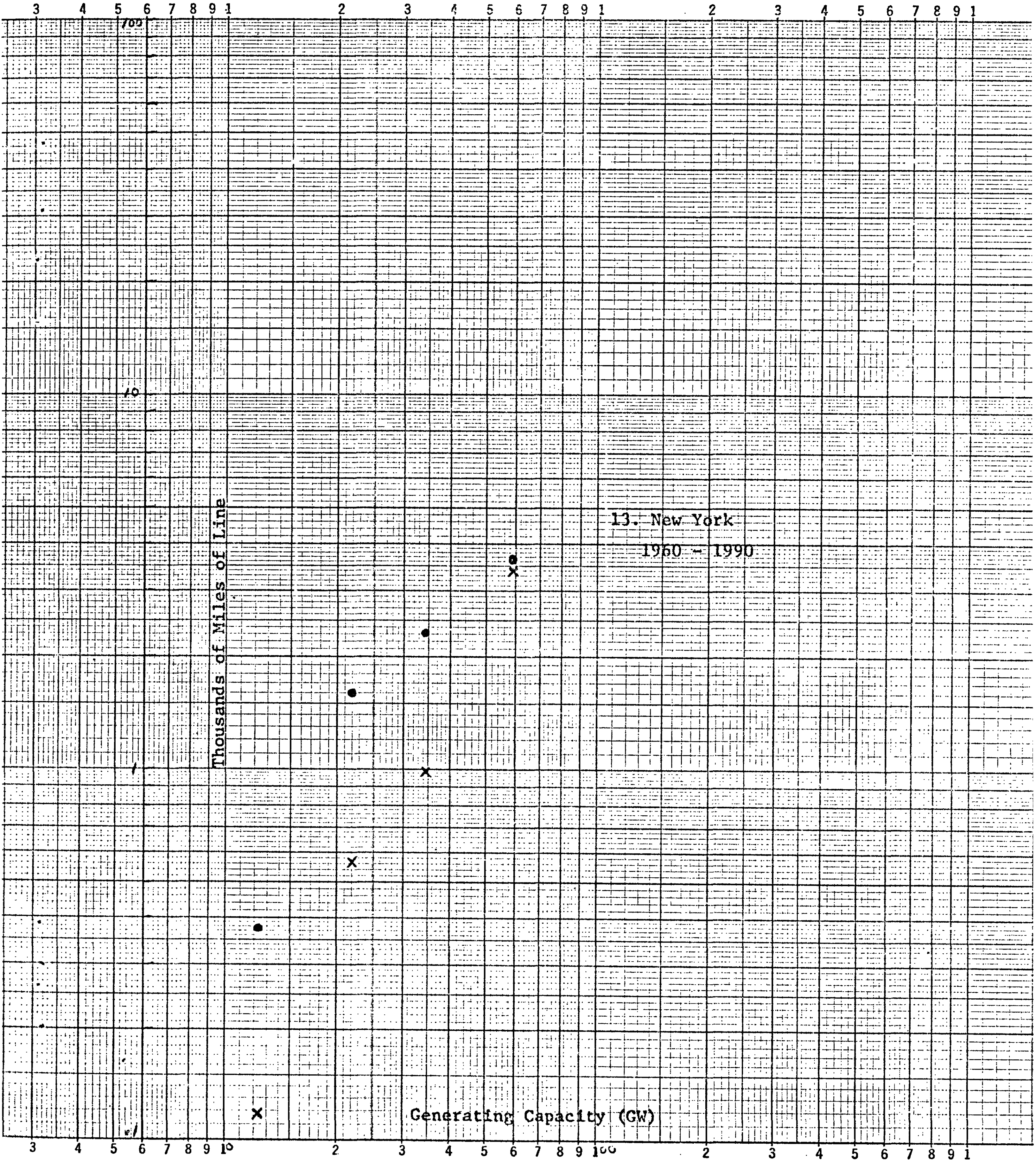


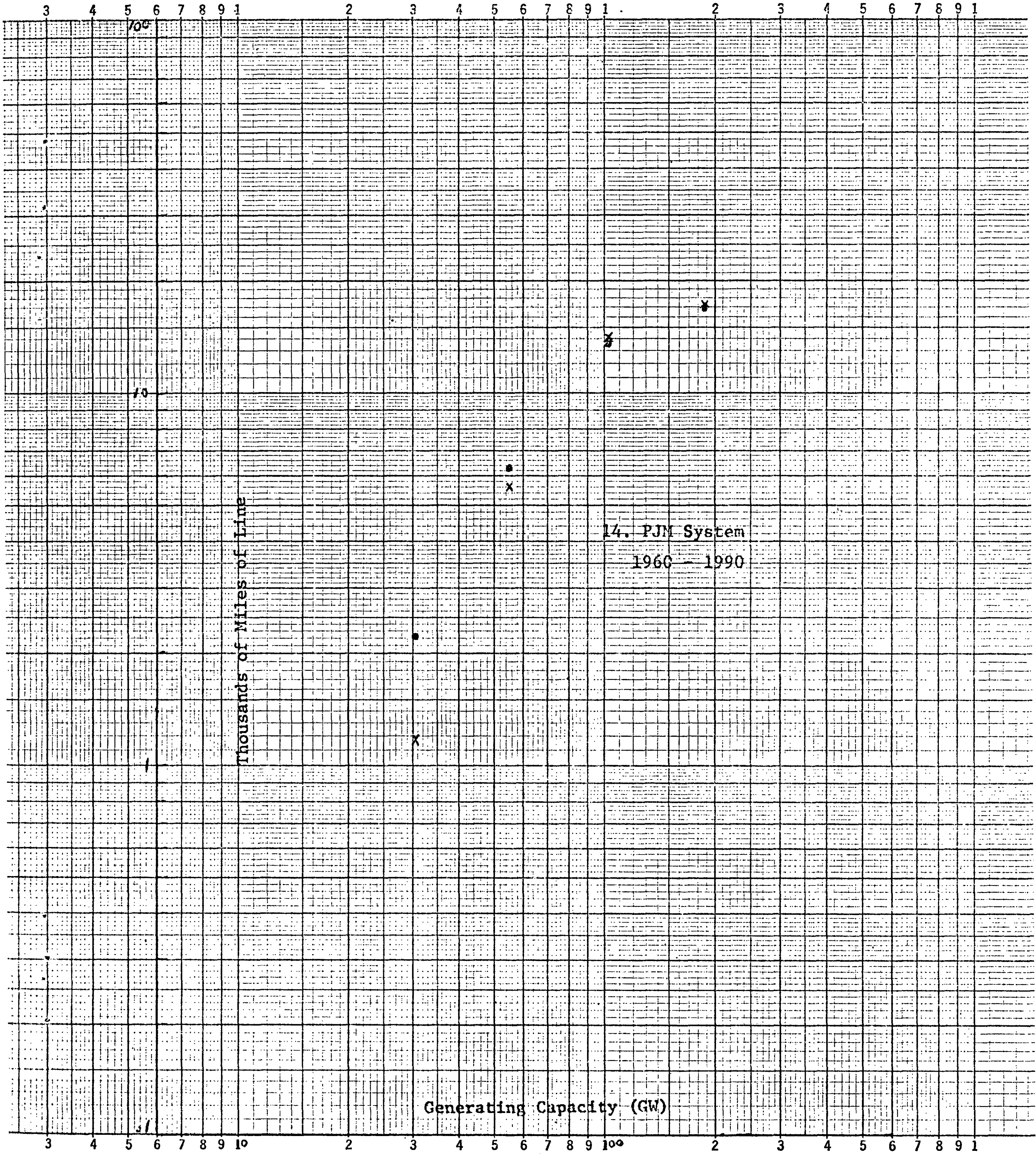

10. Equivalent Miles vs. Generating Capacity 1960 - 1990

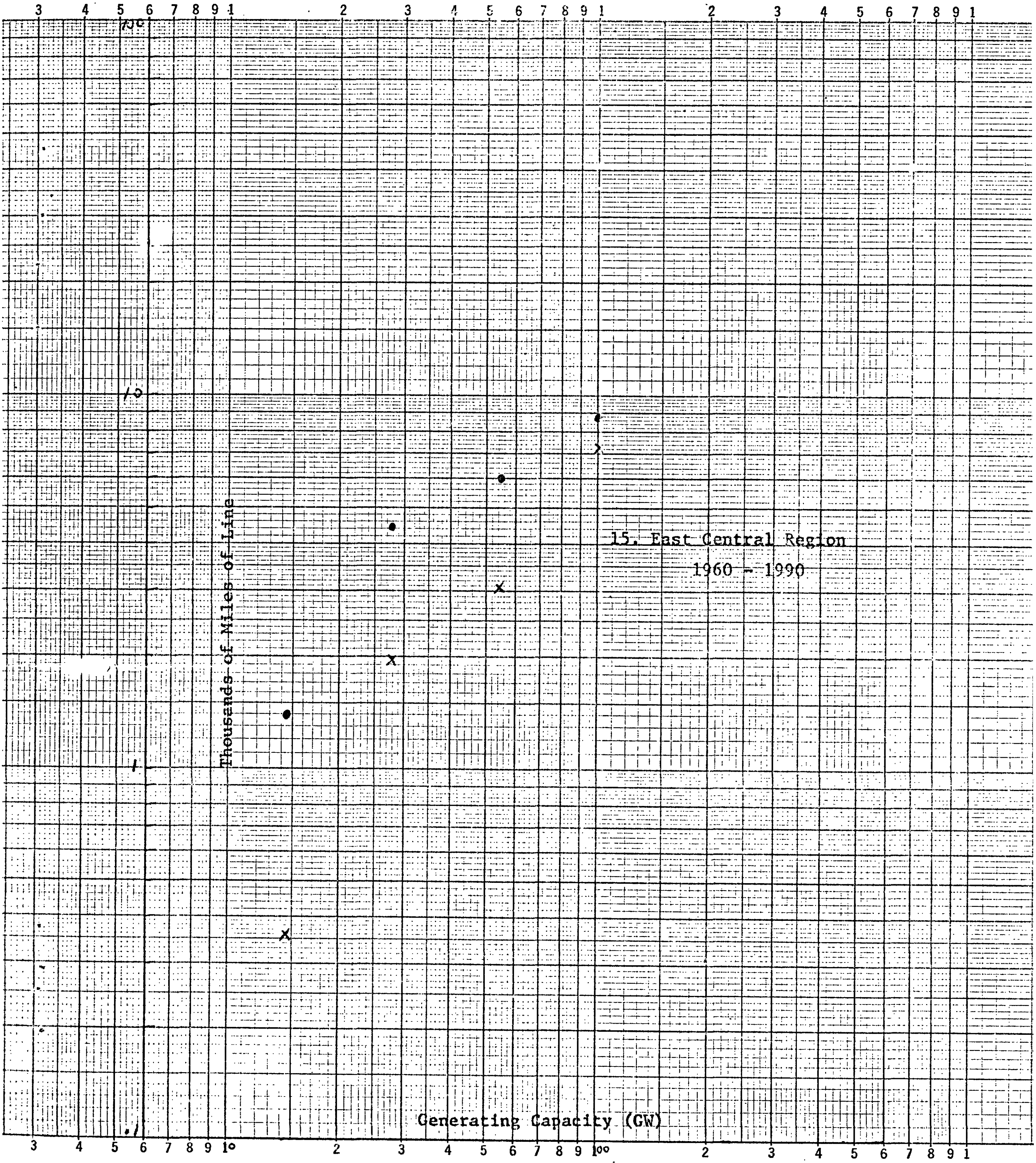


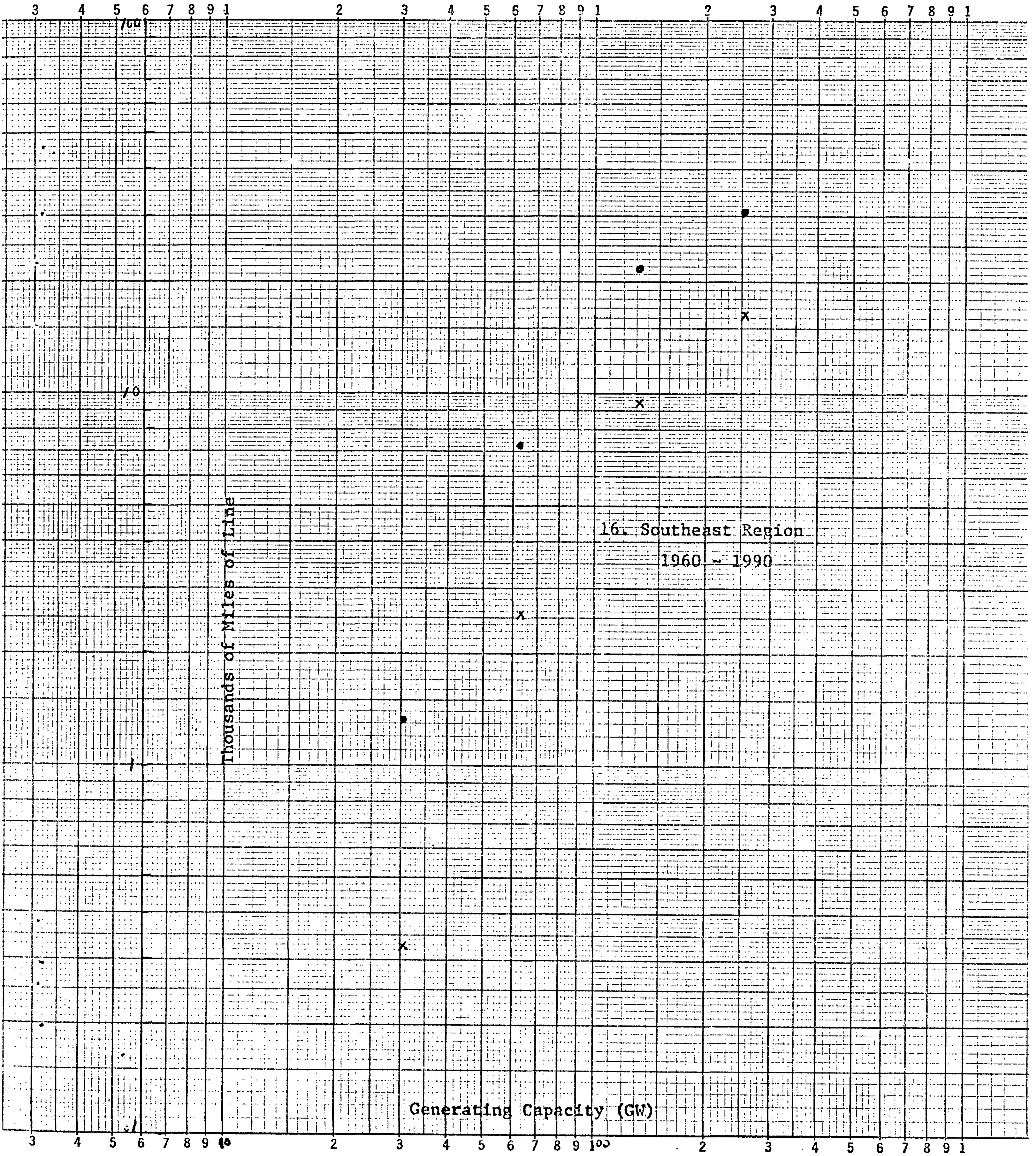
Thousands of Miles of Line
II. Northeast Region
1960 - 1990
Generating Capacity (GW)

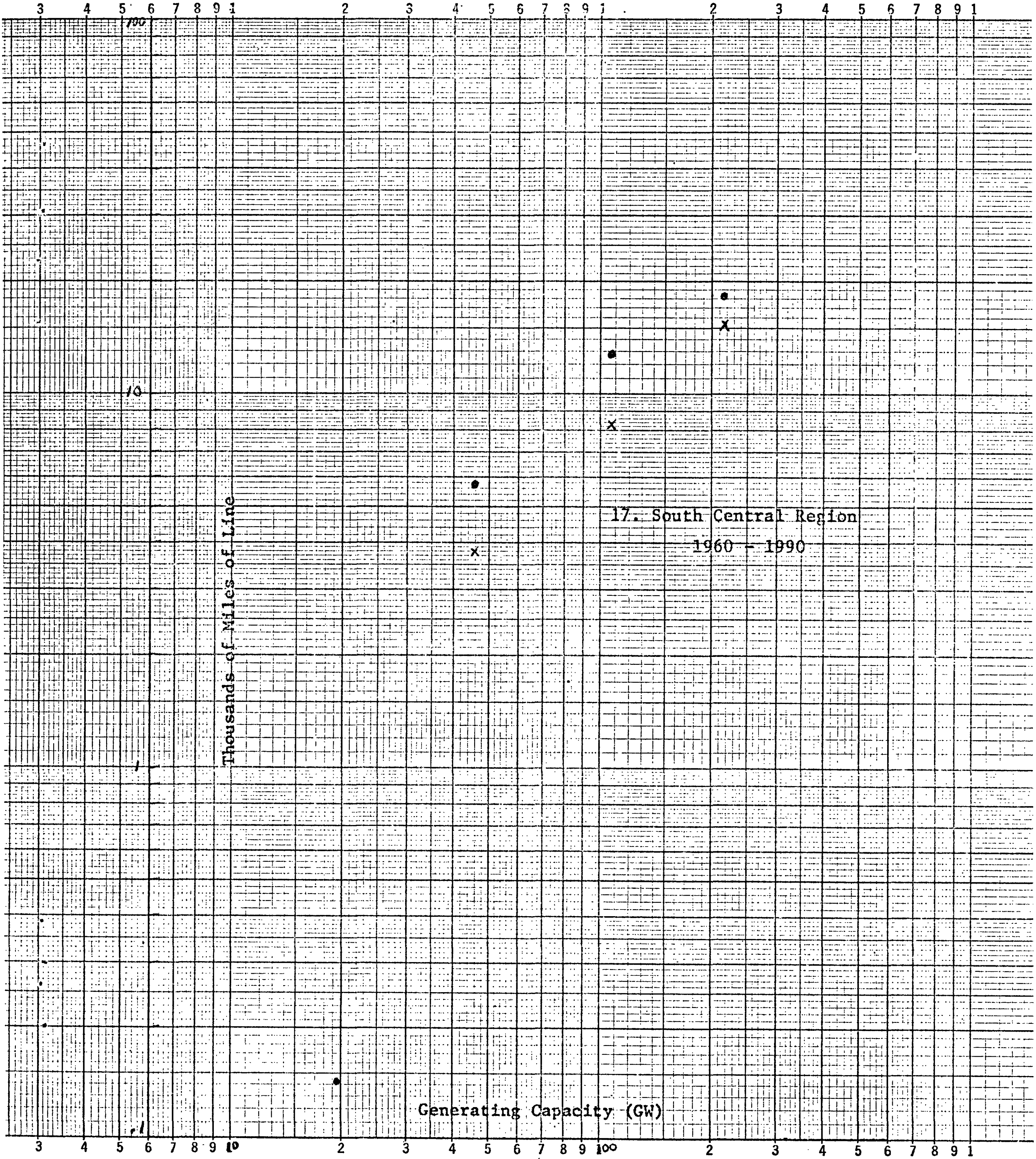

12. New England
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)
100
10
1
10
100

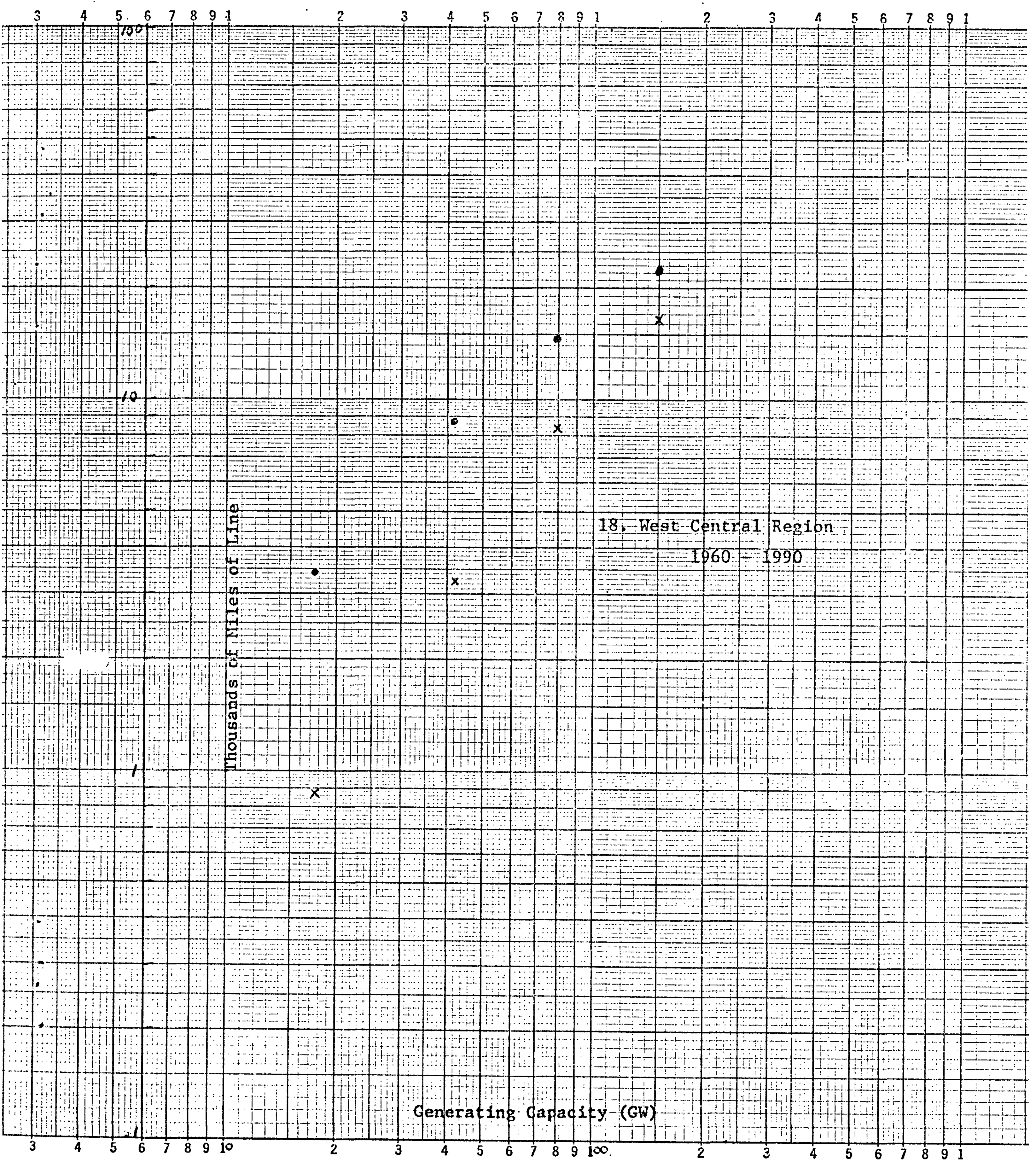

13. New York
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)

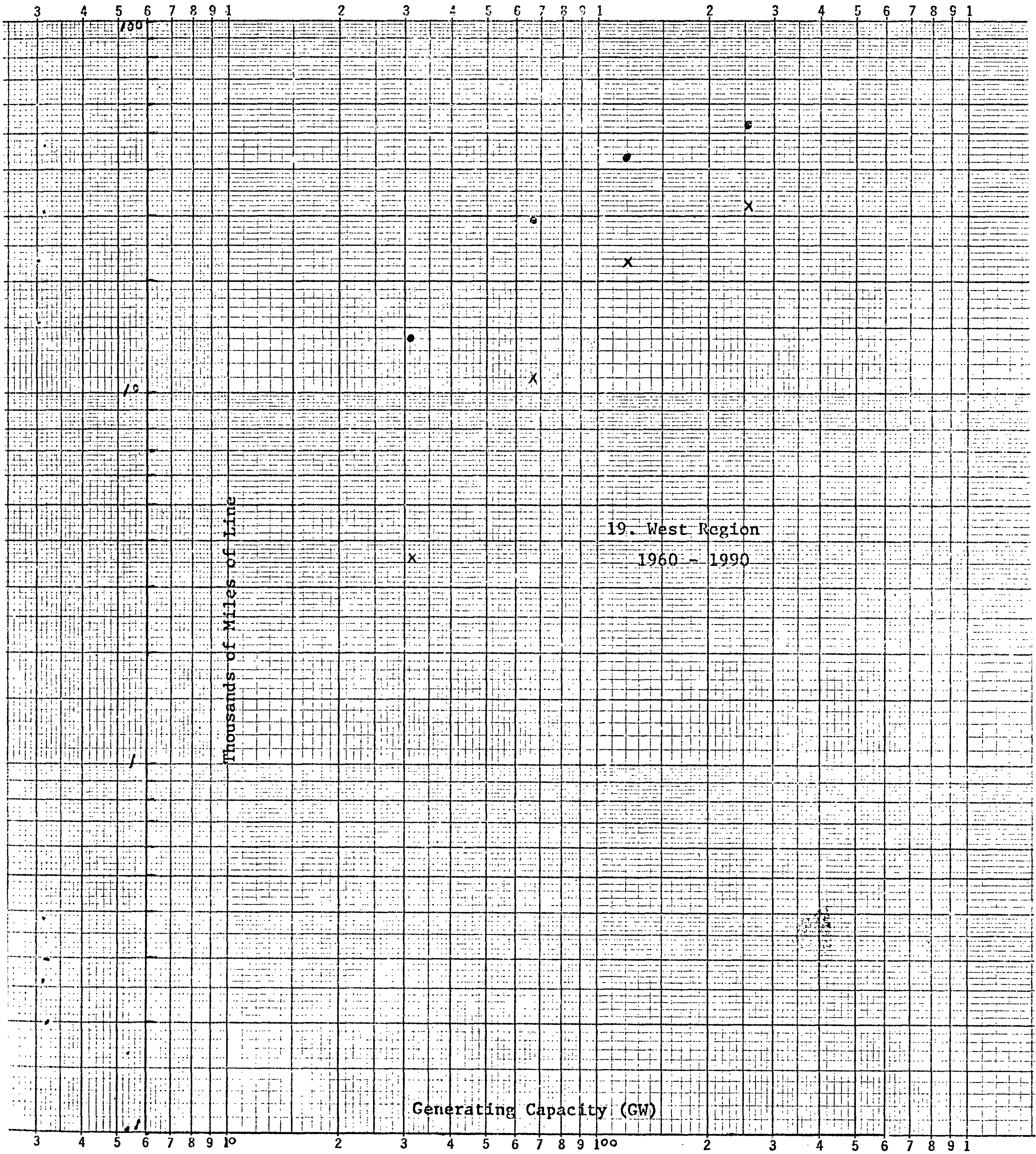

100
10
1
.1
Thousands of Miles of Line
14. PJM System
1960 - 1990
Generating Capacity (GW)
10
100


15. East Central Region
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)
3 4 5 6 7 8 9 10 2 3 4 5 6 7 8 9 100 2 3 4 5 6 7 8 9 1


16. Southeast Region
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)
100
10
1
10
100


17. South Central Region
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)
100
10
1
.1
10
100


18. West Central Region
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)
100
10
1
10
100


19. West Region
1960 - 1990
Thousands of Miles of Line
Generating Capacity (GW)

# APPENDIX IV

## Regression Results

This section contains the results of some of the regressions performed. The results of the final models are given in Chapter 3. Chapter 3's tables also include the predicted or "fitted" values of CM or EM for each region.

Table IV-1 lists all the regressions which were performed in the study.* For ease in reading, the coefficients and constants are left out ($M = e^K A^{a_1} G^{a_2 I}$ becomes $M = A\ G^I$) and abbreviations are used for all variables. (Lists of the abbreviations are given in Table 6, the end of Chapter 2, and Appendix II.)

To perform multiple linear regression, logrithms were taken of both sides of each equation of Table IV-1-A and IV-1-B.

The table entries relating to statistical tests (F-statistic, t-statistic, "$\alpha$", degrees of freedom, standard deviation) are defined in Appendix V.

*Complete results of all regressions listed in Table IV-1 are available from the author or Professor F. C. Schweppe, M.I.T., Room 3-137, Cambridge, Massachusetts.

## Table IV-1

### Regressions Performed

A. Miles = $e^K \prod_i X_i^{a_i}$

M = A I R G
M = A F H G
M = A LC-500 G
M = A LC-1000 G
M = A P G
M = A LC-500 R G METRO H
M = A P LC-1000 R G METRO H
M = A R G METRO H
M = A P LC-1000 R G METRO H AAPL
M = A P LC-1000 R G METRO H TAPL
M = A G METRO
M = A G METRO APL
M = A G METRO AAPL
M = A R G METRO
M = A I G METRO
M = A P R G METRO H AAPL
M = A P I G METRO H AAPL
M = A P I G METRO H APL

<u>By Year</u>

M = A G METRO
M = A G METRO APL
M = A G I
M = A G H
M = A G TPL
M = A G APL
M = A G TAPL
M = A G AAPL
M = G TPL
M = G APL
M = A G

B. Miles = $e^K \left(\prod_i X_i^{a_i}\right) G^{\sum_i X_i b_i}$

M = A G AAPL $G^{LC-500}$ $G^{R}$ $G^{METRO}$ $G^{N}$ $G^{AAPL}$
M = A G TAPL $G^{LC-500}$ $G^{R}$ $G^{METRO}$ $G^{N}$ $G^{TAPL}$
M = A AAPL $G^{R}$ $G^{METRO}$
M = A AAPL $G^{I}$ $G^{METRO}$
M = A AAPL $G^{METRO}$
M = A AAPL G $G^{METRO}$
M = A APL $G^{METRO}$
M = A APL G $G^{METRO}$
M = A $G^{METRO}$

M = A APL $G^{METRO}$
M = A G $G^{METRO}$
M = A $G^{METRO}$

C. Expansion Model for $\frac{M_t - M_o}{M_o}$ ( $= \Delta M/M_o$)

$\Delta M/M_o = \Delta D/D_o + \Delta G/G_o + \Delta^{o} METRO/METRO_o + \Delta H/H_o + \Delta P/P_o + \Delta I/I_o$

$\Delta M/M_o = \Delta D/D_o + \Delta G/G_o + \Delta\, METRO/METRO_o$

$\Delta M/M_o = \Delta D/D_o + \Delta G/G_o$

$\Delta M/M_o = \Delta D/D_o + METRO_o \cdot \Delta G/G_o + \ell n\, G_o \cdot \Delta METRO$

Table VI-1 - continued

$$\Delta M/M_o = \Delta D/D_o + METRO_o \cdot \Delta G/G_o + \Delta G/G_o + \ell n\ G_o \cdot \Delta METRO$$
$$\Delta M/M_o = \Delta D/D_o + METRO_o \cdot \Delta G/G_o + I_o \cdot \Delta G/G_o + \ell n\ G_o \cdot \Delta METRO + \ell n G_o \cdot \Delta I$$
$$\Delta M/M_o = \Delta G/G_o$$
$$\Delta M/M_o = METRO_o \cdot \Delta G/G_o + \ell n\ G_o \cdot \Delta METRO$$
$$\Delta M/M_o = \Delta D/D_o + \Delta G/G_o + \Delta P/P_o + \Delta I/I_o$$
$$\Delta M/M_o = \Delta D/D_o + \Delta METRO/METRO_o + \Delta P/P_o + \Delta I/I_o$$
$$\Delta M/M_o = \Delta D/D_o + \Delta G/G_o + \Delta I/I_o$$

where four values were used for $\Delta D/D_o$:

$(TPL_t - TPL_o)/TPL_o$
$(APL_t - APL_o)/APL_o$
$(AAPL_t - AAPL_o)/AAPL_o$
$(NAPL_t - OAPL_t)/OAPL_t$

TABLE IV-2

$CM = e^{K} A^{a_1} P^{a_2} R^{a_3} G^{a_4} METRO^{a_5} H^{a_6} AAPL^{a_7}$

Degrees of freedom (Df)*: 7/17

F-Statistic/Significance (F)*: 57.67/.01

Average % error (AE): 16%

| Coefficient | Value (V) | Standard Deviation (S) | t-value/$\alpha$ * (t) |
|---|---|---|---|
| K | 1.378 | | |
| $a_1$ | .707 | .113 | 6.26/.001 |
| $a_2$ | - .436 | .302 | 1.44/.2 |
| $a_3$ | -1.219 | .597 | 2.04/.1 |
| $a_4$ | .628 | .231 | 2.72/.05 |
| $a_5$ | .603 | .201 | 2.99/.05 |
| $a_6$ | .0968 | .0873 | 1.11/.3 |
| $a_7$ | .429 | .182 | 2.36/.1 |

* See Appendix V for definition of these terms

TABLE IV-3

$CM = e^K A^{a_1} I^{a_2} R^{a_3} G^{a_4}$

Df: 4/31

F: 33.87/.01

| | V | S | t |
|---|---|---|---|
| K | -2.574 | | |
| $a_1$ | .229 | .141 | 1.62/.2 |
| $a_2$ | - .051 | 1.230 | .04/-- |
| $a_3$ | 1.878 | 1.356 | 1.38/.3 |
| $a_4$ | 1.193 | .141 | 8.44/.01 |

TABLE IV-4

$$EM = e^{K} A^{a_1} I^{a_2} R^{a_3} G^{a_4}$$

Df: 4/31

F: 42.49/.01

| | V | S | t |
|---|---|---|---|
| K | -5.379 | | |
| $a_1$ | .019 | .155 | .12/-- |
| $a_2$ | .400 | 1.346 | .30/.9 |
| $a_3$ | .258 | 1.485 | 1.74/.2 |
| $a_4$ | 1.636 | .155 | 10.57/.01 |

TABLE IV-5

$CM = e^{K} A^{a_1} G^{a_2}$

Df: 2/24

F: 118.32/.01

AE: 31.1%

| | V | S | t |
|---|---|---|---|
| K | -.431 | | |
| $a_1$ | .075 | .108 | .69/.7 |
| $a_2$ | 1.156 | .098 | 11.83/.01 |

TABLE IV-6

$EM = e^K A^{a_1} G^{a_2}$

Df: 2/24

F: 95.60/.01

AE: 41.8%

| | V | S | t |
|---|---|---|---|
| K | -1.711 | | |
| $a_1$ | - .093 | .151 | .62/.7 |
| $a_2$ | 1.562 | .136 | 11.49/.01 |

TABLE IV-7

$CM = e^{K} A^{a_1} G^{a_2 METRO}$

Df: 2/24

F: 116.97/.01

AE: 25.7%

| | V | S | t |
|---|---|---|---|
| K | .830 | | |
| $a_1$ | .835 | .058 | 14.52/.01 |
| $a_2$ | .168 | .022 | 7.62/.02 |

TABLE IV-8

$EM = e^{K} A^{a_1} G^{a_2 METRO}$

Df: 2/24

F: 85.18/.01

AE: 31.5%

| | V | S | t |
|---|---|---|---|
| K | -.032 | | |
| $a_1$ | .867 | .075 | 11.60/.01 |
| $a_2$ | .235 | .029 | 8.19/.02 |

TABLE IV-9

$$CM = e^{K} A^{a_1} G^{a_2} APL^{a_3}$$

| | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|
| Df: | 3/5 | | | 3/4 | | | 3/4 | | |
| F: | 19.20/.01 | | | 18.29/.01 | | | 23.62/.01 | | |
| AE: | 20.4% | | | 21.2% | | | 16.6% | | |
| | V | S | t | V | S | t | V | S | t |
| K | -1.361 | | | -.118 | | | .128 | | |
| $a_1$ | - .070 | .280 | .25/.9 | .384 | .273 | 1.41/.3 | .369 | .318 | 1.16/.5 |
| $a_2$ | 1.136 | .357 | 3.18/.05 | .901 | .434 | 2.08/.2 | .859 | .497 | 1.73/.2 |
| $a_3$ | 1.082 | .515 | 2.10/.2 | .274 | .665 | .37/.9 | -.106 | .464 | .23/.9 |

TABLE IV-10

$$EM = e^{K} A^{a_1} G^{a_2} TPL^{a_3}$$

| | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|
| Df: | 3/5 | | | 3/4 | | | 3/4 | | |
| F: | 18.97/.01 | | | 10.89/.05 | | | 22.43/.01 | | |
| AE: | 22.7% | | | 29.2% | | | 17.0% | | |
| | V | S | t | V | S | t | V | S | t |
| K | -2.226 | | | - .373 | | | - .129 | | |
| $a_1$ | .047 | .224 | .21/.9 | .502 | .361 | 1.39/.3 | .275 | .268 | 1.03/.5 |
| $a_2$ | .030 | .484 | .06/-- | 1.036 | .597 | 1.74/.2 | 1.072 | .323 | 3.32/.05 |
| $a_3$ | 1.165 | .485 | 2.40/.1 | - .150 | .555 | .27/.9 | - .209 | .393 | .53/.7 |

TABLE IV-11

$CM = e^{K} A^{a_1} G^{a_2} H^{a_3}$

| | 1960 | | | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Df: | 3/5 | | | 3/5 | | | 3/5 | | | 3/5 | | |
| F: | 1.42/-- | | | 12.63/.01 | | | 27.57/.01 | | | 29.14/.01 | | |
| AE: | 119.1% | | | 22.9% | | | 19.0% | | | 17.2% | | |
| | V | S | t | V | S | t | V | S | t | V | S | t |
| K | -.905 | | | .367 | | | .143 | | | -.023 | | |
| $a_1$ | .083 | .571 | .15/.9 | .383 | .192 | 1.99/.2 | .420 | .151 | 2.78/.1 | .349 | .156 | 2.24/.1 |
| $a_2$ | 1.462 | 1.149 | 1.27/.3 | .746 | .371 | 2.01/.2 | .890 | .295 | 3.02/.1 | .892 | .287 | 3.11/.1 |
| $a_3$ | .444 | .552 | .80/.5 | .252 | .209 | 1.21/.5 | .169 | .165 | 1.02/.5 | .025 | .122 | .21/.9 |

TABLE IV-12

$EM = e^K A^{a_1} G^{a_2} I^{a_3}$

| | 1960 | | | 1970 | | | 1980 | | | 1990 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Df: | 3/5 | | | 3/5 | | | 3/5 | | | 3/5 | | |
| F: | 1.65/-- | | | 10.61/.05 | | | 22.10/.01 | | | 23.89/.01 | | |
| AE: | 118.2% | | | 28.4% | | | 19.2% | | | 19.1% | | |
| | V | S | t | V | S | t | V | S | t | V | S | t |
| K | 3.108 | | | -3.230 | | | -2.605 | | | -.419 | | |
| $a_1$ | .006 | .549 | .01/-- | .555 | .256 | 2.16/.2 | .573 | .201 | 2.85/.1 | .245 | .154 | 1.60/.3 |
| $a_2$ | 2.039 | 1.230 | 1.66/.2 | .483 | .544 | .89/.5 | .536 | .480 | .91/.5 | .882 | .356 | 2.48/.1 |
| $a_3$ | -3.418 | 3.531 | .97/.5 | 2.054 | 1.738 | 1.18/.5 | 2.011 | 1.216 | 1.65/.2 | .162 | .795 | .20/.9 |

TABLE IV-13

$$CM = e^{K} A^{a_1} P^{a_2} I^{a_3} G^{a_4} METRO^{a_5} H^{a_6} APL^{a_7}$$

Df: 7/17

F: 63.34/.01

AE: 15.1%

| | V | S | t |
|---|---|---|---|
| K | -1.778 | | |
| $a_1$ | .460 | .125 | 3.69/.02 |
| $a_2$ | - .586 | .281 | 2.09/.1 |
| $a_3$ | 1.060 | .497 | 2.13/.1 |
| $a_4$ | 1.091 | .195 | 5.61/.001 |
| $a_5$ | .481 | .180 | 2.67/.05 |
| $a_6$ | .217 | .104 | 2.08/.1 |
| $a_7$ | .392 | .212 | 1.85/.2 |

TABLE IV-14

$EM = e^K A^{a_1} G^{a_2} METRO^{a_3}$

Df: 3/23 (1970, 80, 90 data)

F: 75.93/.01

AE: 29.5%

| | V | S | t |
|---|---|---|---|
| K | -.693 | | |
| $a_1$ | .405 | .130 | 3.12/.1 |
| $a_2$ | .952 | .146 | 6.51/.01 |
| $a_3$ | .191 | .207 | .92/.5 |

$EM = e^K A^{a_1} G^{a_2} METRO^{a_3}$

Df: 3/32 (1960, 70, 80, 90 data)

F: 46.66/.01

AE: 86.6%

| | V | S | t |
|---|---|---|---|
| K | -1.769 | | |
| $a_1$ | .211 | .223 | .95/.5 |
| $a_2$ | 1.465 | .221 | 6.64/.01 |
| $a_3$ | .235 | .378 | .62/.7 |

TABLE IV-15

$CM = e^K A^{a_1} G^{a_2} METRO^{a_3}$

| | | |
|---|---|---|
| Df: | 3/23 | (1970, 80, 90 data) |
| F: | 103.72/.01 | |
| AE: | 22.3% | |

| | V | S | t |
|---|---|---|---|
| K | .370 | | |
| $a_1$ | .610 | .100 | 6.13/.01 |
| $a_2$ | .573 | .112 | 5.10/.02 |
| $a_3$ | .338 | .159 | 2.13/.2 |

---

$EM = e^K A^{a_1} G^{a_2} METRO^{a_3}$

| | | |
|---|---|---|
| Df: | 3/32 | (1960, 70, 80, 90 data) |
| F: | 39.85/.01 | |
| AE: | 71.6% | |

| | V | S | t |
|---|---|---|---|
| K | -.511 | | |
| $a_1$ | .481 | .197 | 2.44/.1 |
| $a_2$ | .961 | .195 | 4.92/.02 |
| $a_3$ | .443 | .335 | 1.32/.3 |

TABLE IV-16

$$EM = e^{K} A^{a_1} G^{a_2} METRO^{a_3} APL^{a_4}$$

Df: 4/20

F: 55.87/.01

AE: 25.8%

| | V | S | t |
|---|---|---|---|
| K | -1.639 | | |
| $a_1$ | .275 | .155 | 1.78/.2 |
| $a_2$ | 1.080 | .176 | 6.15/.01 |
| $a_3$ | .378 | .229 | 1.65/.2 |
| $a_4$ | .540 | .321 | 1.68/.2 |

TABLE IV-17

$$CM = e^{K} A^{a_1} AAPL^{a_2} G^{a_3 R} G^{a_4 METRO}$$

Df: 4/20

F: 62.59/.01

AE: 21.4%

| | V | S | t |
|---|---|---|---|
| K | .299 | | |
| $a_1$ | .780 | .109 | 7.14/.01 |
| $a_2$ | .350 | .203 | 1.75/.2 |
| $a_3$ | -.020 | .072 | .27/.9 |
| $a_4$ | .183 | .045 | 4.06/.02 |

TABLE IV-18

$$CM = e^{K} A^{a_1} APL^{a_2} G^{a_3 METRO}$$

Df: 3/21

F: 73.27/.01

AE: 24%

| | V | S | t |
|---|---|---|---|
| K | .434 | | |
| $a_1$ | .778 | .075 | 10.31/.01 |
| $a_2$ | .248 | .270 | .92/.5 |
| $a_3$ | .195 | .034 | 5.69/.02 |

TABLE IV-19

$$EM = e^{K} A^{a_1} APL^{a_2} G^{a_3 METRO}$$

Df: 3/21

F: 58.63/.01

AE: 31%

| | V | S | t |
|---|---|---|---|
| K | -.543 | | |
| $a_1$ | .785 | .094 | 8.37/.01 |
| $a_2$ | .310 | .336 | .92/.5 |
| $a_3$ | .271 | .043 | 6.36/.01 |

TABLE IV-20

$$\frac{EM_t - EM_o}{EM_o} = K + a_1 \frac{AAPL_t - AAPL_o}{AAPL_o} + a_2 \frac{G_t - G_o}{G_o} + a_3 \frac{P_t - P_o}{P_o} + a_4 \frac{I_t - I_o}{I_o}$$

Df: 4/11

F: 2.73/--

AE in Line Added in last 10 years: 52.7%

AE in Total Line: 18.6%

| | V | S | t |
|---|---|---|---|
| K | - 1.096 | | |
| $a_1$ | - 2.039 | .776 | 2.63/.1 |
| $a_2$ | 1.753 | 1.493 | 1.17/.5 |
| $a_3$ | 2.085 | .820 | 2.54/.1 |
| $a_4$ | -13.243 | 4.645 | 2.85/.05 |

# APPENDIX V

## Determining the Significance of Model Parameters

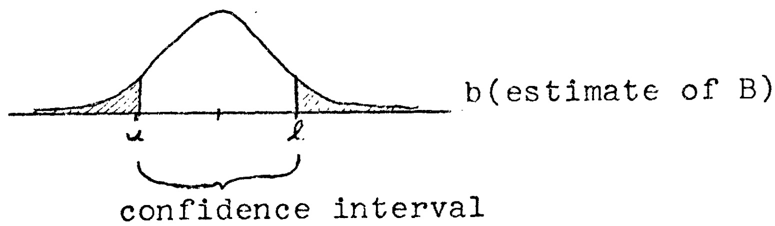
As was mentioned previously, correlation and regression can not be used as proof of cause and effect. However, these techniques are useful in two ways:[1] they may provide further confirmation of a relation that theory tells us should exist, and they may suggest casual relationships not previously suspected (such as when cigarette smoking was found to be correlated with lung cancer). Thus, the "statistical significance" of a parameter is indicative of its correlation with the dependent variable. It is no guarantee of causality.

Two statistics were computed for the parameters estimated by regression: the t-statistic and F-statistic. The t-statistic can be used to compute a confidence interval for the estimated coefficient or to test the null hypothesis. A confidence interval gives the limits within which, with a certain probability, the true value of the parameter to be estimated lies. Thus, a 95% confidence interval for a coefficient, b, (which is an estimate of the true value (B) of the coefficient) is defined by upper and lower limits, u and $\ell$, such that with a probability of .95 the following is true:

$u \leq b \leq \ell$

If the dependent variable is normally distributed, then the regression coefficients, b, have a t-distribution, as shown below:

fig. V-1




To test the null hypothesis, the t-statistic indicates the probability that the null hypothesis will be incorrectly rejected; that is, the probability that a nonzero estimate b of the true value (B) will be made if B is indeed zero. This method of testing the null hypothesis will be described later in this section.

The F-statistic is an alternative method of testing the null hypothesis.[2]

The equations used in the calculation of these statistics and the regression coefficients are given below:[3]

$X_{ij}$ is input data    $i = 1,2 \ldots N$ observations

$j = 1,2,\ldots M$ variables

There are M-1 independent variables and one dependent variable, Y.

The means, $X_j = \dfrac{\sum_{i=1}^{N} X_{ij}}{N}$ are calculated.

The sums of cross products of deviations from the mean are calculated.

$$s_{jk} = \sum_{i=1}^{N} (X_{ij}-\overline{X}_j)\cdot(X_{ik}-\overline{X}_k) - \frac{\sum_{i=1}^{N} (X_{ij}-\overline{X}_j) \sum_{i=1}^{N} (X_{ik}-\overline{X}_k)}{N}$$

The correlation coefficients are

$$r_{jk} = s_{jk}/(\sqrt{s_{jj}} \cdot \sqrt{s_{kk}})$$

The standard deviations of the variables are

$$s_j = \frac{\sqrt{s_{jj}}}{N-1}$$

With "Y" indicating the dependent variable, the regression coefficients are

$$b_j = \left( \sum_{i=1}^{M-1} r_{iY} / r_{ij} \right) \cdot \frac{s_Y}{s_j}$$

The intercept is

$$b_o = \overline{Y} - \sum_{j=1}^{M-1} b_j \cdot \overline{X}_j$$

This gives the fitted equation for $\hat{Y}$, the estimate of Y.

$$\hat{Y} = b_o + \sum_{i=1}^{M-1} b_i X_i$$

To obtain the t- and F-statistics, the correlation coefficient, R, is first computed.

$$R^2 = \sum_{i=1}^{M-1} (r_{iY} \cdot ( \sum_{j=1}^{M-1} r_{jY} / r_{ji} ) )$$

Let $D_{YY}$ = the sum of squares of deviations from the mean for the dependent variable:

$$D_{YY} = \sum_{i=1}^{N} (Y_i - \overline{Y})^2$$

The sum of squares of deviations from regression (SSDR) and the sum of squares attributable to regression(SSAR) are

$$SSAR = R^2 \cdot D_{YY} \qquad \text{and} \qquad SSDR = D_{YY} - SSAR$$

The F-statistic is

$$F = \frac{SSAR/(M-1)}{SSDR/(N-M-2)}$$

with (M-1)/(N-M-2) degrees of freedom.

To obtain the t-statistic, we first define

$$s^2_{y \cdot 1,2, \cdots M-1} = \frac{SSDR}{N-M-2}$$

and

$$D_{jj} = \sum_{i=1}^{N} (X_{ij} - \overline{X}_j)^2.$$

Then the standard deviation of the regression coefficients is given by

$$s_{b_j} = \sqrt{s^2_{y \cdot 1,2 \cdots M-1} / (r_{jj} \cdot D_{jj})}$$

The t-statistic is

$$t_j = b_j / s_{b_j}$$

and it has M-1 degrees of freedom.

The t-distribution is shown below:

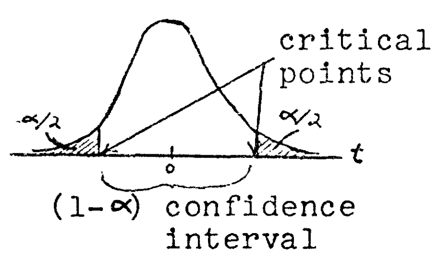



fig. V-2
t-distribution

For the t-distribution to be valid, it is required that the distribution of $Y_i$ (or $Y_i - \overline{Y}$) is normal. The "residual variance" is $s^2 = \frac{1}{N-2} \cdot \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2$ where

$\hat{Y}_i$ = the estimated value of $Y_i$ for each observation. If B = the true value of the regression coefficients and b = their estimated values, then the t-statistic is given by[4]

$$t_j = \frac{b_j - B_j}{\sqrt{s^2 / \sum_{i=1}^{N} X^2_{ij}}}$$

The equation for a 95% confidence interval is[5]

$$B_j = b_j \pm t_{.025} \cdot s / \sqrt{\sum_{i=1}^{N} X_{ij}{}^2}$$

This means that

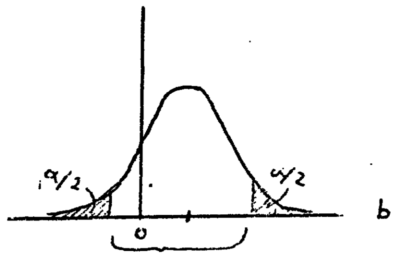
$$Pr(-t_{.025} < \frac{b - B}{\sqrt{s^2 / \sum_{i=1}^{N} X^2_{ij}}} < t_{.025}) = .95$$

where $t_{.025}$ is found in a standard table of t-values. To test the null hypothesis, substitute B=0 into the above equation:

$$t_j = b_j / \sqrt{s^2 / \sum_{i=1}^{N} (X_{ij}{}^2)}$$

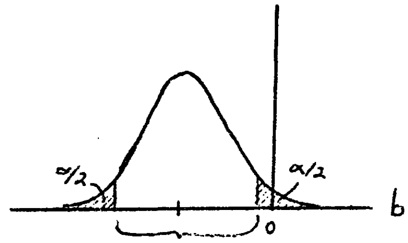
If b, the estimated value of B, is positive, the distribution for b is shown in figure V-3. The midpoint of the distribution is the estimated value b. The standard deviation of the regression coefficient b determines how "spread out" the curve is.

fig. V-3




(1-α) confidence interval

fig. V-4




(1-α) confidence interval

If the point b=o lies outside the confidence interval, that is, in the shaded portion of the curve of figure V-3, then the probability that B is indeed zero (the null hypothesis is true but has been rejected) is given by the shaded area of the curve. Thus, if $\alpha/2 = .025$, then the confidence interval checked is the 95% confidence interval (1-α) and if b, the estimated value of B, is positive, and the point b=0 lies in the shaded portion of the figure V-3, then the probability that B=0 is α/2 or .025. (For a calculated b less than zero, the appropriate curve is figure V-4 above.)

The significant number given after t-statistic in Chapter 3 and Appendix IV of this paper is α. Using α, the probability of a false rejection of the null hypothesis can be calculated. That probability is α/2.

The level of the confidence interval which includes the given regression coefficient, b, is (1-α). Thus, for an α of .1, the probability that the null hypothesis has been falsely rejected is .05 and b lies within the 90% confidence interval.

The F-test similarly tests the hypothesis that $B_1 = B_2 = \ldots = B_{M-1} = 0$. The significance number given after the F-statistic is the probability that this hypothesis has been falsely rejected by assigning non-zero values to the b's.

However, as in all probabilistic functions, it must be remembered that a low t- or F-statistic does not mean that B=0. Rather, it indicates that the hypothesis that B does equal zero <u>can not be rejected</u>. Similarly, a high t- or F-statistic means only that we <u>probably</u> will not wrongly reject the null hypothesis.

# References

Chapter 1

1. Transmission Technical Advisory Committee for the National Power Survey, "The Transmission of Electric Power," p. 89.

Chapter 3

1. IBM , System/360, SSP, Version III

Chapter 4

1. FPC, National Power Survey 1970, p. I-4-26.

Appendix I

1. Transmission Technical Advisory Committee, *op. cit.*, p. 6.
2. Federal Power Commission, National Power Survey 1970.
3. Transmission Technical Advisory Committee, *op. cit.*, p. 11.
4. *Ibid.*, p. 23.
5. *Ibid.*,p. 24.
6. *Ibid.*, p. 11.
7. *Ibid.*, p. 11.
8. *Ibid.*, p. 12.
9. *Ibid*, p. 7.

Appendix II

1. National Electrical Manufacturers Association, "Second Biennial Survey of Power Equipment Requirements of the U.S. Electric Utility Industry".

2. Transmission Technical Advisory Committee, op. cit., p. 23.

Appendix V

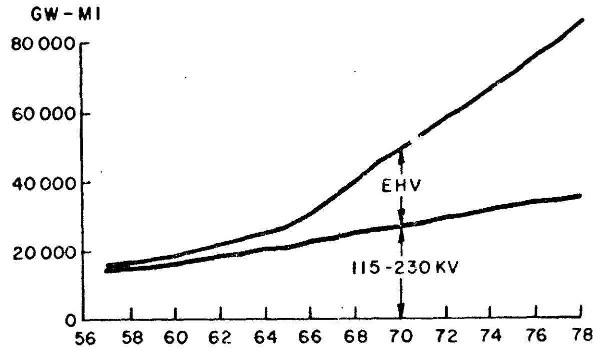
1. Wonnacott, R.J., and Wonnacott, T.H., Econometrics, p. 125.

2. Ibid., p. 119.

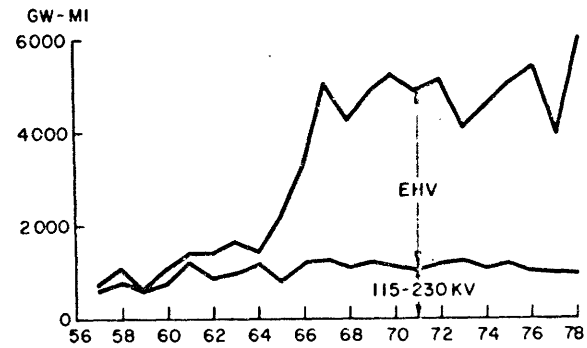
3. IBM, op. cit.

4. Wonnacott and Wonnacott, op. cit. p. 25.

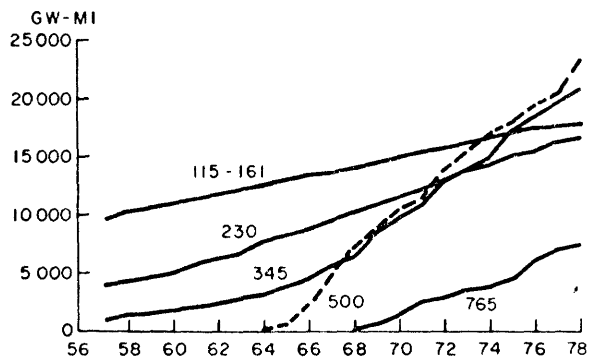
5. Ibid., p. 25.




Transmission Capability




Transmission Capability Additions

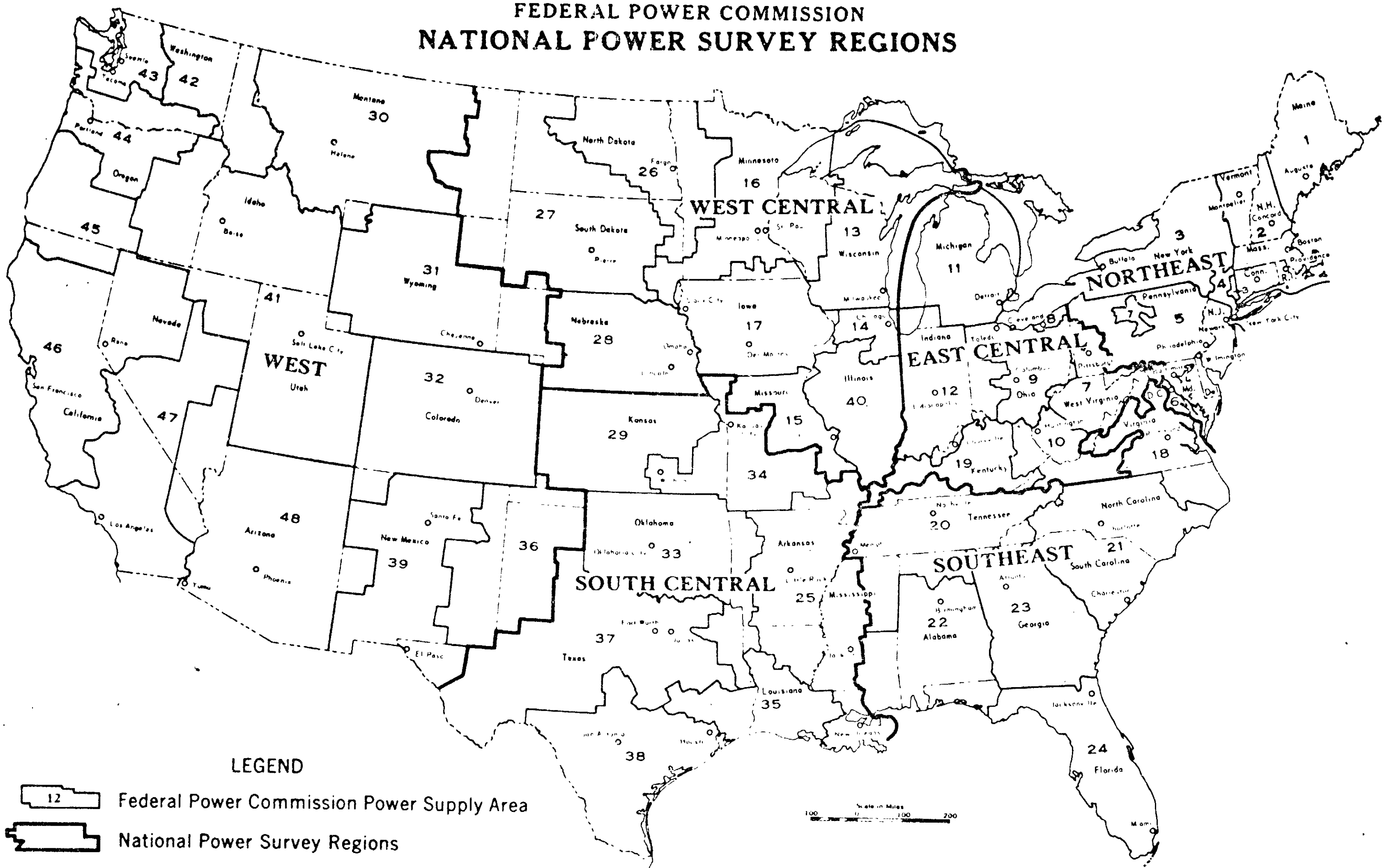



Transmission Capability by Voltage Class

Figure 1. : Growth of EHV Line

Source: NEMA, "Second Biennial Survey,"p. 10 & 12

Figure 2
FEDERAL POWER COMMISSION
# NATIONAL POWER SURVEY REGIONS

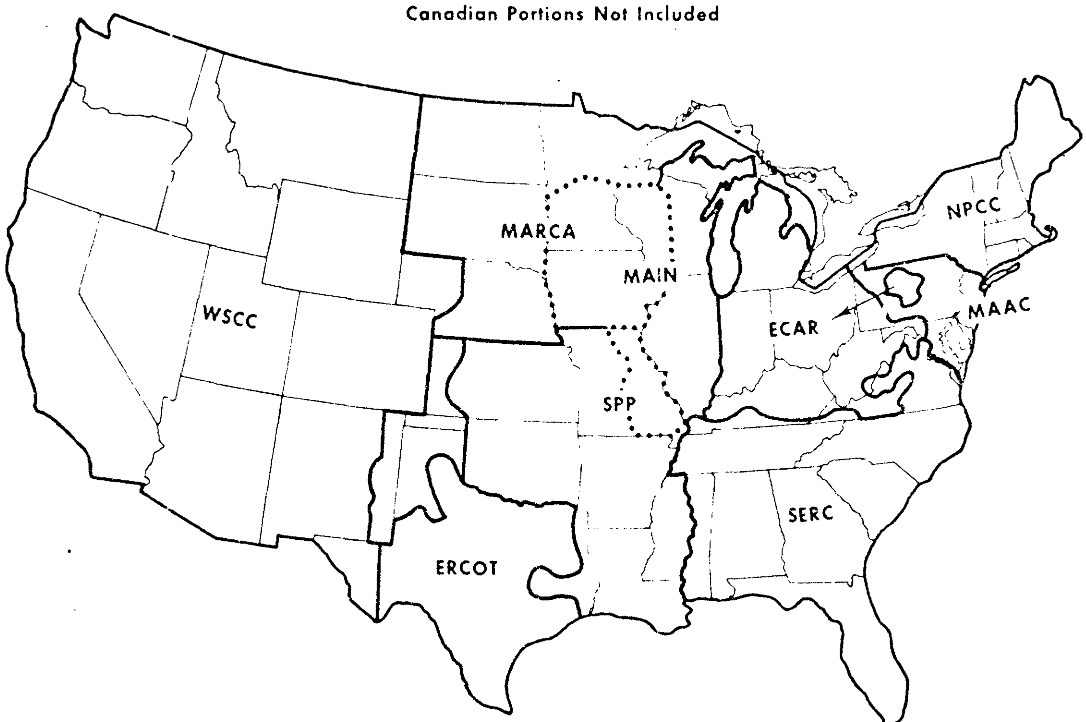



Source: FPC, National Power Survey 1970, p.II-2-2

Figure 3

**NATIONAL ELECTRIC RELIABILITY COUNCIL REGIONS**

Canadian Portions Not Included




| | |
|---|---|
| **WSCC** | Western Systems Coordinating Council |
| **MARCA** | Mid-Continent Area Reliability Coordination Agreement |
| **SPP** | Southwest Power Pool |
| **ERCOT** | Electric Reliability Council Of Texas |
| **MAIN** | Mid-America Interpool Network |
| **NPCC** | Northeast Power Coordinating Council |
| **MAAC** | Mid-Atlantic Area Coordination Group |
| **ECAR** | East Central Area Reliability Coordination Agreement |
| **SERC** | Southeastern Electric Reliability Council |

Source: FPC, National Power Survey 1970, p. I-17-16

Table 1: Some Voltage Class-Dependent Characteristics of Transmission Lines

| Voltage | Loading[1] (MW) | Load Scale[2] Factor | Transmission[1] Distance (mi.) | Structures[1] per mile | Single Circuit[1] Right-of-Way width(yd.) | Single Circuit[1] Right-of-Way acres/mile |
|---|---|---|---|---|---|---|
| 115 KV | 60-100 | .10 | 22 | 19 | 100 | 12 |
| 138 KV | 90-100 | .10 | 30 | 8.5 | 100 | 12 |
| 161 KV | 90-150 | .10 | 32 | 8 | 100 | 12 |
| 230 KV | 150-200 | .25 | 48 | 7.5 | 125 | 15 |
| 345 KV | 300-900 | .60 | 60 | 7 | 125 | 15 |
| 500 KV | 600-3,000 | 1.2 | 66 | 7.8 | 150 | 18 |
| 765 KV | 1,500-4,300 | 2.5 | 60 | 4.5 | 200 | 24 |

Sources: 1 Istvan, Rudyard, "Outputs and Growth in the Electric Utilities"

2 NEMA, "2nd Biennial Survey of Power Equipment Requirements of the U.S. Electric Utility Industry"

Table 2

EHV Transmission Line Mileage

in the United States

| Year | 230 KV | 287 KV | 345 KV | 500 KV | 765 KV | ±400 DC | Total |
|---|---|---|---|---|---|---|---|
| 1940 | 2327 | 647 | | | | | 2974 |
| 1950 | 7383 | 791 | | | | | 8174 |
| 1960 | 18701 | 1024 | 2461 | 13 | | | 22379 |
| 1970 | 40600 | 1020 | 15180 | 7220 | 500 | 850 | 65370 |
| 1980 | 59560 | 870 | 32670 | 20180 | 3540 | 1670 | 118990 |
| 1990 | 67180 | 500 | 47450 | 33400 | 8940 | 1670 | 159200 |

Source: FPC, 1970 National Power Survey, p. I-13-4.

## Table 3

## Costs of EHV Line

Average Total Costs/Mile* in 1964[1]

230 KV: $60,000

345 KV: $77,000

500 KV: $99,000

700 KV: $143,000

*Costs have since risen

Costs/Mile for 1968 in the Northeast Region (in thousands of $)

| New England | Right-of-Way | Line |
|---|---|---|
| 345 KV | $30-$60 | $50-$85 |
| 765 KV | $40-$80 | $200 |
| New York | | |
| 345 KV | $10 | $100-$220 |
| 500 KV | $50 | $200 |
| 765 KV | $75 | $200 |
| P J M | | |
| 500 KV | $10-$100 | $120-$200 |
| 765 KV | $10-$100 | $200 |
| Northeast Region | | |
| 345 KV | $10-$60 | $50-$220 |
| 500 KV | $10-$100 | $120-$200 |
| 765 KV | $10-$100 | $200 |

Sources: 1 FPC, 1964 National Power Survey, Vol. I, pp. 151

2 FPC, 1970 National Power Survey, pp. II-1-63

Table 4

**Actual AC Line Costs, Nominal 500 KV and 700 KV**

| Conductors | Cost per mile: R/W and clearing | Cost per mile: Line Construction | Cost per mile: Total |
|---|---|---|---|
| Eastern area—500 kV: | | | |
| 2—2037 ACSR | $30,700 | $ 80,800 | $111,500 |
| 2—2493 ACAR | 13,500 | 128,500 | 142,000 |
| 2—2049 5005 | 16,700 | 85,800 | 102,500 |
| 3— 971 ACSR | 12,400 | 65,300 | 77,400 |
| 4— 583 ACSR | 10,000 | 95,500 | 105,500 |
| 2—2032 ACSR | 17,000 | 98,000 | [1] 115,000 |
| 2—2490 ACAR | 20,000 | 142,000 | [1] 162,000 |
| 2—2490 ACAR | 59,000 | 272,000 | [1,2] 331,000 |
| 2—2500 ACAR | 22,000 | 118,000 | [1] 140,000 |
| 3— 954 ACSR | 12,000 | 95,000 | [1] 107,000 |
| Central area—500 kV: | | | |
| 3— 954 ACSR | ........ | 84,200 | ........ |
| 3—1024 ACAR | 24,000 | 95,600 | 119,600 |
| Western area—500 kV: | | | |
| 2—1780 ACSR | 7,100 | 72,200 | 79,300 |
| 2—1852 ACSR | ........ | 82,000 | ........ |
| 2—2156 ACSR | 25,000 | 93,900 | 118,900 |
| 2—2156 ACSR | 2,000 | 124,000 | [1,3] 126,000 |
| 700 kV: | | | |
| Average of 735 kV and 765 kV lines | 18,700 | 146,300 | [4] 165,000 |

[1] Latest data—other data prior to 1969.
[2] Line near urban center.
[3] Desert construction.
[4] Includes line sections built over 4-year span.

Source: Transmission Technical Advisory Committee, <u>The Transmission of Electric Power</u>, p. 23

Table 5-A: Data for Models

| REGION | AREA | %METRO | %I | %R | %C | %F | %H | %N | GEN.CAP. |
|---|---|---|---|---|---|---|---|---|---|
| NE60 | 182,056 | 19.6 | 47.0 | 30.6 | 22.4 | 89.0 | 11.0 | 0 | 33,860 |
| NE65 | " | 21.6 | 45.3 | 30.2 | 24.5 | 85.3 | 13.5 | 1.2 | 47,833 |
| NE70 | " | 23.6 | 44.3 | 30.8 | 24.9 | 80.9 | 11.7 | 7.4 | 62,347 |
| NE80 | " | 27.6 | 41.8 | 32.1 | 26.1 | 51.9 | 13.1 | 35.0 | 116,890 |
| NE90 | " | 31.6 | 38.7 | 33.7 | 27.6 | 31.7 | 9.9 | 58.4 | 213,420 |
| ND60 | 66,608 | 7.3 | 46.0 | 36.9 | 17.1 | 82.2 | 17.8 | 0 | 7,055 |
| ND65 | " | 8.1 | 41.6 | 36.6 | 21.7 | 84.0 | 13.9 | 2.1 | 8,905 |
| ND70 | " | 8.8 | 40.1 | 36.8 | 23.1 | 80.9 | 9.7 | 11.4 | 12,652 |
| ND80 | " | 10.3 | 37.7 | 37.1 | 25.2 | 46.3 | 13.7 | 40.0 | 28,000 |
| ND90 | " | 11.8 | 35.3 | 37.3 | 27.4 | 27.0 | 11.2 | 61.8 | 52,100 |
| NY60 | 49,576 | 25.1 | 45.0 | 28.4 | 26.6 | 83.5 | 16.5 | 0 | 12,379 |
| NY65 | " | 28.1 | 41.8 | 28.1 | 30.1 | 75.0 | 23.4 | 1.6 | 16,942 |
| NY70 | " | 31.1 | 39.1 | 30.1 | 30.8 | 70.6 | 18.1 | 11.3 | 21,915 |
| NY80 | " | 37.1 | 35.0 | 32.4 | 32.6 | 40.8 | 20.2 | 39.0 | 34,550 |
| NY90 | " | 43.1 | 31.1 | 34.6 | 34.3 | 25.7 | 16.6 | 57.7 | 59,750 |

Table 5-A continued

| REGION | AREA | %METRO | %I | %R | %C | %F | %H | %N | GEN.CAP. |
|---|---|---|---|---|---|---|---|---|---|
| PJ60 | 65,872 | 28.0 | 52.3 | 29.6 | 18.1 | 95.7 | 4.3 | 0 | 14,426 |
| PJ65 | " | 30.5 | 48.1 | 28.9 | 23.0 | 93.7 | 5.8 | 0.5 | 21,986 |
| PJ70 | " | 33.0 | 49.7 | 28.7 | 21.6 | 90.0 | 7.5 | 2.5 | 27,780 |
| PJ80 | " | 38.0 | 47.7 | 29.8 | 22.5 | 61.8 | 8.2 | 30.0 | 54,340 |
| PJ90 | " | 43.0 | 44.8 | 31.5 | 23.7 | 37.7 | 5.4 | 56.9 | 101,570 |
| EC60 | 200,305 | 14.3 | 66.8 | 22.4 | 10.8 | 96.5 | 3.5 | 0 | 30,389 |
| EC65 | " | 17.6 | 55.1 | 28.6 | 16.3 | 97.1 | 2.7 | 0.2 | 37,906* |
| EC70 | " | 20.9 | 55.6 | 27.9 | 16.5 | 92 | 6 | 2 | 55,215 |
| EC80 | " | 27.5 | 54.9 | 27.6 | 17.5 | 76 | 8 | 16 | 103,560 |
| EC90 | " | 34.1 | 54.1 | 27.8 | 18.1 | 43 | 2 | 55 | 188,635 |
| SE60 | 355,000 | 9.4 | 56.5 | 29.6 | 13.9 | 78.0 | 22.0 | 0 | 30,586 |
| SE65 | " | 11.1 | 50.8 | 33.2 | 16.0 | 81.0 | 19.0 | 0 | 43,863 |
| SE70 | " | 13.6 | 48.0 | 34.5 | 17.5 | 85.2 | 14.8 | 0 | 63,700 |
| SE80 | " | 17.8 | 50.1 | 28.7 | 21.2 | 82.8 | 11.3 | 25.8 | 132,000 |
| SE90 | " | 22.0 | 51.7 | 22.1 | 26.2 | 52.9 | 10.2 | 36.9 | 255,000 |

*1966 value of generating capacity

Table 5-A continued

| REGION | AREA | %METRO | %I | %R | %C | %F | %H | %N | GEN.CAP. |
|---|---|---|---|---|---|---|---|---|---|
| SC60 | 579,856 | 5.9 | 57.5 | 25.3 | 17.2 | 94.6 | 5.4 | 0 | 19,770 |
| SC65 | " | 7.4 | 43.4 | 32.7 | 23.9 | 92.8 | 7.2 | 0 | 31,226 |
| SC70 | " | 8.8 | 45.2 | 32.4 | 22.4 | 95.1 | 4.9 | 0 | 45,904 |
| SC80 | " | 11.7 | 48.6 | 31.8 | 19.6 | 86.8 | 5.6 | 7.6 | 107,735 |
| SC90 | " | 14.6 | 51.7 | 31.7 | 16.6 | 72.5 | 5.7 | 21.8 | 218,905 |
| WC60 | 547,525 | 6.3 | 45.5 | 35.8 | 18.7 | 91.4 | 8.6 | 0 | 17,438 |
| WC65 | " | 7.1 | 44.2 | 35.8 | 20.0 | 88.7 | 10.3 | 1.0 | 28,529 |
| WC70 | " | 7.8 | 41.6 | 38.0 | 20.4 | 83 | 8 | 9 | 42,065 |
| WC80 | " | 9.3 | 41.3 | 38.2 | 20.5 | 58 | 5 | 37 | 79,832 |
| WC90 | " | 10.8 | 40.4 | 39.0 | 20.6 | 40 | 3 | 57 | 151,041 |
| W 60 | 1,187,753 | 9.9 | 44.2 | 35.2 | 20.6 | 42.5 | 57.5 | 0 | 31,390 |
| W 65 | " | 10.8 | 40.0 | 36.3 | 23.7 | 48.5 | 50.1 | 0.4 | 45,568 |
| W 70 | " | 11.6 | 40.2 | 36.5 | 23.3 | 50.7 | 47.5 | 1.8 | 66,700 |
| W 80 | " | 13.3 | 41.2 | 35.8 | 23.0 | 44.2 | 36.4 | 19.4 | 120,436 |
| W 90 | " | 15.0 | 46.2 | 32.6 | 21.2 | 35.3 | 25.1 | 39.6 | 254,000 |

Table 5-B: Data

| REGION | EM | CM | POPULATION (millions) | P | LC-500 | LC-1000 |
|---|---|---|---|---|---|---|
| NE60 | 542 | 2,130 | 41.3 | | | |
| NE70 | 3,124 | 7,452 | 46.9 | 60 | 17 | 9 |
| NE80 | 5,331 | 10,855 | 53.0 | 88 | 32 | 17 |
| NE90 | 13,104 | 15,610 | 60.0 | 105 | 44 | 31 |
| ND60 | 67 | 266 | | | | |
| ND70 | 618 | 1,397 | | 11 | 3 | 1 |
| ND80 | 1,328 | 2,580 | | 21 | 10 | 3 |
| ND90 | 2,467 | 3,370 | | 27 | 17 | 9 |
| NY60 | 117 | 371 | | | | |
| NY70 | 561 | 1,600 | | 24 | 5 | 3 |
| NY80 | 987 | 2,310 | | 30 | 7 | 5 |
| NY90 | 3,411 | 3,595 | | 34 | 9 | 7 |
| PJ60 | 358 | 1,393 | | | | |
| PJ70 | 1,945 | 4,455 | | 25 | 9 | 5 |
| PJ80 | 3,016 | 5,965 | | 37 | 15 | 9 |
| PJ90 | 7,226 | 8,645 | | 44 | 18 | 15 |
| EC60 | 1,189 | 2,221 | 28.3 | | 12 | 5 |
| EC70 | 5,550 | 6,160 | 30.2 | 54 | 33 | 11 |
| EC80 | 13,818 | 13,730 | 36.2 | 91 | 50 | 28 |
| EC90 | 17,058 | 17,030 | 40.9 | 112 | 63 | 44 |
| SE60 | 324 | 1,318 | 27.8 | | | |
| SE70 | 2,529 | 7,160 | 31.4 | 39 | 33 | 16 |
| SE80 | 9,385 | 21,770 | 35.4 | 77 | 62 | 33 |
| SE90 | 16,154 | 30,340 | 40.1 | 105 | 71 | 61 |
| SC60 | 38 | 143 | 21.5 | | | |
| SC70 | 3,778 | 5,700 | 24.4 | 42 | 25 | 9 |
| SC80 | 8,260 | 12,780 | 28.3 | 81 | 46 | 25 |
| SC90 | 15,011 | 18,270 | 33.1 | 116 | 66 | 45 |
| WC60 | 874 | 3,417 | 25.7 | | | |
| WC70 | 3,232 | 8,770 | 28.1 | 25 | 14 | 16 |
| WC80 | 8,384 | 14,780 | 32.4 | 41 | | |
| WC90 | 16,423 | 22,060 | 37.8 | 60 | 19 | 17 |
| W 60 | 3,599 | 13,962 | 27.7 | | | |
| W 70 | 11,996 | 29,280 | 34.7 | 45 | 27 | 15 |
| W 80 | 22,608 | 42,910 | 43.4 | 75 | 36 | 28 |
| W 90 | 31,946 | 51,970 | 52.4 | 103 | 42 | 36 |

Table 5-C: Data

| REGION | TPL | APL | AAPL | NTPL | NAPL | NAAPL | OTPL | OAPL | OAAPL |
|---|---|---|---|---|---|---|---|---|---|
| NE70 | 1,992 | 33 | 33 | | | | | | |
| NE80 | 2,636 | 30 | 55 | 1,556 | 55 | 101 | 1,080 | 18 | 33 |
| NE90 | 2,526 | 24 | 69 | 828 | 49 | 142 | 1,698 | 19 | 55 |
| ND70 | 438 | 40 | 40 | | | | | | |
| ND80 | 891 | 42 | 93 | 573 | 48 | 106 | 198 | 18 | 40 |
| ND90 | 1,212 | 45 | 120 | 471 | 79 | 210 | 741 | 35 | 93 |
| NY70 | 474 | 20 | 20 | | | | | | |
| NY80 | 578 | 19 | 32 | 290 | 49 | 82 | 288 | 12 | 20 |
| NY90 | 534 | 16 | 37 | 126 | 32 | 73 | 408 | 14 | 32 |
| PJ70 | 1,080 | 43 | 43 | | | | | | |
| PJ80 | 467 | 32 | 57 | 693 | 69 | 124 | 594 | 24 | 43 |
| PJ90 | 780 | 18 | 68 | 231 | 33 | 124 | 549 | 15 | 57 |
| EC70 | 1,816 | 34 | 34 | | | | | | |
| EC80 | 2,440 | 27 | 40 | 1,180 | 32 | 47 | 1,260 | 23 | 34 |
| EC90 | 2,260 | 20 | 42 | 553 | 26 | 54 | 1,707 | 19 | 40 |
| SE70 | 1,664 | 42 | 42 | | | | | | |
| SE80 | 2,388 | 31 | 50 | 1,354 | 36 | 58 | 1,034 | 26 | 42 |
| SE90 | 2,100 | 20 | 53 | 619 | 22 | 58 | 1,481 | 19 | 50 |
| SC70 | 2,720 | 63 | 63 | | | | | | |
| SC80 | 3,175 | 39 | 85 | 1,965 | 50 | 108 | 1,210 | 29 | 63 |
| SC90 | 3,405 | 30 | 87 | 1,020 | 31 | 90 | 2,385 | 29 | 85 |
| WC70 | 1,368 | 49 | 49 | | | | | | |
| W 70 | 4,315 | 96 | 96 | | | | | | |
| W 80 | 3,253 | 43 | 103 | 1,346 | 48 | 115 | 1,738 | 40 | 96 |
| W 90 | 4,461 | 43 | 106 | 1,515 | 48 | 118 | 3,115 | 42 | 103 |

Table 5-D: Data

| REGION | % FOSSIL new | % HYDRO new | % NUCLEAR new |
|---|---|---|---|
| NE80 | 18.7 | 14.6 | 66.7 |
| NE90 | 7.3 | 6.2 | 86.5 |
| ND80 | 19.6 | 16.9 | 63.5 |
| ND90 | 4.6 | 8.3 | 87.1 |
| NY80 | -10.8 | 23.7 | 87.1 |
| NY90 | 4.8 | 11.9 | 83.3 |
| PJ80 | 32.4 | 8.8 | 58.8 |
| PJ90 | 10.0 | 2.1 | 87.9 |
| EC80 | 61.2 | 10.1 | 28.7 |
| EC90 | 42.1 | 2.4 | 55.5 |
| SE80 | 42.0 | 8.2 | 49.8 |
| SE90 | 42.3 | 8.9 | 48.8 |
| SC80 | 79.7 | 6.3 | 14.0 |
| SC90 | 58.1 | 5.7 | 36.2 |
| WC80 | 31.4 | 0.1 | 68.5 |
| WC90 | 19.9 | 0.6 | 79.5 |
| W 80 | 37.6 | 25.3 | 37.1 |
| W 90 | 26.2 | 13.5 | 60.3 |

Table 6

Data and Scale Factors Used in Regressions

| Symbol | Variable | Scale Factor |
|---|---|---|
| CM & EM | Miles of Line | $10^{-3}$ |
| Pop | Population | $10^{-6}$ |
| A | Area (sq. miles) | $10^{-5}$ |
| P | Number of Plants 400 MW | $10^{-1}$ |
| LC | Number of Load Centers | $10^{-1}$ |
| I,R & C | % of GWH for Industrial, Residential, or Commercial Users | $10^{-1}$ |
| G | Generating Capacity (MW) | $10^{-4}$ |
| METRO | % Metropolitan Area | $10^{-1}$ |
| F, H & N | % Generating Capacity Fossil Fueled, Hydroelectric or Nuclear | $10^{-1}$ |
| TPL | Total Plant to Load Center Distance (Miles) | $10^{-2}$ |
| APL | Average Plant to Load Center Distance (Miles) | $10^{-1}$ |
| AAPL | Average Adjusted Plant to Load Center Distance (Miles) | $10^{-1}$ |
| NTPL | TPL for Plants Built within Last 10 years | $10^{-2}$ |
| NAPL | APL for Plants Built within Last 10 years | $10^{-1}$ |
| NAAPL | AAPL for Plants Built within Last 10 years | $10^{-1}$ |
| OTPL | TPL for Plants over 10 Years Old | $10^{-2}$ |
| OAPL | APL for Plants over 10 Years Old | $10^{-1}$ |
| OAAPL | AAPL for Plants over 10 Years Old | $10^{-1}$ |
| $F_{new}$, $H_{new}$ & $N_{new}$ | % of Generating Capacity added within the Last 10 Years that is Fossil Fueled, Hydroelectric or Nuclear | $10_{-1}$ |

TABLE 7

COMPARISON OF LINEAR AND EXPONENTIAL PROJECTIONS OF VARIABLE METRO

A. Model: $CM = e^{K} A^{a_1} P^{a_2} P^{a_3} G^{a_4} METRO^{a_5} H^{a_6} AAPL^{a_7}$

Regions and years used: (NE, ND, NY, PJ, EC, SE, SC, W) 70, 80, 90
(WC) 70

Degrees of freedom*: 7/17

| | Linear | Exponential |
|---|---|---|
| F-Statistic/Significance*: | 67.51/.01 | 65.61/.01 |
| Average % Error: | 15.6% | 15.6% |

| Coefficient | Value | Std. Dev. | t-value/α* | Value | Std. Dev. | t-value/α* |
|---|---|---|---|---|---|---|
| K | -1.946 | | | -1.938 | | |
| $a_1$ | .583 | .113 | 5.16/.01 | .594 | .120 | 4.96/.01 |
| $a_2$ | - .471 | .280 | 1.68/.2 | - .449 | .283 | 1.59/.2 |
| $a_3$ | 1.325 | .480 | 2.76/.05 | 1.335 | .486 | 2.75/.05 |
| $a_4$ | .756 | .218 | 3.47/.02 | .685 | .228 | 3.00/.02 |
| $a_5$ | .566 | .188 | 3.01/.05 | .577 | .199 | 2.89/.05 |
| $a_6$ | .223 | .100 | 2.23/.1 | .249 | .099 | 2.53/.05 |
| $a_7$ | .341 | .157 | 2.17/.1 | .358 | .163 | 2.20/.1 |

B. Model: $EM = e^{K} A^{a_1} AAPL^{a_2} G^{a_3 METRO}$

Regions and years used: (NE, ND, NY, PJ, EC, SE, SC, W) 70, 80, 90
(WC) 70

Degrees of freedom*: 3/21

| | Linear | Exponential |
|---|---|---|
| F-Statistic/Significance*: | 89.00/.01 | 91.02/.01 |
| Average % Error: | 23.6% | 22.3% |

| Coefficient | Value | Std. Dev. | t-value/α* | Value | Std. Dev. | t-value/α* |
|---|---|---|---|---|---|---|
| K | -.936 | | | -.869 | | |
| $a_1$ | .728 | .070 | 10.44/.01 | .696 | .069 | 10.13/.01 |
| $a_2$ | .546 | .165 | 3.31/.05 | .572 | .163 | 3.50/.05 |
| $a_3$ | .242 | .023 | 10.43/.01 | .202 | .019 | 10.56/.01 |

*See Appendix V for definition of these terms

TABLE 8

Summary of Models Developed

Model #

1 $CM = e^K A^{a_1} P^{a_2} I^{a_3} G^{a_4} METRO^{a_5} H^{a_6} AAPL^{a_7}$

2 $EM = e^K A^{a_1} G^{a_2} METRO^{a_3} AAPL^{a_4}$

3 $CM = e^K A^{a_1} G^{a_2} APL^{a_3} G^{a_4 METRO}$

4 $CM = e^K A^{a_1} AAPL^{a_2} G^{a_3 I} G^{a_4 METRO}$

5 $CM = e^K A^{a_1} AAPL^{a_2} G^{a_3 METRO}$

6 $EM = e^K A^{a_1} AAPL^{a_2} G^{a_3 METRO}$

7 $EM = e^K A^{a_1} AAPL^{a_2} G^{a_3} G^{a_4 METRO}$

Model #'s

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| K | -1.946 | -1.384 | -.185 | .219 | .304 | -.936 | -.999 |
| $a_1$ | .583 | .502 | .490 | .639 | .756 | .728 | .563 |
| $a_2$ | - .471 | .708 | .424 | .301 | .319 | .546 | .420 |
| $a_3$ | 1.325 | .582 | .510 | .0448 | .172 | .242 | .377 |
| $a_4$ | .756 | .490 | .105 | .125 | | | .163 |
| $a_5$ | .566 | | | | | | |
| $a_6$ | .223 | | | | | | |
| $a_7$ | .341 | | | | | | |
| Average % Error: | 15.6% | 23.2% | 20% | 20% | 22% | 24% | 21% |
| See Also Table: | 3-1 | 3-2 | 3-7 | 3-8 | 3-9 | 3-10 | 3-11 |

BIBLIOGRAPHY

Bonneville Power Administration,U.S. Department of the Interior, Pacific Northwest Economic Base Study for Power Markets, Volume 1: Summary, Bonneville Power Administration.

Bureau of the Census, U.S. Department of Commerce, 1970 Census of Population, Volume 1: Number of Inhabitants, Part 1: U.S. Summary, U.S. Government Printing Office, Washington, D.C., 1971.

Bureau of the Census, U.S. Department of Commerce, 1960 Census of Population, Volume 1: Characteristics of Population, Part 1: Summary, U.S. Government Printing Office, Washington, D.C., 1961.

Draper, N.R. and Smith, H., Applied Regression Analysis, John Wiley and Sons, Inc., New York, 1966.

Edison Electric Institute, Statistical Year Book of the Electric Utility Industry for 1960, for 1961, for 1965, for 1966, for 1970, E.E.I., New York.

Electrical World, 1971 Annual Statistical Report, McGraw Hill, New York.

Electrical World, Directory of Electric Utilities, 75th Edition 1966, 76th Edition 1967, 77th Edition 1968, 78th Edition 1969-70, 79th Edition 1970-71, McGraw Hill, New York.

Federal Power Commission, National Power Survey 1964, U.S. Government Printing Office, Washington, D.C.

Federal Power Commission, National Power Survey 1970, U.S. Government Printing Office, Washington, D.C.

Federal Power Commission, "Principal Electric Facilities", Map series for 1971, U.S. Government Printing Office, Washington, D.C., 1971.

Federal Power Commission, Statistics of Privately Owned Electric Utilities in the United States, 1960 and 1968, U.S. Government Printing Office, Washington, D.C., 1961 and 1969.

Federal Power Commission, Statistics of Publicly Owned Electric Utilities in the United States, 1960, and 1968, U.S. Government Printing Office, Washington, D.C., 1961 and 1969.

Goss, Floyd L., "Transporting Energy by Wire", in Electrical World, Energy for the Seventies, Proceedings of the Utility Executive Conference, Washington, D.C., January 14-15, 1971.

IBM, System/360 Scientific Subroutine Package, Version III, Programmer's Manual, IBM Corp., White Plains, New York, 1970.

Istvan, Rudyard, "Outputs and Growth in the Electric Utilities", B.A. Thesis, Department of Economics, Harvard University, Cambridge, Mass., 1972.

Johnston, J., Econometric Methods, McGraw Hill, New York 1963.

National Electrical Manufacturers Association, "Second Biennial Survey of Power Equipment Requirements of the U.S. Electric Utility Industry, 1969-78", NEMA, New York, 1970.

Transmission Technical Advisory Committee for the National Power Survey, The Transmission of Electric Power, Federal Power Commission, Washington, D.C., 1971.

Wonnacott, R.J., and Wonnacott, T.H., Econometrics, John Wiley and Sons, New York, 1970.