73-26,325

SPURLOCK, Carl Wayne, 1941-
PREDICTIVE MARKETING REGIONS FOR LIQUIFIED PETROLEUM GAS: A CONCEPTUAL RELATIONSHIP OF HEATING DEGREE DAYS AND DEMAND.

The University of Oklahoma, Ph.D., 1972
Geography

University Microfilms, A XEROX Company, Ann Arbor, Michigan

PREDICTIVE MARKETING REGIONS FOR LIQUIFIED PETROLEUM GAS: A CONCEPTUAL RELATIONSHIP OF HEATING DEGREE DAYS AND DEMAND

By

CARL W. SPURLOCK

A DISSERTATION PRESENTED TO THE GRADUATE
COLLEGE OF THE UNIVERSITY OF OKLAHOMA IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF OKLAHOMA
1972

PREDICTIVE MARKETING REGIONS FOR LIQUIFIED PETROLEUM GAS:

A CONCEPTUAL RELATIONSHIP OF HEATING DEGREE DAYS AND DEMAND

A DISSERTATION

APPROVED FOR THE DEPARTMENT OF GEOGRAPHY

By

ACKNOWLEDGMENTS

The completion of this dissertation was made possible by the advice and encouragement of many people in many places. To all of these people, I express my sincere appreciation.

In addition, certain individuals made direct contributions to this study, although at this point it is often difficult to distinguish between my ideas and those contributed by others. To Professor Gary Thompson, Chairman of my Ph.D. Committee, I owe a special thanks. Not only did he make important conceptual contributions, but also he provided encouragement and invested a great deal of time and effort in trying to do something about my unorthodox views on grammar and sentence construction. Professor Jim Bohland made contributions that ranged from the initial encouragement to pursue this topic as a dissertation, to eleventh-hour help with a computer program. Professor Nelson Nunnally originally suggested the methodology that was used, as well as some key ideas that I regard as being especially significant. Professor Bill Southard provided assistance with the marketing aspects of the study. To Professor John Morris, I owe a special debt of gratitude. His assistence and encouragement goes beyond the dissertation to include my entire graduate program at the University of Oklahoma.

Officials of Phillips Petroleum Company, Bartlesville, Oklahoma, were always helpful within the limits of their corporate policy. In particular, Mr. Sheldon Cunningham, of the Marketing and Supply Division, provided valuable information.

I owe the largest debt of all to my wife, Norma. Not only did she make direct contributions in the form of cartographic work and manuscript typing, but her encouragement, devotion, and willingness to sacrifice have been invaluable in completing this study. Norma, along with our son, Michael, have provided me with a continual reminder that writing a dissertation is not the most important thing in the world.

TABLE OF CONTENTS

TABLE OF CONTENTS, continued:

LIST OF TABLES

LIST OF MAPS

Map

Abstract of Dissertation Presented to the Graduate College of the University of Oklahoma in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

PREDICTIVE MARKETING REGIONS FOR LIQUIFIED PETROLEUM GAS: A CONCEPTUAL RELATIONSHIP OF HEATING DEGREE DAYS AND DEMAND

By

Carl W. Spurlock

September, 1972

Chairman: Professor Gary L. Thompson
Major Department: Geography

The research problem was to develop a set of marketing regions that petroleum companies could utilize to achieve a greater degree of success in predicting heating-fuel demand than the regions currently used for this purpose. It was assumed that if the derived regions had less intraregional variability of heating-degree-day occurrence, they would be better predictive regions.

A statistical technique, commonly called "numerical taxonomy" was used to accomplish the regionalization. This technique classified observations on the basis of similarities and, in a series of steps, progressively built up groups (regions). The total space of the study area had been partitioned into 372 weather polygons which served as the

observations to be grouped. Each weather polygon was associated with a single weather station, from which the mean and standard deviation of heating degree days of the ten-year period, 1962-71, had been computed. These two sets of weather variables, which did not covary, served as the similarity attributes used to group the polygons into homogeneous regions. Each weather polygon was located in taxonomic space by its heating-degree-day characteristics and a taxonomic distance measure (D^2) computed. At each step in the grouping procedure, the observations with the smallest D^2 were grouped. Since contiguity in geographical space was required, however, only those observations that were contiguous could be grouped. In addition, each weather polygon was associated with a population value (used as a surrogate for demand). In order for each region to contain an optimal demand level, a population constraint of 1,500,000 was built into the grouping procedure. At each step in the grouping, if a potential combination caused the population limit to be exceeded, the combination was not allowed. Thus, it was assured that no region's population would exceed 1,500,000. In this manner, a set of uniform weather regions that are in a best-fit relationship with demand intensity was derived.

Sixteen predictive marketing regions were delimited in the study area by the taxonomic procedure. Although the size and shape of the derived regions varied widely, thirteen of them had population totals within acceptable limits of the desired demand level. The remaining three were thought to represent weather types marginal to the study area. In many cases, the east-west length of the derived regions far exceeded the north-south width. Since the predictive regions have no distribution function, the lack of compactness is not viewed as an undesirable characteristic. Whether or not the derived regions can be utilized to predict heating-fuel demand more accurately than an existing set of LP-gas marketing regions is the most important criterion for their success.

To determine whether the derived regions were, in fact, more effective in prediction, they were tested against an existing set of LP gas-marketing regions. The objective was to see which set had the smallest sum of within-groups variation of heating-degree-day occurrence weighted by population (used as a surrogate for demand). With this consideration, the propane marketing regions of Phillips Petroleum Company, Bartlesville, Oklahoma, were obtained. Phillips uses twenty-five marketing regions within the four-state area. In order to insure that the test results were consistent, the same geographical area and the same weather stations were used for both sets of regions.

Three hundred and seventy-two weather stations with complete heating-degree-day records served as the test data for the comparison of the two sets of regions. A Total Pre-

diction Index (TPI) of 2,082.49 was computed for Phillips' regions, and a TPI of 737.87 was obtained for the derived set of regions. Thus, the derived regions showed a substantially smaller TPI in spite of consisting of eight fewer regions. The test results gave a strong indication that the derived set of regions would be more effective in predicting heating-fuel demand than those currently being used by Phillips Petroleum Company. Likewise, the results point to the feasibility of using taxonomic techniques to delimit predictive marketing regions.

CHAPTER ONE

THE MARKETING OF LIQUIFIED PETROLEUM GAS

The Need for Forecasting LP Gas Demand: An Introduction

In 1971, sales of liquified petroleum gas* in the United States reached 19.3 billion gallons, about one-half of which was utilized for residential heating (National LP-Gas Association News Release, 1971). In order to effectively serve the growing demand for LP gas as a space-heating fuel, the gas industry must operate an efficient market forecasting and product distribution system. During the past several years, however, petroleum companies which market LP gas have experienced problems in forecasting demand for such heating fuels. The problems have been caused primarily by incorrect estimates of heating fuel consumption for forthcoming heating seasons. The nature of LP gas marketing is such that long-range forecasting of the intensity and spatial variation of product demand is mandatory,

---

*The three types of liquid petroleum gas are propane, butane, and ethane; subsequently the term LP gas will be used to refer to this fuel type.

a fact which becomes more evident when LP gas-marketing channels are viewed from the perspective of the consumer. Local franchised dealers, serving retail to consumers, rely heavily on the capabilities of long-distance pipelines to deliver sufficient fuel in the amounts and at the times required to fulfill daily demands. This requirement considers both the amount previously contracted with a gas supplier, an amount that is based on a long-range estimate of local demand, and the reserve storage capabilities of the supplier. The availability of storage fuels, however, does not mean unlimited delivery capability for several reasons. First, the pipelines used to ship LP gas are classified as common carriers and, as such, transport various types of liquid products for different companies on a first-come, first-served basis. Scheduling pipeline space for LP gas must be made in advance. Hence, there exists a time lag between initial notification to the storage facility and the responding delivery of additional gas. Second, the amount of LP gas that can be delivered is limited by the size of the pipeline as well as the maintenance of a required gas pressure (the gas pressure that keeps LP gas in a liquid form) which limits the amount that may be withdrawn from storage on a per-day basis. Some form of long-range forecast applicable to estimating demand is required in order for the gas supplier, pipelines, and distributors to fulfill their obligations in an efficient and economical manner.

As a consequence of incorrect estimates, periodic oversupply or undersupply of heating fuels prevails at various places throughout the marketing areas. Both situations are highly undesirable. Underprediction for a local area may necessitate gas rationing until additional supplies become available. Large-scale underprediction may require production plants to go into full operation, a procedure which requires added time to call additional, unscheduled personnel needed to operate the plants. The cost of such "peaking gas," may run five times that of previously contracted gas (American Gas Association Monthly, 1971, p. 31). Conversely, if the experienced demand conditions are not as large as have been forecast, the disposal of overdrawn storage gas can become a major economic hardship (American Gas Association Monthly, 1971, p. 31). Unnecessary transportation expenses, decreased sales, and loss of customers occur as a result of incorrect heating fuel demand forecasts.

Incorrect Forecasts

Two major reasons appear to account for the incorrect estimates of heating fuel consumption. The most important reason is fluctuations in weather from year to year. Unexpectedly cold heating seasons will produce increased demands for heating fuels and vice versa. The major parameter for establishing LP gas needs in advance is a long-range weather forecast. A reliable projection of the number of heating degree days* that should occur in each marketing region is essential for estimates of heating fuel consumption.

From the information available, it appears that petroleum companies utilize a variety of methods to generate weather forecasts (American Gas Association Monthly, 1971, p. 29). These methods range from relatively sophisticated predictive models to a reliance on governmental publications such as the Weather Outlook. Since weather occurrence has a random component and cannot be exactly predicted for any place on a long-range basis, these methods all utilize the probabilities of weather occurrence based on long-term "normals." Given the unpredictable nature of weather occurrence, these forecasting methods produce results that are usually within acceptable reliability limits for a single point in each region. However, reliability in projecting weather occurrence is needed not only for a few representative points but also for an entire system of marketing regions.

The second major reason for incorrect estimates of LP gas demand is that data from a single weather station is

---

*Heating degree days are defined by the U.S. Weather Bureau as the difference between a constant (65^{o} F) and the mean daily temperature. For example, if the mean daily temperature were 55^{o} F, ten heating degree days would have been recorded. If the mean daily temperature is 65^{o} F or above, no heating degree days occur.

used to represent large areas. Since it is highly impractical for petroleum companies to attempt a weather prediction for every place that LP gas is sold, large numbers of places have been grouped together into what is referred to as "marketing regions" or "weather regions." One city is selected for each region (usually the city where a regional administrative office is maintained), and climatic data for that city is used to make an estimate of product demand for the entire region. If climatic data from the selected station does not correctly represent all places in that region, faulty demand estimates will result. No matter how accurate the forecast for the selected station, little is accomplished if greatly different weather occurs in another part of the region. To be effective in prediction, some degree of weather homogeneity in the region is a prerequisite. The marketing regions currently used by petroleum companies appear to lack this characteristic. For example, one major petroleum company uses an LP-gas marketing region with Oklahoma City, Oklahoma as the representative city. Oklahoma City, over the past ten years, has averaged 3,369 heating degree days per heating season (July-June). Included within the boundaries of the region are Jefferson, Oklahoma, which averages 3,909 and Ardmore, Oklahoma, averaging 2,630. This variation of 600 to 700 heating degree days makes it readily apparent that Oklahoma City does not correctly represent this region in terms of weather occurrence. However, climatic data from Oklahoma City are utilized to make a long-range, regional forecast of LP-gas sales. This example, one of many that could be cited, points to the lack of weather uniformity in the regions currently being utilized to market LP gas.

Population Distribution As a Factor in Forecasting Demand

A complicating factor, interrelated with the two reasons for incorrect estimates, is the need for higher levels of reliability in predicting fuel consumption for those areas that have greater demand. The spatial distribution of heating fuel users becomes a significant consideration. In areas having high demand intensities (urbanized areas having a large number of gas customers), the margin of error in estimating sales must be minimized. The best method of reducing the margin of error and producing a correct forecast is to have a very small predictive region, since there is a higher probability of uniform weather occurring in a small area. If all marketing regions were very small, their utility would be decreased; and the costs of utilizing an unlimited number of small regions would reach insupportable levels. Conversely, it is not so necessary to have a small marketing region if the magnitude of heating fuel demand in that region is small. A marketing region with a small number of customers cannot justify its existence when administrative cost exceeds potential revenue. Therefore, the areal size of a marketing region must be large enough to include the number of customers that will provide a satisfactory profit level. The final determination of the areal size of predictive marketing regions should reflect a compromise between what is administratively feasible and the requirements for reliable prediction.

The Administrative Function of LP Gas-Marketing Regions

The regional system utilized by petroleum companies to forecast LP gas demand, also serves an administrative function. Each region contains a regional office that is usually located in the highest order city in that region. A Regional Manager supervises each office and acts as a liaison between the parent company and the local franchised dealers who retail LP gas. Prior to a forthcoming heating season (in June or July), the Regional Manager contacts all local dealers in his region and attempts to assess the amount of LP gas each will need for the heating season. The assessment is based largely on the amount of sales from the previous heating season. After obtaining an estimate from each local retail dealer, the Regional Manager submits a total regional estimate to the market department in the company headquarters. Here, each regional estimate of demand is adjusted for expected weather occurrence by using climatic data from the representative weather station of that region. The exact procedure by which regional estimates of demand are adjusted for expected weather occurrence varies among petroleum companies. In any case, the adjusted regional estimates become the basis for preseason production of amount of LP gas that a petroleum company expects to sell during the forthcoming cold season.

Summary

Since the effective marketing of LP gas-heating fuels requires a long-range regional forecast of product demand, problems will occur if the regions are not uniform with respect to weather characteristics. Even if the methods used to forecast weather occurrence at the representative city of each region were entirely accurate, the regions would not be useful in predicting demand if greatly different weather conditions occurred in other parts of the same region. The use of predictive regions that have been arbitrarily delineated, without regard to the elements (latitude, elevation, proximity to water bodies, etc.) that would cause weather differences, has been partially responsible for the marketing problems experienced by petroleum companies.

CHAPTER TWO

APPROACHES TO THE PROBLEM OF REGIONALIZATION

Statement of Problem

The problem is to develop a set of marketing regions that petroleum companies can utilize to achieve a greater degree of success in predicting heating fuel demand than the regions currently used for this purpose. The overall problem can be divided into two interrelated subproblems. First, there is a need to develop a set of homogeneous weather regions. Second, these homogeneous weather regions must be adjusted into predictive marketing regions on the basis of spatial variation in demand. These two subproblems must be considered simultaneously as components of the overall problem.

The problem is basically one of spatial planning in which areas are to be delimited and arranged in an optimum manner. At the center of the conceptual framework for spatial planning is the idea of the region. Thus, the focus of this study is the application of and testing of a method of regionalization for the purpose of increasing the efficience of marketing LP gas.

Regionalization in Geography

The concept of "regions" has been a central theme in modern geography. Much of the present status of geography as a discipline has been built on regional studies. In general, the concern of these studies has been to ask factual questions about the landscape features under observation. The objective was to demonstrate that the character of a portion of the earth's surface was significantly different from other portions. Regional studies of this type described areas in terms of their differences from each other. In fact, in a most scholarly review of the historical development of geography, Hartshorne reports in his classic Nature of Geography that " . . .no universals need be evolved, other than the general law of geography that all its areas are unique (Hartshorne, 1939, p. 468). The concept of areal differentiation had a strong impact on geography in the 1940's.

In recent years, however, scholars in geography have advocated the use of regions as a form of classification. Schaefer (1953) and Bunge (1962) have challenged the "unique" approach to regions. Each suggests that uniqueness is a point of view and not an inherent property of regions. Instead, as reported by Haggett, "the regional concept appears to be one of the most logical and satisfactory ways of organizing geographical information" (1965, p. 241).

The organization of information is a problem common to all sciences. Harvey states "classification is . . . the basic procedure by which we impose some sort of order and coherence upon the vast inflow of information from the real world" (1969, p. 326). Classification is the systematic grouping of objects or events into classes on the basis of properties or relationships they have in common. Since information of concern to geographers is expressed in spatial terms, spatial classification systems are necessary. The regional concept appears to fit exactly this need if it can also be assumed that objects or events have sets of common characteristics and that these sets of common characteristics serve as a proper criterion for assignment to a spatial classification. This is not to say that the use of regions for areal differentiation is invalid. On the contrary, the contributions made to geographical theory by this approach were substantial. The intent of this study is to use regions as a means of areal integration. That is, regions were delineated on the basis of similarities rather than differences.

Although there have been attempts to formulate sets of multiple-criteria regional units, such as the general categories developed by Whittlesey (1954, pp. 32-51), a frequently used technique is the single-purpose region designed to meet specific research needs. Indeed, as Harvey states, "We possess . . . no means of assessing the adequacy or efficiency of a given classification independently of the job it is designed to do" (1969, p. 326). This notion applies to the regions that were generated in this research. They were designed only for a specific purpose and can only be evaluated in relation to that purpose.

Research Hypothesis

Implicit within the problem statement is the assumption that a system of LP gas marketing regions can be delimited and that these regions will be more effective than those currently being used. More "effective" in this case refers to a greater reliability in predicting heating-fuel demand. It is logical to assume that if the derived regions had less intraregional variability of heating-degree-day occurrence, they would be better predictive regions. An indication of whether the derived regions are, in fact, more effective in prediction can be derived by testing against an existing set to see which has the smallest amount of mean within-groups variation of heating-degree-day occurrence. To insure that the test results would be consistent, the same geographical area and the same weather stations are used for both sets of regions. It should be expected that the derived set of regions would have substantially less mean within-groups variation than the regions currently being used.

The research hypothesis tested was stated as follows:

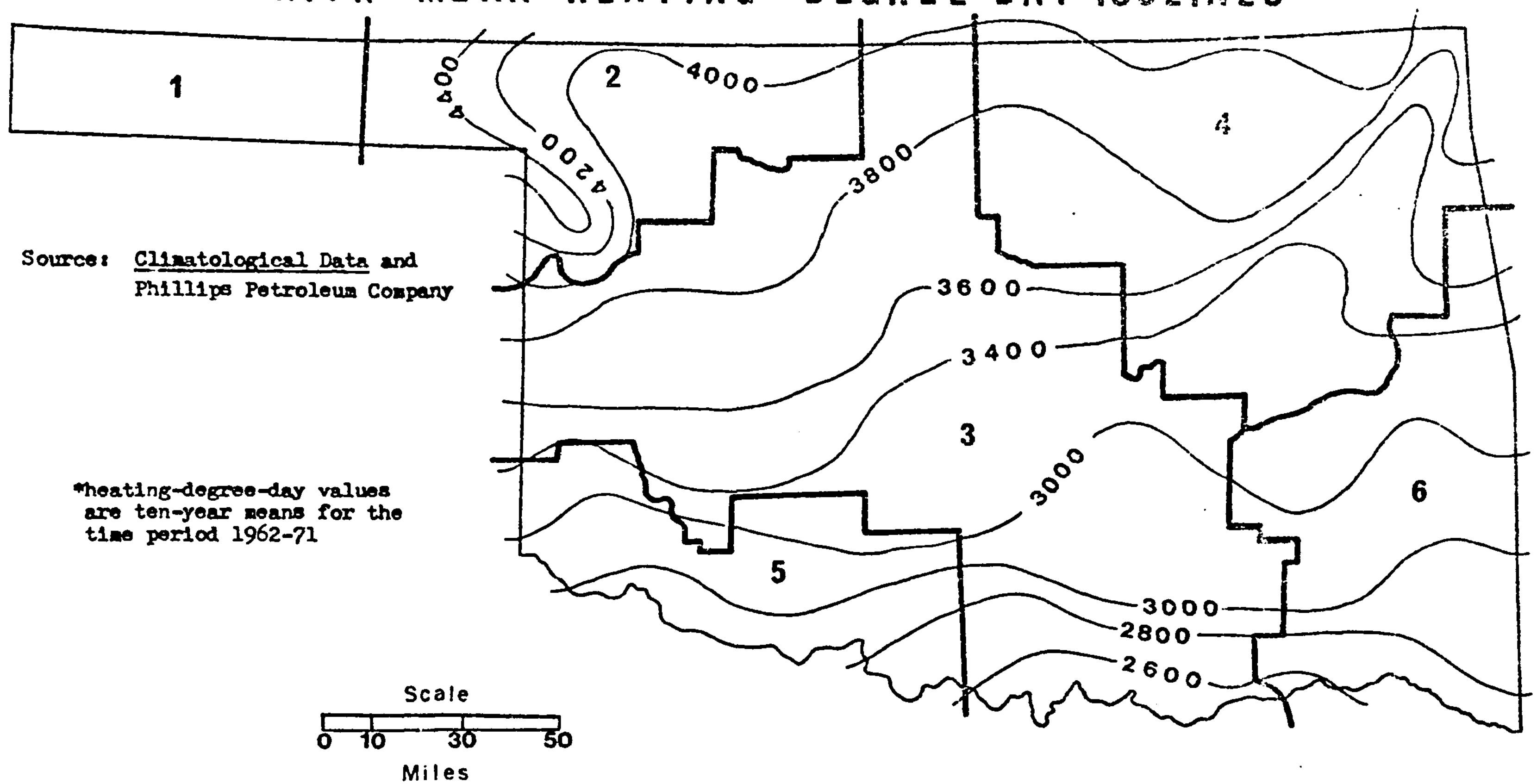
> For the same geographical area and the same weather stations, the derived regions will have less mean within-groups variation than the gas-marketing regions currently being used.

Rationale for the Research Hypothesis

The rationale for the research hypothesis is based upon observations of marketing regions that are currently being used. As an example, the LP gas-marketing regions of Phillips Petroleum Company (Bartlesville, Oklahoma) for the state of Oklahoma are shown by Map 1. In some cases, Phillips' regions (Region 1, for example) are quite small because they are part of larger regions that extend across state boundaries. Region 3, however, is complete. Also shown on Map 1 are heating degree day isolines that range in values from 2,600 to 4,500. It is apparent that the regions have not been delimited on the basis of weather characteristics and are only partially responsive to physical elements that would cause climatic differences. Of particular note is Region 3, which has a heating degree range from less than 2,600 to more than 4,000. Map 1 provides evidence that the regions being used to market LP gas are not as homogeneous as possible in terms of weather occurrence.

The assumption that more effective regions could be derived stems from evidence in the literature of economic geography. Such evidence points to the existence of some very powerful analytical techniques with which classification and regionalization problems can be handled. One form of classification, which is especially pertinent to geographic research, involves the identification or delimitation of regions which

OKLAHOMA MARKETING REGIONS OF PHILLIPS PETROLEUM COMPANY WITH MEAN HEATING-DEGREE-DAY ISOLINES*



Map 1

are relatively homogeneous with respect to a single set of characteristics. This classification is based on distance measures in taxonomic (graph) space, in which the smaller the distance value, the more similar are the observations in question.

It is possible, therefore, in a series of discrete steps to group together observations which are close to one another in taxonomic space. Ward (1963) has operationalized this procedure, and his algorithm has been applied to a number of regionalization problems. Berry, in particular, has utilized this method and has concluded that it provides an "exact solution" (1961, p. 275). Similarly, Bunge (1966, p. 377) has stated:

> Thus, not only are the logical problems of the precisely analogous character of classification and regionalization solved, but the major practical problem of regionalism with computing machines are also conquered. At long last we can replicate our regions.

Other instances of a grouping algorithm being used to delimit uniform regions include Weaver's (1954) crop-combination agricultural regions in the mid-west, Spence's (1968) multifactor uniform regions of British counties based on employment data, and Steiner's (1965) multivariate climatic regions of the United States.

These studies indicate that regions which are acceptably uniform with respect to a set of weather characteristics could be derived. This fact, combined with evidence that LP gas-marketing regions currently in use are not homogeneous weather regions, led to the assumption that the derived regions would have less variation in numbers of heating degree days from year to year, and would be, therefore, more reliable as predictive regions.

Adjusting the Homogeneous Weather Regions to Demand Patterns

Another objective of this study is the adjustment of the derived homogeneous weather regions into effective marketing regions on the basis of demand intensity. The underlying rationale for this step is based on the idea that homogeneous weather regions and effective marketing regions are not necessarily equivalent. There may be geographic similarities, but the structural properties will differ in that weather regions do not take into account the spatial variation of customer concentrations, which generate heating-fuel demands independent of weather occurrence.

If LP gas customers were evenly distributed, then the regional uniformity of weather occurrence would be the only consideration. The central place network of settlement dictates, however, that some areas will have large numbers of customers, while others will have very few. It is a basic principle of marketing that, in many cases, those areas having large customer concentrations should receive special attention because it is in these areas that the highest level of profits is obtainable. In areas having high demand intensities, the conditions for making a correct estimate of LP gas sales should be maximized. The best method of maximizing the conditions for a correct forecast is to have a very small predictive region because of the relatively higher probability of uniform weather occurring in a small area. If all marketing regions were very small, their utility would be decreased, and administrative operating costs for a large number of regions would reach insupportable levels. Obviously, it is not practical to have a small marketing region if the magnitude of heating-fuel demand in that region is small. Such a region would not contribute a pro-rata share to the overhead costs of administering the marketing system. Ideally, a set of regions could be established in which the relationship between the number of customers and intraregional heating-degree-day variation would be inversely correlated. That is, as the number of customers increases, the variation of regional heating-degree-day occurrence should decrease. The ideal situation is approached as the uniform weather regions are adjusted to conform to the spatial distribution of gas users.

Scope of the Research Problem

In scope, the research problem undertaken in this study is a type of economic geographic research that has been catagorized as "marketing geography." The need for such a geography was initially spelled out by a leading practitioner, William Applebaum, who has contended that the objectives of marketing geography should be the delimitation and measurement of markets and the study of the channels of distribution through which goods move from producers to consumers (1954, p. 245). Later, store site evaluation and store location research was recognized as a legitimate concern of marketing geography (Cohen and Applebaum, 1961, p. 95). Above all, it has been recognized that the best place to develop marketing geography is in business, using the laboratory of actual business operations (Berry, 1967, p. 125). In general, the subject matter of this study fits into the traditional framework of marketing geography.

CHAPTER THREE

PROBLEM PARAMETERS AND THE STATISTICAL DESIGN

Introduction

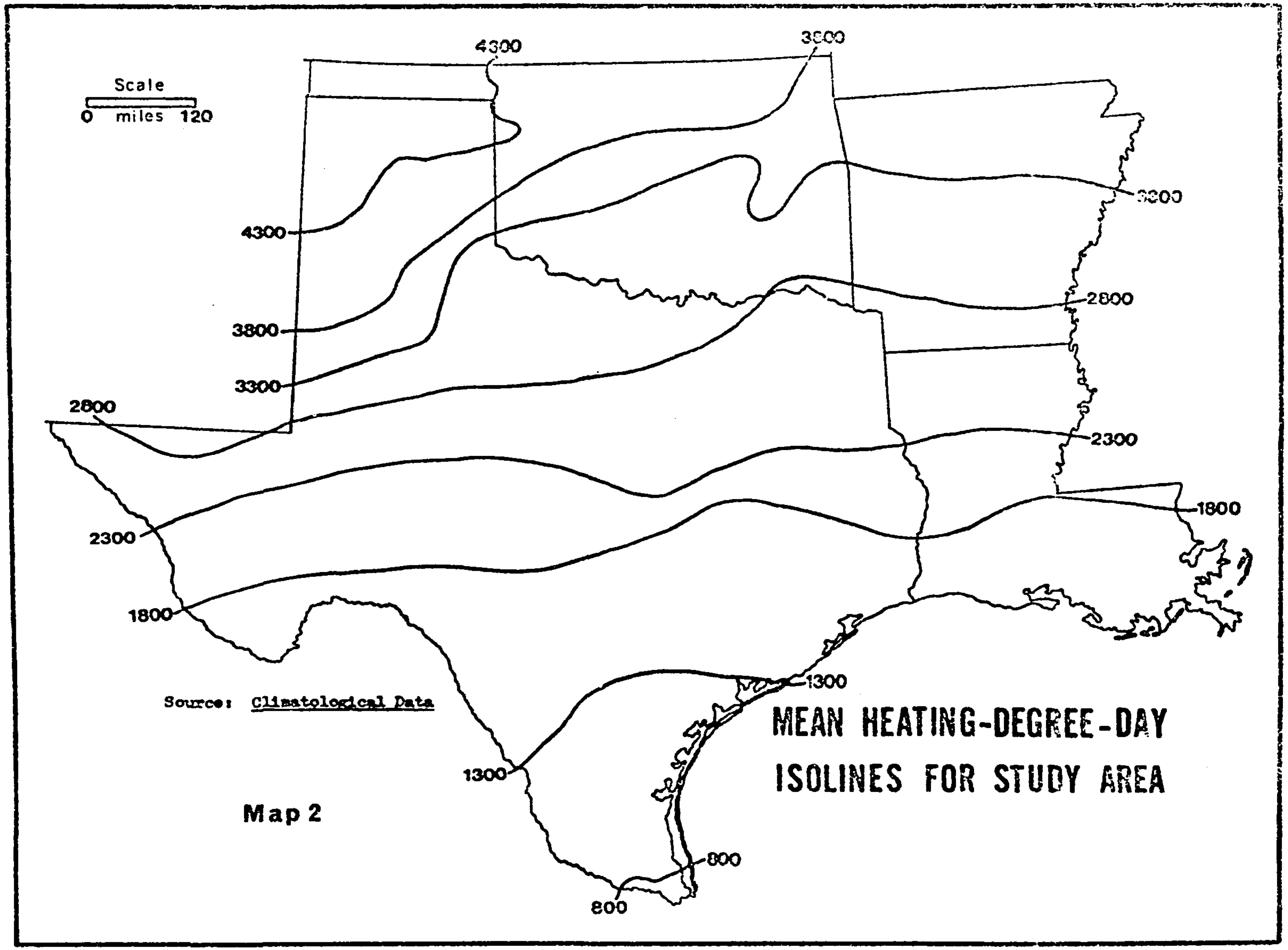
It is the objective of this chapter to describe the problem parameters and the statistical methodology utilized to formulate a solution to the problem of deriving better predictive marketing regions for LP gas. Initially, the focus is upon the study area and its areal extent. A brief rationale for selecting the study area is discussed. Next, the weather variables and the demand constraint that serve as the attributes used to identify the regions within the study area are examined. In addition, careful attention is afforded the method of selecting and collecting the sample data. Of particular concern is the method of regionalizing and the method of testing the derived regions. These two statistical procedures are discussed in detail in the latter sections of this chapter.

Study Area

The study area includes the four south-central states of Arkansas, Louisiana, Oklahoma, and Texas. These states represent an especially troublesome area for marketing gas-heating fuels. Unlike more nothern areas of the United States that can be expected to have cold winters with insignificant variation, average temperatures in southern states may fluctuate widely from day to day and season to season. The spatial variation of heating-degree-day in the study area is shown by Map 2. Isolines drawn at 500 heating-degree-day intervals represent the mean heating-degree-day occurrence for the time period 1962 through 1971. For the study area, a variation of approximately 3,500 heating degree days exists, ranging from 800 in the Lower Rio Grande Valley to more than 4,300 in the High Plains of the Texas Panhandle. Thus, the four states represent an area where an effective marketing strategy for gas-heating fuels is particularly important.

Heating-Degree-Day Variables

Two factors are prerequisite to the delimitation of uniform weather regions needed to predict LP gas consumption. These two factors are related to the geographical pattern of heating-degree-day occurrence--the basic unit for determining heating-fuel consumption. Ideally, the number of heating-degree days should not vary significantly within a marketing region. All weather stations within such a region should closely approximate all other stations in that region in the average (mean) occurrence of heating degree days. Any weather



MEAN HEATING-DEGREE-DAY ISOLINES FOR STUDY AREA

Map 2

Source: Climatological Data

region not possessing such homogeneity will periodically experience an over or undersupply of heating fuels, independent of the accuracy of the heating-degree-day forecast for the representative station. Thus, mean heating degree days and the standard deviation of heating degree days for each selected weather station are used in this study to delineate uniform weather regions. By using the mean and the standard deviation to establish the data matrix for the grouping procedure, both temporal and spatial variations in heating-degree-day occurrence are considered. Mean heating degree days reflect the spatial variations through time, and the standard deviation of heating degree day reflect temporal variations over space. As such, the two types of weather variation that would influence a heating-fuel forecast are both considered.

Population as a Surrogate for Demand

Although it was expected that the delineation of homogeneous weather regions would increase the reliability of estimating gas-heating fuel demands, homogeneous weather regions are not necessarily the most effective marketing regions. Gas users are unequally distributed. For example, two or three counties containing a small number of LP gas customers might be identified as a uniform weather region. Although a region of this type might be highly homogeneous in terms of heating-degree-day occurrence, it could not be considered an effective marketing region as the administrative costs of maintaining a regional office would exceed the profits possible from a limited number of customers. The converse situation of a large city containing a large number of gas users being included in a geographically large region could also occur. As the size of the region increases, it becomes more difficult to achieve regional weather homogeneity and consequently, it becomes more difficult to correctly forecast LP gas consumption. It is, therefore, important that forecasting efficiency be maximized at places having large population concentrations. In the interest of good marketing strategy, it is desirable to prevent the delineation of an overfit or underfit region. This was accomplished by adjusting regional size in relation to the spatial distribution of gas users.

Delineation of marketing regions with desirable size characteristics was achieved by placing a constraint on the grouping procedures. The specification of the constraint involves a decision as to the maximum number of LP gas users to be allowed per region. In making this decision, a knowledge of the threshold population required to provide an acceptable profit level is necessary.* Petroleum companies,

---

*The term "threshold population" is used here in a slightly different sense than in some economic geographic studies. Instead of a population surrounding a single central place, it is used here as the population level needed to provide acceptable profits in a marketing region.

understandably, are reluctant to provide information of this nature. In order to accomplish the goals of this study, an estimation of the threshold population was utilized.

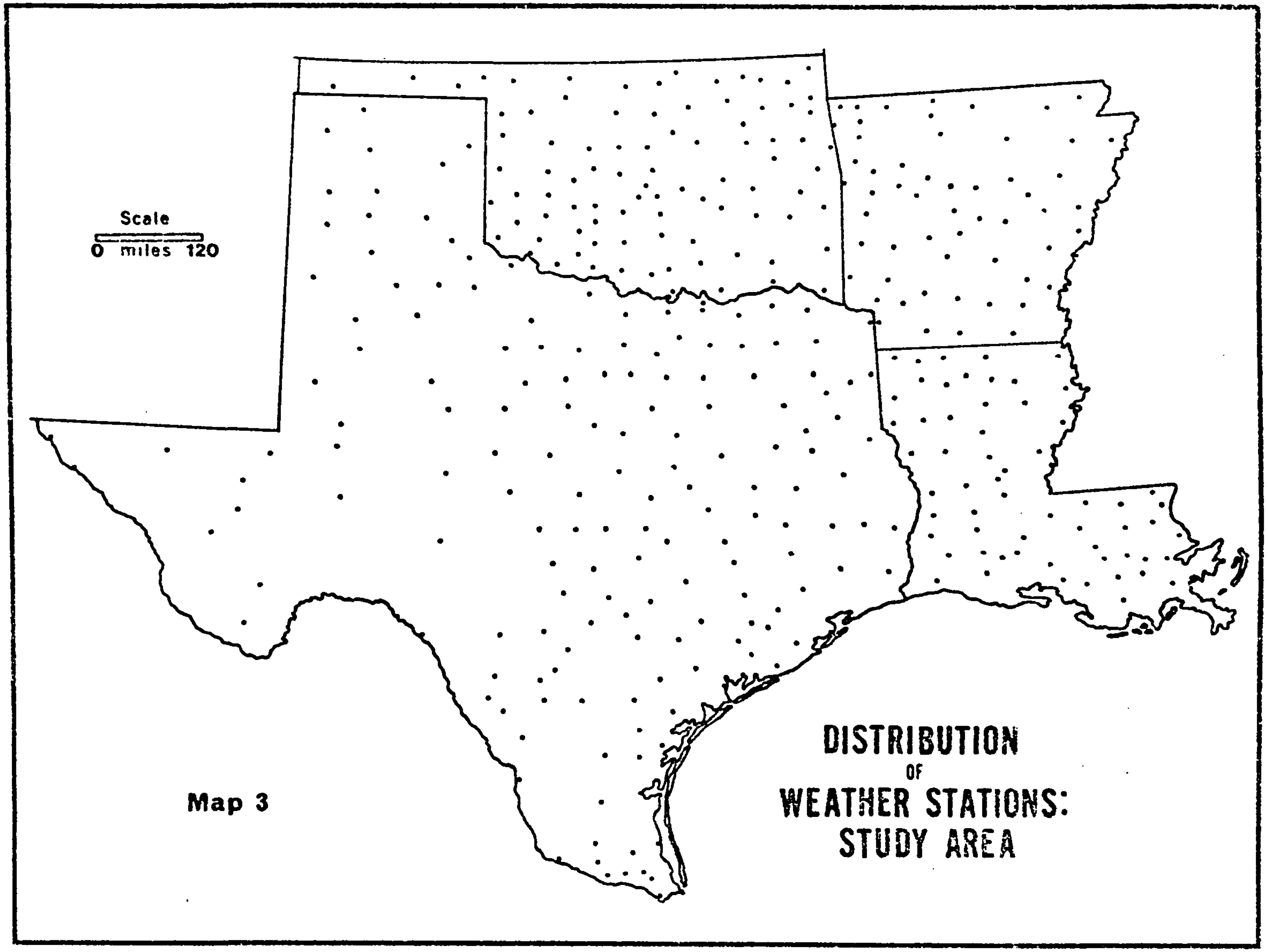
Population by county from the 1970 Census of Population was utilized as a surrogate for demand levels. Although census data of this type do not necessarily represent a specific petroleum company's gas customers, it is apparently the best available in the absence of customer data from company files. The same situation applied in attempting to determine the range of population that would provide a desirable profit level in any homogeneous weather region. Petroleum companies do not allow the publication of profit data. However, officials of one major petroleum company indicated that they would consider any population that fell between the populations of the Dallas, Texas SMSA and the Fort Worth, Texas SMSA as providing a good profit level. This information was used to justify the establishment of a population range that is as applicable to the real situation as possible. The population of the Dallas SMSA is 1,555,950; the population of the Fort Worth SMSA is 762,086. For purposes of this study, 750,000 people in any region was defined as the threshold population while 1,500,000 people was defined as the maximum allowable population in any region. In other words, a constraint was placed on the grouping procedure so that no region would have more than 1,500,000 inhabitants. This procedure insured that all regions identified would have reasonable similar populations, although the areal extent may widely differ. Thus, marketing regions containing densely populated areas (such as the Dallas, Texas area) are geographically small, while sparsely populated areas (such as southwest Texas) are part of a large marketing region. No matter what the physical size, however, a large proportion of the regions should contain about the same population, that is, an amount approaching the arbitrary maximum population limit of 1,500,000 so that regional weather homogeneity is in a best-fit relationship with the number of customers needed to provide acceptable profit levels. Regions derived in this manner should be the most effective possible for forecasting LP gas demand. Again it is emphasized that the population constraint used in this study had to be selected somewhat arbitrarily because population was used as a surrogate for demand. The constraint can easily be modified, however, to fit particular marketing philosophies.

Sample Size and Method of Selection

The ten-year period from 1962 to 1971 was selected as the test period for weather conditions. Two factors prompted the decision to use this time period: (1) since 1962, annual heating-degree-day data for all reporting weather stations have been published for each state by the U.S. Department of Commerce, Weather Bureau, in the July issue of Climatological Data, and (2) recent studies indicate that "normals" based upon periods somewhat shorter than thirty years possess a higher degree of accuracy because they are more sensitive to short-term fluctuations in climate (Court, 1970). Although practically all climatic averages in the United States are based on a thirty-year time period, Court in a study of selected cities across the U.S. found that averages over a time period of between seven and thirteen years were better for predicting weather occurrence. Thirty-year averages are less accurate in prediction because they reflect more than one short-term climatic cycle. In light of Court's findings, ten-year averages of heating degree days have been used in this study.

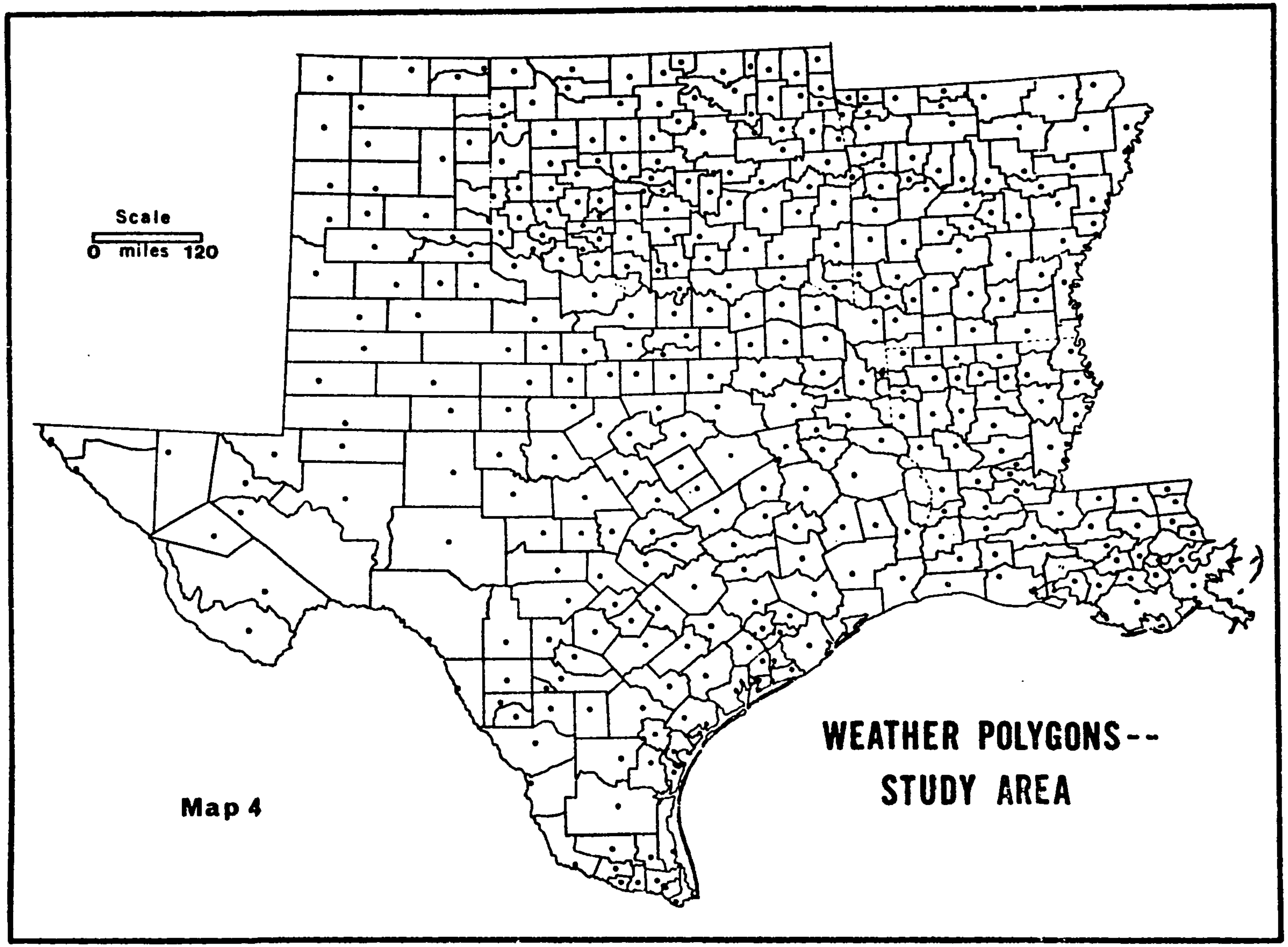
The selection of the weather stations to be utilized was determined by whether or not the station had complete heating-degree-day data for the 1962-71 time period. All stations within the study area with complete heating-degree-day records were used. Any station with missing data was rejected. The total number of weather stations in the study area having complete records was 372. Distributed among the four states of the study area, Arkansas had 56, Louisiana, 59, Oklahoma, 89, and Texas, 168. These weather stations are shown on Map 3 and are listed by name in Appendix A.

The next step was the allocation of the total space of the study area among the 372 sample stations. Although it must be recognized that heating-degree-data from any weather station is only an exact measure at that particular


Scale
0 miles 120
Map 3
DISTRIBUTION OF WEATHER STATIONS: STUDY AREA

place, it is still an indicator of weather occurrence for a limited area surrounding the station. Hence, a method for transforming data locations from points to areas must be used. Thiessen (1911) allocated precipitation to the most efficient distance areas (Thiessen Polygons) around climate stations (points) on the assumption that the area closest to a station is best characterized by the data for that station. The same assumption is utilized in this study for heating degree days; and since the sample weather stations are not uniformly distributed (Map 3), the method used by Thiessen to allocate geographic space can be applied.

However, a second variable must be considered. As described previously, the distribution of heating-fuel users must be taken into account if the derived regions are to be effective as predictive marketing regions. Not only geographic space but also population (used as a surrogate for demand) must be allocated to each weather station. The smallest units for which population statistics are reported are Minor Civil Divisions for Arkansas and Louisiana; and County Census Divisions for Oklahoma and Texas. These divisions have been utilized to assign population to each sample weather station. The method was to assign each population division to the closest weather station so that each weather station would have a central location in the area it represents. The results of this procedure are shown by Map 4. It can be seen that a central location is not always obtained because of the geo-


Scale
0 miles 120
WEATHER POLYGONS--
STUDY AREA
Map 4

metric shape of the county divisions.* These Census Divisions cannot be divided because population totals are reported for county divisions, and the entire area must be assigned to one weather station or another. The sizes of the resultant polygons vary considerably. This is largely a reflection of the central place network of settlement. The population represented by a large weather polygon may be quite small while in most cases small weather polygons represent a relatively large number of people. It must be assumed that the variation in heating degree days within each polygon is minimal.

The 372 weather polygons thus derived serve as the observations to be grouped into a set of relatively homogeneous weather regions. Annual heating degree days for the years 1962-71 for each station were tabulated and the ten-year mean standard deviation computed. These two sets of data serve as the variables used to delimit the uniform regions.

The Nature of Specific Regional Systems

The delimitation of regions is basically a problem in

---

*The state of Oklahoma, for example, established new county census divisions for the Census of Population in 1970. It appears that these new divisions have been arbitrarily delimited mainly on the basis of physical boundaries such as rivers. As such, the new divisions have odd shapes that only partially reflect the central place network of settlement.

classification. The individual observations (weather polygons) must be combined on the basis of similarities (heating-degree-day variables) to form a smaller number of groups. The weather occurrence in each weather polygon may be unique; yet, each weather polygon has characteristics that resemble the characteristics of other places. The objective is to avoid too much uniqueness (if all weather polygons were an individual region) and too much generality (if the whole study area were one region) by assembling information into manageable areal aggregates termed regional systems. The focus of this research is on specific regional systems, which are defined not only by combinations of intrinsic attributes, but also by location (Grigg, 1965). By convention, all parts of a single homogeneous region must be spatially contiguous.

Numerical Taxonomy: A Statistical Design

The two variables, mean heating degree days and the standard deviation of heating degree days, were used to construct a symmetrical matrix which located each observation in graph or taxonomic space. From such a matrix, a measure of observation similarity based on distance (D^2) measures was computed. An important assumption, however, was that these two sets of variables were linearly independent. If they were correlated, the variables contain redundant information and to some degree are measuring the same thing. Hence, the resultant classification would be biased in that some information is given greater weight than other information. Also, if the variables are in any way correlated then the space is nonEuclidean; and in order to calculate the taxonomic distance between the weather polygons the exact structure of an m-dimensional space would have to be defined. However, if the two sets of heating-degree-day variables are not correlated, the taxonomic space is Euclidean and the mathematical distance between observations can be calculated by the Pythagorean Theorum.* In order to determine whether colinearity existed, correlation analysis was performed.

In simple, linear correlation analysis, no distinction between dependent and independent variables is necessary. Instead, the emphasis is on the degree to which the two sets of values covary around their respective means and on the direction of covariation if it exists. The objective is to compute a correlation coefficient (r) which measures the degree of covariation. In general terms, coefficients of between +0.5 and +1 and between -0.5 and -1 are fairly significant, while if values lie between +0.5 and -0.5, then little significant

---

*The Pythagorean Theorum states that in a right triangle the square of the hypotenuse is equal to the sum of the squares of the other two sides.

correlation is to be expected (Gregory, 1963, p. 190). A value of zero indicates that the two sets of data are completely independent of each other and that no correlation exists at all.

A correlation coefficient of +.094 was obtained as the measure of relationship between 372 heating-degree-day means and 372 heating-degree-day standard deviations. This value provides assurance that the two sets of variables do not covary and are, therefore, linearly independent.

From the matrix of distance statistics, the smallest D^2 was identified, and the two corresponding observations grouped. Since spatial contiguity was required, however, the D^2 statistics in the distance matrix were coded positive if observations (weather polygons) were geographically contiguous, negative if they were noncontiguous. The subsequent grouping algorithm was, thus, constrained so that only positive distances were considered in the progressive building up of a spatially contiguous region. In this manner it was assured that no homogeneous weather region would be identified unless all parts were spatially contiguous.

If weather area X was most like weather area Y (and Y was most like X), they were grouped (if spatial contiguity existed). Several methods could have been used to assign the residual individuals in the matrix to groups, once contiguous, reciprocal pairs were identified (Johnston, 1968). The method used in this study was centroid replacement if which the taxonomic distance used was that between the residual observation being grouped and the centroid of the members of the nearest similar group. The centroid is the mean position of the members of the group (sometimes called the center of gravity). Likewise, the distance between two groups was calculated as the distance between two centroids.

At each step of the grouping, it is possible to compute a value for an objective function. For this study, the selection of the groups that were joined at each step follows the grouping algorithm of Ward (1963), in which the criterion is that the increment to the pooled within-groups sum of squares be minimized. The process could be continued through a series of steps until eventually only one group, or region, containing all the observations exists.

Loss of Detail in Regionalizing

At each step in the grouping process, the within-groups ΣD^2 is an objective measure of the amount of detail that is lost. If each individual weather polygon were a region, the sum of the distances squared (ΣD^2) would be a measure of the amount of detail present. As the weather polygons are grouped, the ΣD^2 between-groups declines as the number of groups declines and as detail is lost. Prior to the first grouping, the within-groups ΣD^2 is zero because there are no groups. As grouping proceeds, the ΣD^2 within-groups increases as the number of groups declines and as generality increases. Before grouping begins, 100 percent of detail is retained; after all weather polygons have been grouped into one class, 100 percent of the detail will have been lost.

The nature of this particular problem is such that a large degree of detail is needed in forming predictive regions. It seems reasonable that the more detail that can be retained, the greater will be the success of the derived regions. However, it is economically and administratively infeasible to consider each weather polygon a region. Grouping must, therefore, occur but with the loss of as little of the original detail as possible.

The Problem of Determining the Number of Regions to be Delineated

No analytic solution exists for the problem of deciding how many regions are to be identified (King, 1969, p. 199). This is a key problem because there is no way to know precisely how many homogeneous weather regions should be created. As Berry observes, however, "It is possible to select that level deemed most desirable for a particular problem, and know that an optimal minimum-variance stratification has been achieved. . . ." (Berry. In King, 1969, p. 199).

The most desirable level for this particular problem is the level at which the population of each derived region approaches 1,500,000 people for reasons explained earlier.

Thus, the best set of predictive marketing regions that could be derived would be a set with the population of each region approaching 1,500,000, and with the weather occurrence (heating degree days) in each region as internally uniform as possible. A set of regions with these characteristics was derived by placing a population constraint on the grouping algorithm so that at each step of the grouping, the sume of within-groups population was checked. A potential combination of polygons was not allowed if it caused the population of a region to exceed 1,500,000. Such a constraint was used to prevent the grouping process from continuing iteratively through a series of steps until only one region containing all the observations existed. At some step, no more grouping was possible because any combination that could have been made would have resulted in a region exceeding the population constraint. At this step, the best possible set of predictive marketing regions has been selected. The system of regions identified at this key step was used in this study as the optimum set. Thus, the decision as to the number of regions to be delimited was determined by that grouping step at which no more grouping was possible.

Method of Comparing the Derived Regions With an Existing Set

After the regionalization has been completed, the concern becomes whether or not the derived regions have greater reliability in predicting LP-gas demand than a set of regions currently being used. Logically, if the derived regions have less intraregional variability of heating-degree-day occurrence, they are better predictive regions. An indicator of whether the derived regions are more homogeneous than an existing set can be determined by computing the within-groups sum of squares for both sets of regions and comparing the results to see which is less. If the same weather stations and the same geographical area were used for both sets, the set with the smallest total sum of squares would be the most uniform internally. Regional uniformity is the desirable characteristic of good predictive regions.

In this case, however, simply computing the sum of squares is not the most realistic measure of an effective set of regions that can be obtained. Although calculating the within-groups sum of squares does provide a measurement of error (the squared difference of each observation value from the regional mean), the assumption is that all error (variance) is equally important. In utilizing a region to predict LP gas demand, all prediction error does not have equal significance because of the uneven distribution of customers. Consider two opposite possibilities: first, the error (variance) from the regional mean of a weather polygon with a small population and, second, the error of a weather polygon with a large population. The latter of these two possibilities is much more significant because the regional mean inaccurately represents a larger number of customers. Therefore, to obtain an accurate measure of the effectiveness of a set of regions in predicting demand, the variance of each weather polygon should be weighted by the population of that polygon. This procedure has been used in this study to evaluate the effectiveness of the derived regions. The resulting coefficient is referred to as the Predictive Index. Each regional sum of squares was weighted by population, and the value obtained used as an index of heating-degree-day error per unit of population. The formula for calculating the Predictive Index is

$$\sum_{1 = j}^{K} \sum_{1 = i}^{N_j} P_{ij}(X_{ij} - \overline{X}_j)^2$$

where X_{ij} is the mean heating degree days associated with each weather polygon in a region; $\overline{X}_j$ is the regional mean of heating degree days; and P_{ij} is the population of each weather polygon.

A set of LP gas-marketing regions currently being used by a major petroleum company were obtained in order to make a comparison with the derived set. The method used to compute the weighted mean sum of squares was the same as for the derived regions with one exception. One weather station in each region is used by the petroleum company to make a regional estimate of LP gas demand. Since the regional error (variance) of prediction is the difference between the mean value for one station and all other stations in that region, the mean value of heating degree days for the representative station in each region was substituted for the regional mean in computing the weighted regional sum of squares. The formula is

$$\sum_{1=j}^{K} \sum_{1=i}^{N_j} P_{ij}(X_{ij} - U_{rs})^2$$

where U is the mean heating degree days for the representative station of each region.

The data for the comparison were obtained by using all weather stations in the study area with complete heating-degree-day records, a total of 372. Mean heating degree days for the ten-year period, 1962-71, were computed and served as the variables for the analysis.

The 372 weather stations served as the test component for two sets of gas-marketing regions. Set one was the regions currently being used by a major petroleum company, and set two was the regions derived in this study. The number and the areal extent of these two sets of regions differed, but each set was superimposed upon the 372 weather stations to determine which stations fell into which regions. The strength of this type of comparison lies in the fact that the same geographical area and the same weather stations are being used for both sets of regions. The results determined which set of regions had the smallest amount of weighted within-groups mean sum of squares, and this information served as a framework for evaluating the relative merits of the derived regions.

CHAPTER FOUR

THE DEVELOPMENT OF A SET OF PREDICTIVE MARKETING REGIONS FOR LP GAS

The Method of Regionalizing: An Overview

To recapitulate, the two sets of variables were arranged in a symmetrical matrix which located each weather polygon in graph (taxonomic) space.* Since it has been shown that the means and standard deviations do not covary, the graph space is two-dimensional, and the Pythagorean Theorum was used to compute the taxonomic distance (D^2) between any pair of observations. From the matrix of distance statistics, the criterion for grouping was that the increment to the pooled within-group sum of squared distances be minimized. Once a pair of weather polygons were grouped, their centroid coordinates were replaced into the distance matrix and used to determine the taxonomic distance to any residual observation.

*The grouping program used for computation was written by Robert R. Williams of the University of Oklahoma. A complete description is contained in Appendix B.

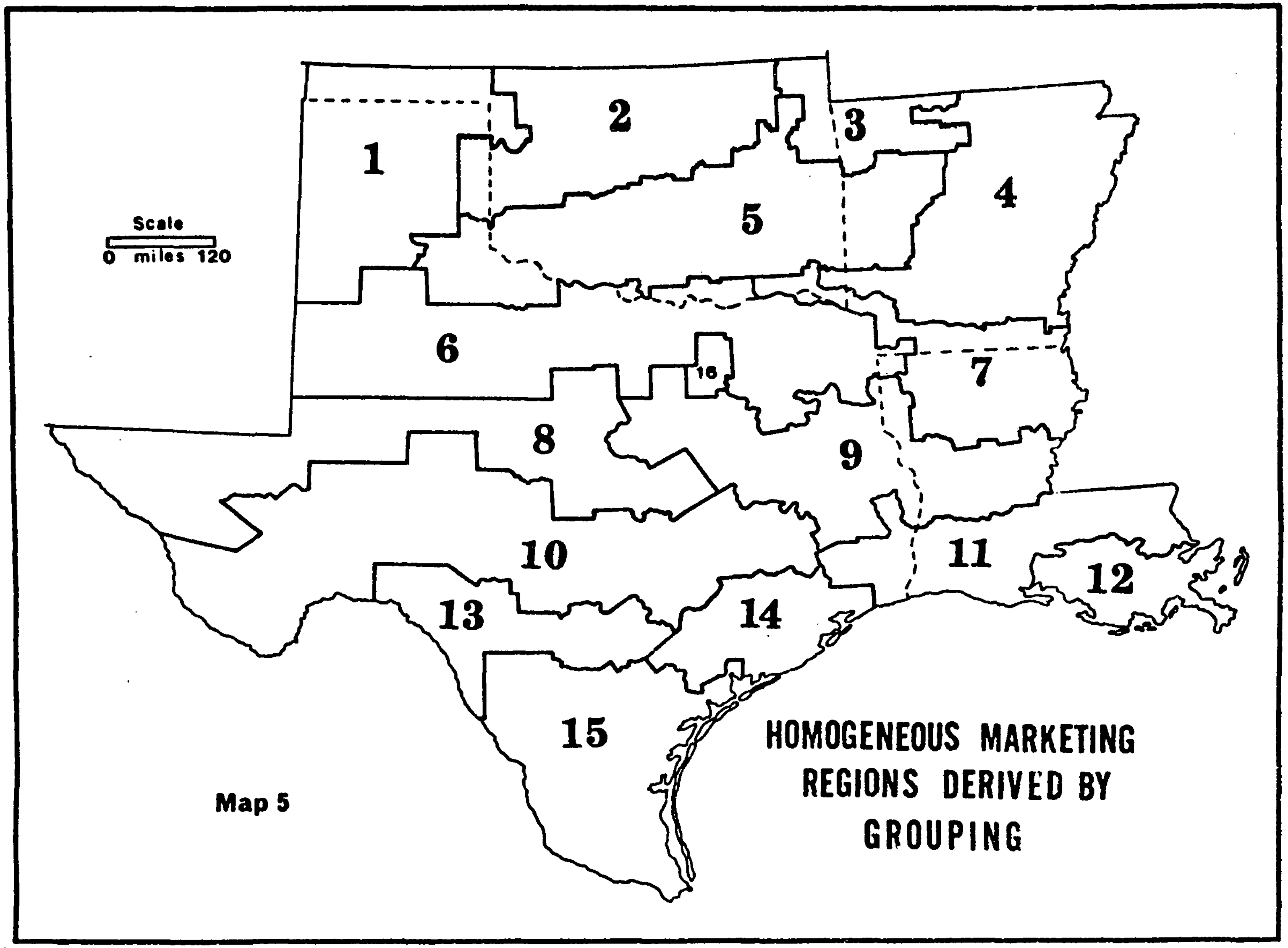
As described in Chapter Three, the Census Population Count of 1970 was used as a surrogate for LP gas customers in the absence of demand data from a petroleum company. An arbitrary value of 750,000 to 1,500,000 was assigned as an optimal number of customers in any region. Thus, a constraint that would not allow the population of any region to exceed a population of 1,500,000 was placed on the grouping process. At each step in the grouping, the pooled within-group population was checked. If the population of a potential addition to a group caused the arbitrary limit to be exceeded, the combination was not allowed. This was an important procedure because it determined the number of uniform weather regions to be delineated. If there had been no population constraint, grouping would have progressed through 371 steps whereupon all weather polygons would have been in a single group or region. However, the population constraint limited the number of grouping steps because the point was reached at which no further combinations could be made without exceeding the population limit. The number of regions identified at the step where grouping stops was used in this study as the best set of predictive marketing regions for LP gas that could be derived.

The Derived Characteristics of Homogeneous Weather Regions

The grouping procedure went through 356 iterative steps before no further grouping was possible. At step 356, the original 372 observations had been collapsed into sixteen regions (Map 5). The total population and the percentage of observations for each region are summarized in Table 1.

As was expected, the sizes and shapes of the derived regions vary widely. For example, consider the size differences between Region 16 and Region 5 (Map 5). Region 16 is very small, being composed of only two weather polygons. On the other hand, Region 5 contains fifty-six of the original weather polygons and extends from the Texas Panhandle, across the entire state of Oklahoma and into approximately one-third of Arkansas. On first impression this tremendous size difference appears to be an undesirable characteristic of the derived set of regions. If the total population in each of the two regions is compared, however, Region 16 contains 1,411,039 people and Region 5, a population of 1,220,390. Both of these figures are within the population range stipulated as optimal for profit levels in marketing LP gas. Thus, the different sizes of the derived regions were necessary in order to develop a set of effective demand forecasting regions. The probability of uniform weather occurring within a region is maximized in relation to the number of customers needed to provide a sufficient demand level.

Of the sixteen regions identified, thirteen have populations greater than 750,000 (Table 1). Three of the derived regions, however, have population totals lower than



Map 5

HOMOGENEOUS MARKETING REGIONS DERIVED BY GROUPING

TABLE 1

THE POPULATION AND PERCENTAGE OF OBSERVATIONS CONTAINED IN EACH OF THE DERIVED REGIONS

Region Identification Number (Map 5)	Number of Weather Stations Contained	Percentage of Original Observations Contained	Total Population (1970)[a]
1	17	4.5	367,283
2	33	8.8	1,376,933
3	15	4.0	277,626
4	32	8.6	1,367,224
5	56	15.0	1,220,399
6	33	8.8	1,379,677
7	22	5.9	655,940
8	21	5.6	1,228,939
9	31	8.3	1,415,757
10	26	7.0	904,552
11	20	5.3	1,479,155
12	12	3.2	1,432,060
13	8	2.1	986,655
14	15	4.0	1,293,921
15	29	7.7	934,151
16	2	0.5	1,411,039
TOTALS	372	99.8	17,731,311

[a]Source: U.S. Census of Population, 1970 (PC-2).

the desirable level. Region 1, with a population of 367,283; Region 3, with 277,526; and Region 7, with a total of 655,940 and stand out as not approaching the desired goal in terms of the number of customers needed to provide a good profit level. This outcome must be expected, in some cases, when the geographical space of a limited area is allocated on the basis of a continuous variable such as weather occurrence. In every case, it can be observed that the regions failing to obtain the desired population abut upon the boundaries of the study area. It seems reasonable that these three regions represent types of weather occurrence that are marginal to the study area. A solution to the problem might be obtained by extending the regions across state boundaries to include more area. Given the defined limits of the study area, however, the regions shown by Map 5 are believed to represent a marked improvement over those drawn by arbitrary methods. This contention can only be substantiated by comparing the derived set of regions against an existing set that is being used currently as marketing regions for LP gas.

Another factor which deserves consideration is the shape of the derived regions. In most cases the east-west length far exceeds the north-south width.* From Map 5,

---

*This outcome must be expected because heat energy from the sun is received at the earth's surface in rays that are vertical in the northern hemisphere only as far as $23\frac{1}{2}^{o}$ North

Regions 6, 8, and 10 are especially good examples of this propensity toward longitudinal dominance. The irregular shapes and the lack of compactness of the derived regions may cause questions to arise as to their effectiveness as regions.

In many economic geographic studies the objective is to derive distribution regions, which characteristically are compact in areal shape. For example, a hexagonal marketing region with all sides an equal distance from a central place has been theorized as the optimal shape for a marketing region within a system of marketing regions (Berry, 1967). The hexagonal marketing region minimizes distances in the regional distribution of products from a number of central places. Thus, transportation time and cost are theoretically equalized in every direction from the central place distributing point.

Latitude. The sun's vertical rays are much more intense than those received at an angle and cause greater warming of the earth's surface. Poleward from $23\frac{1}{2}^{o}$ North, the sun's rays are received at a larger and larger angle, resulting in a general decrease of average temperatures with increasing latitude. As temperature regions are delineated, it is natural to expect strong similarities in an east-west direction and increasing differences in a north-south direction, as is readily apparent in the regions derived in this study. The lack of an even, consistent pattern is due to unequal elevation and the different heating/cooling properties of land and water. The frequency of cold air intrusion from the north also declines gradually towards the south.

The set of regions derived in this study do not function as distribution regions. Delivery of LP gas to local franchised dealers is basically accomplished by long distance pipelines. Local dealers in the same region may be supplied from different points, depending upon their locations with respect to a pipeline terminal. The amount of fuel that will be delivered to each retail dealer is based upon a demand forecast that is made before the heating season begins. The objective is to correctly forecast the amount of LP gas to be needed in each region. Hence, the primary purpose of the regions derived in this study is to serve as predictive marketing regions. Their associated function is that of administrative regions to facilitate communication between the local dealers, the regional manager, and the home office. Thus, the elongated shape of the derived regions may not necessarily be regarded as an undesirable characteristic. Whether or not the derived regions can be utilized to predict heating fuel demand more accurately than an existing set of LP gas marketing regions is the most important criterion for their success.

The Loss of Detail in Regionalizing

One of the goals in delineating predictive marketing regions is the loss of as little of the original detail as possible. In the case of this study, grouping was terminated at step 356. At that step, 57.1 percent of the original detail had been lost. As a result, the weather homogeneity of the derived set of regions is based on 42.9 percent of the original detail.

A methodological question arises as to how the total sum of squared distances was known when the grouping program only ran through 356 steps. In order to obtain this value, the population constraint was removed from the grouping program, which allowed the program to run 371 steps at which time all weather polygons were in one group. The total sum of squared distances obtained by this procedure was used as a constant when computation with a population constraint was made. At each step the ratio of the within groups ΣD^2 and the total sum of squares was calculated. In this manner, the percent loss of detail at each grouping step was obtained with a population constraint in the grouping program.

It is interesting to observe the loss of detail in percentage terms that is sacrificed by having a population constraint on the grouping process. From Table 2, it can be seen that 47.8 percent of the detail was lost at grouping step 356 when there was no population constraint. Thus, 9.3 percent of detail was sacrificed by having a population constraint on the regionalizing process. The within-groups D^2 for both the population constraint and the nonconstraint methods were almost identical all through the grouping process until step 351. At that point, the population constraint

TABLE 2

A COMPARISON OF THE PERCENTAGE OF DETAIL LOST IN GROUPING WITH A POPULATION CONSTRAINT AND A NONCONSTRAINT

Grouping Step	Percentage Lost with Population Constraint	Percentage Lost with No Population Constraint
356	57.1	47.8
355	52.8	46.5
354	48.4	45.4
353	46.9	44.6
352	44.5	42.8
351	42.7	41.9
350	41.7	41.0

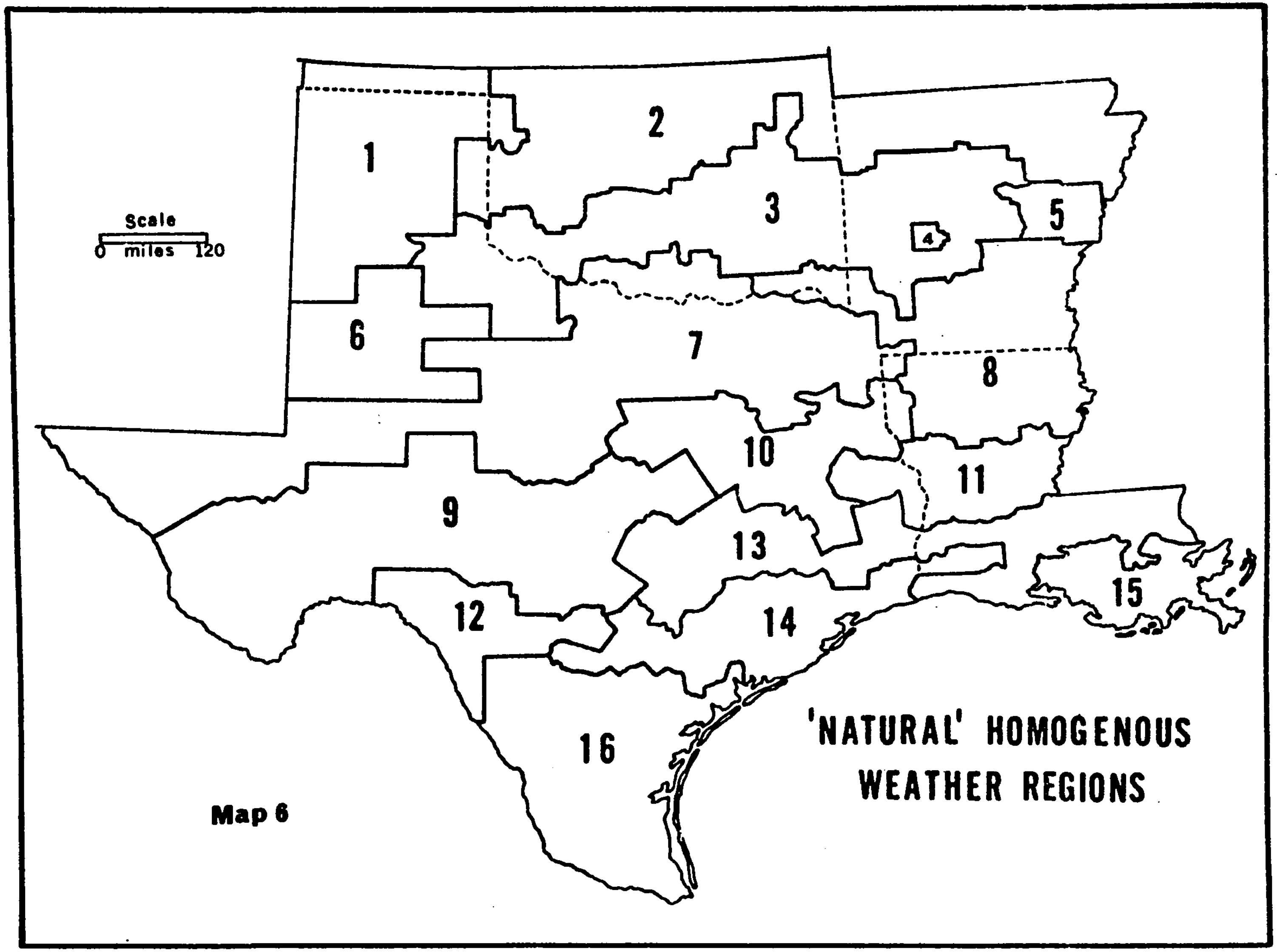
began to take effect as regions began to reach the population limit. The smallest D^2 could not always be used to make a combination; and the within-groups D^2 increased quickly, relative to the increase by the nonconstraint method. The sacrifice of 9.3 percent loss of detail was necessary, however, in order to identify the most effective set of predictive marketing regions as outlined in the purpose of this study.

"Natural" Homogeneous Weather Regions

The question arises as to what changes in regional boundaries might have occurred if no population constraint had been used during the grouping process. Regions derived with no population constraint would have been different in form than the regions derived with a constraint as when the population limit was reached for a region, "forced" grouping would have begun. Evidence of such is provided by the 9.3 percent greater loss of detail in the set of regions derived with a population constraint. Map 6 shows sixteen "natural" homogeneous weather regions derived with no population constraint. These sixteen "natural" regions occurred at grouping step 353 and were compared with the sixteen regions derived with a population constraint. By comparing Map 5 and Map 6, it can be seen that there is a great deal of similarity in the two sets of regions. One anomaly appears on Map 6, which is the weather polygon in central Arkansas that has not been grouped. This polygon is represented by a weather station located in Hot Springs, Arkansas, which apparently has temperature characteristics quite different from any surrounding station.

Summary

The development of a set of predictive marketing regions for LP gas was accomplished by grouping 372 weather polygons into sixteen regions. Mean heating degree days and the standard deviation of heating degree days were the weather


Scale
0 miles 120
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
'NATURAL' HOMOGENOUS
WEATHER REGIONS
Map 6

variables on which the weather uniformity of the regions was based. In addition, a population constraint applied to the grouping insured that no region would have a population that exceeded the arbitrary limit of 1,500,000. As a result the grouping technique produced a set of homogeneous weather regions that are in a best fit relationship with the distribution of population. The percentage of detail lost by having a population constraint was 9.3 percent greater than it would have been with no constraint. The sacrifice of 9.3 percent of detail was necessary, however, in order to delineate the most effective set of predictive marketing regions.

CHAPTER FIVE

EVALUATION OF THE DERIVED REGIONS IN RELATION TO AN EXISTING SET OF LP GAS MARKETING REGIONS

The Purpose of a Classification

The set of uniform weather regions delimited represents a form of spatial classification. There have been many such classifications in geographic analysis. These vary from those based on similarity of traits to those that classify by relationship. Different approaches have been used; some examples are morphological, functional, genetic, or multivariate, the latter incorporating more than one of these approaches. Almost all phenomena with which geographers are concerned, such as land uses, towns, climates, soils, and economies have been classified at one time or another. Yet classification remains, essentially, a means to the end of solving a specific problem. A classification without an objective is of little practical value. There appear to be many occasions when geographers have overlooked this fact. Harvey expresses this idea when he states,

> Classifications have been produced without it ever being quite clear what purposes they are designed for. The geographic literature is replete with complex classifications of towns,

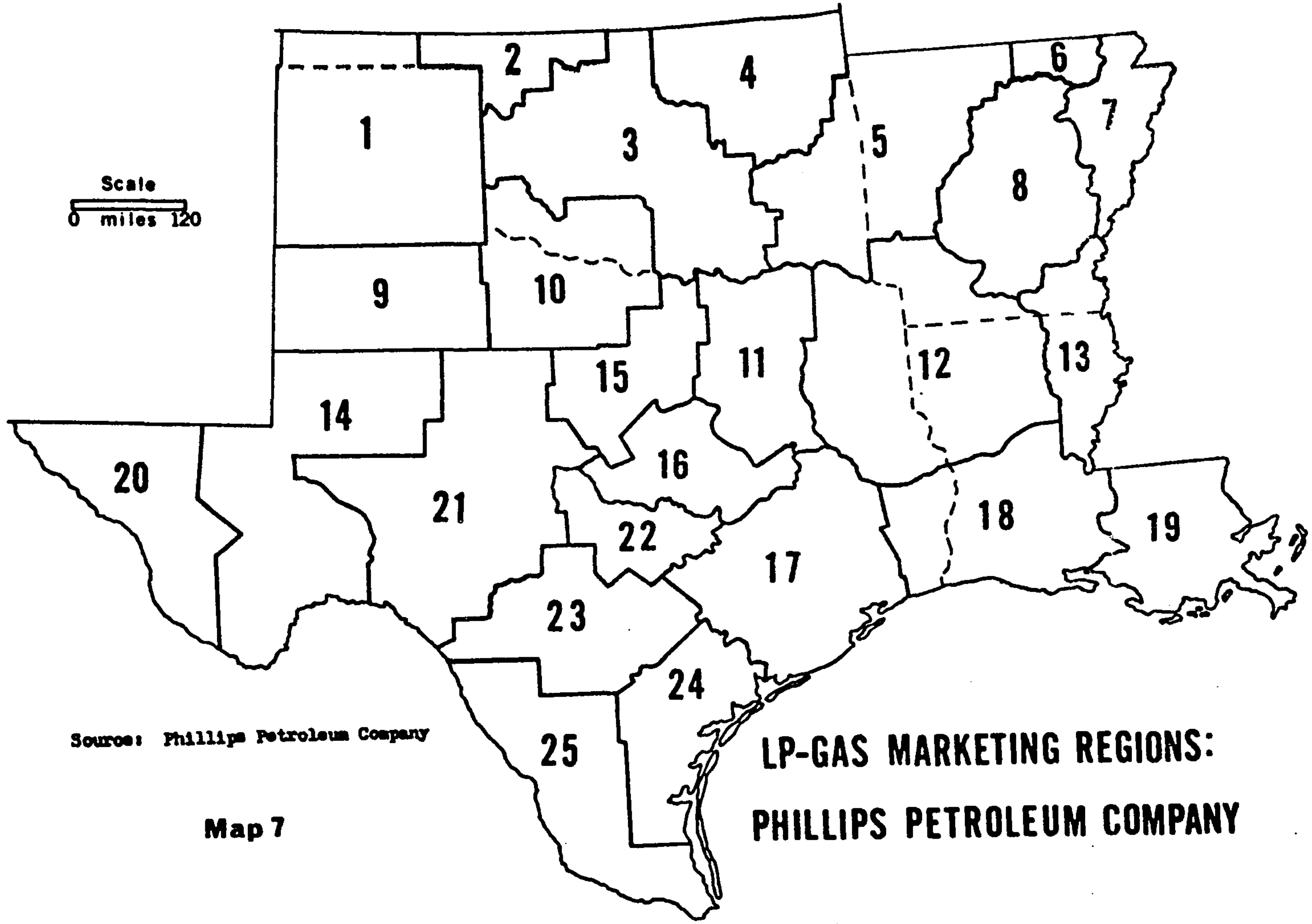
> land uses, climates, regions, morphometric features, and the like, which appear to have been devised with no particular purpose in mind. It is scarcely surprising that many of these classifications have never been used for anything (1969, p. 326).

It is emphasized, therefore, that the spatial classification of heating-degree-day variables accomplished in this study is of little value without a purpose. The adequacy or efficiency of the derived set of uniform weather regions can be assessed only in relation to their purpose. That purpose was to serve as predictive marketing regions for LP gas-heating fuels. The value of the derived set of regions can be demonstrated by a comparison to existing regions now being used for the same purpose.

Arkansas, Louisiana, Oklahoma, and Texas LP gas-marketing regions of Phillips Petroleum Company of Bartlesville, Oklahoma were obtained for comparison. This set of twenty-five marketing regions is shown on Map 7. In some cases, Phillips' regions (Region 6, for example) are quite small because they are part of larger regions that extend across state boundaries. To facilitate the desired comparison, however, the regions were assumed to be only for the four states of the study area.

Comparison of the Two Sets of Regions

The two sets of regions (the derived set and the



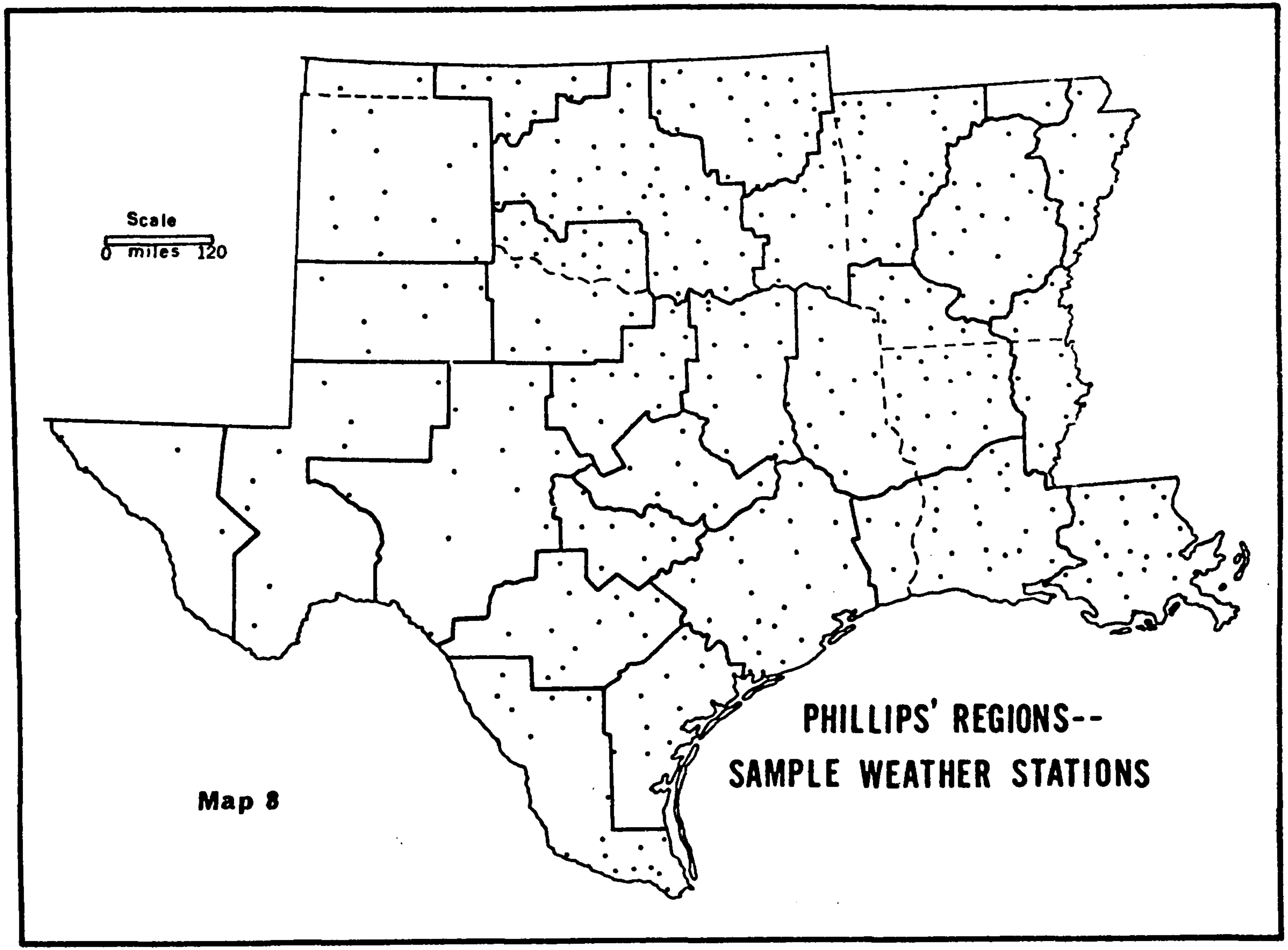
LP-GAS MARKETING REGIONS: PHILLIPS PETROLEUM COMPANY

Source: Phillips Petroleum Company

Map 7

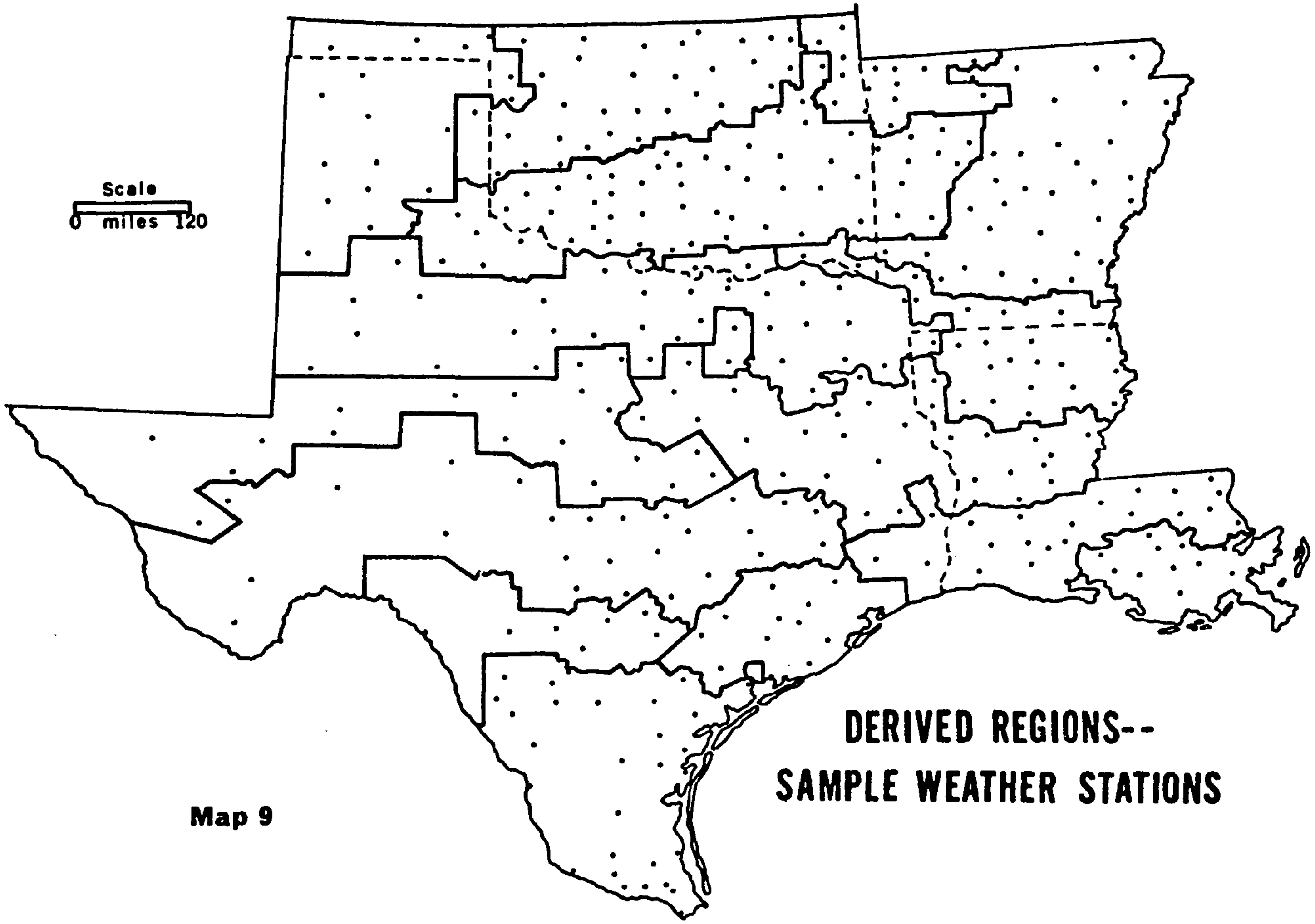
set currently being used by Phillips Petroleum Company) were superimposed upon the 372 weather stations to determine which stations fell into which regions. The resultant pattern for Phillips' regions is shown by Map 8 and for the derived regions by Map 9. It should be noted that Region 6 in the Phillips' set was represented by only one weather station and, therefore, not included in the comparison. Phillips' Region 6 is very small and is an extension of a much larger region centered in southern Missouri. The Phillips' set consisted of twenty-four regions, while sixteen regions constituted the derived set.

Because the key to successful heating-fuel estimates for an entire marketing region is the uniformity of heating-degree-day occurrence within that region, the set of regions with the least intraregional variation of heating degree days will be the most effective. One method of making such a comparison is to compute the mean within-regions sum of squares for both sets of regions. In the case of this study, however, a more realistic measure can be obtained by weighting the within-regions sum of squares by population. The rationale and the formulas for calculating the Prediction Index and the Total Prediction Index have been given in Chapter Three (pp. 38-39). A measure of this type provides an index of prediction error per unit of population for each region and hereafter will be referred to as the Prediction Index. When the Prediction Index of all regions is summed a Total Prediction Index is obtained. If the derived regions have



Map 8

PHILLIPS' REGIONS--
SAMPLE WEATHER STATIONS



Map 9

DERIVED REGIONS--
SAMPLE WEATHER STATIONS

a smaller Total Prediction Index than the set of LP gas-marketing regions currently utilized by Phillips Petroleum Company, they will be more effective in forecasting heating-fuel demand.

The results of the comparison are summarized for the derived regions in Table 3 and for Phillips' Regions in Table 4. Of special note from Tables 3 and 4 are the two Total Prediction Indexes. The Total Prediction Index for the twenty-four regions utilized by Phillips Petroleum Company is 2,082.49, while the Total Prediction Index for the sixteen derived regions is 737.87.

Evaluation of the Results

As a result of the Total Prediction Indexes that were obtained for the two sets of regions, the research hypothesis of this study was accepted. By obtaining a substantially lower Total Prediction Index for the derived set of regions, strong evidence is provided that they are characterized by less intraregional variability of heating-degree-day occurrence than the set of LP gas-marketing regions that are currently being used by Phillips Petroleum Company. The fact that the same geographical area and the same weather stations were used for both sets of regions increases the validity of the results. The results give a strong indication that the derived regions would be more effective in predicting

TABLE 3

PREDICTION INDICES FOR DERIVED REGIONS

Region Number	$\bar{X}$ HDD of Region	Population	Weighted SS_w (in Millions)	Prediction Index (in Thousands)
1	4306.7	367,283	43,347.75	118.02
2	3771.0	1,376,933	78,393.83	56.93
3	3843.1	277,626	10,468.57	37.70
4	3214.5	1,367,224	180,522.30	132.03
5	3237.0	1,220,399	63,175.99	51.76
6	2949.0	1,379,677	141,615.20	102.64
7	2571.3	655,940	8,389.76	12.79
8	2562.4	1,228,939	60,810.71	49.48
9	2225.0	1,415,757	57,189.00	29.86
10	1994.1	904,552	39,656.32	43.84
11	1737.5	1,479,155	33,160.51	22.41
12	1489.9	1,432,060	13,796.80	9.63
13	1513.7	986,655	17,569.79	17.80
14	1335.8	1,293,921	7,025.52	3.06
15	1017.9	934,151	42,141.91	45.11
16	2379.5	1,411,039	6,620.94	4.69
			TOTAL PREDICTION INDEX (in Thousands)	737.87

TABLE 4

PREDICTION INDICES FOR PHILLIPS REGIONS

Number	$\bar{X}$ HDD of Representative Station*	Population	Weighted SS_W (in Millions)	Prediction Index (in Thousands)
1	4101	337,140	49,000.67	145.34
2	4667	54,729	25,365.37	463.47
3	3692	1,143,863	172,518.40	150.82
4	3773	876,509	25,790.73	29.42
5	3413	684,986	124,614.90	181.92
7	3254	368,893	43,725.86	118.53
8	3234	809,347	69,043.61	85.30
9	3493	351,138	26,438.72	75.29
10	2984	412,835	12,105.22	29.32
11	2311	1,706,144	82,255.86	45.54
12	2272	1,296,407	172,225.90	123.33
13	2490	317,254	8,891.02	28.02
14	2680	327,405	26,173.74	79.94
15	2428	925,687	47,584.83	46.39
16	2083	391,047	17,117.76	43.77
17	1350	1,536,850	94,524.92	37.26
18	1616	1,135,858	97,096.23	78.56
19	1422	1,796,631	94,180.40	49.65
20	2636	372,241	1,621.59	4.35
21	2273	346,350	37,635.71	108.66
22	1763	400,403	15,052.93	13.97
23	1653	1,076,909	15,052.93	13.97
24	1038	549,316	14,917.59	27.15
25	638	477,029	36,830.24	77.20
			TOTAL PREDICTION INDEX (in Thousands)	2,082.49

* The representative weather stations are listed by name and region in Table 5, p. 63.

TABLE 5

REPRESENTATIVE WEATHER STATIONS OF PHILLIPS' REGIONS WITH MEAN HEATING DEGREE DAYS, 1962-71

Number	Station Name	$\bar{X}$ HDD 1962-71
1	Amarillo	4101
2	Dodge City, KS	4667
3	Oklahoma City	3692
4	Tulsa	3773
5	Fort Smith	3413
6	Springfield, MO	4280
7	Memphis, TE	3254
8	Little Rock	3234
9	Lubbock	3493
10	Wichita Falls	2984
11	Dallas	2311
12	Shreveport	2272
13	Vicksburg, MI	2490
14	Midland	2680
15	Fort Worth	2428
16	Waco	2083
17	Houston	1350
18	Lake Charles	1616
19	New Orleans	1422
20	El Paso	2636
21	San Angelo	2273
22	Austin	1763
23	San Antonio	1653
24	Corpus Christi	1038
25	Brownsville	638

heating fuel demand than the regions currently used by Phillips for this purpose.

An indication of how much more effective the derived set of regions would be in predicting demand can be obtained in percentage terms by computing the ratio of the Total Prediction Indexes for the two sets of regions. In order to make the two values comparable, the effects of different numbers of regions must be removed. The formula is

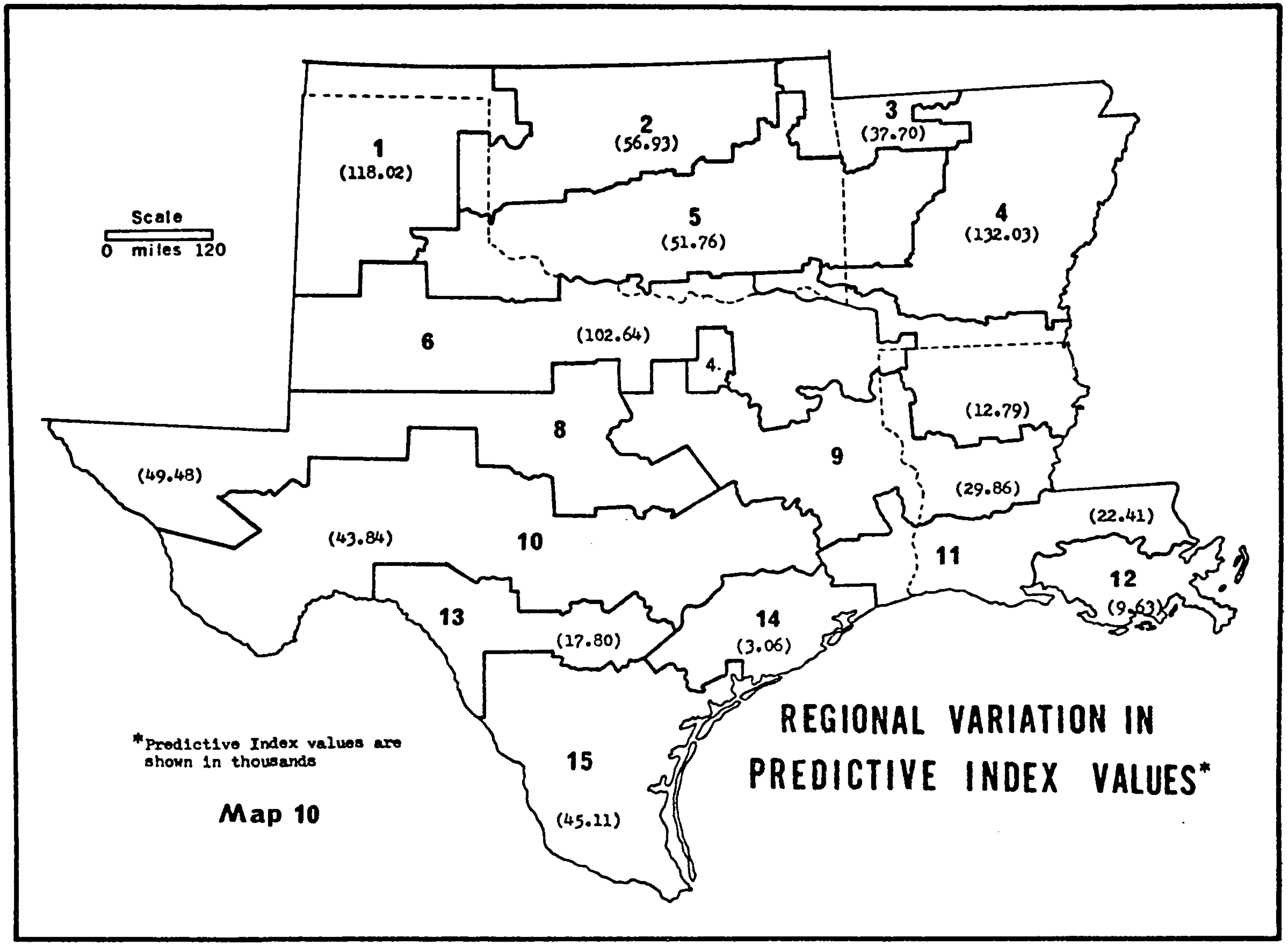
$$\frac{\dfrac{MS_w \text{ (weighted)}}{K_d}}{\dfrac{MS_w \text{ (weighted)}}{K_p}}$$

where K_d is the number of regions in the derived set and K_p is the number of regions in the Phillips' set. The results showed that the derived regions have 53 percent less intraregional variability and would, therefore, be 53 percent more effective in forecasting the demand for LP gas.

An important consideration is that the number of regions in the derived set is less than the number of regions in the set used by Phillips Petroleum Company. Sixteen regions were identified by the regionalizing process as an optimal set, while Phillips uses twenty-four regions for the same geographical area. Thus, the derived set of regions achieved 53 percent less variability in spite of consisting of eight fewer regions. It might be expected that a small Total Prediction Index would result from the use of a large number of highly uniform weather regions. By achieving a substantially smaller Total Prediction Index while using fewer regions provides further evidence that the regions delineated in this study are a more effective set of marketing regions than the set being used by Phillips Petroleum Company. By having eight fewer regions, eight fewer regional offices would have to be maintained. The derived set of regions could be used to cut administrative costs, and, at the same time, be more effective in forecasting demand for LP gas-heating fuels.

The Total Prediction Index obtained for the derived set of regions conceals the variation in homogeneity of the individual regions. For each region, the mean within-region sum of squares (weighted by population) was calculated and termed the Prediction Index. The regional variation in regional Predictive Index values is shown by Map 10. Values range from as low as 3.06 in the case of Region 14, to a high of 132.03 for Region 4, clearly demonstrating that the model does not work equally well at all places. Generally it appears that the smaller, compact regions have less variation, while regions with extreme distances (either north-south or east-west) have greater variation. Regions 1 and 4, for example, have greater north-south distance and also have the highest Prediction Index values.

Although the emphasis of this study is on a total



Map 10

REGIONAL VARIATION IN PREDICTIVE INDEX VALUES* system of regions, it is interesting to compare the individual regions in the two sets. In Table 6, the forty regions have been ranked in order of increasing Prediction Index and identified as belonging to the derived set or Phillips' set. It can be seen from Table 6 that all the individual derived regions do not necessarily have a small Prediction Index. Four of the Phillips' regions are in the first quartile and three of the derived regions are in the fourth quartile.

An important consideration in interpreting Table 6 is that the derived set of regions are not as homogeneous in heating-degree-day occurrence as they could be. Although uniform weather regions provide the important base for developing an effective regional marketing system, they do not take into account the spatial distribution of gas users. Therefore, the weather regions derived in this study have been adjusted into a best-fit relationship with demand intensity. As was discussed in Chapter Four, 9.3 percent of the total detail was sacrificed in order to obtain relatively homogeneous weather regions which would also serve as effective marketing regions. An even smaller Total Prediction Index could have been obtained by removing the population constraint from the grouping process if weather homogeneity had been the only factor considered. Even with the population constraint, however, a substantially smaller Total Prediction Index was obtained for the derived regions than for the set currently being used by Phillips Petroleum Company. The derived set

TABLE 6

INDIVIDUAL REGIONS OF BOTH SETS RANKED BY PREDICTION INDEX

Rank	Prediction Index	Region Source	Region Number
1	3.06	Derived	14
2	4.35	Phillips	20
3	4.69	Derived	16
4	9.63	Derived	12
5	12.79	Derived	7
6	13.97	Phillips	23
7	17.80	Derived	13
8	22.41	Derived	11
9	27.15	Phillips	24
10	28.02	Phillips	13
11	29.32	Phillips	10
12	29.42	Phillips	4
13	29.86	Derived	9
14	37.26	Phillips	17
15	37.70	Derived	3
16	39.20	Phillips	22
17	43.77	Phillips	16
18	43.84	Derived	10
19	45.11	Derived	15
20	45.54	Phillips	11
21	46.39	Phillips	15
22	49.48	Derived	8
23	49.65	Phillips	19
24	51.76	Derived	5
25	56.93	Derived	2
26	75.29	Phillips	9
27	77.20	Phillips	25
28	78.56	Phillips	18
29	79.94	Phillips	14
30	85.30	Phillips	8
31	102.64	Derived	6
32	108.66	Phillips	21
33	118.02	Derived	1
34	118.53	Phillips	7
35	123.33	Phillips	12
36	132.03	Derived	4
37	145.34	Phillips	1
38	150.82	Phillips	3
39	181.92	Phillips	5
40	463.47	Phillips	2

of regions is not only adjusted to customer distribution, but is also relatively homogeneous with respect to heating degree days.

CHAPTER SIX

CONCLUSIONS

This study represents an extension to traditional marketing geography, both in a conceptual and a methodological sense. Traditionally, the emphasis has been on delineating marketing regions as a geographic area in which people would be most likely to trade at a particular town, shopping center, or store. Studies of this type attempted to explain the manner in which geographic space was structured by economic processes. The purpose of this study, however, has been to develop marketing regions, in order to present a strategy that petroleum companies could use to solve the problems that unequal weather occurrence and the unequal distribution of customers created in marketing LP gas as a residential heating fuel. As such, this research is a type of "spatial planning," in which the idea is to organize geographic space in a manner that will best serve human needs. The distinction between explaining how geographic space is utilized in economic activities and organizing geographic space to optimize a particular economic activity is a significant difference in approach. This study, by assuming the latter of these two viewpoints, expands the traditional concepts of marketing geography and points toward what appears to be a fruitful type of research. Intuitively, it appears that the spatial organization theme would have many potential applications within the framework of marketing geography.

No published evidence was found in that the methodology used in this study has been used previously to delimit uniform regions in other marketing geography studies. The results that were obtained in the comparison with an existing set of LP gas marketing regions gave a strong indication of the feasability of using taxonomic techniques to derive uniform regions. Since the delineation of regions is one of the established objectives of marketing geography, numerical taxonomy appears to be a particularly valuable technique.

Further, it is thought that this study has provided some insight into solving the enigmatic question of how many uniform regions should be delimited. As has been described, the taxonomic process groups observations in a series of steps until all observations are in one group. The regionalization technique does not determine the optimum number of groups (regions). Berry (1961) and Bunge (1966) have advocated using the number of groups at which the F ratio was highest. In neither of these cases, however, was a contiguity constraint necessary as in this study. If there were a contiguity constraint on the grouping process, the F ratio would not increase constantly, as the most similar observations might not be spatially contiguous. Also, the grouping step at which the highest F ratio is obtained might not represent the best number of groups for the purpose of the study. Hence, a complete reliance on the F ratio to indicate the best number of regions could lead to a less than optimum solution.

A better solution can be obtained in those cases when uniform regions are derived on the basis of some desirable size characteristic. A constraint can be built into the grouping process so that no group (region) can exceed the specified size characteristic. In this fashion, all groups should approach the desired characteristic limit, and grouping will stop when no further combinations can be made. At this step the best number of groups will have been identified. This procedure was used, as a population total was associated with each weather polygon and a population constraint of 1,500,000 placed on the grouping. As a result, sixteen regions were delimited, of which thirteen had population totals close to the desired figure. This is not to suggest that population is the only characteristic which can be used to determine group size--there are many possible constraints, including the number of observations desired in any group. The decision as to which characteristic to use and the constraining value for that characteristic must be made in keeping with the objectives of the research problem. The results of this study indicate that the constraint method of grouping is an effective way to determine the best number of groups when there is no *a priori* knowledge of the number of groups needed.

In regard to the specific objectives of this study, it can be concluded that the derived set of regions could be more effective in forecasting LP gas-heating fuel demand than the regions currently being used for that purpose by Phillips Petroleum Company. The evidence for this conclusion was provided by demonstrating that the derived regions possess substantially less intraregional heating-degree-day variation per unit of population than the Phillips' regions. Since the accuracy of a regional forecast of LP gas demand is dependent upon regional weather uniformity, any petroleum company using arbitrarily delimited regions will periodically experience problems of oversupply or undersupply. It appears that such problems could be partially alleviated by using the methods described in this study to delimit predictive marketing regions. It is possible, however, to question the accuracy of this conclusion since the Total Prediction Indexes (TPI) for the two sets of regions were calculated differently. The TPI for the derived set of regions was computed by using the regional mean of heating degree days while the TPI for the Phillips' set was computed by using data from one weather station (the representative station of each region). Were Phillips to use the regional mean to forecast demand, the TPI might be quite similar to the derived regions.

To evaluate this possibility, a TPI was calculated for Phillips' regions by using the regional mean of heating degree days--the same procedure used for the regions derived in this study. The results are shown in Table 7. A TPI of 1,638.39 was obtained using the regional mean as compared with a TPI of 2,082.49, using data from one weather station (see Table 4). This decrease of 444.10 is not as substantial as it appears. In Table 7, the Prediction Index (PI) of each Phillips' region is compared with the PI that was computed when heating-degree-day data from the representative station was used (PI_1 - PI_2). It can be seen that most of the decrease (444.10) is accounted for by Region 2, having a PI difference of 412.58. Region 2 is represented by a weather station that is outside the study area (Dodge City, Kansas). Furthermore, all the weather stations used to obtain the variance for this region are in Oklahoma, hence, the extremely high PI of 463.47 obtained in Table 4. If Region 2 were deleted from the results, the difference in TPI's between Table 4 and Table 7 would only be 31.52, a nine percent decrease in the TPI by using the regional mean. Furthermore, as shown by Table 7, for nearly one-half of the individual regions, the Predictive Index was increased when obtained by using the regional mean.

Since Phillips uses the highest order city in each region as the representative weather station and since the spatial variance of heating degree days is being weighted by population to obtain the Predictive Index, it is not surprising that the PI of all individual regions was not substantially

TABLE 7

PREDICTION INDICES FOR PHILLIPS' REGIONS WHEN USING REGIONAL MEAN OF HEATING DEGREE DAYS

Region Number	$\bar{X}$ HDD	Weighted SS_w (in Millions)	PI_2 (in Thousands)	$PI - PI_2$
1	4146.5	48,213.03	143.00	2.34
2	4097.5	2,785.46	50.89	412.58
3	3416.0	75,802.60	66.26	84.56
4	3734.3	26,896.51	28.67	.75
5	3463.9	106,580.20	170.90	11.02
7	3496.3	16,775.76	45.47	70.06
8	3068.8	53,647.68	66.28	19.02
9	3419.5	28,021.45	79.80	+ 4.51
10	3077.2	11,889.18	28.79	.53
11	2591.18	143,194.50	79.28	+ 33.75
12	2555.43	108,245.00	77.51	45.82
13	2594.1	8,060.23	25.40	2.62
14	2547.0	27,401.42	83.69	+ 3.75
15	2681.8	68,927.28	67.20	+ 20.81
16	2107.0	15,653.17	40.02	3.75
17	1564.7	142,106.20	56.01	+ 18.75
18	1850.1	104,198.00	84.31	+ 5.75
19	1639.9	74,446.33	39.25	10.40
20	2966.2	39,654.73	106.52	+102.17
21	2343.0	34,440.98	99.43	9.23
22	2079.4	35,222.04	87.96	+ 48.76
23	1613.3	16,847.48	15.64	+ 1.67
24	1089.3	17,445.81	31.75	+ 4.60
25	988.4	30,638.67	64.22	12.98
			TOTAL PREDICTION INDEX	1,638.39

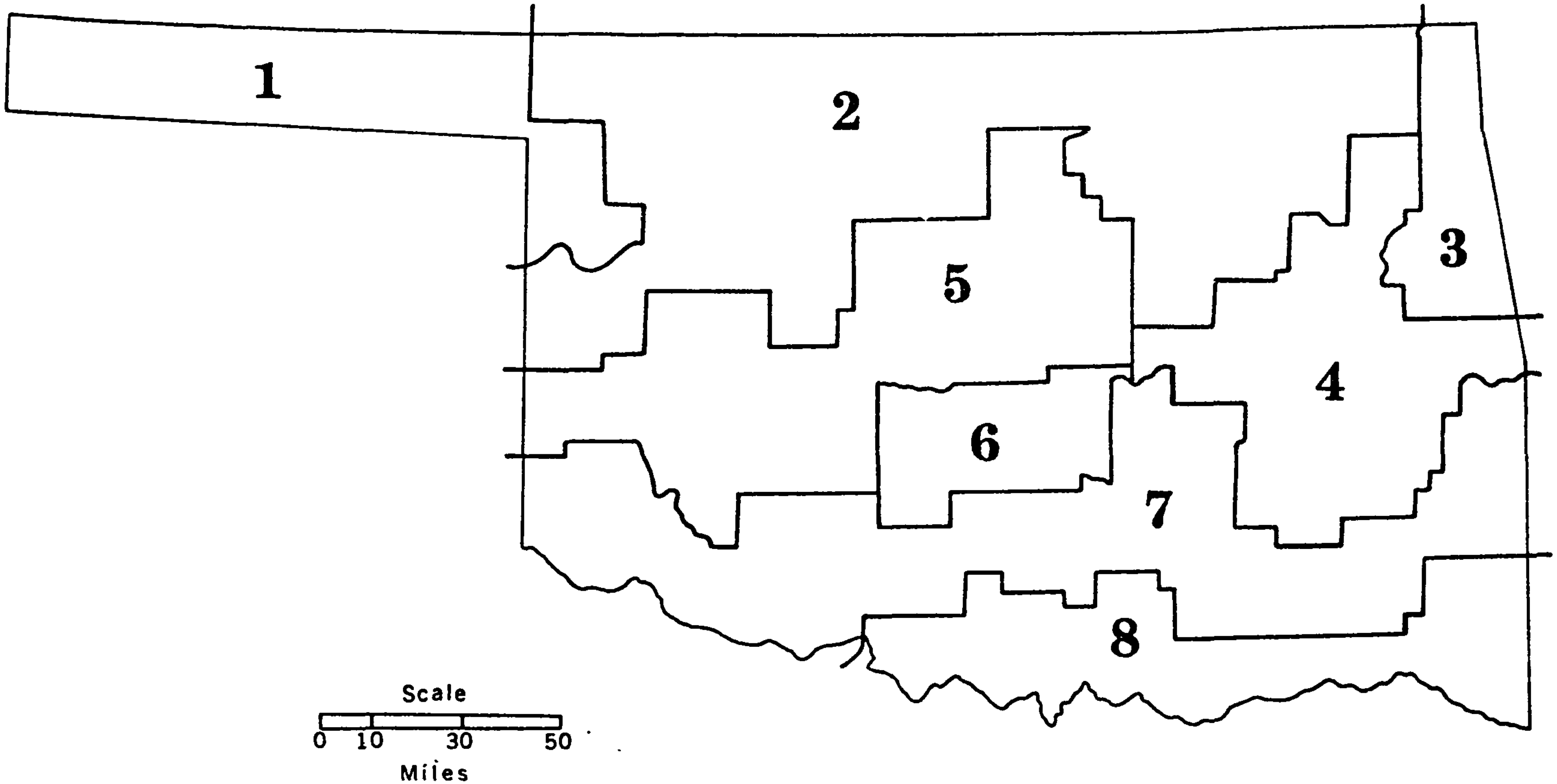
decreased when computed by using the regional mean. In light of the small difference obtained when comparing the two methods, it is doubtful that Phillips could justify the time/cost necessary to obtain heating-degree-data for all weather stations and to compute numerous regional means. The fact that the TPI was not significantly decreased when calculated by using the regional mean further supports the conclusion of this study that the accuracy of a regional forecast of LP gas demand is dependent upon regional weather uniformity. Any petroleum company using arbitrarily delimited regions will periodically experience problems of oversupply or undersupply. It appears that such problems could be partially alleviated by using the methods described in this study to delimit predictive marketing regions.

It should be noted, however, that the taxonomic techniques described and utilized in this study will obtain different results if a different geographical area is used. An example is provided by Maps 10 and 11. Map 10 depicts the regions that were derived for Oklahoma when grouping was accomplished by using eighty-nine weather stations in Oklahoma.* The regions that were derived for Oklahoma when 372 weather stations from the four states of Arkansas, Louisiana, Oklahoma, and Texas were used for the grouping are shown by Map 11. Grouping in both cases was accomplished without the use of a population constraint. Although there are some similarities in the shapes of the two sets of regions, the number of regions

---

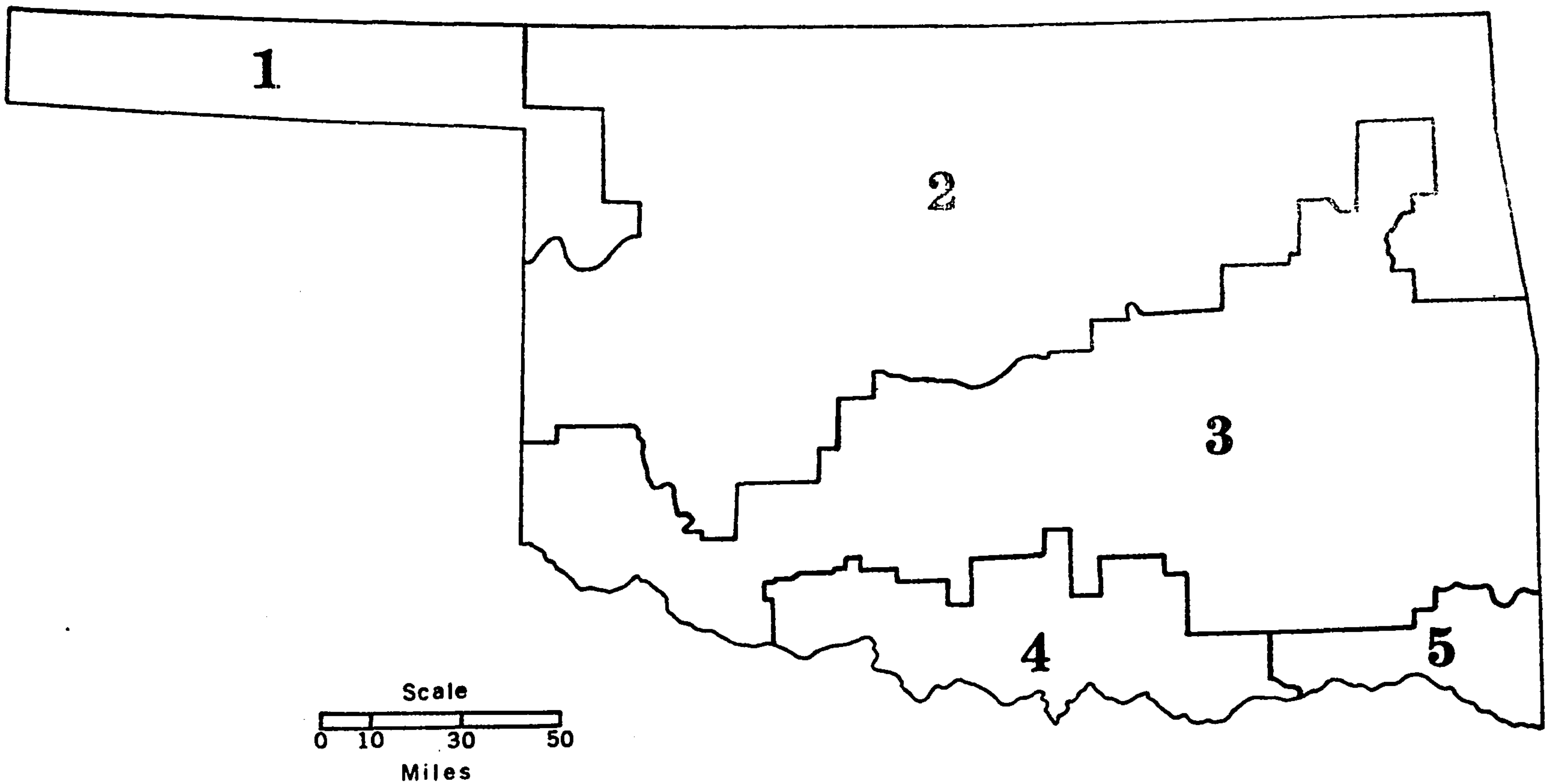
*A population constraint was not used in this case.

REGIONS DERIVED WHEN GROUPING ONLY FOR OKLAHOMA



Map 11

REGIONS FOR OKLAHOMA WHEN GROUPED WITH LARGER AREA



Map 12

derived in each case are different. Eight regions were obtained when grouping only for Oklahoma and five resulted when the grouping was done for four states. A TPI was calculated for each of the two sets based on eighty-nine weather stations in Oklahoma. For the set of regions derived for Oklahoma (Map 10), a TPI of 107.83 was obtained, while a TPI of 150.45 characterized the regions based on data from four states (Map 11). The increased intraregional variability in using a larger geographical area is evident by comparing the two TPI's. As in the case of the two sets of Oklahoma regions, the regions derived in this study were based on a defined study area and will undoubtedly change if a different geographical area is used for grouping.

Although the orientation of this study has been toward the marketing of LP gas by petroleum companies, it is important to emphasize that the methodology utilized has a far wider range of applications. Any product in which demand is influenced by weather occurrence and for which a long-range regional forecast of demand must be made could be substituted for LP gas. Winter clothing, which must be on display in local retail stores in advance of the winter season, is only one of many examples which could be cited. Hence, the system for delineating predictive marketing regions that has been developed in this study could be useful in solving marketing problems for a wide range of products.

APPENDICES

APPENDIX A

SAMPLE WEATHER STATIONS WITH HEATING DEGREE DAY AND POPULATION DATA

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
1	Bentonville	Arkansas	4029	295	24,362
2	Eureka Springs	"	3874	303	13,871
3	Fayetteville AP	"	4178	211	54,397
4	Fayetteville Ex. Sta.	"	3801	223	31,527
5	Gravette	"	4093	230	8,065
6	Harrison AP	"	4092	245	19,073
7	Lead Hill	"	3765	240	7,000
8	Siloan Springs	"	3808	278	18,456
9	Gilbert	"	3842	200	13,575
10	Mountain Home	"	4112	294	29,476
11	Bald Knob	"	3455	261	33,612
12	Batesville	"	3584	170	32,314
13	Corning	"	3783	206	43,536
14	Jonesboro	"	3382	141	99,342
15	Keiser	"	3752	251	110,166
16	Pocohontas	"	3861	180	35,338
17	Searly	"	3251	204	34,613
18	Booneville	"	3218	231	19,284
19	Clarksville	"	3274	195	13,630
20	Dardanelle	"	3293	186	14,208
21	Fort Smith AP	"	3413	150	86,054
22	Fort Smith Water Plant	"	3582	207	7,505
23	Mena	"	3243	200	14,121
24	Ozark	"	3163	175	14,141

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
25	Russellville	Arkansas	3339	140	28,607
26	Subiaco	"	3162	205	3,946
27	Waldron	"	3219	246	10,124
28	Alum Fork	"	3096	239	9,092
29	Benton	"	3120	180	35,806
30	Conway	"	3183	179	34,635
31	Hot Springs	"	2833	176	54,131
32	Little Rock AP	"	3234	220	287,189
33	Malvern	"	3055	151	48,128
34	Morrilton	"	3006	201	19,805
35	Brinkley	"	3509	252	38,949
36	Des Arc	"	2945	199	13,433
37	Helena	"	3421	188	36,493
38	Keo	"	3106	194	26,249
39	Marianna	"	3185	123	45,744
40	St. Charles	"	3254	223	16,033
41	De Queen	"	2848	177	19,705
42	Hope	"	3052	209	15,889
43	Mount. Ida	"	3330	198	9,803
44	Nashville	"	3201	080	15,835
45	Okay	"	2556	160	9,792
46	Texarkana AP	"	2508	122	43,403
47	Camden	"	2718	319	32,219
48	El Dorado AP	"	2848	151	41,413
49	Magnolia	"	2640	173	25,952
50	Prescott	"	2686	247	14,810
51	Warren	"	2812	315	15,380
52	Crossett	"	2580	175	34,143
53	Dumas	"	2827	179	12,193
54	Monticello	"	2891	347	15,157
55	Pine Bluff	"	2638	215	99,178

APPENDIX A, continued:

ID Number	Station Number	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
56	Rohwer	Arkansas	2973	254	23,701
57	Cotton Valley	Louisiana	2653	245	18,312
58	Logansport	"	2298	181	23,234
59	Minden	"	2567	157	49,334
60	Plain Dealing	"	2771	214	32,921
61	Shreveport AP	"	2272	161	236,449
62	Westdale	"	2586	158	16,758
63	Arcadia	"	2615	163	9,012
64	Calhoun	"	2633	159	40,136
65	Chatham	"	2651	187	22,132
66	Farmerville	"	2353	188	16,994
67	Homer	"	2459	157	15,024
68	Monroe	"	2495	181	119,688
69	Ruston	"	2607	151	33,800
70	Winnfield	"	2309	166	21,500
71	Lake Providence	"	2572	143	32,439
72	St. Joseph	"	2408	163	9,732
73	Tallulah	"	2548	150	15,065
74	Winnsboro	"	2485	151	32,982
75	Ashland	"	2700	114	10,850
76	Converse	"	2546	162	20,140
77	Leesville	"	2018	133	51,598
78	Natchitoches	"	1909	144	29,518
79	Alexandria	"	2109	128	18,802
80	Alexandria AP	"	2332	186	91,592
81	Colfax	"	2207	136	22,202
82	Grand Coteau	"	1643	152	87,928
83	Hineston	"	2340	183	8,630
84	Jonesville	"	2395	146	34,347
85	Melville	"	1688	149	41,102
86	Woodsworth	"	2132	145	37,803

APPENDIX A, continued:

ID Number	Station Names	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
87	Amite	Louisiana	1959	162	32,143
88	Baton Rouge AP	"	1767	140	291,564
89	Bogalusa	"	1971	180	31,038
90	Clinton	"	1894	160	22,640
91	Covington	"	1763	205	24,462
92	Franklinton	"	2089	159	16,575
93	Hammond	"	1786	192	75,579
94	Slidell	"	1621	299	33,497
95	Carville	"	1549	140	54,031
96	Franklin	"	1521	239	28,723
97	Lafayette AP	"	1661	141	173,982
98	Morgan City	"	1422	166	41,979
99	New Iberia	"	1570	175	62,927
100	De Quincy	"	1802	174	13,023
101	De Ridder	"	1994	129	24,520
102	Elizabeth	"	2036	157	16,720
103	Hackberry	"	1675	172	7,628
104	Jennings	"	1569	179	33,944
105	Lake Arthur	"	1709	137	9,876
106	Lake Charles	"	1616	222	136,635
107	Oberland	"	1687	176	39,310
108	Diamond	"	1408	207	31,741
109	Donaldsville	"	1488	168	45,665
110	Houma	"	1311	190	110,691
111	New Orleans (Moisant)	"	1631	176	215,974
112	New Orleans (Audubon)	"	1422	155	761,639
113	Paradis	"	1432	171	8,852
114	Reserve	"	1701	177	26,897
115	Schriever	"	1424	168	42,941
116	Arnett	Oklahoma	4429	239	2,737
117	Beaver	"	4396	358	4,189

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
118	Boise City	Oklahoma	4467	291	4,145
119	Buffalo	"	3917	239	5,151
120	Gage AP	"	4315	226	5,878
121	Gate	"	4072	252	2,093
122	Hooker	"	4411	353	16,352
123	Alva	"	4032	237	13,419
124	Blackwell	"	3864	198	44,170
125	Enid	"	3757	212	58,230
126	Jefferson	"	3909	238	8,338
127	Newkirk	"	3976	273	4,621
128	Perry	"	3579	225	13,486
129	Waynoka	"	3778	265	5,725
130	Woodward	"	3841	242	15,537
131	Barnsdall	"	3885	189	19,691
132	Bartlesville	"	3842	201	39,235
133	Bixby	"	3742	293	35,134
134	Claremore	"	4018	284	28,425
135	Jay	"	3858	288	11,740
136	Kansas	"	3676	257	6,027
137	Miami	"	3793	218	29,800
138	Nowata	"	3821	217	9,773
139	Pawhuska	"	3911	259	10,059
140	Ralston	"	3795	229	11,338
141	Spavinaw	"	3472	226	23,302
142	Tulsa AP	"	3773	258	376,697
143	Vinita	"	4052	221	14,722
144	Wagoner	"	3473	239	22,163
145	Clinton	"	3613	221	13,984
146	Cloud Chief	"	3529	191	7,028
147	Cordell	"	3607	190	10,005
148	Erick	"	3611	199	6,486

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
149	Geary	Oklahoma	3637	226	1,299
150	Reydon	"	3849	240	7,306
151	Sayre	"	3648	192	13,416
152	Taloga	"	3968	167	5,656
153	Watonga	"	3802	232	11,471
154	Weatherford	"	3649	180	17,021
155	Antlers	"	3054	159	9,385
156	Hugo	"	2653	173	15,141
157	Idabel	"	2716	176	22,338
158	Poteau	"	3130	174	34,806
159	Smithville	"	3276	198	3,291
160	Wilburton	"	3247	208	8,601
161	Blanchard	"	3287	225	4,730
162	Chickasha Ex. St.	"	3363	232	21,741
163	Cushing	"	3923	252	59,890
164	El Reno	"	3670	207	28,199
165	Guthrie	"	3585	215	19,645
166	Kingfisher	"	3626	257	12,857
167	Norman	"	3302	262	52,285
168	Okemah	"	3340	246	21,839
169	Oklahoma City Penn. Ave.	"	3369	224	429,545
170	Oklahoma City AP	"	3692	270	130,897
171	Purcell	"	3413	194	17,539
172	Seminole	"	3160	198	25,144
173	Shawnee	"	3304	213	41,137
174	Stillwater	"	3761	356	38,040
175	Hanna	"	3280	190	15,764
176	Holdenville	"	3196	218	29,183
177	McAlester AP	"	3347	235	37,521
178	McCurtain	"	3195	133	9,578
179	Muskogee	"	3428	255	78,791

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
180	Sallisaw	Oklahoma	3264	220	28,323
181	Stilwell	"	3669	201	13,405
182	Tahlequah	"	3627	351	23,174
183	Altus	"	3111	185	29,968
184	Anadarko	"	3414	199	12,746
185	Apache	"	3266	213	5,765
186	Chattanooga	"	3158	207	4,268
187	Frederick	"	2913	212	10,845
188	Hobart AP	"	3699	218	7,460
189	Hollis	"	3223	179	9,323
190	Lawton	"	3107	204	97,727
191	Mangum	"	3202	217	7,979
192	Tipton	"	3153	214	2,814
193	Walters	"	3006	266	6,832
194	Wichita Mt. W. R.	"	3427	233	6,149
195	Ada	"	3082	273	12,952
196	Ardmore	"	2630	197	32,750
197	Healdton	"	3000	174	8,475
198	Madill	"	2868	192	33,234
199	Marietta	"	2868	198	5,637
200	Marlow	"	3226	172	34,663
201	Pauls Valley	"	3140	189	24,874
202	Sulphur (Platt)	"	3119	216	18,958
203	Tishomingo W. R.	"	2928	151	7,870
204	Waurika	"	2798	212	17,677
205	Amarillo AP	Texas	4101	206	96,869
206	Canadian	"	3685	204	3,084
207	Canyon	"	3595	231	53,885
208	Crosbyton	"	3565	206	13,286
209	Dalhart AP	"	4526	411	8,794
210	Dumas	"	4587	210	38,503

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
211	Floydada	Texas	3784	328	45,181
212	Hereford	"	4347	232	23,905
213	Lubbock AP	"	3493	267	205,017
214	Miami	"	4377	220	27,916
215	Midland AP	"	2680	184	52,942
216	Midland Plain	"	2548	138	158,393
217	Muleshoe	"	4131	205	31,860
218	Seminole	"	3166	280	28,197
219	Stratford	"	4784	249	19,712
220	Tahoka	"	3334	187	35,858
221	Tulia	"	4048	203	22,651
222	Vega	"	4650	218	2,258
223	Abilene AP	"	2557	212	101,968
224	Ballinger	"	2396	254	13,620
225	Childress AP	"	3317	200	8,924
226	Clarendon	"	3979	180	5,536
227	Haskell	"	2906	164	12,343
228	Matador	"	3098	388	2,178
229	Memphis	"	3324	321	4,606
230	Paducah	"	3045	296	5,415
231	Quanah	"	3480	201	2,215
232	Roscoe	"	2639	227	25,293
233	Seymour	"	3144	155	11,193
234	Snyder	"	3144	176	22,992
235	Wichita Falls AP	"	2984	250	132,837
236	Albany	"	2800	201	19,429
237	Bonham	"	2652	203	22,705
238	Bowie	"	2528	196	19,833
239	Boyd	"	2918	158	23,671
240	Breckenridge	"	2780	205	8,414
241	Bridgeport	"	2806	179	15,333

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
242	Brownwood	Texas	2557	204	35,792
243	Cameron	"	1717	173	25,376
244	Cleburne	"	2292	167	52,137
245	Dallas AP	"	2311	186	1,344,119
246	Dublin	"	2715	218	29,039
247	Evant	"	2021	224	18,248
248	Ft. Worth	"	2428	190	716,317
249	Gainesville	"	3158	157	23,471
250	Georgetown	"	1815	318	23,865
251	Graham	"	2885	178	16,307
252	Greenville	"	2922	152	44,400
253	Hico	"	2214	212	10,153
254	Hillsboro	"	2459	203	32,086
255	Marlin	"	1941	212	17,300
256	McGregor	"	2330	171	141,746
257	McKinney	"	2448	194	66,920
258	Mexia	"	2280	178	31,611
259	Mineral Wells	"	2475	175	28,962
260	Paris	"	2874	185	36,062
261	Pilot Point	"	2918	182	62,120
262	Ranger	"	2696	259	22,182
263	Sherman	"	2978	180	83,225
264	Taylor	"	2032	176	29,287
265	Throckmorton	"	2969	219	2,205
266	Waco AP	"	2083	213	147,553
267	Waxahachie	"	2244	287	72,615
268	Weatherford	"	2854	234	33,888
269	Center	"	2345	168	23,987
270	Centerville	"	2198	161	8,738
271	Clarksville	"	2907	196	13,343
272	College Station AP	"	1811	173	65,866

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
273	Conroe	Texas	1824	155	51,329
274	Crockett	"	2237	156	25,483
275	Dialville	"	2052	169	21,489
276	Gilmer	"	2662	194	28,048
277	Henderson	"	2447	194	47,253
278	Huntsville	"	1898	145	28,993
279	Karnack	"	2240	200	12,778
280	Kirbyville	"	1722	139	29,276
281	Livingston	"	2126	128	17,996
282	Longview	"	2410	164	78,507
283	Lufkin AP	"	1991	152	91,850
284	Madisonville	"	1722	139	11,377
285	Marshall	"	2598	174	48,837
286	Mr. Pleasant	"	2917	163	42,308
287	Palestine	"	2303	188	27,789
288	Rusk	"	2261	192	15,996
289	Sulphur Spgs.	"	2857	156	25,637
290	Texarkana	"	2831	130	74,960
291	Tyler	"	2622	185	114,138
292	Warren	"	1862	170	12,417
293	Wills Point	"	2634	211	51,061
294	Balmorhea	"	2237	173	14,909
295	Chisos Basin	"	2235	244	2,780
296	El Paso AP	"	2636	190	360,234
297	Mt. Locke	"	3364	208	1,527
298	Panther Junction	"	1792	227	9,882
299	Pecos	"	2506	186	14,487
300	Salt Flat	"	3025	186	3,429
301	Wink AP	"	2615	157	22,823
302	Ysleta	"	2840	153	7,051

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
303	Blanco	Texas	2212	181	9,833
304	Boerne	"	2018	229	29,953
305	Burnett	"	1996	229	11,420
306	Del Rio AP	"	1521	201	30,139
307	Eden	"	2147	227	11,508
308	Fredericksburg	"	1969	204	15,038
309	Lampasas	"	2479	186	17,936
310	Llano	"	2259	187	6,979
311	Mason	"	2475	236	3,806
312	McCamey	"	2184	226	15,241
313	San Angelo AP	"	2273	193	76,260
314	Sonora	"	2224	166	13,294
315	Uvalde	"	1615	277	19,361
316	Austin AP	"	1763	203	301,083
317	Beeville	"	1330	190	27,797
318	Brenham	"	1749	177	33,199
319	Chapman Ranch	"	912	154	2,795
320	Corpus Christi	"	859	165	234,596
321	Corpus Christi AP	"	1038	183	59,636
322	Golidad	"	1092	157	16,027
323	Hallettville	"	1345	174	15,259
324	Hondo	"	1664	204	12,175
325	Luling	"	1801	181	32,342
326	Natalia	"	1371	155	22,317
327	New Braunfels	"	1798	231	45,541
328	Nixon	"	1383	203	20,556
329	Refugio	"	1037	136	18,396
330	Robstown	"	1046	154	21,182
331	San Antonio AP	"	1653	188	830,460
332	Sealy	"	1420	154	32,876
333	Sequin	"	1462	197	33,554

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
334	Smithville	Texas	1902	167	35,389
335	Yoakum	"	1255	160	27,670
336	Angleton	"	1471	163	108,312
337	Bay City	"	1319	150	18,241
338	Danevang	"	1216	158	2,778
339	El Campo	"	1252	160	16,984
340	Galveston	"	1229	168	181,999
341	Houston AP	"	1350	171	1,741,912
342	Liberty	"	1647	137	63,010
343	Matagorda	"	1151	145	6,894
344	New Gulf	"	1447	156	8,070
345	Palacios AP	"	1331	146	6,894
346	Pierce	"	1493	156	18,617
347	Pt. Comfort	"	1115	175	19,674
348	Pt. Arthur	"	1481	198	315,943
349	Sugar Land	"	1485	151	45,372
350	Victoria	"	1274	148	62,043
351	Alice	"	1024	163	38,943
352	Carrizo Springs	"	1216	182	6,974
353	Catarina	"	1010	176	2,065
354	Cotulla AP	"	1237	193	4,231
355	Crystal City	"	1227	183	11,370
356	Dilley	"	1481	186	2,892
357	Eagle Pass	"	1441	177	18,093
358	Encinal	"	1321	188	1,350
359	Falfurrias	"	1090	143	20,557
360	Laredo	"	971	246	72,292
361	Pearsall	"	1145	161	12,720
362	Rio Grande City	"	1019	150	11,880
363	Tilden	"	1320	172	1,095
364	Brownsville	"	638	138	79,947

APPENDIX A, continued:

ID Number	Station Name	State	Mean Heating Degree Days	Standard Deviation	Polygon Population
365	Harlingen	Texas	817	127	50,848
366	McAllen	"	803	127	75,006
367	McAllen AP	"	747	138	35,159
368	McCook	"	935	161	12,342
369	Port Mansfield	"	855	108	1,056
370	Raymondville	"	892	159	14,514
371	San Benito	"	704	126	9,573
372	Weslaco	"	639	120	69,234

APPENDIX B

COMPUTER PROGRAM USED FOR REGIONALIZING

```
      DIMENSION LINK(372),X(372),Y(372),POOL(372),LINE(372),KOUNT(372)
      1,DIST(9,372),WSUMSQ(372),SAVE(372)
      INTEGER PTR(372),CONTIG(9,372),LIST(372)
C   INPUT N,XMAX,MOST
C   'N' IS THE NUMBER OF OBSERVATIONS,'XMAX' IS THE MAXIMUM
C   ALLOWABLE WITH-IN GROUP POOLED POPULATION,'MOST' IS THE
C   LARGEST NUMBER OF CONTIGUOUS OBSERVATIONS ASSOCIATED
C   WITH ANY ONE OBSERVATION
      READ(5,150) N,XMAS,MOST
  150 FORMAT(13,2X,F10.0,1X,12)
C   PRINTING THE HEADINGS
      WRITE(6,30)
   30 FORMAT(1H1,50X,'PROGRAM-GROUPING',//,51X,'CARL SPURLOCK',/,45X,
     1'THE UNIVERSITY OF OKLAHOMA',/,45X,'DEPARTMENT OF GEOGRAPHY',//)
      WRITE(6,31) N,XMAX
   31 FORMAT(1X,
     1118('*'),/,5X,'THIS PROGRAM GROUPS N=',13,1X,'OBSERVATIONS'/,5X,
     1'ON THE BASIS OF :',/,10X,'(1) CONTIGUITY',/,10X,'(2) MINIMUM DIST
     ANCE BETWEEN GROUPS',/,10X,'(3) AN ASSIGNED MAXIMUM WITHIN-GROUP PO
     1PULATION=',F9.1)
      WRITE (6,972)
  972 FORMAT(  ///,//8('*'),//,5X,'HERE IS A LISTING OF ALL GROUPS AND
      1THEIR',/,5X,'POPULATIONS AT EACH STEP',///)
C   INPUT THE COORDINATES X,Y AND 'POOL' WHICH
C   IS THE POPULATION ASSOCIATED WITH EACH COORDINATE
      DO 151 K=1,N
      READ(5,152) X(K),Y(K),POOL(K)
  152 FORMAT(5X,F5.0,6X,F4.0,21X,F6.0)
  151 CONTINUE
C   INPUT THE CONTIGUITY MATRIX 'CONTIG'
      READ(5,153) CONTIG
  153 FORMAT(3X,9I3,12X,9I3)
C/ -------------------------------------------------------------
C   REDUCING CONTIG MATRIX
      MT=MOST-1
      DO 650 KR=1,N
      DO 651 MR=1,MOST
  665 KX=CONTIG(MR,KR)
      IF(KX.LE.0) GO TO 650
      IF(KR.LT.KX) GO TO 651
```

```
C   SHIFT
      IF(MR.NE.9) GO TO 659
      CONTIG(9,KR)=0
      GO TO 650
  659 DO 664 LB=MR,MT
  664 CONTIG (LB,KR)=CONTIG(LB+1,KR
      CONTIG(9KR)=0
      GO TO 665
  651 CONTINUE
  650 CONTINUE
C ----------------------------------------------------------------
C   INITIALIZING PTR AND LINK
      DO 1 K=1,N
      PTR(K)=K
      LIST(K)=K
      KOUNT(K)=K
      SAVE(K)=0
    I LINK(K)=0
C ----------------------------------------------------------------
C   CALCULATING DISTANCE MATRIX 'DIST'
      DO 850 L1=1,N
      DO 851 L2=1,MOST
      IF(CONTIG(L2,L1).LE.0) GO TO 850
      L3=CONTIG(L2,L1)
      DIST(2,L1)=(X(L1)-X(L3))**2+(Y(L3))**2
  851 CONTINUE
  850 CONTINUE
C ----------------------------------------------------------------
C   BEGINNING OF MAIN PROGRAM
      ISTEP=0
  300 SMALL=10.0**25
      LOOK=0
C ----------------------------------------------------------------
      DO 100IQ=1,N
      DO 200JP=1,MOST
      I=LIST(IQ)
      J=CONTIG(JP,IQ)
      IF(I.EQ.J) GO TO 200
      IF(J.LE.0) GO TO 100
C ----------------------------------------------------------------
C ----------------------------------------------------------------
C   CHECKING TO SEE IF THE SUM OF THE POOLS OF
C   GROUPS I AND J EXCEEDS THE SPECIFIED CONSTRAINT
  101 SUM=POOL(I)+POOL(J)
      IF(SUM.GT.XMAX) GO TO 200
C ----------------------------------------------------------------
    IS DISTANCE I,J THE SMALLEST?
      IF(DIST(JP,IQ).GE.SMALL) GO TO 200
      SMALL=DIST(JP,IQ)
      ISTOR=I
      JSTOR=J
```

```
      LOOK=1
  200 CONTINUE
  100 CONTINUE
C ----------------------------------------------------------------
   COMBINING GROUPS ISTOR AND JSTOR
C  CHECKING TO SEE IF THE GROUPING PROCESS IS COMPLETE
      IF(LOOK.EQ.0) GO TO 99
      SAVE(ISTOR)=SAVE(ISTOR)+SAVE(JSTOR)+SMALL
C  CHANGING POOLED SUM
      ISTEP=ISTEP+1
      POOL(ISTOR)=POOL(ISTOR)+POOL(JSTOR)
C  CHANGING COORDINATES
      X1=KOUNT(ISTOR)*X(ISTOR)=KOUNT(JSTOR)*X(JSTOR)
      Y1=KOUNT(ISTOR)*Y(ISTOR)+KOUNT(JSTOR)*Y(JSTOR)
      XY=KOUNT(ISTOR)+KOUNT(JSTOR)
      X(ISTOR)=X1/XY
      Y(ISTOR)=Y1/XY
C  CHANGING KOUNT
      KOUNT(ISTOR)=KOUNT(ISTOR)+KOUNT(JSTOR)
C  LINKAGE OF GROUPS ISTOR AND JSTOR
      M=PTR(ISTOR)
    6 LAST=M
      M=LINK(M)
      IF(M.EQ.0) GO TO 5
      GO TO 6
    5 LINK(LAST)=JSTOR
      PTR(JSTOR)=0
C  CHANGING MATRIX CONTIG,DIST,AND LIST
      DO 94 K1=1,N
      IF(LIST(K1).EQ.JSTOR) LIST(K1)=ISTOR
      DO 7 K2=1MOST
      IF(CONTIG(K2,K1).EQ.JSTOR) CONTIG(K2,K1)=ISTOR
      IF(LIST(K1).EQ.ISTOR) GO TO 93
      KN=CONTIG(K2,K1)
      IF(KN.LE.0) GO TO 94
      IF(CONTIG(K2,K1).NE.ISTOR) GO TO 7
   93 IF(LIST(KL).EQ.CONTIG(K2,K1)) GO TO 7
      LI=LIST(K1)
      KN=CONTIG(K2,K1)
      DIST(K2,K1)=(X(LI)-X(KN)**2+(Y(LI)-Y(KN))**2
    7 CONTINUE
   94 CONTINUE
C  WRITING THE GROUPS
      WRITE(6,75) ISTEP
   75 FORMAT(1X,118('*'),/,1X,'STEP',I3)
      WSUMSQ(ISTEP)-0
      DO 50 L=1,N
      IF(PTR(L).EQ.0) GO TO 50
      IF(KOUNT(L).LT.2) GO TO 50
      WSUMSQ(ISTEP)=WSUMSQ(ISTEP)+SAVE(L)
      LL=0
```

```
      NB=KOUNT(L)
      LK=PTR(L)
      DO 70 KB=1,NB
      LL=LL+1
      LINE(LL)=LK
   70 LK=LINK(LK)
      WRITE (6,80) POOL(L),(LINE(KK),KK=1,NB
   80 FORMAT(5X,F10.2,5X,(25I4))
   50 CONTINUE
      GO TO 300
   99 WRITE(6,82)
   82 FORMAT(///,118('*'),//,5X,'HERE IS THE PERCENT DETAIL LOST AT EACH
    1 STEP')
      WRITE(6,81)
   81 FORMAT(///,5X'STEP',5X,'PERCENT',/)
      TOTAL=0
      DO 889 I=1,N
      IF(PTR(I).EQ.0) GO TO 889
      IF(KOUNT(I).LT.2) GO TO 889
      TOTAL=TOTAL+SAVE(I)
  889 CONTINUE
      DO 890 I=1,ISTEP
      WSUMSQ(I)=WSUMSQ(I)/TOTAL*100.0
  890 WRITE(6,891) I,WSUMSQ(I)
  891 FORMAT(5X,I3,F13.5)
      CALL EXIT
      END
```

BIBLIOGRAPHY

Aabund, L. R. (1970), "Ravenous Demand Spells Brighter Days for LP-Gas," The Oil and Gas Journal, Vol. 67, No. 27, pp. 77-81.

Abler, R., Adams, J. S., and Gould, P. (1971), Spatial Organization, Englewood Cliffs, New Jersey: Prentice-Hall.

Applebaum, W. (1954), "Marketing Geography," American Geography: Inventory and Prospect (published for the Association of American Geographers by Syracuse Univ.), pp. 245-51.

Berry, B. J. L. (1958), "A Note Concerning Methods of Classification," Annals of the Association of American Geographers, Vol. 48, pp. 300-303.

Berry, B. J. L. (1961), "A Method for Deriving Multifactor Uniform Regions," Przeglad Geograficzny, Vol. 33, pp. 263-82.

Berry, B. J. L. (1964), "Approaches to Regional Analysis: A Synthesis," Annals of the Association of American Geographers, Vol. 54, pp. 2-11.

Berry, B. J. L. (1967), Geography of Market Centers and Retail Distribution, Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Bunge, W. (1962), "Theoretical Geography," Lund Studies in Geography Series C, Vol. 1.

Bunge, W. (1966A), "Locations are not Unique," Annals of the Association of American Geographers, Vol. 56, p. 377-384.

Bunge, W. (1966B), "Theoretical Geography," Lund Studies in Geography Series C, Number 1 (second edition, Lund).

Cohen, S. B. and Applebaum, W. (1961), "Evaluating Store Sites and Determining Store Rents," Economic Geography, Vol. 37, pp. 95-106.

Court, A. (1970), Seminar given at the University of Oklahoma in the fall of 1970.

Games, P. A. and Klare, F. R. (1967), Elementary Statistics: Data Analysis for the Behavioral Sciences, New York: McGraw-Hill, Inc.

Gregory, S. (1963), Statistical Methods and the Geographer. London: Longmans, Green and Co., Ltd.

Grigg, D. (1965), "The Logic of Regional Systems," Annals of the Association of American Geographers, Vol. 55, No. 3, pp. 465-491.

Haggett, P. (1965), Locational Analysis in Human Geography, New York: St. Martin's Press, Inc.

Hartshorne, R. (1939), The Nature of Geography, Chicago: Rand McNally.

Harvey, D. (1969), Explanation in Geography, London: Edward Arnold Publishers, Ltd.

Johnston, R. F. (1968), "Choice in Classification: The Subjectivity of Objective Methods," Annals of the Association of American Geographers, Vol. 58, No. 3, pp. 575-589.

King, L. J. (1969), Statistical Analysis in Geography, Englewood Cliffs, New Jersey: Prentice-Hall.

Lankford, P. M. (1969), "Regionalization: Theory and Alternative Algorithms," Geographical Analysis, pp. 196-212.

National LP-gas Association News Release (1971).

Schaefer, F. K. (1953), "Exceptionalism in Geography: A Methodological Examination," Annals of the Association of American Geographers, Vol. 43, pp. 226-49.

Spence, N. A. (1968), "A Multifactor Uniform Regionalization of British Counties on the Basis of Employment Data for 1961," Regional Studies, II, pp. 87-104.

Steiner, D. (1965), "A Multivariate Statistical Approach to Climatic Regionalization and Classification," Tijdschrift van het Koinklijk Nederlansch Aardrijkskundig Gerootschap Tweede Reeks, LXXXII: 4, pp. 324-347.

Taylor, P. J. (1969), "The Location Variable in Taxonomy," Geographical Analysis, pp. 181-195.

Thiessen,A. H. (1911), "Precipitation Averages for Large Areas," Monthly Weather Review, Vol. 39, pp. 1082-1084.

U.S. Department of Commerce, Environmental Science Services Administration, Weather Bureau, Average Monthly Weather Outlook.

U.S. Department of Commerce, Weather Bureau, Climatological Data.

Ward, J. H. (1963), "Hierarchical Grouping to Optimize an Objective Junction," Journal American Statistical Association, Vol. 58, No. 1, pp. 236-243.

Weather Forecasting Task Group (1971), "The Importance of Weather Forecasting to the Gas Industry," American Gas Association Monthly, pp. 31-36.

Weaver, J. C. (1954), "Crop-Combination Regions in the Middle West," Geographical Review, Vol. 44, pp. 175-211.

Whittlesey, D. (1954), "The Regional Concept and the Regional Method," American Geography: Inventory and Prospect (published for the Association of American Geographers by Syracuse Univ.), pp. 21-68.