REPORT NO. DOT-TSC-OST-74-26

A SURVEY OF NATIONAL GEOCODING SYSTEMS

Pamela A. Werner

NOVEMBER 1974

FINAL REPORT

Prepared for
U.S. DEPARTMENT OF TRANSPORTATION
OFFICE OF THE SECRETARY
Office of the Assistant Secretary for Policy
Plans and International Affairs
Washington DC 20590

Technical Report Documentation Page

. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
OT-TSC-OST-74-26		

. Title and Subtitle

SURVEY OF NATIONAL GEOCODING SYSTEMS

5. Report Date
November 1974

6. Performing Organization Code

. Author(s)
amela A. Werner

8. Performing Organization Report No.
DOT-TSC-OST-74-26

. Performing Organization Name and Address
rban Systems Laboratory*
lassachusetts Institute of Technology
ambridge MA 02139

10. Work Unit No. (TRAIS)
OP409/R5801

11. Contract or Grant No.
DOT-TSC-692

13. Type of Report and Period Covered
FINAL REPORT
October 1973-June 1974

2. Sponsoring Agency Name and Address
.S. Department of Transportation
ffice of the Secretary
ffice of the Assistant Secretary for Policy
lans and International Affairs
ashington DC 20590

14. Sponsoring Agency Code

5. Supplementary Notes

*Under contract to: U.S. Department of Transportation
Transportation Systems Center
Kendall Square
Cambridge MA 02142

6. Abstract

This document describes major geocoding systems. It is organized into sections that categorize geocoding systems by type. Section 2 deals with systems that are primarily geopolitical in nature and provide general reference coding structures for administrative or other purposes. Section 3 includes those geocoding systems that reference either special significance locations or a combination of geopolitical, geostatistical and special significance locations. The geocoding systems in Section 4 include those that reference areas delineated according to special criteria, such as economic or postal distribution patterns. Finally, Section 5 discusses those systems based on grid networks.

17. Key Words

Geocoding
Geography
Area
Geopolitical
Coding

18. Distribution Statement

For Sale by the Superintendent of Documents
U.S. Government Printing Office
Washington DC 20402

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	356	

orm DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

Because information on national geocoding systems is widely scattered and difficult to assemble, a survey of this size and scope would have been impossible without the help and cooperation of many individuals.

I would like to express my gratitude to the following persons for their assistance: Mr. Dwight W. Briggs, Federal Highway Administration; Mr. John Cannavan, Federal Railroad Administration; Mr. Donald E. Church, formerly of the Bureau of the Census; Mr. Edwin J. Coleman, Bureau of Economic Analysis; Mr. Peter L. Conway, Jr., Association of American Railroads; Mr. Reginald A. Creighton, the Smithsonian Institution; Mr. John P. Fajnor, International Business Machines Corporation; Mr. Joel Frisch, Water Resources Council; Mr. J. G. Fultz, Canada Post; Mr. A. Gomes, International Telephone and Telegraph; Mr. Robert E. Graham, Jr., Bureau of Economic Analysis; Mr. Richard J. Greenfield, U.S. Postal Service; Mr. Robert B. Haden, Federal Railroad Administration; Mr. George Haydu, Statistics Canada; Mr. Robert C. Hennell, American Trucking Associations; Col. William W. Higgins, Department of Defense; Mr. Richard W. Jones, U.S. Postal Service; Mr. Dayton P. Jorgenson, Bureau of the Census; Mr. Christos N. Kyriazi, Economic Development Administration; Mr. Edward Margolin, formerly of the Interstate Commerce Commission; Mr. Morton A. Meyer, Bureau of the Census; Mr. Wallace B. Oliver, General Services Administration; Ms. Beatrice Orris, formerly of the Economic Development

Administration; Mr. Hans W. Roesch, Dun and Bradstreet, Inc.; Mr. Walter L. Schlenker, American National Standards Institute; Mr. Allan H. Schmidt, Harvard University; Mr. Louis Schwalb, the Office of Civil Defense; Ms. Nevaire Serraijan, Defense Intelligence Agency; Mr. Dik W. Twedt, formerly of the U.S. Postal Service; Ms. Nell Vetter, Federal Aviation Administration; and Mr. Ralph E. Waters, American Telephone and Telegraph Company.

A special debt for support and help of many kinds is owed to Mr. Alan E. Pisarski, U.S. Department of Transportation, Mr. Robert E. Barraclough, Transportation Consultant, and Mr. H. W. Bruck of the Massachusetts Institute of Technology.

I would also like to express my profound gratitude to Ms. Susan M. Wallberg for helping me prepare the many illustrations in this survey.

CONTENTS

CONTENTS

LIST OF ILLUSTRATIONS

Figure

LIST OF ILLUSTRATIONS (Continued)

Figure

LIST OF ILLUSTRATIONS (Continued)

Figure

LIST OF ILLUSTRATIONS (Continued)

Figure

LIST OF ILLUSTRATIONS (Continued)

Figure

1. INTRODUCTION

The two major structural elements of all geographic coding systems are a concept of areal division, classification or definition and some form of coding logic. In recent years, overt emphasis has been given to the automated aspects of geocoding logic and data storage, retrieval and display in large, geographically referenced information systems. This has resulted in a popular tendency to assume that a legitimate geocoding system must be computer based and requires a fairly sophisticated coding structure. Broadly conceived, however, systematic geographic coding has had a long and diverse history which includes many classification rather than coding oriented systems of geographic reference.

Geographic coding had its genesis in ancient times with the emergence of cartography and geography. As early as 200 B.C., Eratosthenes had calculated the circumference of the earth within 200 miles and had devised a grid of latitude and longitude lines on which he plotted the locations, with coordinate references, of seas, lands, mountains, rivers and towns. Somewhat later, Strabo compiled the Geographia which represented the most ambitious attempt, before the sixteenth century, to catalogue and locate all of the place names in the known world.

For centuries thereafter, geographers, cartographers, explorers, and surveyors were devising geographic reference systems whereby areas were subdivided and coded for purposes of classification, administration or data collection. Current examples in the field of physical geography are the World Climatic Region System developed originally by Köppen (later simplified by G. T. Trewartha), and the World Natural Vegetation System developed by A. W. Küchler.

In the United States, the first large scale geographic coding system had its beginnings in 1785 in a program designed to subdivide and provide unique geographic identification for the vast tracts of publicly owned land. This public land survey was based on a grid system which divided areas of six square miles, known as townships, into forty acre lots, called quarters. Each of these was identified by an alpha numeric designation. The system was used for administration, survey and transfer of title under the public land laws of the United States and was utilized, at least in part, by all states except the thirteen original ones and six others.

As the major source of basic demographic statistics in the United States, the U.S. Bureau of the Census played an important role in the further development of geocoding systems. Since its establishment as a permanent agency in 1902, the Bureau of the Census has attempted to provide a flexible geographic framework for the tabulation of data that was designed specifically to meet the needs of Census data users in both the public and private sectors. Census tracts, for example, were originated by Dr. Walter Laidlaw of New York City in 1906. At his request, the Bureau of the Census defined tracts and tabulated data from the 1910 Census by tract for New York and seven other cities. State economic areas, urbanized areas, standard metropolitan statistical areas, central business districts and numerous other geostatistical units are further examples of geographic areas defined and coded by the Bureau of the Census in order to provide tabulations and aggregations of data that met the needs of Census users.

During the twentieth century, the development of information systems in general, the significant increase in the collection of geographically referenced data, and the automation of data processing led to widespread use of geographic coding. Thus, whether the functional definition of geographic units or a flexible, machine readable coding logic is the primary objective in the design of any given geographic reference system, such systems have become integral to a wide range of information and therefore, of concern to a broad spectrum of users.

1.1 URBAN AND NATIONAL GEOCODING

The requirement for a continuing metropolitan transportation planning process that was made statutory in 1962 created an increased demand upon the Bureau of the Census to provide small area data and expanded data user services. This, in turn, led to the initiation of the Census Use Study in 1966. The primary objective of the Census Use Study was to improve methods for relating census data with local agency data at a fine geographic scale. As a result, the Bureau of the Census developed a highly generalized urban geocoding system and was instrumental in promoting establishment of urban geographic base files which could be used as tools in relating the various available data sets to a single framework of geographic reference. By 1971, a fairly standardized universe of urban geographic base files utilizing the Dual

Independent Map Encoding (DIME) system, based on block face coding, had been implemented in almost all standard metropolitan statistical areas. A battery of DIME related computer programs, including UNIMATCH, GRIDS, CRAM, and DACS, had been developed and were available for use.

In the United States, comparable incentive for the development of a generalized national geographic base file has not existed. Thus, national geocoding is characterized by specialization and multiplicity of structure. Each system tends to be "tailor made" for a specific purpose, these ranging from regional planning, commodity flow analysis, market analysis, transportation planning, shipment routing, plant and terminal location, freight rating, museum cataloguing, and credit accounting to simple clerical indexing of administrative records.

In addition to a relative absence of coordination among the various macro-level geographic base files, national geocoding lacks one very important feature which has been an essential element in the emergence of a unified approach to metropolitan geocoding. This essential feature is the availability of a "common denominator" coding unit. Micro-level geographic base files, such as DIME and GRDSR[1], provide maximum geographic manipulation capabilities because they are based upon the encoding of block face, a primary and virtually universal unit of urban geography. The city block is one of the smallest (second only to parcel units), most standard and relatively permanent urban areal units. By encoding each side of a block individually, it is possible to provide

---

[1]Geographically Reference Data Storage and Retrieval. A Canadian urban geocoding system (See Appendix B).

a very flexible, fine grained geographic coding system for an entire metropolitan area. Block face is the most viable common denominator unit for urban geographic base files because, in general, all special areas within a city are made up of whole block face units. With such common denominator coding, data sets can be aggregated or disaggregated to conform to any number of special area boundaries, e.g., school districts, traffic zones, police precincts, etc. Thus, only one geographic coding system is required to meet the varied demands of many users.

There is no analogous common denominator unit for national geographic base files. A nation's geography encompasses a far more heterogeneous mix of land use patterns, natural areas, governmental entities and other spatial orderings than that which characterizes metropolitan geomorphology. While many national geocoding systems are based on county units or units compatible with county boundaries, there are important exceptions such as ZIP code zones, Congressional Districts and others. These do not necessarily aggregate to the county level. Below the county level, there are numerous obstacles to interfacing the various national geographic coding systems. There is a lack of definition, both semantic and geographic, of subcounty units. There is the problem of variation between urban and rural land use, settlement patterns, and population densities which create great disparities in the size of the spatial units coded both within a single system and among the various systems. And finally, the relative instability of subcounty units and frequent changes affecting this universe pose other problems primarily in the areas of update and maintenance.

The contrast between urban geographic base files and national geographic base files is one of singular versatility opposed to multiple inflexibility.

1.2 CHARACTERISTICS OF NATIONAL GEOCODING SYSTEMS

Aside from the two basic structural elements which all geocoding systems have in common--a concept of areal division or classification and a coding logic--there are numerous characteristics by which these systems could be classified into related groupings. The type of geographic definition and spatial relationships reflected in the system, the type of geographic units coded, the scale, grain and flexibility of each system, and the code assignment process provide a number of taxonomic criteria. However, due to the formidable number and variety of national geocoding systems, no single set of criteria defines mutually exclusive categories applicable to all national geographic coding systems. Some systems are similar in scale or grain but differ in coding logic; other systems have the same coding logic but are utilized for completely different purposes; and some systems encompass a number of subsystems each having a separate structure.

One of the major difficulties in classifying geocoding systems has been semantic. In the words of one geographer, "There is a clearly felt need to clarify terms in the technology and institute more precise phraseology--terms that are more fundamental and independent of the narrow trade names derived from various approaches and computer program acronyms."[1] Beyond trade names and acronyms, a mundane but typical example of the semantic difficulties which have plagued national geocoding are the various usages of the two words, "place" and "point". For many months the American National Standards Institute (ANSI) debated the meaning of the word place. The committee appointed to deal with this issue finally concluded that no formal definition or criterion could be formulated to the satisfaction of all committee members. Many other agencies such as the Department of Defense and the Bureau of the Census have had similar difficulties formulating a definition for the word place. Thus the term place, so frequently used in any discussion of national geocoding systems, has no exact, generally accepted meaning.

The word point is no less promiscuously used, according to perspective and purpose. To a member of the ANSI point location committee, it is an imaginary, nondimensional dot on the surface of the earth which represents a cartesian, state plane or universal transverse mercator coordinate fix. To a member of the American Trucking Associations or the Association of American Railroads, a point represents a siding, a loading dock or waybill station, in short, an entry in the Standard Point Location Code file. In the Census of Transportation, it represents the geographic centroid of a municipality which is one of approximately 5,700 such "Key Points" in the United States.

---

[1]C. E. Barb, "Geocoding Systems, Their Philosophy and Utility," an unpublished dissertation discussion paper, University of Washington, Seattle, n.d., p.1.

These two terms--place and point--as well as others, such as location, site and area, have developed multiple and ambiguous meanings. They are the chameleons of geography, changing color according to context, specifically user dependent and, therefore, major contributors to the state of confusion which many consider characteristic of national geocoding.

1.2.1 Geographic Definiton. A fundamental distinguishing element among the complex variety of geocoding systems is that between geographic identification and geographic definition.[1] Coding systems which consist of a set of unique designations for undefined or implicitly defined locations, such as ANSI place codes or ZIP codes, are geoidentifying systems. Coding systems which provide explicit boundary delineations for the units coded, such as the state plane coordinate systems, are geodefining systems.

A further distinction in regard to this fundamental classification of geocoding systems is between the two types of geoidentifiers. While none of the geographic identification systems provide for explicit boundary definition of the units coded, some provide at least implicit boundary definition, others provide no definition at all. Roger Tomlinson refers to the former as external indices which "identify the data as belonging to a particular defined geographic area . . . and must be used

---

[1]The concept of geodefinition and geoidentification was suggested by Betsy Schumacker in a paper prepared for the National Geocoding Conference, December 1, 1971. See the National Geocoding Conference Proceedings (Washington, D.C.: U.S. Department of Transportation, 1972).

in conjunction with a map."[1] Federal Information Processing Standard (FIPS) codes for the counties and county equivalents of the United States, for example, comprise an external index. These county codes refer to a set of areal units which have explicit boundaries, but the code itself does not define them. Boundaries for the referents of the ANSI place code, on the other hand, have not been established and therefore, these codes do not provide even an implicit areal definition of the units coded.

1.2.2 Spatial Relationships. A second set of fundamental cross-classifications for geocoding systems are the distinctions between nominal, ordinal, and cardinal geographic references.[2] Strictly defined, a nominal geographic code is the name of a location. However, for the purpose of classifying geographic coding systems, nominal codes include names, name abbreviations, mnemonic truncation of names, and certain numerical codes. In this case, the word nominal derives meaning not from its root word, name, but from describing a designation, alpha or numeric, which does not indicate or imply a spatial relation between itself and other similar designations.

Nominal geocodes do not indicate any spatial relationships among the units coded. In contrast, ordinal and cardinal geocodes do.

---

[1] R. F. Tomlinson, "Introduction," Geographic Data Handling, Symposium Edition, Vol. 1, R. F. Tomlinson, ed. A publication of the International Geographical Union Commission on Geographical Data Sensing and Processing for the UNESCO/IGU Second Symposium on Geographical Information Systems, Ottawa, Canada, August, 1972.

[2] This set of coding classifications was suggested by C. E. Barb in an unpublished dissertation discussion paper, op. cit. The author has altered the application of the taxonomy slightly.

Ordinal geocodes indicate the relative positions of coded units within a spatially related system. A hierarchically structured geocoding system such as the one employed by the Bureau of the Census, for example, indicates that a certain census tract is located within a certain minor civil division which is located in a certain county within a certain state. Cardinal geocodes indicate the absolute positions of coded units within a spatially related and incrementally scaled system. The network of meridians and parallels, known as latitude and longitude, is the best example of a cardinal geocoding system.

	GEOGRAPHIC IDENTIFICATION	GEOGRAPHIC DEFINITION
NOMINAL	Location Designation	Boundary Description
ORDINAL	Hierarchical/Grid Code	Local Coordinates
CARDINAL	Coordinate Related Code	Global Coordinates

1.2.3 <u>Type of Geographic Units Coded</u>. National geocoding systems may also be classified according to the type of units coded. Most of the systems surveyed in this paper are concerned with the spatial relationships created by the settlement patterns and land uses in the United States.

This includes geopolitical units such as counties and states; geo-administrative units such as enumeration districts or customs districts; geostatistical units such as standard metropolitan statistical areas or urbanized areas; and, geoeconomic units such as ZIP marketing areas or major retail centers. A few systems identify physical geographic features such as drainage basins or areas of homogeneous soil type in the United States. Other systems, including most of the grid and coordinate geocoding systems, are unrelated to either physical geographic features or human geography.

1.2.4 Scale, Grain and Flexibility. The scale and grain of geographic units coded are also significant characteristics of geocoding systems. They are the most variable characteristics of the systems surveyed in this paper and defy simple categorization. A geocoding system, such as PICADAD, may encode a set of small scale point locations with a universe of only 5,700 entities, while another system, such as the Standard Point Location Code, based on similar scale point locations has a universe of 110,000 coded entities. Scale and grain may vary independently and widely, ranging from theoretically nondimensional points with an infinite universe of entities to eight, nine, or ten large regional areas.

The flexibility of a geocoding system and the ability to disaggregate or aggregate the geographic units to which data is coded is dependent upon all of the other system characteristics. Maximum flexibility is achieved in systems that have small, mutually exclusive, geodefined units with a fine grain. Coordinate geographic coding systems such as latitude and longitude and the state plane coordinate systems are obviously by far the most flexible systems. Many tiered hierarchical systems such as the Census geocoding system or the National Location Code are not as manipulable as coordinate systems, but they do have several alternative levels of aggregation and disaggregation. Geocoded place lists such as those developed by the American National Standards Institute, Dun and Bradstreet, and the General Services Administration have very limited flexibility because the basic geographic unit within these systems lacks geodefinition and there are, at maximum, only two or three possible levels of aggregation.

1.2.5 Code Assignment. Among the national geocoding systems surveyed in this paper there are four basic coding methods and several combinations thereof. The four basic coding methods are: 1) mnemonic name abbreviations, such as airport codes or passenger railroad station codes, which are designed to provide easily recognizable and easily memorized codes facilitating human usage of the coding system; 2) code designations assigned to units sequenced alphabetically by name, such as in the Dun and Bradstreet geocoding system, which provide the simplest form of look up tables for encoding and decoding of data; 3) code designations assigned to units sequenced in some geographic order, such as in the Standard Point Location Code, which are designed to reflect the geographic proximity of the units coded; and, 4) grid and coordinate code assignment, such as in the military reference grid system or latitude and longitude, which provide for area and distance calculations.

A Survey of National Geocoding Systems is intended to be a reference volume containing general descriptions of the major national geocoding systems employed by public and private agencies in the United States. The material included in this survey is limited to those geographic reference systems having nationwide application[1] and no overt attempt has been made to analyze, compare or judge any of the coding systems surveyed.

As previously noted, no single set of criteria defines mutually exclusive categories applicable to all national geographic coding systems. Therefore, this paper has been organized in four major sections which represent a compromise taxonomy. Section 2, Geopolitical and General Location Reference Codes, includes descriptions of those geocoding systems that are primarily geopolitical in nature and provide general reference coding structures for administrative or other purposes. Section 3, Geopolitical and Specialized Location Reference Codes, includes those geocoding systems that reference either special significance locations (e.g., transportation related locations, military installations, etc.), or a combination of geopolitical, geostatistical, and special significance locations. Section 4, Specialized Areal Unit Reference Codes, includes those geocoding systems which reference areas delineated according to specialized criteria, such as economic patterns or postal distribution patterns. Section 5, Grid and Coordinate Codes, includes those geocoding systems based on grid networks.

---

[1]This excludes a number of recently developed regional and statewide geographic base files designed for land use inventory and analysis, many of which utilize a combination of standard data collection methods and remote sensor data. For a comprehensive review of geocoding activity in this field refer to Geographic Data Handling, Vols. 1 and 2, R. F. Tomlinson, ed., op. cit.

2. GEOPOLITICAL AND GENERAL LOCATION REFERENCE CODES

2.1 FEDERAL INFORMATION PROCESSING STANDARDS

Public Law 89-306, "The Brooks Bill" authorizes the establishment of mandatory standards within government data processing systems and a voluntary standards program for nongovernment agencies. It is the responsibility of the National Bureau of Standards working with the cooperation of other government agencies, the American National Standards Institute (ANSI), and the International Organization of Standardization (ISO) to implement the standards programs and to maintain a register of Federal Information Processing Standards (FIPS). This register is the official source within the Federal Government for information pertaining to the approval, implementation and maintenance of standards for all aspects of data processing, including hardware, software, applications, and data standards. Each of the FIPS standards must be approved by the Office of Management and Budget (OMB), is maintained by a designated federal agency, and is published by the National Bureau of Standards (NBS).

Five of the standards adopted and promulgated by the National Bureau of Standards pertain to geographic coding. There is a FIPS standard code for:

- States and Outlying Areas of the United States (FIPS PUB 5)
- Counties and County Equivalents of the States of the United States (FIPS PUB 6)

- Metropolitan Statistical Areas (FIPS PUB 8)

- Congressional Districts of the United States (FIPS PUB 9)

- Countries, Dependencies, and Areas of Special Sovereignty (FIPS PUB 10)[1]

2.1.1 States and Outlying Areas of the United States

This publication contains names, abbreviations and codes for representing the Fifty States, the District of Columbia and fourteen outlying areas,[2] all of which are considered first order subdivisions of the United States. The abbreviations adopted to represent the outlying areas, however, are derived from the FIPS standard representations of countries, dependencies, and areas of special sovereignty and are not necessarily the same as the abbreviations employed by the U.S. Postal Service for outlying areas.

The code assigned to each entity listed in FIPS PUB 5 is a two digit number. State codes range from 01 through 56, and they are assigned to the states in alphabetical order. Although the code

---

[1]The scope of this paper is limited to coverage of national geocoding systems and therefore the FIPS standard codes for countries, dependencies, and areas of special sovereignty are not reviewed in this section. Refer to Appendix C, International Geocoding Systems, for further information on this FIPS publication.

[2]The fourteen outlying areas included in FIPS PUB 5 are: American Samoa, Canal Zone, Canton and Enderbury Islands, Guam, Johnston Atoll, Midway Islands, Puerto Rico, Ryukyu Islands (Southern), Swan Islands, Trust Territories of the Pacific Islands, U.S. Miscellaneous Caribbean Islands, U.S. Miscellaneous Pacific Islands, Virgin Islands and Wake Island.

sequence is not gapped to allow for general expansion of the universe, codes have been reserved for possible future use in identifying American Samoa (03), Canal Zone (07), Guam (14), Puerto Rico (43), and the Virgin Islands (52). Outlying areas, including those for which a state code has been reserved, are assigned a two digit number ranging from 60 through 79. Unlike the abbreviations assigned to the outlying areas, the code representations for outlying areas are not derived from another source and are not part of another geocoding standard.

2.1.2 Counties and County Equivalents of the States of the United States

This publication provides names and codes for representing the counties of the fifty States or the county equivalents thereof. Counties are defined as the first order subdivision of each state regardless of their local designations, e.g., parish, borough, district, etc. In addition to the 3,068 entities which qualify under this definition as counties, there are 74 recognized county equivalents. These include Washington, D.C., the 42 independent cities located in Maryland, Missouri, Nevada, and Virginia, the 29 Census Divisions of Alaska, that portion of Yellowstone National Park located in Montana, and a consolidated government in Georgia.

The code representation for each of these 3,142 entities is a three digit number intended for use in combination with either the two character state abbreviation or the two digit state code to provide unique identification of each entity. In general, counties and county equivalents are sequenced alphabetically by state and assigned a serial odd number. Even numbers are reserved for possible future use in the event of universe expansion or change. County equivalents in Alaska, however, are assigned codes ranging from 010 to 290 with a ten integer gap between each assignment, and independent cities are ordered separately at the end of the county listing for a state. Codes for independent cities range from 510 to 840, generally with a ten integer separation (See Figure 2.1).

2.1.3 Standard Metropolitan Statistical Areas

This publication contains names and codes for representing Standard Metropolitan Statistical Areas (SMSAs). The general concept of an SMSA is one of an integrated socioeconomic area containing a large population nucleus. The unit was developed to meet the need for the presentation of general purpose statistics by agencies of the government. With the exception of the SMSAs located in New England all SMSAs consist of a county or group of contiguous counties at least one city or twin cities of 50,000 inhabitants or more. In addition to the county or counties containing such a population nucleus, contiguous counties are included in an SMSA if they are metropolitan in character and are socially and economically integrated with the central city. New England SMSAs consist of townships instead of entire counties. These units are defined, designated, titled, and assigned codes by the Office of Management and Budget (OMB) in conjunction with the Federal Committee on Standard Metropolitan Statistical

STATE NAME:	MARYLAND
STATE ABBREVIATION:	MD
STATE CODE:	24

CODE	COUNTY NAME
001	Allegany
003	Anne Arundel
005	Baltimore
009	Calvert
011	Caroline
013	Carroll
015	Cecil
017	Charles
019	Dorchester
021	Frederick
023	Garrett
025	Harford
027	Howard
029	Kent
031	Montgomery
033	Prince Georges
035	Queen Annes
037	St Marys
039	Somerset
041	Talbot
043	Washington
045	Wicomico
047	Worcester

CODE	INDEPENDENT CITY
510	Baltimore City

Source: The National Bureau of Standards

Figure 2.1 The Federal Information Processing Standard For Counties and County Equivalents

Areas.[1]

The SMSA names and codes adopted by NBS are those originally established by OMB. Although a majority of SMSA titles consist of the name of the largest city within each area, an SMSA title may contain up to three city names representing the three largest cities in an area or as in the case of some recently retitled SMSAs, non-city designations. The Nassau-Suffolk SMSA and the Northeast Pennsylvania SMSA are examples of SMSA titles which do not contain the name of a city. The SMSA code is a four digit number ranging from 0040 to 9320 assigned to the areas in alphabetic sequence. This FIPS publication must be updated frequently because SMSAs are a relatively unstable population of entities. New SMSAs are designated almost every year, county components of existing SMSAs are sometimes rearranged, and, occasionally, either the title or the code of existing SMSAs may change.

2.1.4 Congressional Districts of the United States

This publication contains the specifications for representing the Congressional Districts of the United States, but does not contain a full list of districts and their codes. Congressional Districts are defined as any of the districts into which a state is divided for the purpose of electing representatives to the House of

[1]Refer to Section 4.3 for specific criteria applicable to Standard Metropolitan Statistical Areas.

Representatives of the United States Congress. A two digit number is used to represent each district which is identified in the current edition of the Congressional Directory. The First Congressional District is represented as "01", etc. In states where representatives are designated "at large", the Congressional District is coded as "00". Data systems concerned with the identification of Congressional Districts in more than one state require that the district codes must be used in conjunction with a state code prefix in order to provide unique identification. The number of representatives, and therefore the number of Congressional Districts in a state, may change as a result of the decennial Census of Population and Housing, and the boundaries of the districts may be altered through legislative action by states. Therefore, where data systems require continuity of identification or relationship between prior or subsequent Congressional Districts, the code may also be used in conjunction with a two digit number representing the particular congress. The code "250292", for example, identifies the Second Congressional District of Massachusetts during the 92nd Congress.

Since the National Bureau of Standards first adopted a FIPS standard code for states and for counties of the United States (effective January 1, 1970) these representations have been incorporated in the data processing systems of many major government agencies. The Bureau of the Census, the Bureau of Economic Analysis (BEA), and the Economic Development Administration (EDA) are examples of agencies which maintain large geographically referenced data bases and are currently employing the FIPS standard representations for states and counties (See Sections 3.1, 4.4, and 4.5). In a few instances the FIPS standard has been altered to meet the needs of a particular agency. BEA, for example, does not recognize independent cities as county equivalents and occasionally combines several counties, for example, the five comprising New York City into one unit. In these cases the FIPS standard cannot be strictly applied. Several other federal agencies are unable to employ the FIPS standard for data processing because it is unsuited for coding multilevel, county cluster structures such as the geographic system maintained by the Office of Preparedness (formerly the Office of Emergency Preparedness). Nevertheless, the Office of Preparedness has added the FIPS standard representations for states and counties to its files as a cross reference (See Section 3.2).

The National Bureau of Standards intends to increase the number of geocoding standards in the FIPS Register. Working in conjunction with the American National Standards Institute and representatives from other federal agencies, NBS is evaluating several other standard geographic representations including codes for cities, towns and places, codes for point locations, and codes for continents and water areas of the world.[1]

---

[1]A complete description of these and all other programs in which FIPS is involved is contained in the "Federal Information Processing Standards Index", FIPS PUB 12.

References:

U.S. Department of Commerce, National Bureau of Standards. Federal Information Processing Standards Publications. FIPS PUB Nos. 0, 5, 6, 8, 9, 10, and 12. Washington, D.C.: U.S. Government Printing Office, 1968 - 1974.

2.2 GENERAL SERVICES ADMINISTRATION

The primary responsibility of the General Services Administration (GSA) is the management of government property. In fulfilling this responsibility GSA has, since it was established in 1949, maintained a worldwide inventory of real property owned or leased by the Federal Government. It was for use on the records of this inventory that GSA originally developed a set of geographical references for states, counties, and cities of the United States and for continents, countries, and cities of the world. As the GSA grew into what "might well be termed a conglomerate"[1] and as its activities expanded to include the management of government wide automated data processing resources, this system of geographic coding was more universally applied. It developed into an unofficial federal standard, the forerunner of a Federal Information Processing Standard.

When the FIPS standard code for states and counties of the United States became effective on January 1, 1970, the General Services Administration replaced its original state and county coding system with those codes prescribed by the National Bureau of Standards. This change did not constitute a radical overhaul of the system. The

---

[1]Office of Federal Register, National Archives and Records Service, General Services Administration, U.S. Government Organization Manual 1972/1973, (Washington, D.C.: U.S. Government Printing Office, July 1972), p. 448. The current responsibility of the GSA is defined as the management of government property and records, including construction and operation of buildings, procurement and distribution of supplies, utilization and disposal of property, transportation, traffic and communications management of government wide automated data processing resources.

original GSA state code was a two digit number which, although it differed from the FIPS state code in every instance, was readily commutable. The original GSA county codes were assigned in exactly the same manner as the FIPS standard county codes - three digit serial odd numbers matched to county names, sequenced alphabetically by state. Therefore, a majority of the existing county codes remained unchanged. The only prominent alteration of the GSA county coding structure which occurred as a result of the transfer to FIPS specifications was the addition of county equivalents in Alaska and independent cities in Maryland, Missouri, Nevada, and Virginia.

Although the original state and county codes of the General Services Administration have been superseded by the FIPS standard, the GSA city coding system, for which the FIPS Register has no official standard, remains unaltered. This set of geographical references includes codes for cities, towns, villages, places, and military installations. In fact, any location which qualifies as a "populated area having a recognized entity or geographical boundary"[1] and which is listed in the Rand McNally Commercial Atlas and Marketing Guide is is eligible to receive a GSA city code. All of the accepted entities are sequenced alphabetically by state and assigned a four digit number (See Figure 2.2). When these codes were first established, intervals of ten integers were maintained between code numbers. Gapping

---

[1]Office of Finance, General Services Administration, Worldwide Geographical Location Codes, (Washington, D.C.: U.S. Government Printing Office, February 1972), p. I.

NAME	CODES		
	STATE	CITY	COUNTY
Maryland	24		
Aberdeen		0010	025
Aberdeen Prov Grnd		0015	025
Accident		0020	023
Andrews AFB		0025	033
Annapolis		0030	003
Arnold		0032	003
Army Chem Center		0035	025
Army Med Center		0036	031
Avenel		0038	031
Bainbridge		0040	015
Baltimore		0050	510
Barclay		0060	035
Barnesville		0070	031
Barton		0080	001
Beantown		0085	031
Bel Air		0090	025
Beltsville		0100	023
Bengies		0103	005
Berlin		0110	047
Berwyn Heights		0120	033
Bethesda		0130	031
Betterton		0140	029
Bladensburg		0150	033
Boonsboro		0160	043
Bowie		0170	033

Source: General Services Administration

Figure 2.2 The Geographical Location Codes Of The General Services Admininstration the code sequence in this manner allowed for the preservation of alphabetically ordered place names even as the universe expanded over the years. Gradually, however, this margin for expansion is being exhausted and the GSA city coding system is currently nearing saturation at several points. 33,000 locations are included in this system.

The structure of the GSA city code resembles that of other major, general purpose place identifiers. Both IBM and Dun and Bradstreet employ the same basic two digit state and four digit place code configuration to provide unique location identification. It is a simple two level hierarchy with individual entities ordered alphabetically at each level. However, the GSA coding system does manifest a degree of specialization not shared by either the IBM or Dun and Bradstreet city coding systems. The one characteristic of the system which reflects the extent of this specialization is the uneven assignments of GSA city codes. All U.S. military installations automatically receive a GSA city code while other code assignments must be specifically requested of, and then approved, by the Office of Finance at GSA.

The effect of this code assignment procedure is most exaggerated at the international scale and accounts for the fact that in the Soviet Union only Leningrad and Moscow have received GSA city codes while there are approximately 125 city codes within the Republic of Vietnam.[1]

---

[1]Refer to Appendix C for a description of the international geocodes of the General Services Administration.

Allocation of GSA city codes in the United States is far less erratic. However, the effect of a government property oriented code assignment procedure is still discernible. There is, for example, an unusually large number of GSA city code assignments in Alaska and an unusually small number of GSA city code assignments in Maine.[1]

Generically, the set of geographic codes maintained by the General Services Administration is a member of a family of geocoding systems which includes the systems employed by IBM and Dun and Bradstreet. These systems have three basic elements: a two digit state code, a three digit county code, and a four digit place code. Although the specific numbers assigned to each set of geographic units varies from system to system, the basic coding logic is identical and at the state and county level even the universe of entities coded is relatively standard. It is at the place level that these systems differ significantly according to the applications for which the system is designed and therefore the universe coded.

References:

General Services Administration. Office of Finance. Worldwide Geographical Location Codes. Washington, D.C.: U.S. Government Printing Office, 1972.

---

[1]For four sample states, the General Services Administration's city coding pattern compares with the Dun and Bradstreet city coding pattern as follows:

	Arizona	Alaska	New York	Maine
Number of GSA Codes	290	434	1,648	352
Number of DUN Codes	325	254	2,467	563

2.3 INTERNATIONAL BUSINESS MACHINES

During the 1950's, as the number of IBM computers in use by private industry and public agencies began to proliferate, the International Business Machines Corporation made available to the user of its equipment a numerical code for states, counties and cities in the United States. These geocodes were published and distributed by IBM in order to assist users in establishing the machine readable geographic references necessary for direct mail operations, market research, sales analyses, territory assignment and transportation routing.

In the original geocoding manual published by IBM three digits were allotted for the coding of counties, and three digits were allotted for the coding of cities. This was an ample number of digits at the time, permitting unique location identification with sufficient space for filling in new additions so that an alphanumeric sequence could be maintained. Subsequent to the 1960 Census of Housing and Population, however, the three digit city code became saturated as the number of urban places having a population of 2,500 or more increased substantially. Consequently, IBM established a new, four digit, city code and published these in a manual of geographic codes. Thus, post 1960 IBM geocodes are structured in exactly the same manner as the codes maintained by the General Services Administration (See Section 2.2) and Dun and Bradstreet (See Section 2.4).

The full IBM numerical code for states, counties and cities is an eleven digit number. The first two digits indicate state, ranging from 01 for Alabama to 51 for Hawaii (the District of Columbia is treated as a state and has been assigned state code 08). The next three digits identify a county. The counties of each state are sequenced alphabetically and then assigned an odd number beginning at 001. The Population Group Code occupies the following two digit positions. Population figures are divided into seven major groups and then further subdivided into three of four minor groups as follows:

POPULATION GROUP	CODE	POPULATION GROUP	CODE
1,000,000 or more	11		
750,000 to 999,999	12		
500,000 to 749,999	13	40,000 to 49,999	51
		35,000 to 39,999	52
400,000 to 499,999	21	30,000 to 34,999	53
300,000 to 399,999	22	25,000 to 29,999	54
250,000 to 299,999	23		
200,000 to 249,999	24	20,000 to 24,999	61
		15,000 to 19,999	62
175,000 to 199,999	31	10,000 to 14,999	63
150,000 to 174,999	32	5,000 to 9,999	64
125,000 to 149,999	33		
100,000 to 124,999	34	4,000 to 4,999	71
		3,500 to 3,999	72
80,000 to 99,999	41	3,000 to 3,499	73
70,000 to 79,999	42	2,500 to 2,999	74
60,000 to 69,999	43		
50,000 to 59,999	44		

The last four digits of the IBM code designate a city. All cities in each state are sequenced alphabetically and then assigned a serial number at intervals of ten beginning at 0010.

The IBM geocode manual contains a unique state/county/city code for every place having a population of 2,500 or more based on the 1960 Census of Housing and Population. Thus, the universe of 5,463 place (city) entities coded by IBM is not specialized. The criteria for inclusion in this geocoding system is population rather than special logistical significance to a particular user, and the population criteria applied conform to those dictated by the Bureau of the Census. Very few of the United States military installations, for example, which are significant to and coded by the Department of Defense or even the General Services Administration are included in the IBM coding manual.

In addition to the eleven digit number assigned as unique geo-identification to each place listed in the IBM code manual, every unincorporated place is assigned the suffix letter "U", and all places which had been assigned codes prior to 1960 are identified by the old three digit city code as well as the current four digit designation (See Figure 2.3). If a city extends into two counties, and each sector has 2,500 or more people, then both counties are shown beside the city name; if one sector has a population under 2,500, the county in which that sector is located is omitted. In either case, the city population group code is based on the combined population of both sectors.

	STATE	COUNTY	OLD CITY CODE	POP GROUP	NEW CITY CODE
Aberdeen	19	025	004	64	0010
Annapolis	19	003	010	61	0020
Arbutus	19	005		61	0030 U
Baltimore	19	007	020	12	0040
Bel Air	19	025	024	71	0050
Bethesda	19	031		44	0060 U
Bladensburg	19	033	027	73	0070
Brentwood	19	033	028	72	0080
Brunswick	19	021	030	72	0090
Cambridge	19	019	040	63	0100
Capitol Heights	19	033	041	73	0110
Carrollton	19	013		73	0120
Catonsville	19	005		52	0130 U
Chestertown	19	029	050	72	0140
Cheverly	19	033	051	64	0150
Cockeysville	19	005		74	0160 U
College Park	19	033	055	62	0170
Crisfield	19	039	060	72	0180
Cumberland	19	001	070	53	0190
District Heights	19	033		64	0200
Dundalk	19	005		41	0210 U
Easton	19	041	100	64	0220
Elkton	19	015	110	64	0230
Essex	19	005		52	0240 U
Forest Heights	19	033		72	0250

Source: International Business Machines Corporation

Figure 2.3 IBM City Codes

In view of the code standardization activities of FIPS (See Section 2.5), IBM has decided not to republish the manual of geographic codes, and is referring users to FIPS publications for state and county codes. However, until a place code standard is adopted and authorized by FIPS, IBM will continue to make available to users the current edition of the coding manual which has not and will not be updated to reflect the 1970 population data from the Bureau of the Census.

References:

International Business Machines Corporation. Data Processing Division. Numerical Code for States, Counties and Cities of the United States. GC 20-8073-0. White Plains, New York: IBM Technical Publications Department, 1961.

International Business Machines Corporation. "IBM Installation Newsletter." Issue No. 70-22. White Plains, New York: IBM Technical Publication Department, November 6, 1970.

2.4 DUN AND BRADSTREET

Since it was founded in 1841 as "The Mercantile Agency," Dun and Bradstreet, Inc. has earned the reputation of "a mysterious warehouse of statistics."[1] As an organization which compiles and interprets vast amounts of credit data, Dun and Bradstreet has, in fact, developed one of the largest business information systems in the world, and this system includes a set of geographic codes for nations of the world, and codes for states, counties, Standard Metropolitan Statistical Areas (SMSAs), and cities of the United States.[2]

The Dun and Bradstreet national code is a three digit number randomly assigned to the nations of the world. The state code is a two digit number assigned to the 50 United States and the District of Columbia in alphabetic sequence beginning with 01 and enumerated at intervals of two (i.e., Alabama is 01, Alaska is 03, Arizona is 05, etc.). The only even number assigned as a state code is for the District of Columbia (state code 16). The county code is a three digit number assigned to counties in alphabetic sequence by state (See

---

[1] J. Wilson Newman, "Dun and Bradstreet: Established in 1841 for the Promotion and Protection of Trade," A Newman Society Address, October 25, 1956.

[2] Dun and Bradstreet files also cover approximately 8,000 locations in Canada. The geocodes are assigned to the Canadian geopoltical units equivalent to states, counties and cities in the same manner as the codes are assigned for locations in the United States. They are not identical to the Standard Geographic Classification Codes established by Statistics Canada (Appendix B). However, Dun and Bradstreet does employ the six digit Canadian Postal Zone code established by the Canadian Post Office Department.

Figure 2.4). These code numbers are generated by a semi-random enumeration process designed so that the county codes are always of increasing numeric magnitude by alphabetic sequence within a state.[1] The SMSA code is a three digit number assigned to each area in alphabetic sequence by SMSA title beginning at 010. These SMSA codes are gapped by three digits in order to allow for the inclusion of new SMSAs within the proper alphabetic sequence.

The city code is a four digit numeric code assigned to cities in alphabetical order within a state. These city code numbers are generated and assigned in the same manner as the Dun and Bradstreet county codes.

The complete Dun and Bradstreet geographic code file which includes approximately 44,000 locations in the United States contains the following items: A 13 character alphabetic abbreviation of the city name; a 3 character alphabetic abbreviation of the state name; the two digit Dun and Bradstreet state code; the 4 digit Dun and Bradstreet city code; the 3 digit Dun and Bradstreet county code; a

---

[1]The first position of the county code is assigned so that roughly the same number of county codes within a state begin with the same numeric character. In Montana where there are 57 counties, for example, the first digit is assigned to counties in blocks of 5 or 6. In New Hampshire where there are only 10 counties, each county code begins with a different number. The second and third digits of this county code are various regular or irregular number series. For example, the last digit number for counties in Texas is one of the number series 3,6,9,2,5,8,1,4,7,0 which is repeated to the end of the listing of counties in that order. This procedure varies from state to state. The number series may be irregularly repeated, it may be only one number (e.g., the county codes for South Dakota all end in the number 4) or it may be an alternating pattern of only two numbers as it is for counties in Colorado.

STATE MONTANA (51)

County	Code	County	Code
Beaverhead	024	McCone	500
Big Horn	041	Meagher	517
Blaine	058	Mineral	534
Broadwater	075	Missoula	551
Carbon	092	Musselshell	568
Carter	109	Park	585
Cascade	126	Petroleum	602
Chouteau	143	Phillips	619
Custer	160	Pondera	636
Daniels	177	Powder River	653
Dawson	194	Powell	670
Deer Lodge	211	Prairie	687
Fallon	228	Ravalli	704
Fergus	245	Richland	721
Flathead	262	Roosevelt	738
Gallatin	279	Rosebud	755
Garfield	296	Sanders	772
Glacier	313	Sheridan	789
Golden Vallen	330	Silver Bow	806
Granite	347	Stillwater	823
Hill	364	Sweet Grass	840
Jefferson	381	Teton	857
Judith Basin	398	Toole	874
Lake	415	Treasure	891
Lewis and Clark	432	Valley	908
Liberty	449	Wheatland	925
Lincoln	466	Wibaux	942
Madison	483	Yellowstone	959
		Yellowstone N P	976

Source: Dun and Bradstreet, Inc.

Figure 2.4 Dun and Bradstreet County Codes one digit geographic check code; the 5 digit United States ZIP Code; the 3 digit Dun and Bradstreet reporting office designation; and a cross reference indication for cities for which no ZIP Code is listed. A sample printout of this file is reproduced in Figure 2.5 and a sample of the data which is associated with these codes in the Dun's Market Identifiers (DMI) magentic tapes is provided in Figure 2.6.

While the number assignment process for Dun and Bradstreet county and city codes is relatively complex and unique among the geocoding systems reviewed in this report, the basic code structure is identical with that of GSA (See Section 2.2) and IBM (See Section 2.3) geocodes which also identify numerically a place within a county within a state. Thus, all three of these systems have the same degree of flexibility and the same specificity of geoidentification.

The only significant differences among these systems is in the nature and size of the universe coded. The universe of entities coded by GSA is much larger than the universe of places assigned codes by IBM, however the IBM and the Dun and Bradstreet systems are more alike in that these two sets of coded places are not specialized. They reflect general population density rather than the distribution of government property or some other code assignment bias.

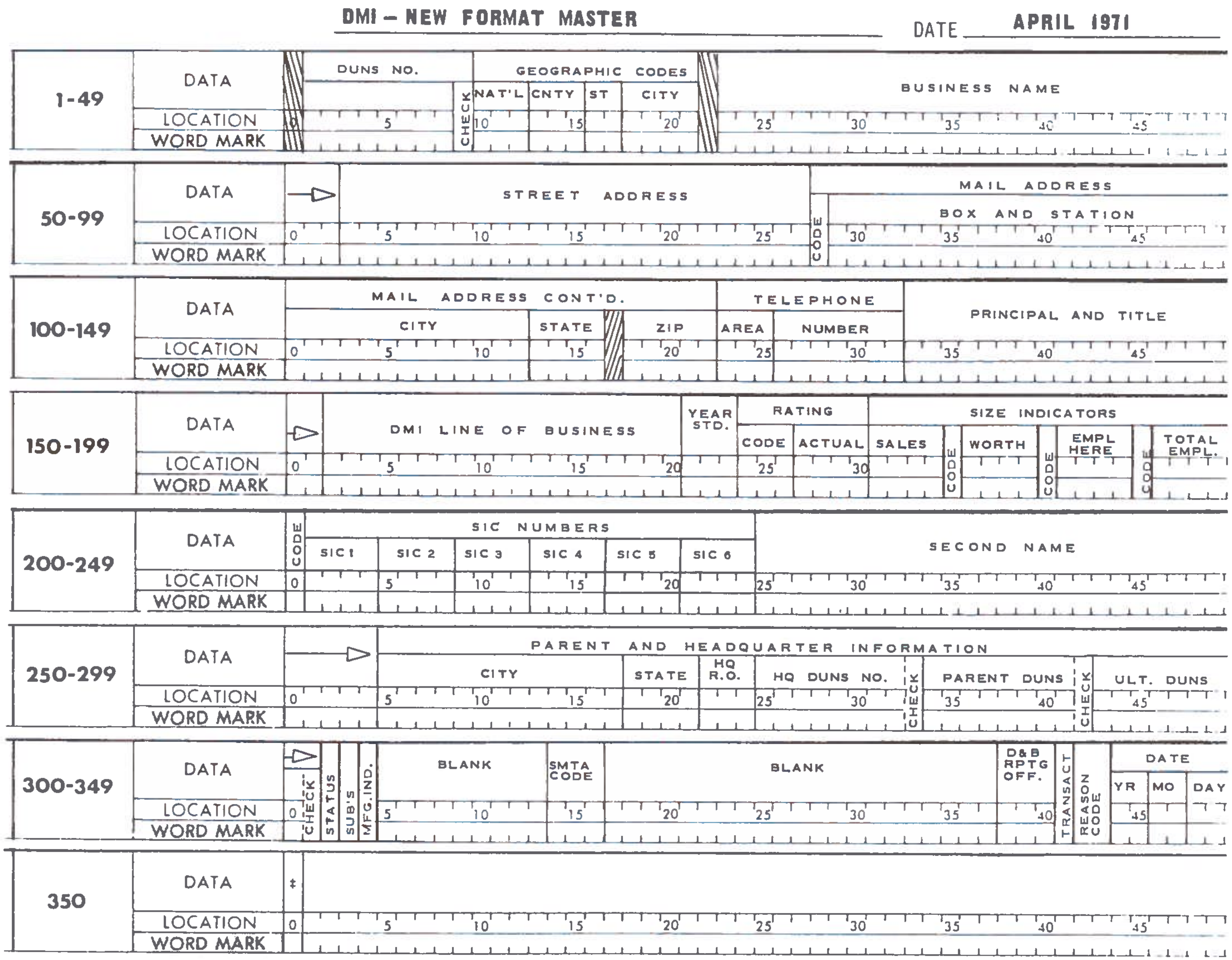
References:

Dun and Bradstreet Inc. Marketing Services Division. County Code Book. Standard Metropolitan Statistical Areas Code Book. and City Code Book. New York: Dun and Bradstreet, Inc., 1971.

TOWN NAME	STATE ABBREVIATION	STATE CODE	CITY CODE	COUNTY CODE	ZIP CODE	TOWN TO SEE IF NO ZIP CODE LISTED
Chancellor	Ala	01	1663	458	36316	
Chapmen	Ala	01	1674	122	36015	
Chase	Ala	01	1685	654	35745	
Chastang	Ala	01	1686	710	36517	
Chatom	Ala	01	1696	934	36518	
Chavies	Ala	01	1709	374		See Rainsville Ala
Chelsea	Ala	01	1722	850	35043	
Chepultepec	Ala	01	1735	094		See Allgood Ala
Cherokee	Ala	01	1748	262	35616	
Chestnut	Ala	01	1752	724	36425	
Chickasaw	Ala	01	1762	710	36611	
Childersburg	Ala	01	1773	878	35044	
China Grove	Ala	01	1786	794		See Linwood Ala
Choccolocco	Ala	01	1799	136	36254	
Choctaw	Ala	01	1812	192	36905	
Chrysler	Ala	01	1818	724	36520	
Chunchula	Ala	01	1825	710	36521	
Citronelle	Ala	01	1839	710	36522	
Claiborne	Ala	01	1850	724	36434	
Clairmont	Ala	01	1861	878		See Talladega Ala

Source: Dun and Bradstreet, Inc.

Figure 2.5 Dun and Bradstreet Geographic File



Source: Dun and Bradstreet, Inc.

Figure 2.6 Dun's Market Identifiers

Dun and Bradstreet, Inc. Marketing Services Division. Magnetic Tape Description. New York: Dun and Bradstreet, Inc., 1971.

2.5 AMERICAN NATIONAL STANDARDS INSTITUTE

The American National Standards Institute (ANSI) is a nonprofit organization comprised of trade associations, private companies, professional societies and government agencies which serves as the national clearinghouse and coordinating agency for voluntary standards in the United States. Since its inception in 1918, over 2800 American National Standards have been authorized and published.[1] Of these, more than one third were submitted by individual participating organizations that had developed standards through their own procedures, and supplied evidence of consensus in support of such standards. The remaining American National Standards result from the work carried on by the complex of ANSI committees. The American National Standards Institute works in close cooperation with the National Bureau of Standards (NBS) and is the United States member body of the International Organization of Standardization. (ISO).

ANSI Committee X3L84 which is one of the 58 subcommittees designated under the reorganization of the Standards Committee for Computers and Information Processing (X3) in 1969 is specifically charged with the development of data standards for representing geographical and geopolitical entities. Currently, X3L84 activity includes study

---

[1]ANSI was originally organized as the American Engineering Standards Committee (AESC) in 1918 by five engineering societies. In 1928, AESC was reorganized as the American Standards Association (ASA) and then renamed the United States of America Standards Institute in 1966. The present name, American National Standards Institute, was adopted in 1969.

projects on:

- Codes for cities, towns, and places of the United States
- Coordinate point locations in the United States
- A standard representation for mailing and shipping addresses
- Codes for countries of the world and their national subdivisions
- Codes for foreign cities
- Codes for continents and water areas

Previously this committee formally approved and adopted the FIPS codes for designating states and counties of the United States (See Section 2.1).

2.5.1 Place Codes

In recent years the place code study project has been one of the most active concerns of ANSI X3L84 and work on this standard is nearing completion. The place code developed by the study project is a five digit number assigned in alphabetic sequence to each named place within a state. Number gapping between assigned codes allows for the addition of names in proper alphabetic sequence. Within each state the code assignment begins with 00100 and ends with a code no higher than 89999. This will provide unused numbers at the end of the code sequence for the special needs of individual users. The ANSI Place Code may be represented as 01-10510, in which 01 is the standard FIPS state code and 10510 is the unique code for a place within that state. Alternatively the form AL-10510 may be used, employing the alpha designation of the state code rather than the numeric standard. This two digit state/five digit place code combination is a major deviation from the place code structure (two digit state/three digit county/four digit place) employed in the older systems maintained by Dun and Bradstreet, the General Services Administration, and IBM (See Sections 2.2 through 2.4). While the ANSI Place Code is two digits shorter than the other codes and therefore requires less space on a file record, it is a less flexible code in that the entities coded in this system cannot be sorted by county.

The ANSI Place Code covers approximately 130,000 named populated places and related entities. The list of places is based primarily on an amalgamation of entities contained in the Rand McNally Commercial Atlas and the Standard Point Location Code (See Section 3.6). Because of the varied character of the named places included in this list, no single definition or criterion can be stated as a qualification for inclusion in the ANSI Place Code System. Instead, ANSI has established eighteen categories of entities and an entity fulfilling any one of these definitions may be included in the system. In general the ANSI Place Code covers named populated cities, towns, villages, and similar populated communities, whether or not they are organized as municipal governments; and several categories of named entities that are similar to these in one or more important respects, including: 1) scattered rural communities; 2) important military and naval installations; 3) townships in the states where such units have governmental powers; 4) Indian reservations; 5) national and state parks; and 6) named places that form parts of other places as defined. Also included are named places with no permanent residents having significance for transportation, industrial or commercial purposes, such as unpopulated railroad points, airports, and shopping centers (See Figure 2.7). Thus, the universe of places coded in the ANSI file is much larger than the universe of any other place code system. It includes both the generalized universe of populated places and the places of more specialized significance.

> The common characteristic of these varied types of entity is that all of them are recognized as named places by a significant segment of the public. In other words, for each category of entity there is an important group of users who would expect to find it included in a standard place code. The reverse is also true--there are users who will not wish certain categories of entity included for their special purposes. Therefore the categorization of entity by type is a very important adjunct to use of the code, since it permits users to select those types of entity which fit their own particular conception of 'populated places.' [1]

In conjunction with the basic place code there are three additional code features in the ANSI system: 1) a class designator which is a single letter mnemonic code serving to class the individual entity into one of the eighteen place categories; 2) a nineteenth class which is designated X and is comprised of the places with alternate and former names; and 3) an "inclusion" indicator which is assigned to places that meet the criteria for individual codes but can also be parts of other places uniquely coded.

---

[1]Richard L. Forstall, Chairman, Place Codes Subgroup, ANSI Task Group X3L84, "The American National Standard Code for Named Populated Places and Related Entities," The National Geocoding Conference, Proceedings, (Washington, D.C.: U.S. Department of Transportation, 1972), p. 111.3.

CATEGORY	DESIGNATION	COUNT (THOUSANDS)	COMPLETENESS
Incorporated Places	C	19.0	*
Unincorporated Places	U	42.7	**
Seasonally Populated Places	K	1.0	***
Rural Communities	R	22.0	***
Military Installations Wholly or Largely Within Incorporated Places	L	0.1	**
Military Installations Wholly or Largely Outside Incorporated Places	M	0.4	**
Unpopulated Industrial Points	F	0.3	**
Shopping Centers (not parts of other places)	G	0.2	**
Unpopulated Transport Points	S	9.2	*
Airports	A	0.5	**
Indian Reservations	D	0.2	*
National and State Parks	N	1.0	*
Places Which are Parts of Incorporated Places	P	12.0	***
Places Which are Parts of Populated Unincorporated Places	Q	3.3	***
Urban Townships	T	0.3	*
Townships Associated With a Locality of Identical Name	V	5.6	*
Townships Not Associated With a Locality of Identical Name	W	13.5	*
Townships Wholly Comprised Within an Incorporated Place	Y	0.5	*
		131.9	TOTAL

* Exhaustive: Original file should include every entity meeting the definition, and future maintenance should add any new or corrected entities that meet the definition	** Complete: The criteria for inclusion in the original file are intended to include all the more important entities of the class; additional entities should be added if (a) they meet the criteria and (b) there is a demonstrated need for users of the standard to recognize them separately	*** Not Complete: Application of the criteria should result in including most or all of the significant entities; additional entities can be added that meet the criteria if there is a demonstrated need for users of the standard to recognize them separately

Source: The American National Standards Institute

Figure 2.7 ANSI Categories for Named Populated Places and Related Entities

2.5.2 Point Codes

Another important activity of X3L84 has been the evaluation of codes for point locations in the United States. A task force was assigned to "study available systems of coding geographic point locations and to recommend a standard which is practical for use and is currently available."[1] After careful deliberation, this task force recommended latitude and longitude (See Section 5.1), the Universal Transverse Mercator Grid (See Section 5.2), and the State Plane Coordinate System (See Section 5.5) as the three most suitable systems for coordinate point location standards.

The ANSI task force on point location codes notes that all three of the approved coordinate systems meet the precision requirements of a point location code.[2] All are mathematically compatible and interconvertible. However, each system has specific disadvantages. Lati-

---

[1]Henry D. Walker, "Report on Work of the Subcommittee for Codes for Point Locations in United States," a draft report presented to ANSI Task Group X3L84, September 15, 1971.

[2]The most rigorous requirement is the location of a survey point, measured to the nearest 100th of a foot.

tude and longitude is the most universally accepted, but difficult to work with and altitude conversion must be taken into consideration. The Universal Transverse Mercator Grid (UTM) is measured in metric units and there are large distortions in certain parts of the UTM zones. The State Plane Coordinate System encompasses 127 zones with separate origins and is based on two different map projections, making conversion from zone to zone an arduous process. Therefore, the ANSI committee on point location codes concluded that latitude and longitude is the best system for measuring great distances; the Universal Transverse Mercator system is most suitable for coding locations within the same UTM zone; and the State Plane Coordinate System is most effective for local use. X3L84 will continue work on this study project and an enlarged subcommittee will be appointed to draft a code for the interchange of geographic point location information based on the current recommendations of that subcommittee.

2.5.3 Representations for Mailing and Shipping Addresses

The current status of ANSI's study project on the standard representation for mailing and shipping addresses is somewhat ambiguous. In a memo to this study group, a representative of the United States Postal Service (USPS, also a member of the study group) suggested:

> A practical machine-oriented but human-readable address standard would have wide application not merely in the communications and transportation industries, but in all information systems that carry address as part of the data base...Considering the number and volume of existing address files, however, a standardized orthography will

> not be widely adopted until it is accompanied by a 'standardizer,' i.e., a computer software package that will accept non-standard addresses inputted from magnetic tape files and convert them to standard spelling and format, outputted into a new, 'standard' file...at economic throughput rates... Therefore, we propose that Project 95 (representation of mailing and shipping addresses) be temporily suspended until the ongoing USPS effort to develop a universal standardizer reaches the point where we have something concrete to demonstrate to the committee.[1]

While this position was generally agreed upon, members of the committee felt that some work could proceed with the identification of address elements and even with the specification of their structure while a standardizer was being developed. Therefore, the current activity of this ANSI task force has been limited to a compilation of a list of address element in priority order.

The progress of ANSI task groups concerned with geocoding standards is often very slow and arduous. This is due in part to the time constraints imposed upon committee members who have many other commitments associated with their normal work and due to the difficulty involved in reaching a consensus among rather diverse interests on often controversial items.

References:

U.S. Department of Commerce. National Bureau of Standards. Federal Information Processing Standards Index. FIPS PUB 12. Washington, D.C.: Government Printing Office, 1970.

---

[1]Memo from Mr. George W. Proffitt of the United States Postal Service appended to the Meeting Minutes of ANSI Task Group X3L84, February 21, 1974.

Forstall, Richard L. "The American National Standard Code for Named Populated Places and Related Entities," The National Geocoding Conference Proceedings. Washington, D.C.: U.S. Department of Transportation, 1972.

Walker, Henry D. "Report on Work of the Subcommittee for Codes for Point Locations in the United States." A draft report presented to ANSI Task Group X3L84. September 15, 1971.

ANSI Task Group X3L84. "Meeting Minutes," February 21, 1974.

3. GEOPOLITICAL AND SPECIALIZED LOCATION REFERENCE CODES

3.1 BUREAU OF THE CENSUS

Since the First Decennial Census of 1790 when population totals for all 292 counties and most larger cities of the sixteen United States were published in fifty-six pages until 1900, Census data were tabulated almost exclusively for geopolitical areas with legally defined boundaries. However, after the Bureau of the Census was established as a permanent agency in 1902, statistically, administratively, and functionally defined areas also became a part of the Census geographic system. Currently, the Bureau collects and tabulates data for over forty different spatial units included in one or more of the eight continuous national Censuses (See Figure 3.1).[1] The geocoding system maintained by the Bureau in order to tabulate data for these various political, statistical, administrative and functional areas is the most comprehensive national system of geographic codes.

---

[1]The eight continuous national Censuses are: the Census of Housing and Population, the Census of Manufactures, the Census of Agriculture, the Census of Mineral Industry, the Census of Governments, the Census of Business, the Census of Construction and the Census of Transportation.

	Population	Manufactures	Agriculture	Mineral Industries	Governments	Business	Housing	Construction	Transportation
1790	●								
1800	●								
1810	●	●							
1820		●							
1830									
1840	●	●	●	●					
1850	●	●	●	●	●				
1860	●	●	●	●	●				
1870	●	●	●	●	●				
1880	●	●	●	●	●				
1890	●	●	●	●	●				
1900	●	●	●						
1902				●					
1905		●			●				
1910	●	●	●	●					
1912					●				
1915		●							
1920	●	●	●	●					
1921		●							
1922					●				
1923		●							
1925		●	●						
1927		●							
1929		●							
1930	●		●	●		●		●	
1931		●							
1932					●				
1933		●				●			
1935		●	●			●		●	
1937		●							
1939		●							
1940	●		●	●		●	●	●	
1942					●				
1945			●						
1947		●							
1948						●			
1950	●		●				●		
1954		●		●		●			
1955			●						
1957					●				
1958		●		●		●			
1960	●		●				●		
1962					●				
1963		●		●		●			●
1965			●						
1967		●		●	●	●		●	●
1970	●		●				●		
1971						●			
1972				●	●			●	●

Source: A. Ross Eckler, The Bureau of the Census

Figure 3.1 Major National Censuses

3.1.1 Definition of Geographic Areas[1]

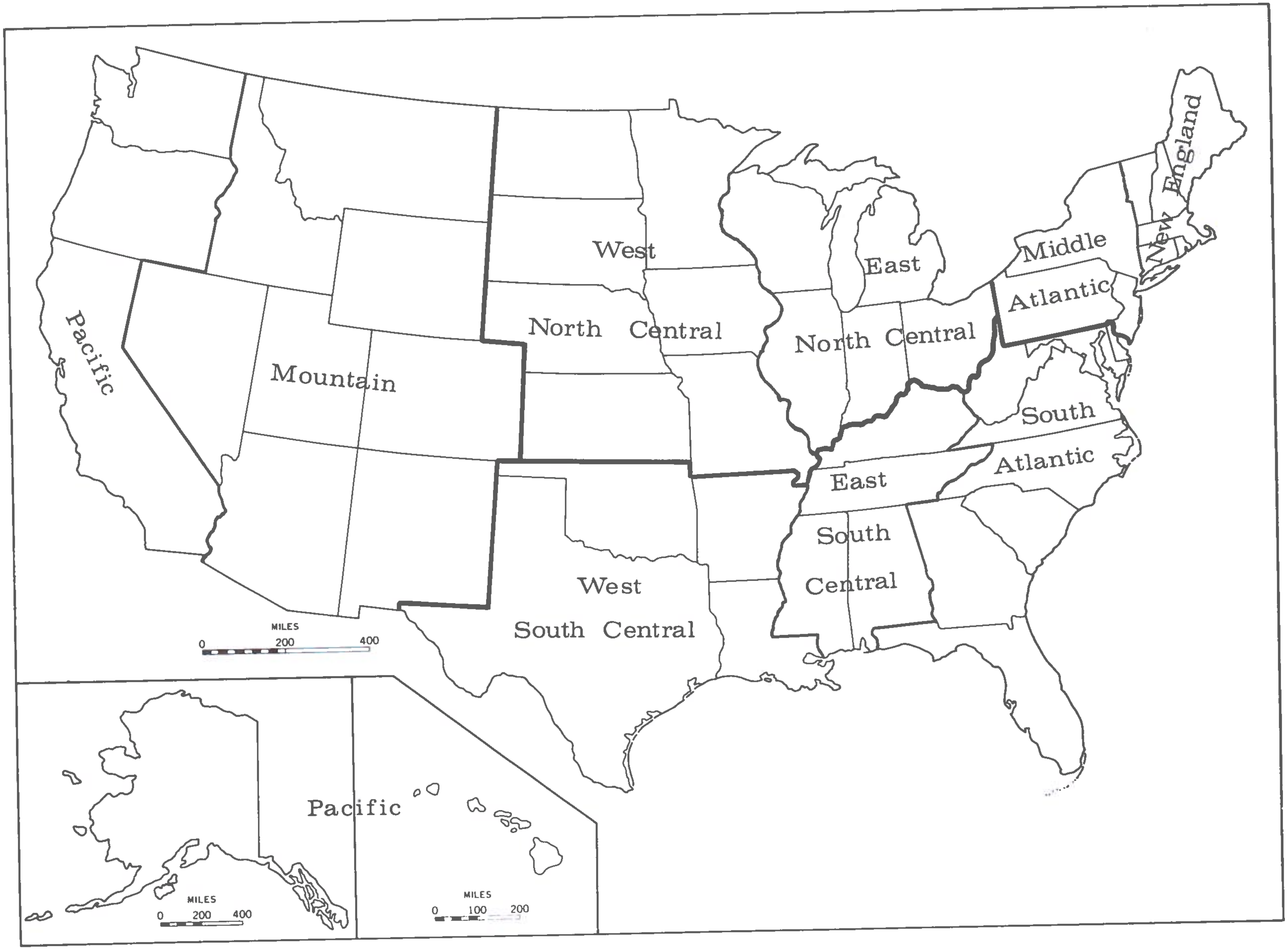
1. Regions. Regions are large, geographically contiguous state aggregates (with the exception of one region including Alaska and Hawaii) which are the first order Census subdivisions of the United States. There are four regions: Northeast, North Central, South, and West (See Figure 3.2). Although Census data are tabulated by these units there are no codes or abbreviations for the regions.

2. Geographic Divisions. Geographic divisions are areas composed of contiguous states with the extracontinental states of Alaska and Hawaii included in the Pacific Division. There are nine geographic divisions, two or more of which constitute a region (See Figure 3.2). As of 1970, the one digit division code has been discontinued.

3. States. The codes assigned to these first order political subdivisions of the United States are two digit numbers assigned in sequence to the fifty states and the District of Columbia listed alphabetically. These state codes conform to the FIPS standard designations (See Section 2.1). Prior to 1970, the Census state code was a two digit number composed of the one digit division number and an

---

[1] This section is extracted from the following sources: U.S. Department of Commerce, Bureau of the Census, *1970 Census Users' Guide* Part 1 (Washington, D.C.: U.S. Government Printing Office, 1970). U.S. Department of Commerce, Bureau of the Census, "Introduction to Small Area Geographic Subdivisions for Which the U.S. Bureau of the Census Collects and Tabulates Data," prepared by Valerie McFarland and Ann D. Casey of the Data User Services Office (Washington, D.C.: U.S. Bureau of the Census, 1972).



Source: Bureau of the Census

Figure 3.2 Census Regions and Divisions

additional digit designating a state within a division (See Figure 3.3).

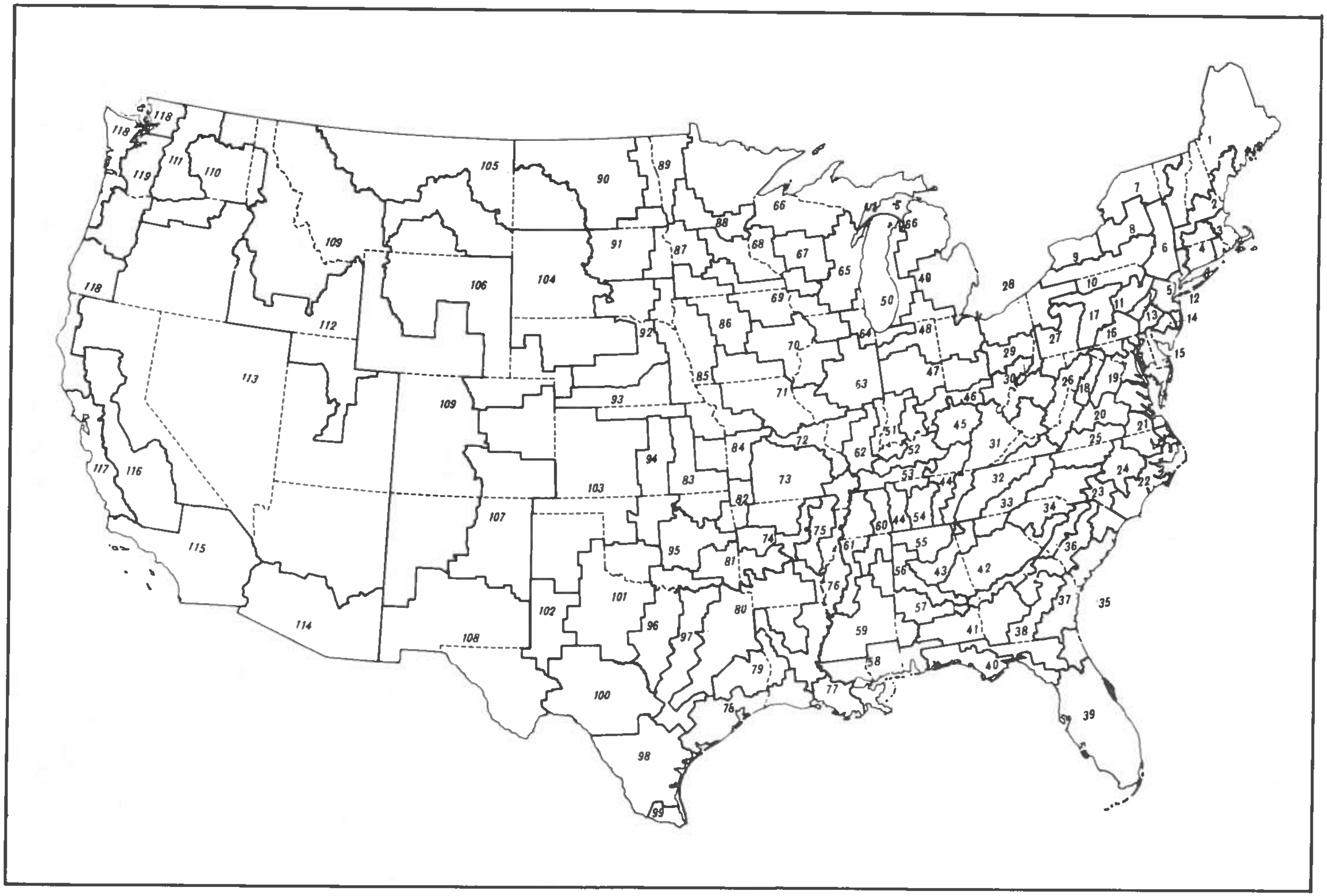
4. Economic Subregions (ESRs). The 121 economic subregions are interstate clusters of areas which are closely related in terms of their economic and social characteristics. They are aggregates of state economic areas (SEAs) which are, in turn, county aggregates. A three digit numeric code is assigned to each ESR (See Figure 3.4).

5. State Economic Areas (SEAs). State economic areas consist of a single county or a group of counties within a state which are relatively homogeneous with respect to economic and social characteristics. The delineation of these 509 SEAs involved the compilation of data on population, manufacturing, transportation, topography, climate, natural resources and agriculture. As many as sixty-four statistical indices were employed simultaneously to determine areas of socioeconomic homogeneity. Standard Metropolitan Statistical Areas (See Item 13) of one million or more population automatically constitute whole SEAs except in New England where SEAs include the non-SMSA remainder of counties containing an SMSA (in New England SMSAs are groups of townships rather than whole counties) and in cases where SMSA boundaries cross state lines, thereby necessitating the designation of two separate SEAs. State economic areas are identified by a two digit number code (non-SMSA SEAs) or a one digit alphabetic code

NORTHEAST		NORTH CENTRAL	
NEW ENGLAND	MIDDLE ATLANTIC	EAST NORTH CENTRAL	WEST NORTH CENTRAL
11 Maine 12 New Hampshire 13 Vermont 14 Massachusetts 15 Rhode Island 16 Connecticut	21 New York 22 New Jersey 23 Pennsylvania	31 Ohio 32 Indiana 33 Illinois 34 Michigan 35 Wisconsin	41 Minnesota 42 Iowa 43 Missouri 44 North Dakota 45 South Dakota 46 Nebraska 47 Kansas

SOUTH			WEST	
SOUTH ATLANTIC	EAST SOUTH CENTRAL	WEST SOUTH CENTRAL	MOUNTAIN	PACIFIC
51 Delaware 52 Maryland 53 D. C. 54 Virginia 55 West Virginia 56 North Carolina 57 South Carolina 58 Georgia 59 Florida	61 Kentucky 62 Tennessee 63 Alabama 64 Mississippi	71 Arkansas 72 Louisiana 73 Oklahoma 74 Texas	81 Montana 82 Idaho 83 Wyoming 84 Colorado 85 New Mexico 86 Arizona 87 Utah 88 Nevada	91 Washington 92 Oregon 93 California 01 Alaska 02 Hawaii

Source: Bureau of the Census



Source: Bogue and Beale, Economic Areas of the United States

Figure 3.4 Economic Subregions

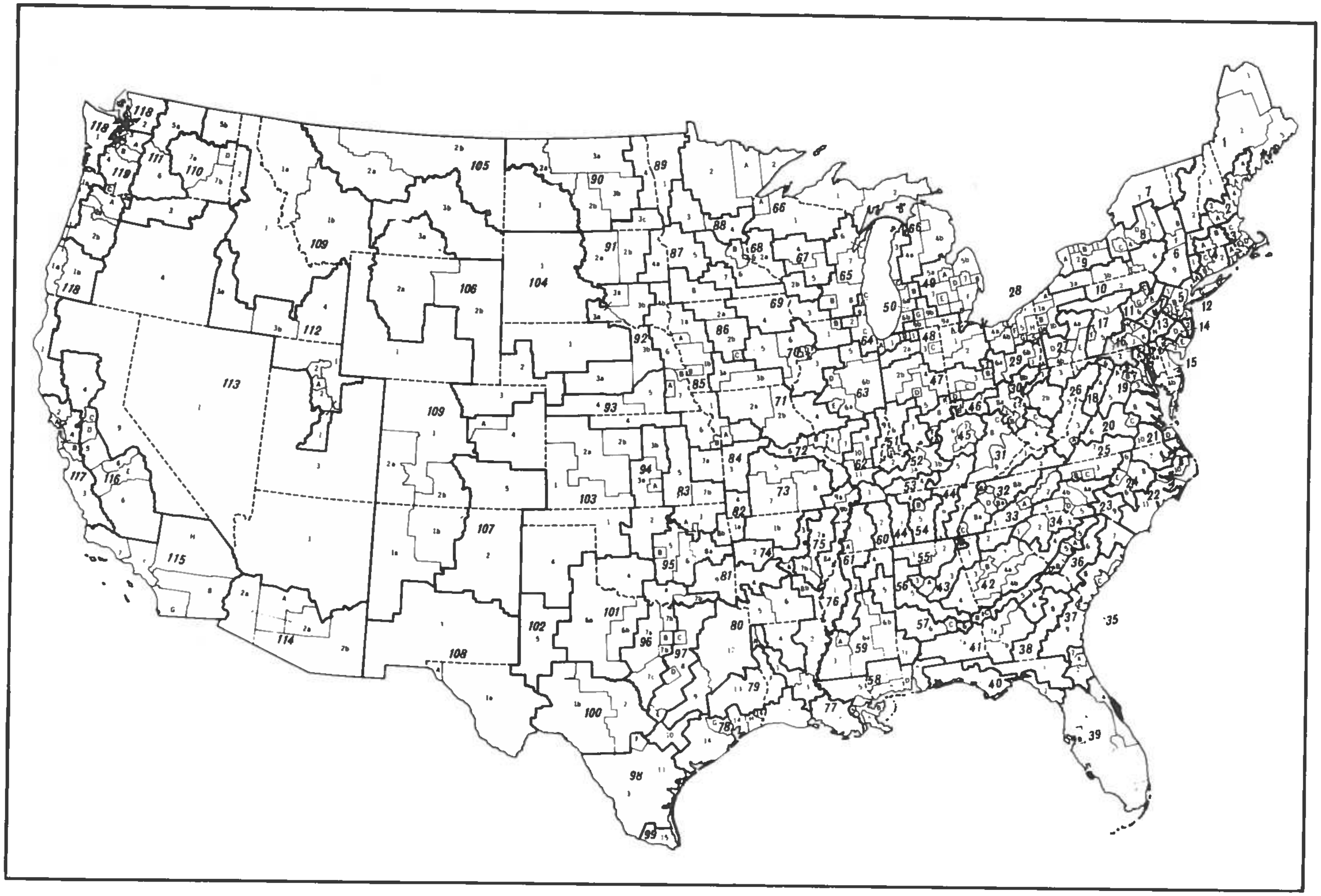
(SMSA SEAs) assigned sequentially within a state (See Figure 3.5).[1]

6. Counties. Counties and county equivalents are defined as the primary political and administrative subdivisions of the fifty states. County equivalents include: 29 Census Divisions in Alaska; 64 Parishes in Louisiana; Baltimore City in Maryland; St. Louis City in Missouri; Carson City in Nevada; the 39 Independent Cities in Virginia; the Consolidated Government of Columbus in Georgia; and that part of Yellowstone National Park located in Montana. The District of Columbia is also considered a county equivalent. There are approximately 3,134 counties and county equivalents in the United States varying considerably in size and population. According to the 1970 Census, Upper Yukon, Alaska, is the largest county with an area of 84,142 square miles; New York County, New York (coextensive with the borough of Manhattan) is the smallest with 23 square miles. The greatest number of people, over seven million, reside in Los Angeles County, California, while only 164 people live in Love County, Texas. County data are provided for nearly all Censuses.

A three digit numeric code unique within state is assigned to each county. The codes used are those defined in the Federal Information Processing Standards Publications (FIPSPUB No. 6, See Section 2.1). On the 1970 Census of Housing and Population summary tapes,

---

[1]For a more detailed description of the development and criteria for economic subregions and state economic areas refer to: Donald J. Bogue, and Calvin L. Beale. Economic Areas of the United States. New York: The Free Press of Glencoe, Inc., 1961.



Source: Bogue and Beale, Economic Areas of the United States

Figure 3.5 State Economic Areas

each geographic identification record carries two codes for county - the 1970 county code and the 1970 county of tabulation code. These two codes usually agree. However, when the 1970 county of tabulation code differs from the 1970 county code, a record had been assigned an incorrect 1970 county code, and the 1970 county of tabulation code represents a correction which should be used in the aggregation of records. A one digit code, known as the central county code, indicates those counties that contain central cities of SMSAs. The code "0" in this field indicates all those counties which contain SMSA central cities and the code "1" indicates all counties which do not.

7. Minor Civil Divisions (MCDs). Minor civil divisions are the primary political and administrative subdivisions of counties or county equivalents as established by state law. Although almost two thirds of all MCDs are townships, an MCD may also consist of towns, precincts, or other independent municipalities. While separate enumeration was made for some county subdivisions even in the First Decennial Census of 1790, it wasn't until 1870 that population totals were published for county subdivisions in all states. In 1870 MCDs were described as "civil divisions less than counties." In 1880 the term "minor civil division" was adopted and has been used ever since. In 1970 there were over 28,000 MCDs ranging in size from townships of less than 100 people tc a few cases where the county itself is the MCD. Each MCD or equivalent area (census county divisions) is assigned a three digit numeric code in alphabetic sequence within a county. Gaps of five

digits were originally allowed for the inclusion of new units in the proper alphabetic sequence.

8. Census County Divisions (CCDs). In states where minor civil divisions are not satisfactory statistical units, either because the areas have lost their original significance, are too small, have frequent boundary changes, or have indefinite boundaries, the Bureau of the Census in cooperation with state and local authorities has established relatively permanent, statistically relevant county subdivisions designated as census county divisions. In delineating these county subdivisions, consideration is given to recognizing the trade pattern or service areas of principal settlements and in some cases to major land use or physiographic differences. The boundaries generally follow physical features, such as transportation lines, streams or ridges. The larger incorporated places are recognized as individual census county divisions, and the boundaries of these divisions change when annexations occur. CCDs were first established in the state of Washington for use in the 1950 Censuses. Between 1950 and 1960, the program was extended to seventeen more states in which the pattern of MCDs was not well suited for statistical purposes. During the last decade CCDs have been added in three other states.[1] All counties in these states have been subdivided into census county divisions. In reviewing the existing

---

[1]The 21 states with census county divisions are: Alabama, Arizona, California, Colorado, Delaware, Florida, Georgia, Hawaii, Idaho, Kentucky, Montana, New Mexico, North Dakota, Oklahoma, Oregon, South Carolina, Tennessee, Texas, Utah, Washington and Wyoming.

CCDs for the 1970 Census, some revisions were made to improve their usefulness. For example, most of the counties with small populations which were single CCDs in 1960 have been divided into two CCDs in order to provide more statistical detail. The Census of Housing and Population is the only one for which census county division data is provided. The total number of CCDs and MCDs is approximately 36,000. They are assigned a three digit numeric code and the boundaries of both CCDs and MCDs are updated annually by a boundary and annexation survey conducted by the Bureau of the Census.

9. Places. Both incorporated and unincorporated concentrations of population generally termed places are recognized by the Bureau of the Census. A four digit numeric code is assigned to each place in alphabetic sequence within a state. Place codes are unique within states but place boundaries may cross county, MCD, or CCD lines. The codes are gapped at intervals of five digits to permit insertion of codes for additional places in the proper alphabetic sequence. In addition to the place code itself there are two other place associated codes. The place description code is a one digit number which identifies place by type. Places are classified and coded as follows:

1 - central city (SMSA)

2 - central city (Urbanized Area)

3 - central city (SMSA and Urbanized Area)

4 - other incorporated place

5 - unincorporated place

6 - not a place

Places are also qualified by a two digit place size code which identifies one of sixteen population groups into which a place falls on the basis of the 1970 Census. The place size codes are:

00 - under 200	08 - 10,000 to 19,999
01 - 200 to 499	09 - 20,000 to 24,999
02 - 500 to 999	10 - 25,000 to 49,999
03 - 1,000 to 1,499	11 - 50,000 to 99,999
04 - 1,500 to 1,999	12 - 100,000 to 249,999
05 - 2,000 to 2,499	13 - 250,000 to 499,999
06 - 2,500 to 4,999	14 - 500,000 to 999,999
07 - 5,000 to 9,999	15 - 1,000,000 or more.....

9.1 Incorporated Places. These are political units incorporated as cities, towns, villages, or boroughs (with the exception of towns in New England, New York or Wisconsin and boroughs in Alaska which are not recognized as incorporated places). Statistics for incorporated places of all types and sizes have been published in the population and housing Census report since 1870. In 1970 there were more than 18,500 incorporated places.

9.2 Unincorporated Places. An unincorporated place is a closely settled population center without legally defined corporate limits or municipal powers. Each unincorporated place has a definite residential nucleus, and boundaries are drawn to include, insofar as possible, all the surrounding closely settled areas. In the 1940 Census of Housing and Population, data for unincorporated places were collected for some areas in the United States. Beginning with the 1950 Census, data have been collected for unincorporated places nationwide. In 1970, there was a significant increase in the number of unincorporated places defined and reported as a result of increased participation in their definition by state highway agencies and the local census tract committees. The decision to recognize unincorporated places of 5,000 or more inhabitants within urbanized areas (See Item 15) also increased the number of unincorporated places reported. Due to this substantial increase, data from the 1970 Census of Housing and Population will be published only for those unincorporated places with 1,000 inhabitants or more.

10. Governmental Units. In reports published by the Census of Governments, statistics are tabulated for types of governments rather than for types of places. To be reported as a government, an entity must possess all three of the following attributes: 1) existence as an organized entity as evidenced by the presence of some form of organization and the possession of some corporate powers; 2) governmental character, as indicated where officers of the entity are popularly elected or appointed by public officials with responsibility entailed to the public; and 3) substantial autonomy in fiscal administrative functions. There were 78,269 governmental units in the United States at the beginning of 1972. The 15,781 local school districts and 23,885 special districts accounted for a little over one half of this total. The

remainder includes the federal government, the 50 state governments, 3,044 counties, 18,517 municipalities, and 16,991 townships. The 1972 count of local governments is based on a directory card listing which shows for each governmental unit and dependent school system its name, county location, mailing address, and selected characteristics--including the population of counties, municipalities, and townships and the enrollment of school districts and dependent school systems. This directory card record was first prepared in 1951-52 and has since been successively updated.

10.1 Municipalities. For purposes of Census classification, a municipality is a political subdivision within which a municipal corporation has been established to provide general local government for a specific population concentration in a defined area. A municipality may be legally termed a city, village, borough (except in Alaska), or town (except in New England States, New York and Wisconsin).[1] The concept of municipalities corresponds generally to the "incorporated places" that are recognized in Census Bureau reporting of population and housing statistics subject to an important qualification--the count of municipalities in this report excludes places which are reported as currently governmentally inactive.

---

[1]In Alaska, the term "borough" corresponds to units classed as county governments and in New England, New York and Wisconsin, the term "town" relates to an area subdivision which (although it may be legally termed a municipal corporation and have a similar governmental organization) has no necessary relationship to a concentration of population and thus corresponds to townships in other states.

10.2 Townships. The term townships applies to 16,991 organized governments, located in 21 states. This category includes governmental units officially designated as towns in the six New England states, New York, and Wisconsin and some "plantations" in Maine and "locations" in New Hampshire as well as governments called townships in other areas. As distinguished from municipalities, which are created to serve specific population concentrations, townships exist to serve inhabitants of areas defined without regard to population concentrations. Excluded from this count of township governments are unorganized township areas, townships coextensive with cities where the city governments have absorbed the township functions, and townships known to have ceased to perform governmental functions.

10.3 School Districts. Of the 17,238 school systems in the United States in early 1972, only the 15,781 that are independent school districts are recognized as separate units of government. The other 1,457 dependent school systems are regarded as agencies of other governments--county, municipality, township, or state--and are excluded from the count of governmental units. Because of the variety of state legislative provisions for the administration and operation of public schools, marked diversity is found in school organization throughout the United States. In 27 states, responsibility for public schools rests solely with school districts which are independent governmental units. In another three states--California, Indiana, and Ohio--all school systems that provide education through grade 12 are independent governments.

However, each of these states has an institution of higher education operated by a city or county government. A mixed situation is found in 15 states, with public schools that provide elementary and secondary education operated in some areas by independent school districts and elsewhere by some other type of government. In the District of Columbia and in five states (Alaska, Hawaii, Maryland, North Carolina, and Virginia) there are no independent school districts, and all public schools are administered by systems that are agencies of county or city governments, or of the state.

10.4 Special Districts. Special districts make up the most varied area of local government. With the exception of Alaska, these units are found in every state and in the District of Columbia. Most special districts are established to perform a single function but some are authorized by their enabling legislation to provide several kinds of services. More than one fourth of all special districts are concerned with natural resources; fire protection districts constitute one sixth of the total number; and housing authorities and urban water supply districts each one tenth. A variety of functions account for the remaining number of special districts. Most of the units recognized as multiple function in nature involve some combination of urban water supply with other services, most commonly sewerage services. A total of more than 3,000 special districts are concerned with urban water supply either as the sole function or as one of a combination of functions.

11. Congressional Districts. These areas are defined by state legislatures for the purpose of electing congressmen to the United States House of Representatives and are subject to change, based on population counts, after each Decennial Census. Almost all of the fifty states revised their congressional districts boundaries in 1971 and 1972 as a result of the 1970 Census of Housing and Population. Due to recent decisions of the United States Supreme Court requiring unprecedentedly close equality of population in congressional districts, states may be required to change the district boundaries more frequently. Congressional districts are identified by a two digit numeric code which corresponds to the number assigned in state legislation, except that the code 01 is used to identify areas in which members of congress are elected at large, rather than by district. These congressional district codes also conform to the FIPS standard (See Section 2.1). The Censuses of Population and Housing are the only Censuses for which statistics are tabulated by congressional districts. In 1970 there were 435 congressional districts.

12. Wards. Wards are political subdivisions of cities used for voting and representation purposes. Population totals for wards were reported for some cities as early as 1800. These areas are reported in the Census of Housing and Population tabulations in cities of 10,000 or more which have provided boundary information; however the statistical significance of wards is negligible, primarily because of the frequent boundary changes. The ward code is a two digit number

generally assigned by each city.

13. Standard Metropolitan Statistical Areas (SMSAs). Except in the New England states, a standard metropolitan statistical area consists of a county or group of contiguous counties containing at least one city of 50,000 inhabitants or more, or "twin cities" with a combined population of at least 50,000. In addition to the county, or counties containing such a city or cities, contiguous counties are included in an SMSA if they are metropolitan in character and are socially and economically integrated with the central city.[1] In the New England states, SMSAs consist of towns and cities instead of counties. Each SMSA must include at least one central city and there is no limit to the number of adjacent counties included in the SMSA as long as they are integrated with the central city. SMSAs may also cross state boundaries. The concept of the SMSA was developed for the 1950 Census in order to present general purpose statistics. Originally these units were called standard metropolitan areas. Later the term statistical was added in order to avoid misunderstanding concerning the nature of the areas so recognized. Standard metropolitan statistical areas are delineated by the Office of Management and Budget (OMB) with the advice of the Federal Committee on Standard Metropolitan Statistical Areas, which is composed of representatives of the major federal statistical agencies. The definitions are reviewed from time to time

---

[1]A detailed list of the exact criteria for SMSAs is included in Section 4.3.

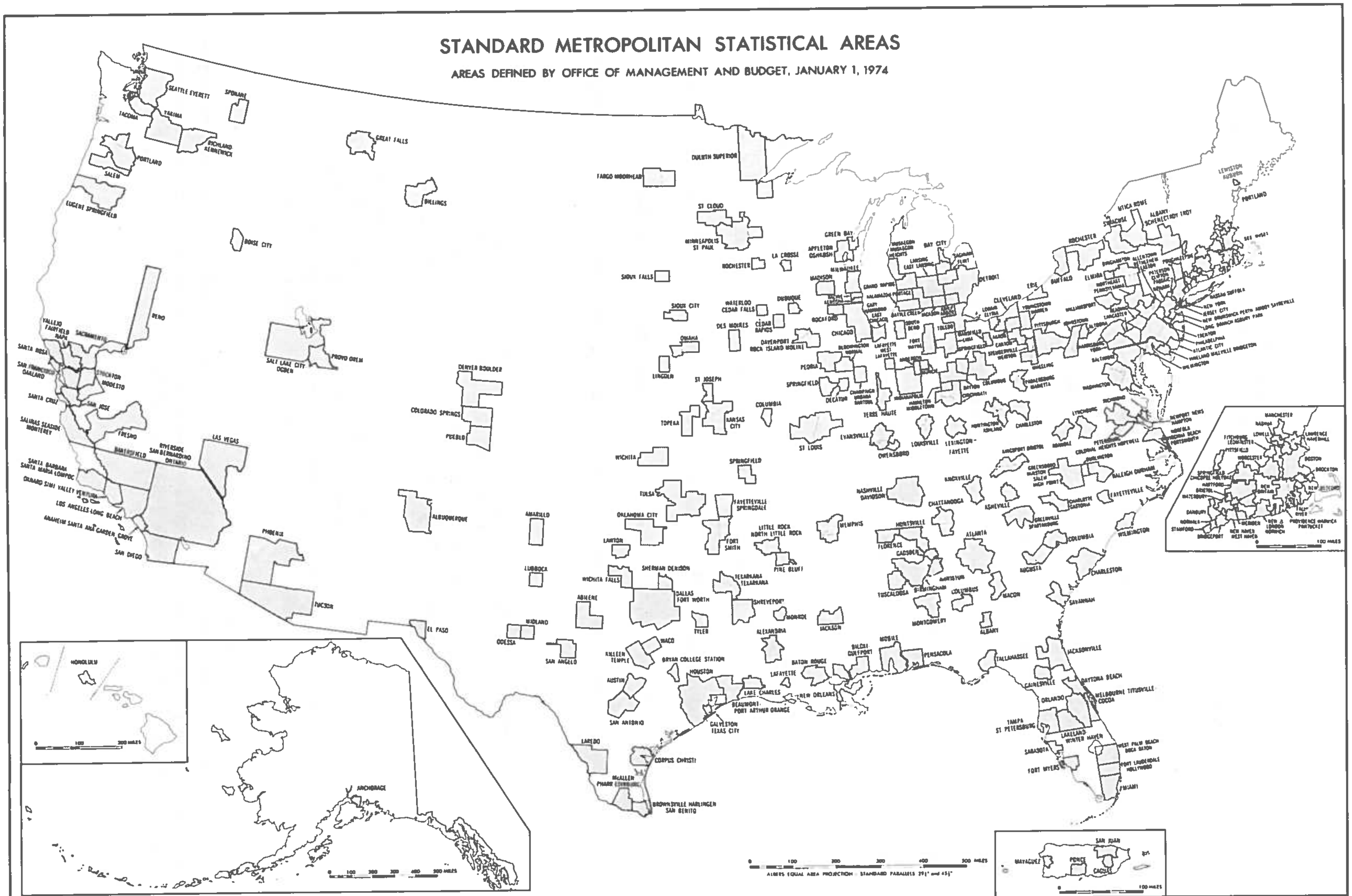
and revisions are made on the basis of more recent data. As of May 1973 there were 263 SMSAs in the United States and 4 additional designated SMSAs in Puerto Rico (See Figure 3.6).[1] Each of these areas is assigned a four digit number in alphabetic sequence by title conforming to the FIPS standard for SMSA codes.[2] This series of numbers was originally gapped at intervals of forty. The population living in SMSAs is designated as the metropolitan population. This classification is further subdivided into population "inside central city or cities" and population "outside central city or cities." The population living outside SMSAs constitutes the nonmetropolitan population.

14. Standard Consolidated Areas (SCAs). In view of the special importance of the metropolitan complexes around New York and Chicago, several contiguous SMSAs and additional counties that do not meet the formal SMSA integration criteria but do have other strong interrelationships have been combined into the New York-Northeastern New Jersey and the Chicago-Northwestern Indiana Standard Consolidated Areas. In Census tabulations, a one digit alphabetic code is assigned to each of these SCAs. They are:

---

[1]Every one of the 50 United States except Wyoming and Vermont contain at least one SMSA.

[2]For the 1972 Economic Censuses, the Bureau of the Census employed a three digit truncation of the four digit FIPS SMSA code. Since the last digit of all but one of the FIPS SMSA codes assigned at that time was zero, this truncation does not actually alter the standard codes. After 1972 FIPS added two more SMSA codes with final digits other than zero.



Source: Bureau of the Census

Figure 3.6 Standard Metropolitan Statistical Areas

A. New York SMSA, Newark SMSA, Jersey City SMSA, Paterson - Clifton - Passaic SMSA, and Middlesex and Somerset Counties in New Jersey

B. Chicago SMSA and Gary-Hammond-East Chicago SMSA in Indiana

The concept of standard consolidated areas was introduced in the 1960 Census of Housing and Population. Although the definition of the two SCAs in terms of SMSAs has been altered due to a shuffle of SMSA definitions between the 1960 and 1970 Censuses, the component counties have remained the same--these are the counties that comprised the 1950 SMSA of New York and the 1950 SMSA of Chicago (with two additional counties).[1]

15. Urbanized Areas. The major objective of the Census Bureau in delineating urbanized areas is to provide a clear separation of urban and rural population in the vicinity of larger cities. An urbanized area consists of a central city, or cities, and surrounding densely settled territory. The specific criteria for delineation of an urbanized area are as follows:

1. A concentrated nucleus of population consisting of:
   a. a central city of 50,000 inhabitants or more in 1960, in a special census conducted by the Census Bureau since 1960, or in the 1970 Census; or

   b. twin cities, i.e., cities with contiguous boundaries and constituting for general social and economic purposes a single community with a combined population of at least 50,000 and with the smaller of the twin cities having a population of at least 15,000.

---

[1]It is assumed that the 1970 definition of the New York-Northeastern New Jersey - Standard Consolidated Area will be altered to reflect the alteration of the New York SMSA in 1971.

2. Surrounding closely settled territory, including:
   a. incorporated places of 2,500 inhabitants or more

   b. incorporated places with fewer than 2,500 inhabitants, provided that each has a closely settled area of 100 housing units or more.

   c. small parcels of land normally less than one square mile in area having a population density of 1,000 inhabitants or more per square mile. The areas of large nonresidential tracts devoted to such urban land uses as railyards, airports, factories, parks, golf courses, and cemetaries are excluded in computing population density.

   d. other similar small areas in unincorporated territory with lower population density provided that they serve
      - to eliminate enclaves or
      - to close indentations of one mile or less across an open end, or
      - to link outlying enumeration districts of qualifying density that are not more than 1 1/2 miles from the main body of the urbanized area.

Urbanized areas which were first delineated in 1950, differ from standard metropolitan statistical areas (SMSAs) principally in excluding the rural portions of the counties composing the SMSAs and excluding those towns which are separated by a strip of rural territory from the densely populated fringe around the central city. All persons residing in an urbanized area are classified as urban. The urbanized area population is sometimes divided into those in the "central city (or cities)" and those in the remainder of the area or the "urban fringe." Because urbanized areas are defined on the basis of population distribution at the time of the census, the boundaries are not permanent. A four digit numeric code assigned to urbanized areas in alphabetic sequence provides unique identification for each of the 248 urbanized areas in the United

States.

16. Central Business Districts (CBDs). A statistical area based on census tracts, the central business district is the downtown retail trade area of a city. As defined by the Census Bureau, the CBD is an area of very high land valuation characterized by a high concentration of retail business offices, theaters, hotels, and service businesses, and with a high traffic flow. CBDs consist of one or more whole census tracts and have been defined only in cities with a population of 100,000 or more. The purpose of defining the CBD is to provide a basis for comparing changes in business activity in the CBD with changes in the remainder of the metropolitan area or of the central city. Central business districts were identified for the first time in the 1954 Census of Business. The number of CBDs for which data are published increased from 95 in 1954 to about 150 in 1972. A one digit code, known as the central business district indication, denotes what tracted areas are to be tabulated as part of a CBD.

17. Major Retail Centers (MRCs). Major retail centers are concentrations of retail stores located inside an SMSA but not in the central business district of the chief city of the SMSA. An SMSA may have more than one MRC. To be considered an MRC, a shopping area has to contain at least one major general merchandise store--usually a department store--and have $5 million or more in retail sales annually. MRC's include not only the planned suburban shopping centers but also the

older "string" street and neighborhood developments which meet the prerequisites. Frequently the boundaries of a single MRC include stores located within a planned shopping center as well as adjacent stores outside the planned portion. In general, the boundaries of the MRCs have been established to include all adjacent blocks containing at least one store in the general merchandise, apparel, or furniture and appliance groups of stores. The MRC, like the CBD, was developed to measure the changes in shopping habits from the downtown business districts to outlying suburban shopping centers. Major retail centers were identified for the first time in the 1958 Census of Business. The number of MRCs for which data are published has increased from 472 in 1958 to almost 2,000 in 1972. The Census of Business is the only source of statistics for major retail centers.

18. <u>Census Tracts</u>. Census tracts are small, relatively permanent areas into which large cities and adjacent areas have been divided for statistical purposes. Tracts are designed to be relatively uniform with respect to population characteristics, economic status, and living conditions. The average tract has a little over 4,000 residents. Tract boundaries are established cooperatively by a local census committee and the Bureau of the Census according to the standards that dictate limitations on population size, specify the best type of visible boundaries, and indicate the type of homogeneity required. Geographic shape or area size of tracts is of relatively minor importance. Tract boundaries are established with the intention of being maintained over

a long time so that statistical comparisons can be made from year to year and from census to census. However, occasional changes may be made in tract boundaries due to physical changes in the street patterns of cities caused by highway construction, park development, urban renewal programs, etc. The concept of census tracts was originated by Dr. Walter Laidlaw in New York City in 1906. At his request, the Bureau of the Census tabulated census tract data from the 1910 Census for New York City and seven other cities. The program attracted users slowly in the early decades of its development but by 1950 the census tract system was well established and widely used. For the 1970 census all SMSAs recognized at the time of the census are completely tracted and include about 32,000 census tracts. Over 100 counties, cities, or parts of counties outside SMSAs are also tracted and include about 2,600 census tracts. In many more cities than before, census tracts have been modified to disregard city limits thereby reducing the need for future tract changes as city limits expand. However, census tracts have always and will continue to respect county boundaries. Each census tract is assigned a six digit numeric identification number. The first four digits are the "basic" code and the last two are the "suffix" code. The suffix is only used when necessary to identify two or more tracts formed from a former single tract. Basic tract codes can range from 1 to 9,999, and the suffix codes range from 00 to 95. Suffix code 99 is reserved for crews of ships. The number is always unique within county, usually unique within SMSA, and in a few instances, unique within state. Census tract data from the Census of Housing and

Population are published in a series of reports, one for each SMSA.

19. <u>Enumeration Districts (EDs)</u>. Enumeration districts are small administrative areas containing on the average 250 housing units. They are defined by the Bureau of the Census and used for the collection and tabulation of population and housing data. Two administrative factors play a part in determining the geographic definition: 1) the estimated population size of the ED should constitute an adequate enumerator workload; and 2) the ED will never cross the boundary of a city, township, or other area (except census blocks) for which data are to be tabulated. EDs are delineated by outlining them on the base maps. In 1970, there were approximately 250,000 EDs. For the 1970 Census of Population and Housing, where approximately three fifths of the population were enumerated by a mail-out/mail-back canvass, EDs similar to those of earlier censuses were used for the collection and tabulation of data only for the conventional enumeration areas and for portions of the mail-out/mail-back SMSAs not covered by the Address Coding Guide. There were about 750 people in each ED in the conventional enumeration areas and about 1,500 people in the SMSA areas where EDs were used. A four digit numeric code (ED basic code) is assigned sequentially to each ED within a county, and in some instances within a District Office Territory. A one digit alphabetic suffix code is used to indicate splits of original EDs. Another one digit code, commonly called ED type, identifies an ED as being in one of the following areas: Address Coding Guide (0), Prelist (1), and Conventional

(non-mail) (2). EDs in Address Coding Guide areas are called block groups. Although data for enumeration districts have not been published in the printed reports of the censuses, data for enumeration districts and block groups are available on the 1970 First Count Summary Tape.[1]

20. Block Groups. Block groups are tabulation areas, defined by the Census Bureau, which were used in the census-by-mail areas. The designation "block group" was new in 1970. A block group is a combination of contiguous blocks having a combined average population of about 1,000. Block groups are approximately equal in area (discounting parks, cemeteries, railroad yards, industrial plants, rural areas, etc.); they are subdivisions of census tracts which simplify numbering and data control. Block groups are typically defined without regard to the boundaries of political or administrative areas, such as cities, minor civil divisions, congressional districts, etc. When a block group straddles one or more of these boundaries, data for those parts in different areas are tabulated separately. Each block group is identified by the first digit of the three digit block number. Block group "1" will contain any block in range 101 - 199, block group "2" will range 201 - 299, etc. However, normally only the first few numbers in

---

[1]From the 1970 Census, six counts of population and housing data are available. The first three counts relate to the subject items collected on a 100 percent basis in the Census. The second three counts contain data collected on a sample basis in the Census. The data so tabulated are available--subject to suppression of certain detail where necessary to protect confidentiality--on magnetic computer tape at the cost of preparing the copy.

a range are used. For purposes of providing small area population and housing data, they are equivalent to enumeration districts within the mail-out/mail-back areas where Address Coding Guides have been prepared.

21. Block Numbering Areas. In untracted areas where city blocks are tabulated on a contract basis, blocks have been numbered in block numbering areas which are identified by census tract type numbers ranging from 9400.00 to 9999.00. Block numbering areas are unique within county boundaries and usually contain a population of about 4,000 people. Data are not tabulated for these areas.

22. City Block. A city block is normally a well-defined rectangular piece of land bounded by street and roads. However, it may be irregular in shape or bounded by railroad tracks, streams, or other features. Blocks may not cross census tract boundaries but may cross other boundaries such as city limits. Beginning with the 1940 Census of Housing, block statistics were published for all cities with a population of 50,000 or more at the time of the last decennial or special census. For the 1970 Census, the city block program was expanded to include approximately 1,800,000 blocks. Blocks were identified and data tabulated and published for all cities of 50,000 inhabitants or more in an official census prior to 1970 and for the urbanized areas of these cities. In addition, the Bureau also collected and published data for blocks in over 900 additional areas on a contractual basis where the local area paid for the processing and tabulation. A three digit

identification code number is assigned to each block. The first of this code is always one or greater. Block numbers are unique within each census tract (See Figures 3.7 and 3.8 illustrating Hierarchy of Census Geographic Units).

22.1 Block Face. A block face is the side of a city block; a segment of the periphery of a block; or a segment of a cul-de-sac running into a block. Block faces can be identified by using the Address Coding Guide and grouped to any specifications at request for a special tabulation.

23. Geographic Description Codes.

23.1 Urban/Rural Population. Urban population comprise all persons living in:

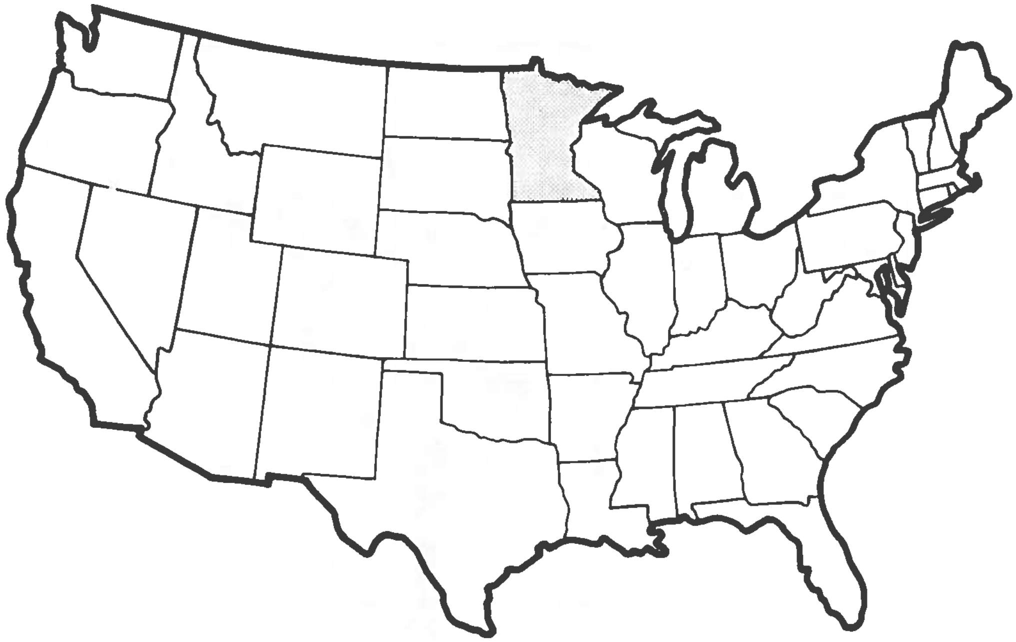
1. Places of 2,500 inhabitants or more incorporated as cities, boroughs, villages, and towns (except boroughs in Alaska and towns in New England, New York, and Wisconsin).

2. The densely settled urban fringe, whether incorporated or unincorporated, or urbanized areas.

3. Unincorporated places of 2,500 or more inhabitants.

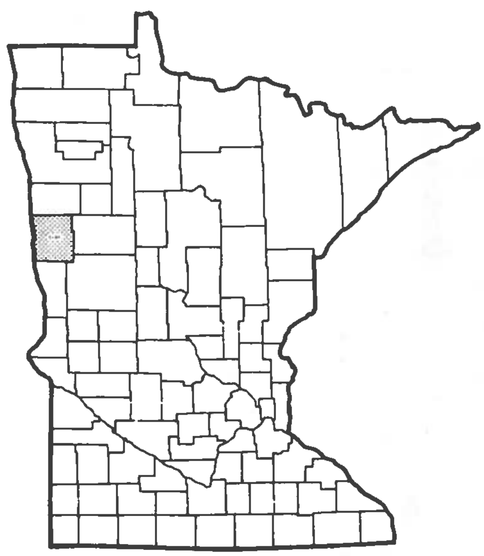
Rural areas are those remaining areas not falling into one of these categories. The Bureau of the Census uses a one digit code on the summary tapes to classify enumeration districts as urban, rural. The urban/rural code designations are as follows: 0 = urban, and 1 - rural.

23.2 Universal Area Code (UAC). All central cities of SMSAs, select towns, and all counties and central business districts in the



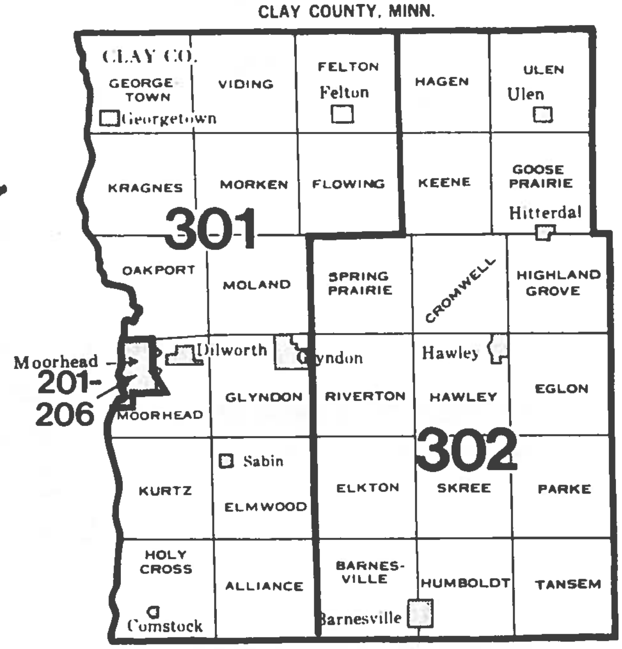
Map of the United States highlighting the State of Minnesota

MINNESOTA



County Outline Map for Minnesota highlighting Clay County

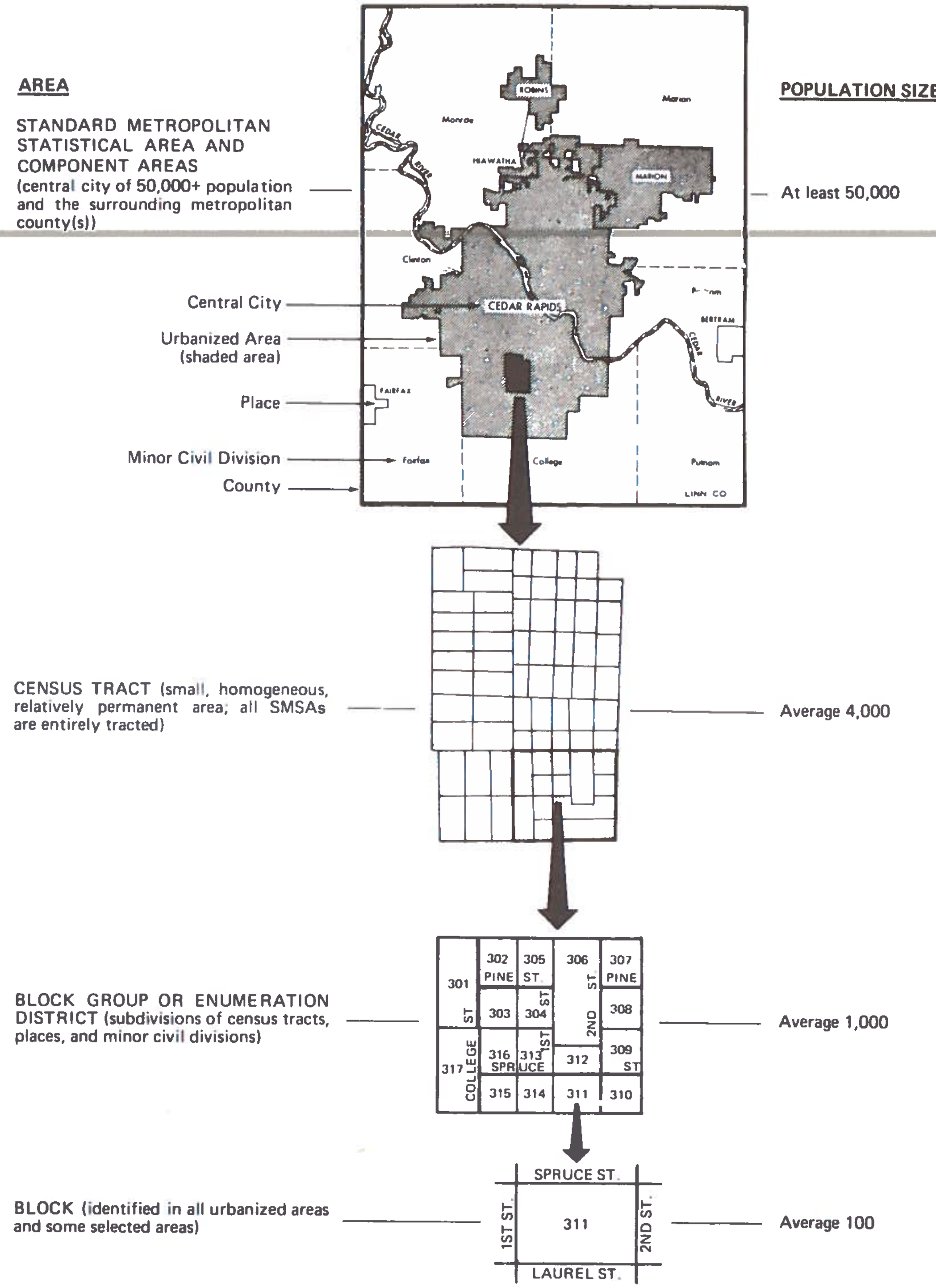
CLAY COUNTY, MINN.



Map of Clay County by Township; highlighting Places and indicating Census Tracts 201-206, 301, and 302.

Source: Bureau of the Census

Figure 3.7 Hierarchy of Census Geographic Units, I



Source: Bureau of the Census

Figure 3.8 Hierarchy of Census Geographic Units, II

United States are assigned a five digit numeric universal area code. The levels are: county (1); town (New England) (2); city (3); and central business district (4).

24. Special Purpose Districts.

24.1 Foreign Trade Statistical Areas. Statistics on United States imports and exports are published for many different areas. Information is shown for foreign countries, foreign ports, Puerto Rico, U.S. possessions (Virgin Islands, Wake Island, Guam, and American Samoa), U.S. coastal districts, U.S. ports (including Great Lakes ports), and for combinations of trading areas (See Section 4.7).

24.2 Water Locations. These are areas established to provide tabulations useful in analyzing the population growth of SMSAs near coasts, lakes, and rivers. Water locations first appeared in the 1960 Population Census Report entitled "Standard Metropolitan Statistical Areas," PC(3)-ID.

24.3 Industrial Water Usage Regions. Twenty of these units defined by a federal interagency committee are recognized in a subject report from the Census of Manufactures, "Water Use in Manufacturing." Each region is a combination of counties grouped to reflect major water drainage basins.

24.4 Fishing Regions. Statistics for these ten regions as defined by the Department of the Interior are printed in the report from the 1963 Census of Commercial Fisheries.

24.5 Petroleum Regions. Statistics for these eight regions are presented in a report from the Census of Business. They are defined by the Departments of Defense and the Interior and by the Executive Office of the President.

24.6 Lumber Industry Regions. Statistics for these ten regions are presented in the annual Current Industrial Report, "Lumber Production and Mill Stocks."

24.7 Regional Marketing Areas. Statistics for these areas are tabulated for brick and structural clay tile (except surfacing tile). They appear in the monthly Current Industrial Report, "Clay Construction Products."

24.8 Oil and Gas Districts. These regions are located in Louisiana, Texas, and New Mexico. In Louisiana, they are composed of parishes and in New Mexico and Texas they are composed of counties. Statistics for the seventeen districts are presented in the reports on petroleum and natural gas industries in the Census of Mineral Industries.

24.9 Standard Location Areas (SLAs). This is an area defined by the Office of Civil Defense. Special Housing and Population Census tabulations are produced for use in the Civil Defense Damage Assessment Program. SLAs consist of census tracts in tracted areas (tracts lying both inside and outside cities of 50,000 or more are regarded as split tracts and are treated as two separate SLAs); wards in untracted cities of 25,000 or more where wards are identified; groups of enumeration districts (averaging 5,000 population) in cities of 25,000 or more with neither tracts or wards; urban places of 2,500 to 25,000 outside tracted areas; or MCDs and CCDs, grouped where necessary to attain a minimum population of 2,000 (See Section 3.2).

24.10 Production Areas and Market Areas. Production areas and market areas are geographic units especially defined for use in the Census of Transportation. A production area consists of a single SMSA or cluster of SMSAs selected to represent relatively compact geographic concentrations of industrial activity. These units are used to reference the origin of commodity movements without disclosing the activities of any company or individual. The production area code is a two digit number ranging from 01 to 25. Market areas are geographic units created to provide more destination detail than possible for the origin detail provided by the production area system. The production areas which are major market as well as production areas and an additional number of SMSA complexes with concentrated marketing activities comprise the full set of market areas (See Section 3.7).

Market areas are coded from 31 to 55.

24.11 ZIP Code Areas. ZIP code areas are special units established by the U.S. Postal Service for directing and sorting mail. They are identified by five digit numeric codes. The first three digits indicate a major city or sectional distribution center; the last two digits signify a specific post office's delivery area within the center (See Section 4.1). ZIP code areas do not coincide with Census areas and change according to postal requirements. They are not mutually exclusive areas and their boundaries do not necessarily follow clearly identifiable physical features. Since ZIP code areas were developed within the last decade, the 1970 Census is the first census to provide data by ZIP code areas. These data are available only in the form of the Fifth Count Summary Tapes. These data are estimates and are of a lower order of precision than other Census information.

3.1.2 Geographic Files

MEDList and GACI. The Bureau of the Census has prepared two geographic index files for the 1970 Census of Housing and Population. The Master Enumeration District List (MEDList), which is the successor of the 1960 Geographic Identification Scheme, is a file associating names of enumeration districts, block groups, and minor civil divisions or census county divisions for a state with their geographic identification codes as well as their population and housing counts. In areas where address coding guides have been developed, the smallest unit on the

MEDList is the block group. Enumeration districts are the smallest unit in all other areas (Figure 3.9).

Another version of the Census MEDList has been prepared by the Office of Civil Defense (Figure 3.10). This file contains the appropriate census codes, the 1970 population and housing counts, and the latitude and longitude of the centroid of the enumeration districts or block groups listed.

A third file called the Geographic Area Code Index (GACI) contains geographic codes for each state, county, minor civil division or census county division, and place in the United States, thus providing a means of associating area names with the geographic codes on census summary tapes. Essentially GACI is an abbreviated or truncated version of MEDList and it does not contain either population or housing counts (Figure 3.11).

<u>ACG/DIME Geographic Base Files</u>. In 1970, the Bureau of the Census expanded those areas canvassed by mail in the Nineteenth Decennial Census of Population and Housing. For the 145 SMSAs canvassed by mail, a system called the address coding guide (ACG) was employed to code individual addresses on the mailing list to specific geographic areas for tabulation. The addresses coding guide consists of an inventory of block faces with their associated street names and address ranges together with codes identifying specified geographic areas within which the block faces lie. Each record in the file refers to a segment of a street by identifying the ZIP code, the

STATE 1970	STATE 1960	FIPS COUNTY	COUNTY OF TABULATION	CENTRAL COUNTY	MINOR CIVIL DIVISION CENSUS CIVIL DIVISION	PLACE CODE	PLACE DESCRIPTION	PLACE SIZE	STANDARD CONSOLIDATED AREA	STANDARD METROPOLITAN STATISTICAL AREA	URBANIZED AREA	TRACTED AREA	UNIVERSAL AREA PREFIX	UNIVERSAL AREA CODE	STATE ECONOMIC AREA	ECONOMIC SUBREGION	CENTRAL BUSINESS DISTRICT	AREA NAME	TRACT BASIC	TRACT SUFFIX	BLOCK GROUP	ENUMERTION DISTRICT CODE	ENUMERTION DISTRICT SUFFIX	URBAN/RURAL	WARD	CONGRESSIONAL DISTRICT
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
26	34	121		1														Muskegon								
26	34	121		1	090													Ravenna TWP								
26	34	121		1	090	2275	4	02		5320		5320	1	34062	06	050	1	Ravenna Village	0029			0442		1		09
26	34	121		1	095		7			5320		5320	1	34062	06	050	1	Remainder of MCD	0029			0443		1		09
26	34	121		1	095													Roosevelt Park City								
26	34	121		1	095	2345	4	06		5320	5320	5320	1	34062	06	050		Roosevelt Park City								
																	1		0022		1			0		09
																	1				2			0		09
																	1				3			0		09
26	34	121		1	110		7			5320		5320	1	34062	06	050		White Hall TWP								
																	1		0030			0404		1		09
																	1		0040			0405		1		09
																	1		0037			0406		1		09
																	1		0038			0407		1		09
																	1		0036		1			0		09
																	1				2			0		09

Source: Bureau of the Census

Figure 3.9 The Master Enumeration District List (MEDList)

State - Maine	Code 23	Total Population - 993663	Housing Units - 397169
County - Androscoggin	Code 001	Total Population - 91279	Housing Units - 31194

MCD	Area Name	Tract	Resident Population	Housing Units	Latitude	Longitude
005	Auburn City		24151	8403	44-05-40N	070-14-20W
005	Auburn	0101	2198	992	44-05-39N	070-13-41W
005	Auburn	0102	4152	1427	44-08-01N	070-14-05W
005	Auburn	0103	3484	1345	44-06-00N	070-14-01W
005	Auburn	0104	2266	684	44-05-08N	070-14-31W
005	Auburn	0105	2618	855	44-05-04N	070-13-22W
005	Auburn	0106	2945	837	44-03-19N	070-13-45W
005	Auburn	0107	2848	831	44-04-19N	070-16-05W
005	Auburn	0108	3640	1432	44-06-15N	070-15-02W
010	Durham Town		1264	376	43-57-44N	070-06-57W
015	Greene Town		1772	758	44-11-25N	070-08-58W
020	Leeds Town		1031	354	44-17-41N	070-07-30W

Source: Office of Civil Defense

Figure 3.10 The X-Y MEDList

STATE						PLACE							UNIVERSAL AREA				
1970	1960	FIPS COUNTY	COUNTY OF TABULATION	CENTRAL COUNTY	MINOR CIVIL DIVISION CENSUS CIVIL DIVISION	CODE	DESCRIPTION	SIZE	STANDARD CONSOLIDATED AREA	STANDARD METROPOLITAN STATISTICAL AREA	URBANIZED AREA	TRACTED AREA	PREFIX	CODE	STATE ECONOMIC AREA	ECONOMIC SUBREGION	AREA NAME
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
08	84																Colorado
08	84	001															Adams
08	84	001	001	1	005					2080		2080	1	84005	A	103	Brigthton Division
08	84	001	001	1	005	0130	4			2080		2080	1	84005	A	103	Brighton
08	84	001	001	1	005	0943	4			2080	2080	2080	1	84005	A	103	North Glenn
08	84	001	001	1	008					2080		2080	1	84005	A	103	Commerce City Division
08	84	001	001	1	008	0245	4			2080	2080	2080	1	84005	A	103	Commerce City
08	84	001	001	1	008	0325	5			2080	2080	2080	1	84005	A	103	Derby (U)
08	84	001	001	1	015					2080		2080	1	84005	A	103	East Adams Division
08	84	001	001	1	015	0075	4	01		2080		2080	1	84005	A	103	Bennett
08	84	001	001	1	020					2080	2080	2080	1	84005	A	103	North Auroda Division
08	84	001	001	1	020	0055	4			2080	2080	2080	1	84006	A	103	Auroda

Source: Bureau of the Census

Figure 3.11 The Geographic Area Code Index (CACI) street name and the first and last address associated with that segment. Each record relates to odd numbered or even numbered addresses, never to both.

The series of geographic identifiers contained on an ACG record include: state, county, minor civil division or census county division, place, ZIP code, census tract, street (name of street, street direction, street type and serial number), address range, block number, SMSA, district office (Bureau of the Census administrative code), area code, local code or census serial number, ward, annex (area annexed between 1960 and 1970), congressional district, post office, and serial number. Most of these identifiers and their significance have been defined previously. The others are:

Address Range. The address range consists of the lowest and highest addresses of a range of addresses on a block face. Usually, the potential address range is recorded. An address range of zero may appear when a block face contains no addresses or the potential range is unknown. High and low addresses may be the same if there is only one address on a block face.

Area Code. During the preparation of address coding guides, a three digit numeric code, known as the area code was devised solely to provide an identification combining MCD and place codes. The area code is assigned to MCDs or CCDs alphabetically within county, and to all places within each MCD or CCD. Numbers were assigned at intervals of five, beginning with 005, (except Cook County, Illinois and Allegheny County, Pennsylvania, where numbers were assigned at intervals of four, beginning with 004) to provide for insertion of new places and changes in alphabetic listing of MCD/CCDs. The minor civil division/place code combination would have required seven digits if the normal codes were used.

Serial Numbers. The serial number is a five digit identifier of a single record in the address coding guide for an urbanized area. The serial number is unique within SMSA. A suffix to the serial number known as a check digit, is mathematically derived from the serial number and used to detect errors in transcribing or punching serial numbers. A typical use of the serial number is to identify records that are to be changed.

Street Code. The street code is a five digit numeric code for each street name in the address coding guide and is unique within postal finance areas. The postal finance area number used by the U.S. Postal Service consists of two digit state code and the last five digits of the seven digit postal data code. Postal data codes must be used with the street codes to distinguish among identical street codes in different postal finance areas.

Address coding guides were originally developed and pretested for use in the 1970 Census of Population and Housing in New Haven, Connecticut during the Census Use Study.[1] They were designed for use in conjunction with another geocoding system called Dual Independent Map Encoding (DIME). A DIME file is composed of segment records with a segment being defined as a length of street or other feature (a river, railroad track, municipal boundary, etc.) between two distinct vertices or nodes. The process of DIME file creation essentially involves translation of urban geographic information such as street patterns from maps and other sources into a form that can be read and manipulated by computers. While an ACG is constructed on a block face basis a DIME file is constructed on a street segment basis. Each ACG record contains the appropriate codes for one side of a street. Each DIME segment record contains the appropriate codes for both sides of a street between two nodes.

---

[1]The Bureau of the Census in conjunction with officials of the New Haven SMSA initiated the Census Use Study in 1967. The objectives of this program were to assess the needs and interests of census data users and to improve the system for relating census data with local agency data at a fine geographic scale.

An ACG/DIME Geographic Base File (GBF) contains all of the information contained in an address coding guide--street name, address ranges, census block numbers, census tracts number and other geopolitical codes--plus node designations and left/right orientation. The nodes are numbered and assigned coordinate readings. Coordinates in the ACG/DIME Geographic Base File can be expressed as geodetic coordinates (degrees of longitude and latitude carried to four decimal places), as State Plane coordinates and as "map miles" from an arbitrary point. The nodes are placed at sharp curves in streets or other features so that such curves can be adequately described by a series of straight line segments when plotted by computer. The GBF system contains all meaningful non-street features such as rivers, municipal boundaries, shorelines, and railroad tracks. The combination of ACG and DIME provide a flexible geographic base file for urban information system development with computer mapping and spatial analysis capabilities.

Additional DIME related programs (i.e., address matching program, edit, update, and augment programs) have been developed in the Census Use Study's Washington headquarters and in the other field offices in Southern California and Indianapolis (See Figure 3.12).

At its present stage of development the ACG/DIME geographic base file system is limited to urban area applications.[1] The address coding guide and the dual independent map encoding systems are depen-

---

[1]Currently the only statewide ACG/DIME Geographic Base File is for Rhode Island.

PROGRAM	FUNCTION
ADMATCH (Address Matcher)	An address matching system that provides the capability of geographically coding computer readable records containing street addresses. The system compares the addresses on input data records (after standardization with a preprocessor) with the address ranges in a reference file. A "match" occurs when the street names are judged identical or equivalent and when the address falls within the defined range. Any or all geographic codes from the reference file may be attached to the matched data records.
UNIMATCH (Universal Matcher)	An improved matching system that has many capabilities not available in ADMATCH such as the ability to handle building names, street intersections, and non-address matching. It is a generalized record linkage system which will compile, assemble, and execute a file matching system tailored to the specific needs of the user.
Grid Related Information Display System (GRIDS)	A computer mapping system developed for producing character printed maps from detailed data. GRIDS has a flexible user oriented language and has several mapping options available.
DIME Area Centroid System (DACS)	A flexible computer system for locating centroids and calculating areas (square feet or acres) for blocks, block groups, census tracts, etc., from DIME files. Centroid location is required primarily as input to GRIDS and other map generating computer packages.
Network Allocation of Population to Shelter (NAPS)	A system developed by the Census Use Study for the Office of Civil Defense. It assigns the population to their closest fallout shelter up to the shelter capacity limit. It uses the DIME file to determine the shortest path to the shelter and uses the Third Count Census Summary Tape to establish population demand.
Computer Resource Allocation Model (CRAM)	A computer tool, based on the DIME file, that is designed for use in facilities planning (schools, service agencies, recreational facilities). It is a generalized facility location system that allocates demand among the set of available facilities according to their capacity to supply services, street accessibility, and access time. CRAM is an outgrowth of the more limited NAPS system.

Source: Bureau of the Census, Census Use Study: DIME Workshops, May 1973.

Figure 3.12 DIME Related Software Packages

PROGRAM	FUNCTION
CREATE	In areas where a GBF does not exist, CREATE establishes a file from locally coded information. In areas where a GBF already exists this program permits a large number of new records to be added to the file.
FIXDIME II	This program: 1) edits locally prepared correction inputs to the GBF for completeness and consistency; and 2) inserts the accepted corrections into the file. FIXDIME 2 does not include the insertion of x-y coordinates. FIXDIME C includes x-y coordinate information.
ADDEDIT-L	ADDEDIT-L consists of two programs (a preprocessor and an edit program) separated by a sort program. Among the options available are: 1) flagging only selected errors; and 2) complete listing of all segments of a feature which has a flagged record or listing of flagged record and segments on either side only.
TOPOEDIT	This program edits network features to determine their validity. There are several options including an option to edit only records in certain tracts. TOPOEDIT consists of two programs (PRESORT and TOPOEDIT) separated by a system sort.
FIXCORD	This program inserts missing or corrected x-y coordinate values. The process involves inputting distance (measured clerically) between points of known values to those nodes lacking coordinate values. FIXCORD calculates node values in each of three coordinate systems (state plane, latitude and longitude, and map set miles) used in GBF/DIME files.
UPDIME	This is a comprehensive program, designed for large computers (180K or larger of usable core). It includes many of the functions of FIXDIME II, FIXCORD, and TOPOEDIT. It does not include the functions of ADDEDIT.

Source: Bureau of the Census, _Present Status and Future Prospects of the Census Bureau's GBF/DIME CUE Program_, April 1974.

Figure 3.13 CUE: Correction, Update, and Extension Program dent upon having information with a structured, city type, address. Non-structured or oddly structured addresses such as intersection addresses, major generator addresses, names of places, and rural federal delivery addresses are not suitable for ACG files. However, the Bureau of the Census is conducting a continuation program for ACG/DIME. This program will focus on the extension of city type addresses in rural areas, the provisions for non-structured address coding and the development of GBF maintenance/update procedures for the over 200 SMSAs which now have ACG/DIME geographic base files (See Figure 3.13).

References:

U.S. Department of Commerce. Bureau of the Census. _1970 Census Users Guide, Part 1_. Washington, D.C.: U.S. Government Printing Office, 1970.

U.S. Department of Commerce. Bureau of the Census. Census Use Study. _The DIME Geocoding System: Census Use Study Report No. 4_. Washington, D.C.: U.S. Government Printing Office, 1970.

U.S. Department of Commerce. Bureau of the Census. Geography Division. _Use of Address Coding Guides in Geographic Coding: Case Studies_. Conference Proceedings, November 19-20, Wichita, Kansas. Washington, D.C.: U.S. Government Printing Office, 1970.

U.S. Department of Commerce. Bureau of the Census. Geography Division. _Geographic Base Files: Plans, Progress, and Prospects_. Conference Proceedings, April 1-2, Jacksonville, Florida. Washington, D.C.: U.S. Government Printing Office, 1971.

U.S. Department of Commerce. Bureau of the Census. Geography Division. _Geographic Base File System: Uses, Maintenance, Problem Solving_. Conference Proceedings, November 16-17, Arlington, Texas. Washington, D.C.: U.S. Government Printing Office, 1971.

U.S. Department of Commerce. Bureau of the Census. Geography Division. "Introduction to Small Area Geographic Subdivisions for Which the U.S. Bureau of the Census Collects and Tabulates Data." Prepared by Valerie McFarland and Ann D. Casey of the Data User Services Office. Washington, D.C.: Geography Division, U.S. Bureau of the Census, 1972.

Eckler, A. Ross. _The Bureau of the Census_. New York: Praeger Publishers, 1972. (Mr. Eckler was Director of the Bureau of the Census, 1965-1969).

U.S. Department of Commerce. Bureau of the Census. Geography Division. _Geographic Base File System: Establishing a Continuing Program_. Conference Proceedings, January 18-19, Seattle, Washington. Washington, D.C.: U.S. Government Printing Office, 1973.

U.S. Department of Commerce. Bureau of the Census. Census Use Study. _DIME Workshops: An Interim Report_. Washington, D.C.: U.S. Government Printing Office, 1973.

U.S. Department of Commerce. Bureau of the Census. Geography Division. _Geographic Base File Systems: A Forward Look_. Conference Proceedings, April 16-17, Boston, Massachusetts. Washington, D.C.: U.S. Government Printing Office, 1974.

3.2 NATIONAL LOCATION CODE

The National Location Code (NLC) is a geocoding system designed to provide adequate small area coverage of the United States for the purpose of coding the geographic aspects of national damage assessment and resource evaluation programs. It is maintained by the Office of Preparedness, U.S. General Services Administration (formerly the Office of Emergency Preparedness) and has been in use since 1956.[1] The full NLC is composed of three code segments: region, state, area and county (RSAC) are identified in the first seven digits; standard location area (SLA) is designated by the next four digit serial number; and the final digits are geodetic coordinate readings to the nearest second for the estimated center of population or the approximate geographic center of the SLA.

3.2.1 Region/State/Area/County (RSAC)

In 1970 the RSAC portion of the National Location Code was revised to reflect the geopolitical boundary changes that occurred since 1960 and to conform with established federal geocoding standards.

---

[1]The first National Location Code (NLC) was developed in 1956 by the Stanford Research Institute as part of a damage assessment system for the Federal Civil Defense Administration. The National Resource Evaluation Center adopted NLC and also used the code in its damage assessment and resource evaluation programs. In 1960, the Office of Civil Defense (OCD) and the Office of Emergency Preparedness (OEP) adopted a revision of NLC that was prepared by the Geography Division of the Bureau of the Census. In 1970, NLC was revised once again by OEP. However, OCD elected to abandon NLC in favor of the Bureau of the Census MEDList codes (See Section 3.1).

The first order geopolitical subdivision of the revised RSAC code is the standard federal region (See Section 4.3) designated by a single digit, either 1 through 9 or A for the tenth region, in the first position of the code (See Figure 3.14). The second digit identifies states within federal regions. The District of Columbia and outlying areas under the jurisdiction of the United States are treated as states in the assignment of codes. Specifically, Puerto Rico and the Virgin Islands are coded as states in federal region 2; the District of Columbia is coded as a state in region 3; the Canal Zone is coded as a state in region 4; and American Samoa and Guam are coded as states in region 9. Within each federal region, states (including the District of Columbia) are coded in alphabetic sequence of state name. Outlying areas assigned to federal regions 2,4, and 9 are assigned codes in alphabetic sequence following the states. Except for the Virgin Islands, all coding is numeric (the letter A is used for the Virgin Islands). The two character state abbreviation shown in the name field of the new code conforms to the Federal Information Processing Standard for state abbreviations.

The one digit entry following state code distinguishes between the SMSA and non-SMSA portions of a state (or state equivalent as noted above). A "1" is used to identify the SMSA portion of a state. A "2" identifies the non-SMSA SEA part of a state (See Section 2.1 Item 5). The closest thing to a "type of area" designation in the old RSAC Code was the use of asterisks (*) to distinguish among SMSA, "special" and "residual" areas recognized in the old code.

Source: Office of Emergency Preparedness

Figure 3.14 National Location Code Regions

Individual SMSAs and non-SMSAs are identified in the next three positions. In the case of SMSAs, a unique code is used for each SMSA so that records for resources located in SMSAs may be brought together, if desired, regardless of existing state and regional boundaries. In order to conform as closely as possible to the newly established FIPS codes, the first three digits of the FIPS four position SMSA codes are used as the NLC SMSA Code. With only three exceptions, the first three digits of the FIPS SMSA code are unique and the fourth position contains zero. Thus the NLC three digit truncation does not actually alter the standard SMSA code. Individual SEAs within the non-SMSA portions of states have been assigned three digit numeric codes beginning with 001. The assignment of codes is unique only within state. The sequence used for the assignment of codes within states reflects the arrangement of SEAs in Bogue and Beal's _Economic Areas of the United States_. The sequence is generally from west to east beginning in the northwest corner of the non-SMSA portion of a state (See Figure 3.5).

The final digit of the new RSAC code is a county designation which identifies counties within their SMSA or non-SMSA SEA portion of a state. Where states are subdivided into units other than counties such as Parishes in Louisiana, Divisions in Alaska or Municipios in Puerto Rico, these areas or divisions are treated as county equivalents. Independent cities are also treated as counties in the assignment of county codes. Within SMSAs and SEAs, counties (or their equivalents) are assigned numeric codes 1 through 9 in alphabetic sequence of name. Where more than nine counties are included in an area, alphabetic characters beginning with A are used following the numeric code entries. To avoid possible confusion with the numerals one and zero, the letters L and O are not used (See Figure 3.15).

The four digit region/state/area/county coding structure is, like the nine digit state/county/place codes, a common configuration. The Interstate Commerce Commission's Waybill Statistics were coded in a similar manner for many years and the Bureau of Public Roads used such a code in processing a nationwide highway travel study. The major advantage of an RSAC coding structure is the intermediate level of aggregation called area which consists of county clusters.

3.2.2 Standard Location Areas (SLAs)

The next four digits of the National Location Code comprise the SLA serial number which identifies an individual SLA within a county. In counties tracted by the Bureau of the Census, each tract is established as a separate SLA. Tracts lying both inside and outside of cities with 50,000 people or more are treated as two tracts, and SLAs are established for each portion of the "split" tract. This procedure guarantees that SLA and city boundaries will conform for cities of 50,000 people or more.

In untracted cities of over 25,000 people for which ward data are available, wards are used as separate SLAs. In cities of 25,000 or more inhabitants which are not divided into either tracts or wards,

Volume II - Region, State, Area and County Codes
Part A - Alphabetic Sequence of States and Counties

State and County Name	RSAC Code					FIPS Code		
	Region	State	Type	Area	County	State	County	SMSA
Allen County, Kansas	7	2	2	009	1	20	001	0000
Anderson County, Kansas	7	2	2	009	2	20	003	0000
Atchison County, Kansas	7	2	2	008	1	20	005	0000
Barber County, Kansas	7	2	2	001	1	20	007	0000
Barton County, Kansas	7	2	2	002	1	20	009	0000
Bourbon County, Kansas	7	2	2	009	3	20	011	0000
Brown County, Kansas	7	2	2	008	2	20	013	0000
Butler County, Kansas	7	2	2	904	1	20	015	9040
Chase County, Kansas	7	2	2	007	1	20	017	0000
Chautauqua County, Kansas	7	2	2	007	2	20	019	0000
Cherokee County, Kansas	7	2	2	010	1	20	021	0000
Cheyenne County, Kansas	7	2	2	002	2	20	023	0000
Clark County, Kansas	7	2	2	001	2	20	025	0000
Clay County, Kansas	7	2	2	005	1	20	027	0000
Cloud County, Kansas	7	2	2	005	2	20	029	0000
Coffey County, Kansas	7	2	2	009	4	20	031	0000
Comanche County, Kansas	7	2	2	001	3	20	033	0000
Cowley County, Kansas	7	2	2	007	3	20	035	0000
Crawford County, Kansas	7	2	2	010	2	20	037	0000
Decatur County, Kansas	7	2	2	002	3	20	039	0000

Source: Office of Emergency Preparedness

Figure 3.15 National Location Code Region, State, Area, and County Designations

SLAs are defined as clusters of enumeration districts. The ED clusters average 5,000 residents each. Urban places of 2,500 to 25,000 people outside of tracted areas are established as separate SLAs. Although tracts, wards and urban places of 2,500 inhabitants are considered important enough to be treated as separate SLAs, thinly populated minor civil divisions (MCDs) and census county divisions (CCDs) in untracted rural areas are considered excessive detail and some consolidation of these divisions is necessary. To hold the number of SLAs to an acceptable figure (i.e., approximately 43,000), contiguous MCDs and CCDs are grouped (providing their land area does not exceed 500 square miles) to form SLAs with populations of not less than 2,000.

Serial numbers are assigned to SLAs sequenced in one of the following three ways: 1) SLAs in rural counties are placed in MCD and CCD alphabetic sequence with SLAs for places of 2,500 to 25,000 inserted in the sequence just after the MCD or CCD in which they are located; 2) SLAs in tracted counties are first grouped by places of over 50,000 residents, then by places of 25,000, and finally, by balance of county. Within places and balance of county, SLAs are sequenced by tract number; 3) SLAs in partially tracted counties are grouped first by tracted area and then by untracted balance of county. Within the tracted area, SLAs are sequenced by tract number. SLAs in the untracted balance of county follow this sequence in MCD or CCD alphabetical sequence.

The four digit serial number used to identify SLAs is actually a contraction of the five digit serial number used to identify the small areas included in the census tabulation programs. The first four digits designate the SLA in which the small area is located. The fifth digit or suffix identifies the small areas within an SLA. If a small area is a complete SLA, the fifth position of the SLA serial number is zero filled. If, on the other hand, an SLA is a grouping of two or more small areas, the fifth position reflects in general the alphabetic sequence of small areas within the SLA.

3.2.3 Geographic Coordinates

In 1960, the absolute geographical position of each SLA was identified by both geodetic (latitude and longitude) and Universal Transverse Mercator (UTM) coordinates. Geodetic coordinate readings were to the nearest second, and UTM coordinates were recorded to the nearest 100 meters. For each SLA, the coordinate readings represent the estimated center of population or the approximate geographic center of the SLA. However, as of the 1970 Census of Housing and Population, only latitude and longitude coordinates will be designated to each SLA.

References:

U.S. Department of Commerce, Bureau of the Census. National Location Code: OCP-OEP Region 2. FG-D-3 1/2. Prepared for the Office of Civil Defense, Department of Defense, and the National Resource Evaluation Center, Office of Emergency Planning. Washington, D.C.: U.S. Government Printing Office, 1962.

Office of Emergency Preparedness. Geographic Codes for Region, State, and County. ISG-111. Washington, D.C.: Office of Emergency Preparedness, 1971.

3.3 DEPARTMENT OF DEFENSE

Because of the magnitude of its involvement in automatic data processing, the Department of Defense (DOD) has generated literally thousands of data elements and codes for use in DOD information systems and also undertaken "an aggressive program to discipline the data language and its related codes" so that any potential communication barrier between systems due to variable data vocabularies can be avoided.[1] The primary product of this Data Elements and Data Codes Standardization Program is a manual containing all DOD standard data elements and their related features, including data chains. This manual provides a single, authoritative reference source for all of DOD's information processing vocabulary and is used by all Department of Defense components concerned with development, use and maintenance of data systems.

Among the many codes maintained in the DOD Manual for Standard Data Elements there are nine geographic and geopolitical coding systems along with a standard for expressing longitude and latitude to the decisecond. In accordance with the procedures set forth in the DOD data elements standardization program the Assigned Responsible Agency (ARA) for these geographic codes is the Defense Intelligence

---

[1]U.S. Department of Defense, Office of the Assistant Secretary of Defense (Comptroller), _Department of Defense Manual for Standard Data Elements_, 5000.12 M (Washington, D.C.: Assistant Secretary of Defense, 1970), p. I. The Data Elements and Data Codes Standardization Program is authorized by DOD Directive 5000.11, December 7, 1964 and under the responsibility of the Assistant Secretary of Defense (Comptroller).

Agency (DIA).[1]

The DOD manual contains coding systems for:

International Geopolitical Entities
- Countries of the World
- States/Provinces of Countries of the World

National Geopolitical Entities
- States of the United States
- Counties of the United States
- Standard Metropolitan Statistical Areas of the United States
- Congressional Districts of the United States

Geographic Entities
- Celestial Bodies of the Solar System
- Divisions of the World (Continents, Seas, and Oceans)
- Water Bodies of the World
- Latitude and Longitude (degrees, minutes, seconds, milleseconds, deciseconds, and terrestrial hemisphere)

All of the DOD codes for national geopolitical entities contained in the manual conform for FIPS standards. In DOD information systems, states are identified by the two digit numeric codes published in FIPS PUB 5; counties are identified by the three digit numeric code published in FIPS PUB 6 and used in conjunction with the state codes for unique identification; SMSAs are identified by the four digit numeric code authorized by the Office of Management and Budget and published in FIPS PUB 8; and, Congressional Districts are identified by the combination four digit state/district code published in FIPS PUB 9. The only variances from the FIPS standard are DOD's use of a variable three or four character state abbreviation (as opposed to the two character FIPS standard) and the use of a two digit alpha

---

[1]Refer to Appendix C, International Geocoding Systems, for a description of the geocoding activities of the Defense Intelligence Agency.

numeric designation beginning A1 through Zi for each representative at large of a state.

In addition to the geographic and geopolitical codes contained in the manual, DOD has developed a Geographic File (GEOFILE) of Geolocation (GEOLOC) codes for locations or places of military significance. The places included in this file range from missile early warning stations to water system filter plants identified within data systems of the Worldwide Military Command and Control Systems and those other systems within the Department of Defense that use GEOLOC codes. As it is explained in geolocation code request documents:

> An attempt to prescribe a code to represent a location/place raises the question as to what constitutes a location/place. There are no generally acceptable definitions available which can be relied upon for guidance. Consequently. . . there is an element of arbitrariness involved in determining what is to be assigned a GEOLOC. This factor of arbitrariness also applies to the determination of what items are to be included in a list of authorized installation types.[1]

The record layout for GEOFILE is illustrated in Figure 3.16. Each record contains the following items:

> GEOLOC. A four character alpha numeric entry which represents the specified geographic location or place of military significance. The geolocation codes for water areas are four character combined codes, ØØ preceding two alphabetic characters.
>
> GEONA. A maximum of 17 alpha characters which identify with a distinguishable and meaningful name the

---

[1] U.S. Department of Defense, Office of the Assistant Secretary of Defense (Comptroller), "Specified Geolocation Code Request," JCS PUB 6, Volume II, Part 6, Chapter 5 (Washington, D.C.: Assistant Secretary of Defense, 1972), p. 6-5-16.

GEOLOC	GEONA	GEODE	CRTCD	CRTCD NAME	LPR	POINT	GEOPE	GEOPE NAME
AABA	A Ludi	AFD	VS	Viets	5D	161600N1071400E	002	Thua Thien
AABF	A Ro	AFD	VS	Viets	5D	153600N1072900E	003	Quang Nam
AABK	A Shau	AFD	VS	Viets	5D	160700N1072000E	002	Thua Thien
AABP	A Tou	AST	VS	Viets	5D	154900N1072100E	003	Quang Nam
AABR	Abadan	APT	IR	Iran	7E	302158N0481356E		
AABS	Aabenraa	CTY	DA	Denmk	4J	550200N0092600E		
AACA	Aachen	AIN	GE	Germy	4K	504600N0060600E		
AADG	Aalen	CTY	GE	Germy	4K	485000N0100600E		
AADT	Aarhus	PRT	DA	Denmk	4J	560600N0101800E		
AADU	Aalesund	PRT	NO	Norwy	4T	592249N0104709E		
AAEL	Abadan	PRT	IR	Iran	7E	302000N0481700E		
AAFK	Abadhiri	CTY	JA	Japan	5H	424800N1414000E		
AAGE	Abbotsford	APT	CA	Canad	3X	490126N1222158F		
AAKR	Aberdeen	PRT	UK	Unite	4L	570900N0793000W		
AALS	Aberdeen	PRT	53	Wash	3B	465900N1234900W		
AAPP	Abidjan	ADM	IV	Ivyco	7A	051900N0035800W		
AARR	Abilene	RTC	48	Texas	3C	322600N0994500W		
AAUT	Abingdon	RTC	51	Va	3A	384900N0770800W		

Source: Department of Defense

Figure 3.16 Department of Defense Geolocation Code specifed geographic location.[1]

GEODE. A three character alphabetic entry which identifies the type of installation present at the location represented by a GEOLOC. There are approximately 70 authorized installation codes.

CRTCD. A two character alpha numeric entry which represents a country or water body of the world, or a two digit number representing a state of the United States.

CRTCD Name. A maximum of 5 alpha characters which provide a distinguishable and meaningful abbreviation of a country, water body or state name.

LPRCO. A two character alpha numeric code designated to meet the needs of the military logistics community. This Logistics Planning and Reporting Code divides the world into land areas and ocean areas.

POINT. The 15 digit alpha numeric characters which indicate the intersecting lines of latitude and longitude in seconds that determine the coordinate point of a geographic location.

GEOPC. A three digit number which provides for the coding of political subdivisions of countries of the world (provinces or states) or a county of a state in the United States (Currently this entry is required only for locations in the Republic of Vietnam).

GEOPC Name. A maximum of 14 alpha characters which identify with a distinguishable and meaningful name the political subdivisions of countries of the world or a county of a state of the United States (Also currently required only for location in the Republic of Vietnam.)

---

[1]The spelling for geolocation names are taken from one of the following sources: Locations in the United States - The National Atlas published by the Department of the Interior; locations outside of the United States - listings of Official Standard Names by Country coordinated and published by the Central Intelligence Agency (commonly called the CIA gazetteer); and for airfields - DOD flight information publications published by the Defense Mapping Agency.

There are approximately 27,000 GEOLOC codes assigned throughout the world. Unlike most other place or location codes, GEOLOC is not hierarchically structured and therefore independent of any other state, county, or province designations. Furthermore, the codes assigned are completely nonsignificant, having no mnemonic implications or sequence logic. Eliminating all the unacceptable four letter combinations ("dirty words"), GEOLOC can be used to uniquely identify over 400,000 geographic locations in four characters as opposed to the minimum of six digits required in most other place codes. This shortest practical code concept has proved to be most cost effective in meeting the total data requirements of DOD data systems.

In the near future DOD intends to develop a conversion and cross reference file for GEOLOC and geographic location codes maintained by the General Services Administration (GSA). First the Air Force "Cities of the World" file will be matched with the GEOLOC file and all unmatched Air Force cities will be added. Then the GSA tape will be matched on location name with the expanded GEOLOC file. Any locations in the GSA tape which do not match a GEOLOC will be assigned a DOD four character code and added to the file. GEOLOC codes not contained in the GSA file will be researched and assignment of a GSA code will be requested for that location.

Although the DOD GEOFILE contains many codes for locations designated simply as general geopolitical entities such as towns and cities, the universe of locations included in this file is clearly specialized with the overwhelming majority of GEOLOC codes identifying military installations ranging from ammunition storage depots to navigational aids. Much like the city codes of the General Services Administration (See Section 2.2), the relatively large number of GEOLOC assignments in the Republic of Vietnam (over 700 locations) reflects the extent of this location specialization.

The GEOLOC and all the other geographic and geopolitical codes in the DOD manual of standard data elements are used extensively in most of the data systems within the DOD which require geographic references.

References:

U.S. Department of Defense. Office of the Assistant Secretary of Defense (Comptroller). Department of Defense Manual for Standard Data Elements. 5000.12 M. Washington, D.C.: Assistant Secretary of Defense, 1970.

U.S. Department of Defense. Office of the Assistant Secretary of Defense (Comptroller). "Specified Geolocation Code Request." JCS PUB 6, Volume II, Part 6, Chapter 5. Washington, D.C.: Assistant Secretary of Defense, 1972.

3.4 WATERBORNE COMMERCE STATISTICS

The U. S. Army Corps of Engineers maintains a clearinghouse for information on the flow of waterborne commerce called the Waterborne Commerce Statistics Center in New Orleans, Louisiana. The center is responsible for compiling, tabulating, printing and distributing waterborne commerce statistics. In fulfilling these tasks, the Corps of Engineers collects, codes and processes data recorded on vessel operation reports. The data which are coded from these vessel operation reports include: vessel identification, vessel type, flag, port, dock, commodity, operator, service, alternative routes and traffic. In this system, the geographic and descriptive codes are completely integrated. Of all the codes used in compiling Waterborne Commerce Statistics, only the vessel code, the port code, the dock code, and the route code are geographic identifiers.

The vessel code consists of seven digits. The first two digits, called the flag code, indicate United States or foreign registry. If a vessel is registered in the United States, the flag code identifies the district of registry. These districts are defined by the by the U. S. Army Corps of Engineers (Figure 3.17). The remaining five digits of the vessel code is the vessel registration number. This is a number assigned to each vessel by the Waterborne Commerce Statistics Center. It remains with the vessel until it is sunk, abandoned, or sold foreign.

The port code is a five digit number. The first digit indicates

Code	District	Code	District
01-02	New England Division	28	Rock Island, Ill.
03	New York, N. Y.	29	St. Louis, Mo.
07	Philadelphia, Pa.	30	Memphis, Tenn.
08	San Juan, P. R.	31	Vicksburg, Miss.
09	Baltimore, Md.	32	New Orleans, La.
10	Washington, D. C.	33	Galveston, Tex.
11	Norfolk, Va.	34	Little Rock, Ark.
12	Wilmington, N. C.	35	Kansas City, Mo.
13	Charleston, S. C.	36	Seattle, Wash.
14	Savannah, Ga.	37	Portland, Ore.
15	Jacksonville, Fla.	38	Alaska
16	Mobile, Ala.	39	San Francisco, Cal.
17	Nashville, Tenn.	40	Sacramento, Cal.
18	Louisville, Ky.	41	Los Angeles, Cal.
19	Ohio River Division	42	Honolulu, Haw.
20	Huntington, W. Va.	43	Omaha, Neb.
21	Pittsburgh, Pa.	44	Walla Walla, Wash.
22	Buffalo, N. Y.	45	Tulsa, Okla.
23	Detroit, Mich.	47	U. S. Maritime Adm. Vessels
24	Duluth, Minn.	48	U. S. Defense Craft
25	Milwaukee, Wis.	49	U. S. Vessels (temporary nos.)
26	Chicago, Ill.	50	Canadian
27	St. Paul, Minn.	60	Other Foreign

Source: U.S. Army Corps of Engineers

Figure 3.17 Waterborne Commerce Statistics District Codes major geographical areas of the United States; the second digit indicates the district within the major area; the third, fourth and fifth digits signify individual port, section within a port, waterway or mile station.

The dock code consists of three digits which designate the dock, wharf, or pier within a port. The first digit of the dock code indicates tenths of a mile from a base point within the port, the second digit identifies the wharf or dock at that point, and the third digit consists of the following codes:

Code	Location
1	right bank (downstream)
2	left bank (downstream)
3	right bank tributary
4	left bank tributary
5	island
6	lock, dam or bridge

Port and dock codes are assigned by either the base line, or the mileage method. In the assignment of base line codes, the Great Lakes and coasts are scaled off, and numbers progressively assigned to the ports and localities in rough geographic order, beginning in the New England Division (certain inland waterways have been assigned dock codes according to the base line method, starting at 000 on the left bank at the mouth and scaling up the left bank and down the right bank to 999 on the right bank at the mouth). The mileage code is assigned to a port or dock according to its location in miles above a zero point which is usually the mouth of the river. This method is used in assigning codes to localities on most rivers and canals.

The two digit waterway codes listed in Figure 3.18 are used to identify certain waterways which are alternative routes available to the vessel on a voyage. The route used by the vessel and alternate routes must be stated by the respondent on the vessel operations reports.

References:

U.S. Department of the Army. Office of the Chief of Engineers. "Reports and Statistics: Waterborne Commerce Statistics." Regulation No. 335-2-1, ENGCW-OM. New Orleans, Louisiana: U.S. Army Corps of Engineers, 1970.

Code	Waterway, Channel, or Lock	Code	Waterway, Channel, or Lock
	ATLANTIC COAST		GULF COAST
02	Annisquam Canal, Mass	30	Gulf Intracoastal Waterway at Apalachee Bay, Fla.
04	Pollock Rip Shoals, Nantucket Sound, Mass.	31	The Rigolets, La.
07	Cape Cod, Mass.	32	Calcasieu River and Pass, La.
10	Port of New York: Lower Entrance Channels	34	Atchafalaya River Bar Channel, La.
12	Port of New York: Long Island Sound Entrance	35	Sabine Pass, Texas
16	Chesapeake and Delaware Canal	36	Ship Island Bar Channel, Miss.
18	Knapps Narrows, Md.	37	Horn Island Pass, Miss.
20	Atlantic Intracoastal Waterway	38	Mobile Bay Bar Channel, Ala.
23	Waterway connecting Lookout Light and Back Sound, N. C.	39	Pensacola Entrance Channel, Fla.
24	Waterway from Norfolk to Beaufort Sound (Dismal Swamp)		
25	Waterway from Pamlico Sound to Beaufort Harbor via Core Sound		
26	Okeechobee Cross-Florida Waterway		GREAT LAKES
		60	Superior Entry, Wis.
		61	Duluth Canal, Minn.
		62	Keweenaw Waterway, Mich.
		64	Sturgeon Bay and Lake Michigan Canal, Wis.
		65	Chicago Harbor (Entrance), Ill.
	MISSISSIPPI RIVER SYSTEM	66	Calumet-Sag Channel, Ill.
		67	Erie Barge Canal (Tonawanda), N. Y.
		68	Oswego Canal, N. Y.
40	Mouth of the Mississippi River: Southwest Pass	69	Black Rock Lock (Niagara River), N. Y.
42	Mouth of the Mississippi River: South Pass		
46	Mouth of the Mississippi River: Other Passes		
47	Mississippi River - Gulf Outlet Entrance Channel, La.		PACIFIC COAST
50	Harvey Lock, La.		
51	Algiers Lock, La.	72	Sacramento River Barge Lock, Cal.
53	Atchafalaya River Route, La.	74	Multnomah Channel, Ore.
54	Port Allen Lock, La.	76	Waterway connecting Port Townsend and Oak Bay, Wash
57	Kentucky Lock, Ky.	78	Swinomish Channel, Wash.
58	Barkley Lock, Ky.	80	Wrangell Narrows, Alaska

Source: U.S. Army Corps of Engineers

Figure 3.18 Waterborne Commerce Statistics Waterway Codes

3.5 FEDERAL AVIATION ADMINISTRATION

A wide variety of operational codes are used in the National Airspace System (NAS). They include codes for special procedures, navigational aids, airspace elements, military routes, aircraft types, aircraft companies, and several sets of location identifiers. These codes, when assigned, are authorized for use on charts, flight information publications, written records, NAS computers, and other communication media. They comprise the vocabulary of a very specialized language.

In view of the fact that the increasing use of codes in the National Airspace System may result in duplication and serious operational problems in an automated environment, the Flight Services Division of the Federal Aviation Administration (FAA) has developed an overall, centrally controlled plan for all air traffic operational codes. This plan provides a unique format for each type of requirement. Each series of codes in the plan has the capacity to provide non-duplicating codes for many years and there are provisions for reserving blocks of codes for future, but presently unknown requirements.

The implementation of this plan is especially important in the assignment of location identifiers. The NAS computer program which correlates radar data with flight plan and geographic information and then generates the radar map display for the controller requires that the geographic definitions of locations be refined to a much higher degree of accuracy than is possible by present manual means. Also the use of such automated data processing systems has required that a great number of additional locations be identified for the internal use of the computer flight plan processing program. Thus one of the major thrusts of the Flight Service Division plan is to simplify flight plan filing and processing, minimize encoding and decoding of data, and make the system of location identifiers more efficient. To this end the Cartographic and Technical Standards Branch of the Federal Aviation Administration has revised the coding structure of certain location identifiers and is compiling an automated master file for the operational coding system.

A summary of the location identifiers maintained by FAA is presented in Figure 3.19. There are six major domestic codes, one international code and a series of Canadian codes that are required in air traffic service and related activities. A complete listing of these location identifiers is compiled and updated three times a year by the FAA.

The first code listed in Figure 3.19, a three character alpha designation, is assigned to those airports, heliports and seaplane bases in the United States on which there is established a manned FAA air traffic control facility, an air terminal navigational aid (NAVAID) within the airport boundaries, or which receive Department of Defense airlift service or scheduled route air carrier service directly into the airport. The three letter location identifier is formulated based on the name of the location it is to represent, if

LOCATION	CODE
Major Airports	LLL
Navigational Aids (NAVAIDS)	LLL
Minor Airports	LLNN
Area Navigation Waypoints (RNAV)	LLLLL
Intersection and Distance Measuring Equipment (DME) Fixes	LLLLL
Computer Coordination Fixes	NNNNN
International Civil Aviation Organization (ICAO Identifiers)	LLLL
Canadian Airports and Navigational Aids	LL,LLL, NL,LNL,LLN

L = letter character

N = number character

Source: Federal Aviation Administration

Figure 3.19 Location Identifiers Used in the National Airspace System practicable. For example, Boston's Logan Airport is coded BOS and the San Francisco International Airport is coded SFO. A code with the initial letter of the location name is not always assignable, however, because of the ever decreasing number of suitable codes available. Choice of code is limited also by restrictions on the use of the "K", "N", "Q", and "W" blocks of codes which have been allocated for other purposes.

United States airports, heliports and seaplane bases not meeting the requirements for an identifier in the three letter series are assigned a code for minor airports provided they average at least two general aviation inbound flight plans daily. Formerly, these minor airport designators were a variable combination of two numbers and a letter. However, they will be changed to a standard letter-letter-number-number configuration.

Effective July 30, 1973, the area navigation waypoint designations for charted route waypoints were simplified. The existing two letter, two number, one letter codes were replaced by five letter pronouncable names. This provides a single five letter name that serves as a computer code, the assigned identifier and the way point name. The five letter name/code will also be expanded to include several types of coordination fixes which are assigned primarily to non compulsory reporting points. Computer coordination fixes will, however, retain the five number designations assigned within the block of numbers appropriate to the altitude.

The four letter international location identifier is established by the International Civil Aviation Organization and covers roughly the same type of locations--airports, navigational aids, charted way-points, etc.--as the domestic codes and designations of the FAA. This code is employed in all international telecommunications.

The Canadian location identifiers are variable two and three digit alpha numeric combinations authorized by the Canadian Ministry of Transport. These designations are used in all communications, notices to airmen, and flight plans between facilities of the Federal Aviation Administration and Canadian facilities.

Weather offices and stations in the United States may be assigned codes in any of the series of location identifiers, depending upon type or combination of station. Generally, however, weather service stations in the United States are assigned three letter codes. Other types of aviation weather reporting stations are listed with two letter, two number identifiers.

In addition to the location identifiers for airports the FAA code manual lists the code of the tie-in facility and control center. The tie-in facility is the telecommunications facility that handles flight plan messages for the listed airport. The identifier for the Air Route Traffic Control Center of jurisdiction for each airport is included to assist in routing instrument flight rules messages.

References:

U.S. Department of Transportation. Federal Aviation Administration. Air Traffic Service. "Air Traffic Operational Coding System." Order 7350.2B. Washington, D.C.: Federal Aviation Administration, U.S. Department of Transportation, 1974.

U.S. Department of Transportation. Federal Aviation Administration. Air Traffic Service. Location Identifiers. Order 7350.4. Washington, D.C.: U.S. Government Printing Office, January 1974.

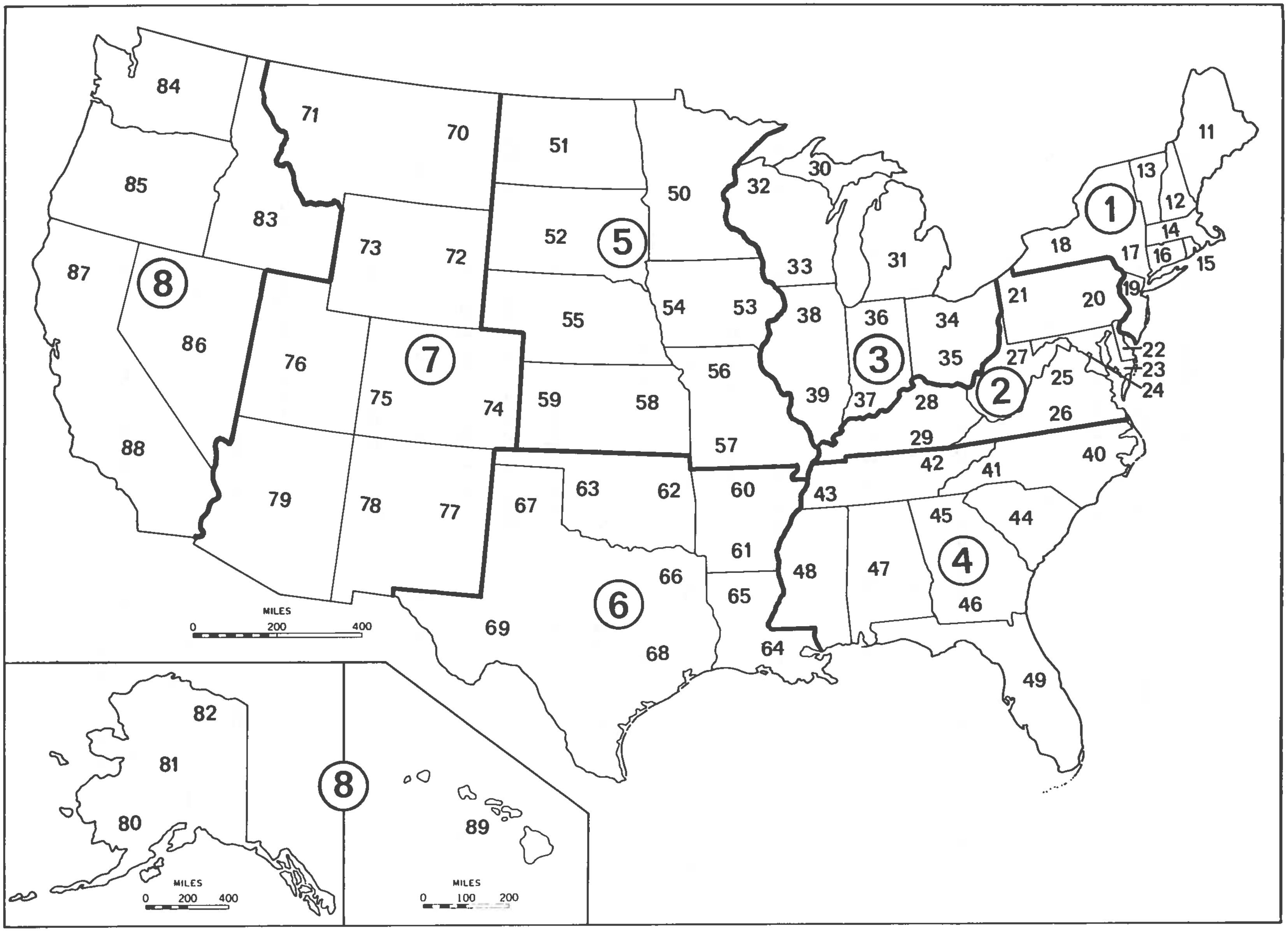
3.6 STANDARD POINT LOCATION CODE

The Standard Point Location Code (SPLC) is a geocoding system designed for transportation purposes which originated as a joint project of the American Trucking Associations (ATA) and the Association of American Railroads (AAR) in 1965. The initial code was compatible for both the truck and rail industries, but over time the ATA version and the AAR version of SPLC developed separately. Although both files were maintained in accord with the general principles of the code, the lack of a central controlling registry led to variations between the universe of points coded. Conceptually, SPLC is based on a system of nesting areal units. Starting at the top of this hierarchy, the United States is divided into eight regions which are composed of contiguous clusters of states. All of these regions are subdivided into ten sections each of which may be a state or a portion of a state (See Figure 3.20). Sections are subdivided into groups of contiguous counties or county equivalents;[1] and these units are further subdivided into counties or county sections.

According to the SPLC coding logic: each region can accommodate 100,000 coded locations; each state or state section can accommodate 10,000 locations; and, each county or county section can accomodate 100 locations. Therefore, code assignment requires that each county

---

[1]The term county equivalent applies to recognized political entities equivalent to county subdivisions such as independent cities of Maryland, Missouri, Nevada and Virginia, parishes in Louisiana and Census divisions in Alaska.



Source: Transportation Data Coordinating Committee

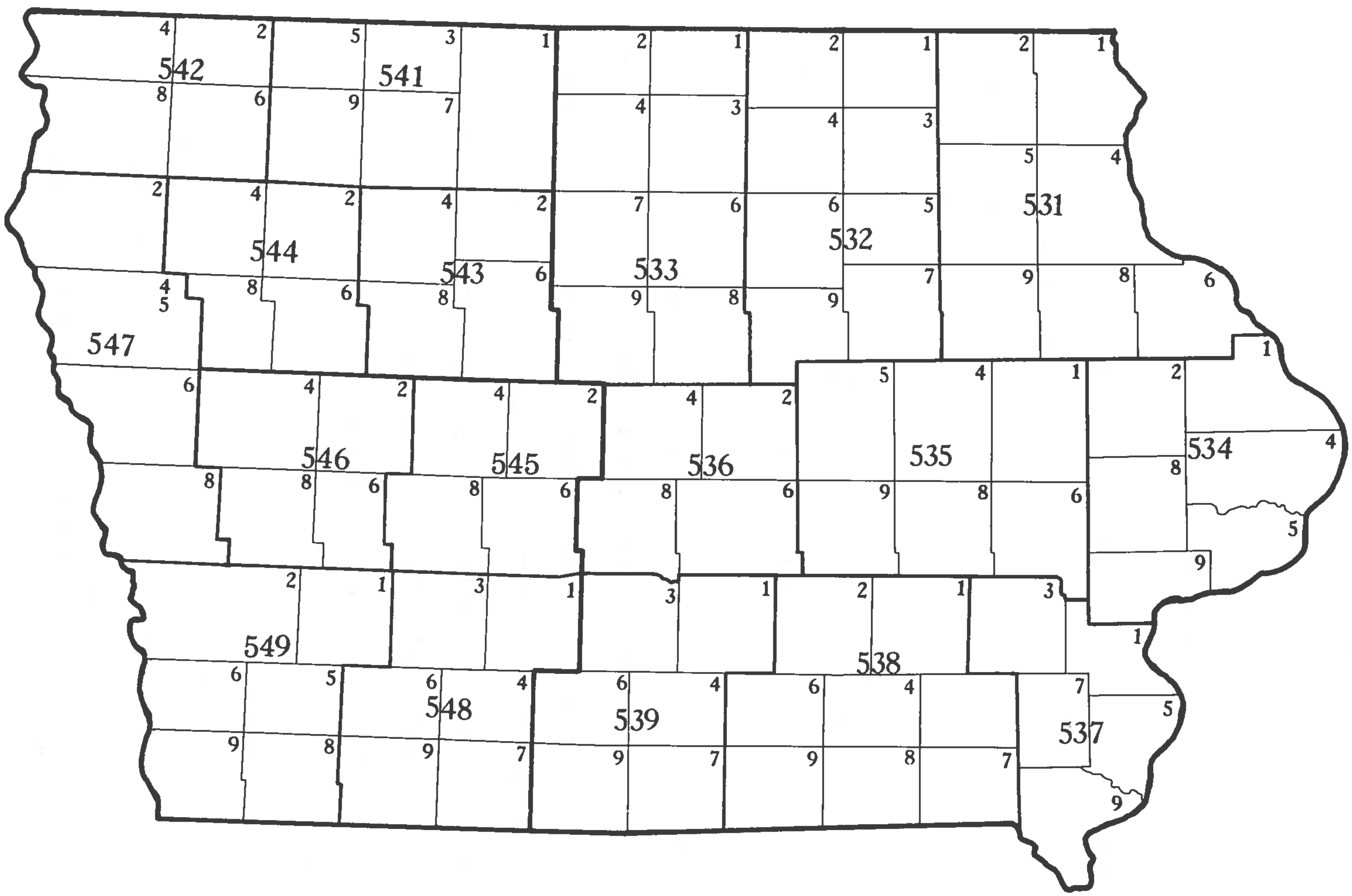
Figure 3.20 Standard Point Location Code: Regions, States, and State Sections or county section be subdivided into as many smaller areal units as necessary to accommodate 100 possible locations in each. As needed, counties of less than 900 square miles or county sections of counties which exceed 900 square miles are divided into nine parts not exceeding a maximum of 100 square miles each. This level of areal unit corresponds roughly with the existing minor civil divisions. In turn, each of these units which do not exceed 100 square miles is partitioned into smaller units not exceeding 10 square miles each, if necessary.

The code assignment plan arranges each digit level by starting at the northwest corner of the previous level area and progressing east to west and north to south in ascending order (Figure 3.21). The first digit of an SPLC code number identifies a region; the second digit identifies a section of that region (state or portion of a state); the third digit identifies a group of counties within that section; the fourth digit identifies a county or county section within that group of counties; the fifth digit identifies an areal unit not exceeding 100 square miles within that county or county section (roughly an MCD); and the sixth digit identifies a 10 square mile unit.

In summary, the basic logic for the nesting geographical units and code assignment of SPLC is as follows:

-first digit = region = clusters of contiguous states = 100,000 locations

-second digit = section = state or portion of state = 10,000 locations



Source: Transportation Data Coordinating Committee

Figure 3.21 Standard Point Location Code: Coding Sequence for State Sections and Counties in Iowa

-third digit = county group = clusters of contiguous counties = 1,000 locations

-fourth digit - county/county section = a county or portion of a county = 100 locations

-fifth digit = minor civil divisions = 100 square mile units = 10 locations

-sixth digit = portion of MCD = 10 square mile unit = 1 location.

The locations which are coded in the ATA version of SPLC are truck carrier tariff points. These are defined as the area (approximately 19 square miles) within a radius of 2 1/2 miles from the center point which may be a post office, city center, cross roads, etc.[1] These locations are coded to the smallest areal unit of the SPLC geographic hierarchy (the 10 square mile unit) in which most of the location is situated.

In practice, the actual configuration and assignment of standard point location codes may vary substantially from the conceptual scheme described above. For the coding of points within metropolitan areas, for example, ATA has modified the SPLC structure as follows: major cities are identified at the three or four digit level and places within the city are coded by replacing the zeros following the three or four digit level with a specific metropolitan code.

Examples:

New York, N.Y................................ 178<u>000</u>

Brooklyn, N.Y................................ 178<u>700</u>

Greenpoint, Brooklyn, N.Y.................... 178<u>721</u>

---

[1]These units are defined in the Interstate Commerce Commission document Docket MC-37.

The specific metropolitan code which is employed as the last two or three digits of SPLC in the major cities depends upon the user. Locations in the city of Los Angeles, for example, are coded to a utility rate zone or a ZIP code zone. If a place is coded to ZIP areas, the fourth digit of the SPLC number will be zero (or the same number for the entire city since the first of the last three ZIP code digits is the same within a major city) and the fifth and sixth digits will be the last two numbers of the ZIP code for the area in which the place is located. If a place is coded to utility rate zone, the last three digits of the SPLC number will be the three digit rate zone code for the zone in which the place is located.

The ATA and AAR versions of SPLC differ in several important features. Many of the locations included in the ATA files are of no interest to the AAR and these have been deleted in the AAR version of SPLC. Conversely, there are locations important to the railroad transportation system which are not coded in the ATA version of SPLC and these have been added. The SPLC file utilized by AAR and printed in the Freight Station Accounting Code Directory (FSAC) contains roughly 60,000 locations in the United States, Canada and Mexico. The SPLC directory developed by ATA and published by the National Motor Freight Traffic Associations (NMFTA) contains approximately 110,000 locations in the United States and 15,000 locations in Canada.

There is also a significant difference in the nature of the locations coded in each file and the use of the two SPLC files. The locations or points coded in the AAR version consist of railroad stations, junctions, sidings or other rail facilities and the file is not used for rating purposes. The locations coded in the ATA version of SPLC are tariff points as described above.

160 manufacturers and 194 carriers who are concerned with coding the geographic aspects of transportation systems have subscribed to the ATA version of SPLC. Roughly 100 shippers and 350 carriers have subscribed to the FASC/SPLC and there are approximately 3,000 copies of this file in use. Of the subscribers who have adopted SPLC, a few of these users have tailored this geocoding system somewhat to suit their particular needs.

As a result of the confusion created by the existence of two major, differing versions of SPLC, the National Motor Freight Traffic Association proposed the creation of a single master file including both versions of the SPLC. The master file would be updated and errors or variations could be resolved. With the help of the U. S. Department of Transportation and support from ATA and AAR this master file has been developed (See Figure 3.22). The file is monitored by a central registry and a 17 man SPLC policy council has been established to deal with such issues as code structure changes, adjustments to existing code designations, the maintenance of historical records, the problem of code saturation, and the overall coordination of the standard point location code.

AAR File Record Layout

SPLC	Station	State	Blank	County Name	Railroad Number	Freight Station Accounting Code	Uniform Alpha Code	Control Field	Blank
6	14	2	1	14	3	5	4	4	1

ATA File Record Layout

SPLC	Finder Code	Unique Code	Change Code	Line Number	SPLC	Blank	Point Name	Blank	County Name	State Code	Blank	Date
6	8	2	1	3	6	2	40	2	20	2	3	5

SPLC Master File Record Layout

Record Code	SPLC	Record Type	Year	Day	Record Number	Point Name	County Name	Multi-County Code	State	Major/Minor Point Code
1	6	2	2	3	2	29	20	1	2	1

Rail Point Code	Truck Point Code	Future Use Code	Obsolete Code	Trans Code	FIPS Code State/County	Future Use Code
1	1	1	1	1	5	21

Source: Association of American Railroads, American Trucking Assocations, and National Motor Freight Traffice Association, Inc.

Figure 3.22 Standard Point Location Code

References:

Transportation Data Coordinating Committee. Standard Point Location Code Congress: Addresses and Panel Discussion. Washington, D.C.: Transportation Data Coordinating Committee, 1970.

Conversations with Mr. Peter L. Conway, Jr., of the Association of American Railroads and Mr. Robert L. Hennell of the American Trucking Associations held on January 11, 1972.

National Motor Freight Traffic Association, Inc., "Statement of Work for the Creation and Maintenance of a Standard Point Location Code Master File." Washington, D.C.: National Motor Freight Traffic Association, Inc., 1973.

Standard Point Location Code Policy Council. "Meeting Minutes," January 1973 through October 1973.

3.7 CENSUS OF TRANSPORTATION: PICADAD

PICADAD is a computerized system developed by the Transportation Division of the Bureau of the Census for automating the geographic aspects of intercity transport. It is designed primarily for tabulations and analyses of movement or spatial relationships between two or more places in the United States. This includes shipments from plant to market, location of distribution points with respect to production and market areas, length of haul, origins and destinations of passenger travel, etc. PICADAD also provides an efficient method for coding and programming tabulations for states, counties and other geopolitical areas. PI stands for Place Identification. CA stands for the Characteristics and Area of the place identified. And, DAD stands for the computer Distance and Direction.

PICADAD was initially developed for use in small scale surveys, because customary mileage block and other standard geographic coding systems were prohibitively expensive and too slow for tabulations and analyses involving travel or traffic flows. In 1963 the system was programmed for large computers in order to tabulate over one million shipping records that were input for the first Census of Transportation.

There are three basic elements within the PICADAD system. These are: the PICADAD place file, the key point reference file, and the complex of computer programs which performs the tabulations and distance calculations. The place file is a listing of all places included in the PICADAD system. Places located in the 48 conterminous states and the District of Columbia are arranged in nominal alphabetic sequence by states in alphabetic order.[1] Each of the 37,000 places in this file is associated with one of the 5,700 key points listed in the key point reference file. While the place file may be expanded indefinitely the universe of key points is fixed.[2]

In addition to a key point serial number, each place in the PICADAD place file is identified by the following additional geographic designators (See Figure 3.23):

- The Federal Information Processing Standard for state (both alpha and numeric), county, and SMSA codes as published in FIPSPUB five, six and eight (See Section 2.1).
- The 1960 Census codes for divisions, states and counties; and the three digit Census version of the FIPS SMSA codes used in the 1972 economic censuses (See Section 3.1).
- The name of each place up to twenty characters long and a six digit state name abbreviation.
- The two digit freight rate territory code established by the Interstate Commerce Commission (See Section 4.8).
- The three digit economic area code established by the Bureau of Economic Analysis (formerly the Office of Business Economics, OBE). (See Section 4.5).
- The three digit transportation zone code which is employed in the national transportation networks of the U.S. Department of Transportation (See Section 3.9).

---

[1]Alaska and Hawaii lie beyond the limits of the plane surface used for arithmetic coordinates, and are treated by a subroutine when needed in specific analysis.

[2]The Census programs for transportation statistics so far have not required a more detailed identification of places and therefore the universe of key points has remained constant. However, if each place required unique identification, this could be done readily by adding digits to the key point code. The use of subcodes would identify each place without forcing any change in the other parts of the system.

FIPS State Code	Key Point Number	Alpha State Code	FIPS County Code	Census Division/State Code	Census County Code
2	4	2	3	2	3

State Abbreviation	Place Name	North-South Coordinates	East-West Coordinates	Census SMSA Code	Production Area Code
6	20	4	4	3	2

Freight Rate Territory Code	Market Area Code	OBE Economic Area Code	FIPS SMSA Code	DOT Transportation Zone Code
2	2	3	4	3

Source: Bureau of the Census, Transportation Division

Figure 3.23 PICADAD Place File

- A two digit production area code and a two digit market area code developed for the PICADAD system by the Transportation Division of the Bureau of the Census.

- Four digit north-south plane coordinates and four digit east-west plane coordinates for the approximate geographic center of each place employed as a key point.

Production area and market area codes are the only geographic identifiers unique to the PICADAD system. Production areas consist of a single SMSA or cluster of SMSAs selected to represent relatively compact concentrations of manufacturing activities. Market areas consist of a single SMSA or cluster of SMSAs which represent areas of concentrated marketing activities. These units are employed in order to geographically reference commodity movements without disclosing the activities of any single company or individual. There are 27 designated production areas (coded 01 through 27) and 23 market areas (coded 31 through 55). If a place is not located within a production area it is identified as either "in an SMSA but not in a production area" or "not in an SMSA." Similarly places not located in a market area are identified as "in an SMSA but not in a market area" or "not in an SMSA." (See Figure 3.24).

The second element of this system, PICADAD's key point reference file, contains a fixed set of about 5,700 "key points" selected to represent all places in the place file (See Figure 3.25). The prime reason for creating a fixed set of key points is to gain the benefits of "canned" computer routines, and still retain the ability to introduce new places into the place file whenever needed for coding or place identification purposes. Key points are selected from the

PRODUCTION AREAS

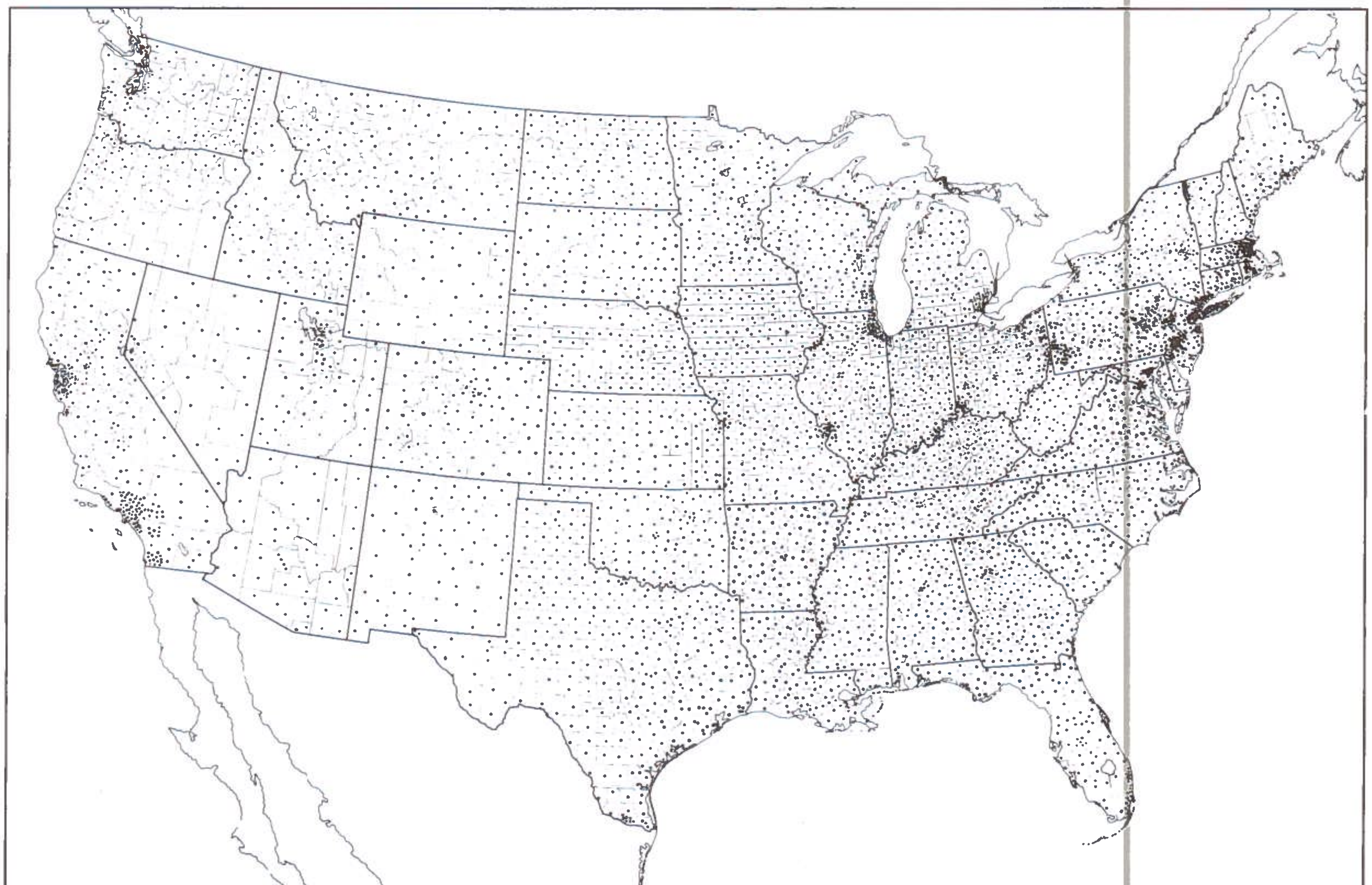
Code	SMSAs Included
1	Boston, Worcester, Providence-Warwick-Pawtucket, Brockton, Lawrence-Haverhill, Lowell
2	Hartford, New Britain, Meriden, Waterbury, New Haven, Bridgeport, Norwalk, Stamford, Springfield-Chicopee-Holyoke
3	New York
4	Newark, Jersey City, Paterson-Clifton-Passaic, Middlesex County, Somerset County
5	Philadelphia, Wilmington, Trenton
6	Baltimore
7	Allentown-Bethlehem-Easton, Reading
8	Harrisburg, Lancaster, York
9	Syracuse, Utica-Rome, Albany-Schenectady-Troy
10	Buffalo, Rochester
11	Cleveland, Akron, Canton, Loraine-Elyria, Youngstown-Warren, Erie
12	Pittsburgh, Steubenville-Weirton, Wheeling
13	Detroit, Flint, Toledo, Ann Arbor
14	Cincinnati, Dayton, Hamilton-Middletown, Springfield
15	Chicago, Gary-Hammond-East Chicago
16	Milwaukee, Kenosha, Racine
17	Minneapolis-St. Paul
18	St. Louis
19	Atlanta
20	Dallas, Fort Worth
21	Houston, Beaumont-Port Arthur-Orange, Galveston-Texas City
22	Denver
23	Seattle-Everett, Tacoma
24	San Francisco-Oakland, Vallejo-Napa, San Jose
25	Los Angeles-Long Beach, Anaheim-Santa Ana-Garden Grove, Riverside-San Bernardino-Ontario
26	Indianapolis, Muncie, Anderson
27	Kansas City, St. Joseph, Topeka

MARKET AREAS

Code	SMSAs Included
31	Scranton, Wilkes-Barre-Hazleton, Binghamton
32	Washington, D.C.
33	Newport News-Hampton, Norfolk-Portsmouth
34	Columbus (Ohio)
35	Grand Rapids, Muskegon-Muskegon Heights
37	Louisville
38	Nashville
39	Memphis
40	Augusta, Columbia
41	Fort Lauderdale-Hollywood, Miami, West Palm Beach
42	Birmingham, Tuscaloosa
43	Tampa-St. Petersburg
44	Mobile, Pensacola
45	New Orleans
46	Omaha, Lincoln
48	Oklahoma City, Tulsa
49	San Antonio, Austin
50	Salt Lake, Provo-Orem, Ogden
51	Phoenix, Tucson
52	Portland (Oregon)
53	Sacramento, Stockton
54	Fresno, Bakersfield
55	San Diego
	Additional Designations
	In an SMSA but not in a production area
	Not in an SMSA
	In an SMSA but not in a market area

Source: Bureau of the Census, Transportation Division

Figure 3.24 Production and Market Areas for the 1972 Census of Transportation



Source: Bureau of the Census, Transportation Division

Figure 3.25 PICADAD Key Points

total universe of places on the following basis:

- One key point, generally the county seat, was selected in each county;

- Additional key points were selected in counties having large ground areas. The objective is to hold the maximum distance between each pair of points to not more than 50 miles, except in sparsely populated areas;

- To assure reliable measurement of distances between all major cities in the United States, each city or town with a population of 5,000 or more in the 1960 Census was included in the key point file without regard to its geographic relationship to other nearby places.

All other places contained in the place file are "keyed" to the nearest key point location by the PI four digit serial number. Key point serial numbers were originally assigned in an ordered fashion which distinguished the low number SMSAs from the high number non-SMSA places, however, expansion of the file has made it impossible to maintain this systematic assignment of numbers. The PICADAD key point code is therefore a semi-random serial number. Coordinates of longitude and latitude were assigned to the key points in one of two ways. Those key points which appear in C. A. Whitten's table of plane coordinates for selected places were assigned the longitude and latitude recorded in that listing.[1] This table covered approximately 10 percent of the key points. The key points which did not appear in C. A. Whitten's table were assigned coordinates scaled from a Rand McNally Atlas and mathematically converted to plane coordinates

---

[1]C. A. Whitten, Air Line Distances Between Cities in the United States, Special Publication No. 238 (Washington, D.C.: U.S. Coast and Geodetic Survey, 1947). pp. 238-241.

within the same axes as established by C. A. Whitten. The north-south coordinates range from slightly more than zero at the southern tip of the United States to about 2,000 at the northern boundary. To avoid possible clerical or programming errors, the east-west coordinates were set in a higher range--from about 3,000 at the east coast to 7,000 on the Pacific coast.

The straight-line (air-line or great circle distances) mileage between any two key points is calculated as the hypotenuse of the right triangle formed by the coordinates of those key points. Specifically, the computations are:

	North Coordinates		West Coordinates
From New York..............	1,298		3,862
To Chicago.................	1,256		4,570
Legs of the triangle...	+42		-708
Computed miles (i.e., hypotenuse)	$(42)^2$	+	$(708)^2$
		503,028	
		709 miles	

A more accurate measurement of the straight-line mileage is 713 miles. The 4 mile difference is caused by simplifying the PICADAD computations to merely the square root of the sums of the squares of the two legs of the triangle, without making adjustments for the extent to which the Lambert projection fails to achieve a perfect plane. The computational error for any given pair of points is a maximum $\pm$ 25 miles. In high density traffic belts, the error is trivial and reaches maximum only at the fringes where little traffic moves. Adjustments can be programmed in the computer, but the improvement in accuracy is not considered to be worth the added cost and increased machine time for statistical tabulations.

Although none of the census studies has involved the computation of direction of flow, this directional aspect can be programmed on the basis of signs of the differences between two coordinates. In addition to the computation of distances between two specific key points, the longitude and latitude readings associated with each key point in the file permits the structuring of key points in grid order. By pairing the digits of the north-south and east-west coordinate numbers, the key points can be grouped into grid zones. In this manner, key points may be identified strictly according to geographic proximity as well as parts of various, larger geopolitical units (See Figure 3.26).

References:

Church, Donald E. "PICADAD: A System for Machine Processing of Geographic and Distance Factors in Transportation and Marketing Data." Washington, D.C.: Transportation Division, U.S. Bureau of the Census, August 1970.

U.S. Department of Commerce. U.S. Bureau of the Census. Transportation Division. _PICADAD: A Computer File of Places with Geographic Characteristics and a Straight-Line Distance Computation Method._ Public Use Tape Users' Manual. Washington, D.C.: Transportation Division, U.S. Bureau of the Census, March 1974.

State	County	NS Coordinate	EW Coordinate	Grid	Place
16	003	1365	3832	13 38 63 52	Watertown, Conn.
16	003	1365	3839	13 38 63 59	Washington, Conn.
21	014	1361	3874	13 38 67 14	Arlington, N. Y.
21	014	1362	3875	13 38 67 25	Poughkeepsie, N. Y.
21	014	1363	3874	13 38 67 34	Fairview, N. Y.
21	014	1366	3871	13 38 67 61	Pleasant Valley, N. Y.
16	002	1375	3818	13 38 71 58	New Britain, Conn.
16	002	1379	3812	13 38 71 92	Nethersfield, Conn.
16	002	1370	3823	13 38 72 03	Southington, Conn.
16	002	1372	3826	13 38 72 26	Bristol, Conn.
16	002	1375	3822	13 38 72 52	Plainville, Conn.
16	003	1378	3840	13 38 74 80	Torrington, Conn.
21	056	1376	3882	13 38 78 62	Kingston, N. Y.
16	002	1386	3866	13 38 80 66	Manchester, Conn.
16	002	1380	3818	13 38 81 08	Newington, Conn.
16	002	1382	3818	13 38 81 28	West Hartford, Conn.
16	002	1383	3814	13 38 81 34	Hartford, Conn.
16	002	1384	3812	13 38 81 42	East Hartford, Conn.
16	003	1388	3836	13 38 83 86	Winsted, Conn.

Source: Bureau of the Census, Transportation Division

Figure 3.26 PICADAD Place File in Grid Order

3.8 FEDERAL HIGHWAY ADMINISTRATION

In the early 1960's the Federal Highway Administration (FHWA) initiated the development of an ambitious nationwide highway travel model to simulate the flow and volume of highway traffic in the continental United States.[1] One of the most important prerequisites for the implementation of this model and for the realistic determination of national travel patterns is a network which adequately describes at a relatively fine grained geographic level the spatial relationships involved. Over several years the FHWA developed such a spiderweb network for the nationwide highway travel project and established a large geographic reference file in support of the model.

Counties are the smallest unit of area to which nationwide trip origins and destinations are assigned in the FHWA spiderweb network. From past experience, it was concluded that any units larger than counties would not produce meaningful results. Furthermore, a considerable amount of information necessary to the model are readily available by counties from the Bureau of the Census and other agencies. The network is constructed with a node at the center of population of each county and each county centroid is connected by link segments

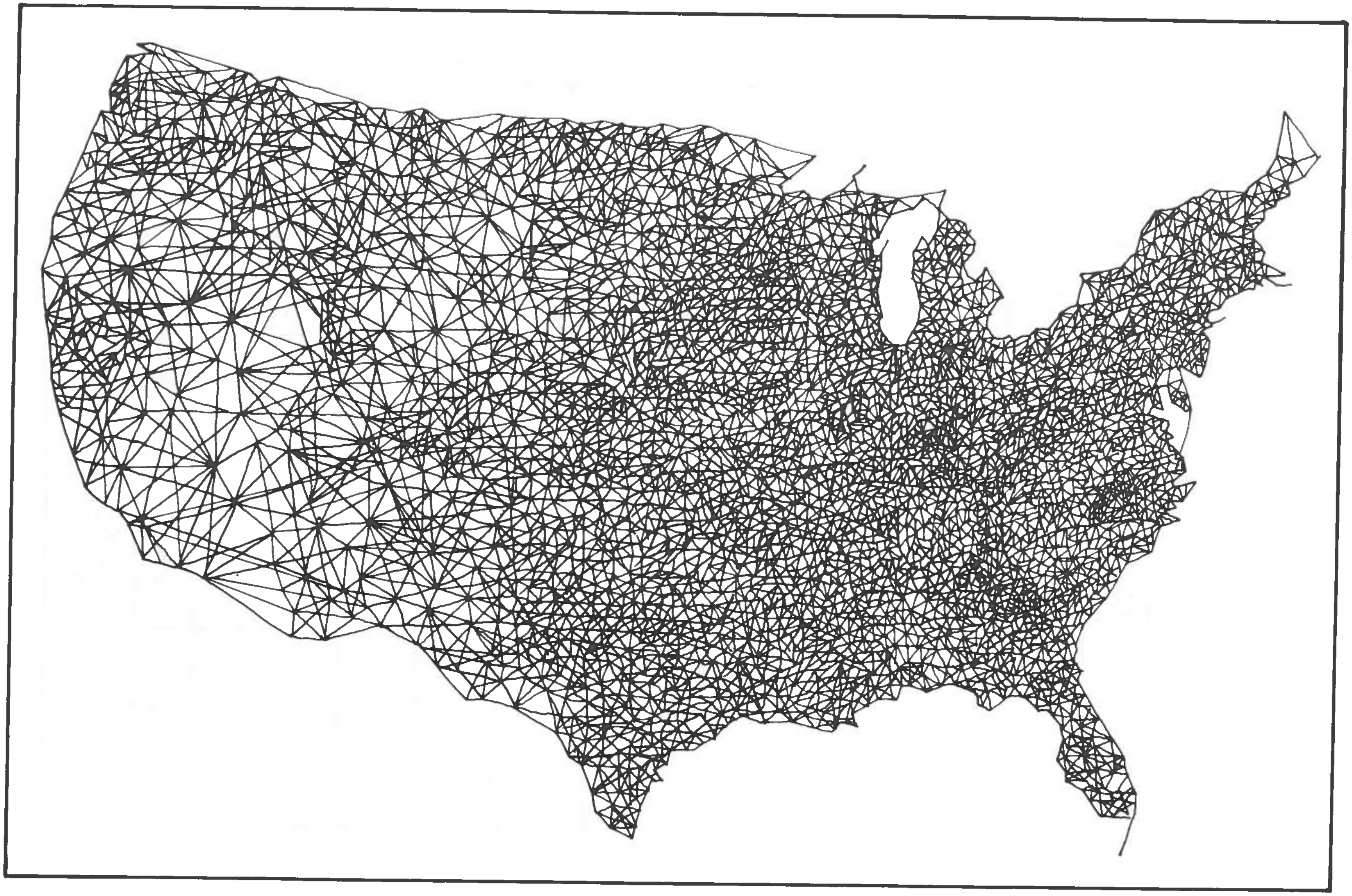
---

[1]The nationwide highway travel model was actually initiated under the auspices of the Bureau of Public Roads, U. S. Department of Commerce. In 1967 when the Department of Transportation was established, the Bureau of Public Roads was reorganized as the Federal Highway Administration and the FHWA continued work on the nationwide highway travel project. To avoid confusion only the current designation for this agency is used in this section.

to nodes in eight surrounding counties. Due to the variation in size, shape, and location of the individual counties, it is not always possible to connect each county with exactly eight other counties. In 211 cases it is necessary to accomodate more than 8 links from each county. This is accomplished by inserting an additional "dummy" node at these locations which is connected to the county centroid with a link of zero distance. In other cases, particularly around the border of the continental United States, there is no need for eight connectors. In total, the network consists of 3,076 county centroids with an additional 211 dummy nodes connected by a system of 11,487 two way links. Thus, on the average, each node has almost exactly 7 links (6.99) per node (See Figure 3.27).

The network is 440,600 miles long as compared to the federal aid primary system of 242,000 miles. The state primary system of 460,000 miles, and the total federal aid system of 850,000 miles. It is slightly more direct between major cities than the existing system -- on the average it is about 13 percent less than the mileage between cities shown in the Rand McNally atlas.

The geographic reference file developed by FHWA in support of this network model was dubbed the ZIMDECK. The file contains a nationwide listing of counties associated with four different county codes, coordinate information for the approximated population center of each county, the county area in square miles, the population of each county and the number of vehicles registered within each county (See Figure 3.28). Both the five digit state/county code maintained



Source: Bureau of Public Roads

Figure 3.27 The Nationwide Highway Travel Network

McGraw Hill St/Co Code	State Abbreviation	County Name	FHWA Sequence Code	FHWA Geographic Code	IBM St/Co Code	Square Mile Area	Latitude	Longitude	X Coordinate	Y Coordinate	Population	Vehicles
22002	ME	Aroostook	0001	1001	18003	6805	467	83	329879	222986	1060	330
22011	ME	Piscataqui	0002	1010	18021	3948	454	92	328399	213109	174	69
22013	ME	Somerset	0003	1011	18025	3948	449	98	326588	208959	397	159
22004	ME	Franklin	0004	1012	18007	1715	448	103	324430	207606	201	80
22010	ME	Penobscot	0005	1020	18019	3408	451	85	332269	212132	1263	448
22015	ME	Washington	0006	1021	18029	2553	450	73	338066	213259	329	136
22005	ME	Hancock	0007	1022	18009	1542	446	83	334237	209131	323	142
22014	ME	Waldo	0008	1030	18027	734	445	91	330677	207306	226	97
22007	ME	Knox	0009	1031	18013	362	441	91	331482	204667	286	120
22008	ME	Lincoln	0010	1032	18015	457	441	96	329108	203950	185	80
22012	ME	Sagadahoc	0011	1033	18023	257	439	98	328549	202345	228	93
22006	ME	Kennebec	0012	1034	18011	865	444	97	328042	205792	892	332
22001	ME	Androscogg	0013	1035	18001	478	441	102	326254	203107	863	306
22003	ME	Cumberland	0014	1036	18005	855	438	103	326355	200983	1828	723
22009	ME	Oxford	0015	1037	18017	2085	444	106	323777	204545	443	180
22016	ME	York	0016	1038	18031	1000	435	106	325489	198581	994	397

Source: Federal Highway Administration

Figure 3.28 Geographic Reference File for the Federal Highway Administration's Nationwide Travel Network by McGraw Hill (currently obsolete) and the five digit IBM state/county code (See Section 2.3) are included on the file to facilitate cross referencing the FHWA file with files outside the administration. There is also a four digit FHWA county sequence serial number assigned to each county in the order in which counties are listed in the ZIMDECK and another FHWA code known as the geographic code which is assigned to each county by contiguous county groups within states, within state group regions. This last code was devised by FHWA in order to provide an indication of geographic relationship among counties rather than the alphabetic sequence reflected by the assignment of both the McGraw Hill and IBM state/county codes. The FHWA geographic county codes permit data aggregation by county clusters within state and by regional clusters of states.

The latitude and longitude of the approximated population center in each county are listed with one decimal point implied (it is necessary to add 600 to longitude values listed in ZIMDECK) and plane X-Y coordinates to hundredths of miles are also listed for the population center. The population and vehicle registration figures are expressed in hundreds and based on 1960 data.

Using the distance along each link in the FHWA spiderweb network as the primary measure of spatial separation, assuming the trip generating values observed in Washington, D.C. and assuming that the attraction of an area is proportional to the population of the area with some adjustments for social and recreational trips, the FHWA developed trip equations to determine the number of trips from one

one county to another. The vehicle trips from each county to all other counties were computed and loaded on the nationwide spider-web for the first study completed in December of 1965.

It was found that in general the model counts are low. This is illustrated by the fact that the counts and assignments agree within 10 percent in 12 states; the assignment is more than 10 percent high in three states, but more than 10 percent low in 33 states. The following conclusions concerning national highway travel patterns were drawn from the first study employing the FHWA network model:

- vehicle ownership must be one of the most important factors in trip generation. Trips in the New York City - New Jersey area were overestimated by more than 3 1/2 times. Approximately 80 percent of the families in Manhattan own no vehicle which is four times the national average at 20 percent.

- population appears to be a reasonable indicator of trip attractions except for recreational travel where it is significantly deficient. Receipts for hotels and motels seem to provide better correlation but probably measure an effect rather than a cause.

- trips between two areas, the centers of which are 20 miles apart, will contain a preponderance of trips less than 20 miles long. The basic trip generation rate should therefore be increased.

- the distance function for the attenuation of trips less than 20 miles long appears to be more complex than distance raised to a constant power.

- travel in large areas such as the nation appear to follow definable principles better than in smaller urban areas probably because the mixture of activities in a zone is more homogeneous.

In the late 1960's further refinements were made to the model and techniques had been developed in plotting networks and trees and inserting numbers on each link denoting volume. Since the volumes on the nationwide network vary from 1 vehicle to 1.5 million vehicles, varying the width of the lines plotted according to volume would produce an indiscernable pattern. Therefore the volumes were stratified roughly by type of facility and plotted in different colors. Volumes above 25,000 were assumed to require more than 4 lane divided highways; 4,000 to 25,000 were assumed to require a high type 2 lane highway; 400 to 1,000 an intermediate type facility and less than 400 a low type facility. The link width within each volume group then reflected the proportion of the load within the limits of the group.

Also in connection with the nationwide highway travel project the boundaries of all counties in the geographic reference file were digitized and FHWA developed an automated county mapping capability.

Although the FHWA nationwide highway network was used in several subsequent studies, eventually it was decided that assignments of origins and destinations to single county zones provided an unnecessarily large network with too fine a geographic grain. Therefore the single county zones were aggregated into 533 county cluster zones and the networks reduced from 440,000 route miles to 110,000 route miles which includes all interstate highways and principal arterials as defined in the FHWA's functional classification manual (See Section 3.9).

References:

Brokke, Glen E. "Nationwide Highway Travel." A paper prepared for the Western Association of State Highway Officials Highway Planning Conference, June 9, 1966, Santa Fe, New Mexico. Washington, D.C.: Bureau of Public Raods, U.S. Department of Commerce, 1966.

3.9 NATIONAL TRANSPORTATION NETWORKS

The Office of Systems Analysis and Information (OSAI), U.S. Department of Transportation has developed five modal networks which are geocoded to a single, nationwide system of transportation zones. Each of these networks--highway, rail, waterway, air, and pipeline--is topologically described as a set of nodes (intersections) and links or segments (lines directly connecting two nodes) representing the configuration of major transportation routes in the United States.[1] Every possible combination of two nodes and the link directly connecting them is identified, located and characterized by a series of codes. The networks and the codes which characterize each link vary from mode to mode, but all five systems are associated with a single set of geographic units called transportation zones.

Transportation Zones. There are 533 OSAI transportation zones which provide a degree of geographic detail suitable for transportation analyses on a national scale. Each standard metropolitan statistical area (SMSA) in the fifty United States constitutes a transportation zone. The remaining area is divided into groups of contiguous counties each of which constitutes a non SMSA transportation zone. In addition

---

[1]The OSAI networks are structured on the Urban Planning System 360 program battery. For a complete description of this system refer to: Urban Transportation Planning: General Information and Introduction to System/360. U.S. Department of Transportation, Federal Highway Administration, June, 1970.

the twenty United States ports ranking highest in foreign trade are designated as separate zones. Zone numbers 001 through 490 are assigned to the transportation zones of the continental United States ranging from the Northeast to the Southwest and numbers 514 through 533 are assigned to the twenty ports. The intervening numbers are assigned to outlying connections or dummy links depending upon the mode of transportation for which the network is designed (See Figure 3.29). Within each of the transportation zones a centroid is designated at either the approximate geographic center of the zone or at the activity/population center of the zone.

Each of the modal networks is superimposed on this system of transportation zones. All of the links within a zone are connected to the centroid of that zone and all of the nodes within a zone are associated with that zone number.

Highway Network. As a direct descendant of the FHWA nationwide highway travel network which was initiated in the early 1960's by the Bureau of Public Roads (See Section 3.8), the OSAI highway network is the oldest and most fully developed of the five transportation networks. It consists of 110,000 route miles, which include all interstate highways and principal arterials as defined in the Federal Highway Administration's functional classification manual. There are approximately 2,500 nodes numbered sequentially beginning at 600 (the numbering starts at 600 in order to avoid confusion and readily distinguish node numbers from centroid numbers). In addition to the

NETWORK	ZONES 1-490	ZONES 491-496	ZONES 497-505	ZONES 506-509	ZONE 510	ZONE 511	ZONE 512	ZONE 513	ZONES 514-533
Highway	Continental United States	Canadian Connections	Canadian Connections	Mexican Connections	Mexican Connection	Mexican Connection	Mexican Connection	Mexican Connection	Twenty Highest U.S. Ports Based on Foreign Tonnage
Rail		Canadian Connections	Canadian Connections	Mexican Connections	Mexican Connection	Mexican Connection	Mexican Connection	Mexican Connection	
Waterway		Dummy	Alaska	Hawaii	Pacific Islands	Puerto Rico And Virgin Islands	Mouth of St Lawrence River	Dummy	
Air		Canadian Connections	Alaska	Hawaii	Mexican Connection	Mexican Connection	Mexican Connection	Mexican Connection	
Oil		Dummy	Dummy	Dummy	Canadian Producing Area	Canadian Producing Area	Canadian Producing Area	Dummy	

Source: John P. Fajnor, IBM, Gaithersburg, Maryland

Figure 3.29 OSAI Transportation Zones

distance, the time or speed, the turn penalty code and the transportation zone code which are associated with every link in the network, each link is associated with a two digit state code which is assigned to the continental states in alphabetic sequence, a two digit geographic unit code ranging from 1 to 78 which is assigned to states or state sections of homogeneous topography, and a one digit terrain code which is assigned in conjunction with these geographic unit designations indicating 1) level terrain, 2) rolling terrain or 3) mountainous terrain. There is also a one digit physical type code (1-5) which indicates the number of divided or undivided lanes running along the link. And, finally, each node is assigned a set of plane coordinates which were digitized from an Albert's Equal Area map projection. All of this information is recorded as the traffic assignment node-link data for the system/360 highway network (See Figure 3.30).

The OASI highway network has been employed in modelling various motor vehicle flow studies. Most recently it was used to test the effect of approximately 13,000 additional miles of proposed highways on the overall flow of highway traffic in the United States. This project was undertaken in conjunction with the 1972 and 1973 national highway acts.

Rail Network, The OSAI rail network consists of the principal class 1 railroads in the United States as designated by the Interstate Commerce Commission. There are approximately 2,600 nodes numbered sequentially beginning at 600. The traffic assignment node-link data

NETWORK	NO. OF ZONES	NO. OF ZONES	NETWORK ELEMENTS	GEOCODES	ADDITIONAL CODES
Highway	533	2,500	Interstate and Principal Arterials	State Code Plane Coordinates OSAI Zone Number	Link Distance Time or Speed Turn Penalty Code Geographic Unit Code Terrain Code Physical Type Code
Rail	533	2,600	Principal Class I Railroad	Division/State Code Plane Coordinates OSAI Zone Number	Link Distance Time or Speed Track Type Code Number of Tracks Signal System Code Railroad Company
Waterway	533	1,800	All Inland Waterways	Water Resource Region Code OSAI Zone Number	Link Distance Time or Speed Waterway Code Waterway Type Code Number of Locks Size of Locks
Air	533	101	Major Airports	State Code OSAI Zone Number	Link Distance Time or Speed Airport Access Airport Type Code
Crude Oil Pipeline	533	1,300	All Crude Oil Pipelines	OSAI Zone Number	Link Distance Time or Speed Size of Pipeline
Refined Products Pipeline	533	1,100	All Refined Products Pipelines	OSAI Zone Number	Link Distance Time or Speed Size of Pipeline

Source: John P. Fajnor, IBM, Gaithersburg, Maryland

Figure 3.30 OSAI Transportation Networks recorded for this network include the standard items for each link in the OSAI transportation networks--distance, time or speed, transportation zone number--and the following additional information: 1) a two digit division/state code which was devised by the Bureau of the Census (See Section 2.1); 2) a one digit link type code which indicates whether the railroad track is a main branch line; 3) a two digit alphabetic abbreviation of the name of the railroad company in control of the tracks; 4) the number of tracks; and 5) a one digit code indicating which of the three types of signal systems is in operation along the link. On this system it is possible to code items 3 through 5 for a maximum of four railroad lines running on the same link. Although both the OSAI rail network and the OSAI highway network have been digitized the plane coordinates assigned to the two networks are not compatible.

Until recently the rail network had not been employed in any major transportation analysis. However, in accordance with section 204 of the Regional Rail Reorganization Act of 1973 (P.L. 93 - 236) the Federal Rail Administration initiated a research project concerning the restructuring and reorganization of the railroad system in the Midwest and Northeast regions of the United States. In this study this OSAI rail network was used in the determination of recommendations for the improvement of transportation service and the financial viability of the railroad companies involved.[1]

---

[1]U.S. Department of Transportation, Rail Service in the Midwest and Northeast Region, Volumes 1 and 2 (Washington, D.C.: U.S. Department of Transportation, February 1974).

Waterway Network. All of the inland waterways (including the Great Lakes-St. Lawrence Seaway system) and major coastal routes in the United States are described in the OSAI waterway network. There are approximately 1,800 nodes in this system, and they are numbered sequentially beginning at 600. In addition to the distance, time or speed, and transportation zone codes each link is characterized by the following designations: 1) a one digit link type code ranging from 1 to 5 which indicates whether the waterway link is greater than 9 feet deep, less than 9 feet deep, an ocean or deep water route, a centroid connection, or a dummy connection; 2) the number of locks operating along the warerway link; 3) a two digit code ranging from 1 to 44 which indicates the size of the locks; 4) a two digit code assigned by the Water Resources Council which locates the link within one of the 20 water resource regions (See Section 4.2); and 5) a two digit waterway code ranging from 01 to 55 which identifies the specific river, canal, harbor or bay on which the link is located. This waterway network has not been digitized and has not been used in other than the original OSAI study for which it was created.

Air and Pipeline Networks. The OSAI air network and pipline networks have never been fully developed due to practical difficulties. Although two pipeline networks were developed (one for crude and the other for refined petroleum products) these are relatively unrealistic configurations because real pipeline networks consist of many disconnected independent lines. In the OSAI pipeline network these independent lines are linked together due to the continuity requirements of traffic assignment type networks. A true air network was never developed for OSAI because time and budget limitations precluded development of a network consisting of more than zone to zone air distance and shipment data.

References:

U.S. Department of Transportation, Federal Highway Administration. Urban Transportation Planning: General Information and Introduction to System/360. Washington, D.C.: Federal Highway Administration, U.S. Department of Transportation, June, 1970.

Fajnor, John P. International Business Machines Corporation, Gaithersburg, Maryland. Conversations and correspondence, 1974.

3.10 FEDERAL RAILROAD ADMINISTRATION

Because of the extreme financial difficulties that have plagued passenger rail service and rail freight service particularly in the Northeast and Midwest, the federal government has been directing a restructuring of rail transportation in the United States over the past five years.[1] The establishment of the National Railroad Passenger Corporation in 1970 and the subsequent initiation of Amtrak service constituted the first major changes. Passage of the Regional Rail Reorganization Act in 1973 and the creation of the United States Railway Association are further developments in this direction. The Federal Railroad Administration (FRA) whose general function is to consolidate government support of rail transportation activity in the United States along with Amtrak and the U. S. Railway Association has become involved in various aspects of the rail reorganization programs. FRA is responsible for providing a unified and unifying national policy for rail transportation, conducting research and development programs in support of improved rail transportation and analyzing the future requirements for rail and ground transportation in the United States.

Two of the geographically referenced information systems maintained by the FRA in support of rail reorganization analyses and the

---

[1]While not all individual railroad companies suffered major economic difficulties during the 50's and 60's the railroad industry has, in general, declined and in several cases, most notably the Penn Central, gone bankrupt. Refer to U. S. Congress, Senate, Failing Railroads, Hearings before the Senate Committee on Commerce on S.4011, S. 4014, and S. 4016, 91st Congress, 2nd Session, 1970 or U. S. Congress. Senate, Railroad Industry Overview 1971, Hearings before the Subcommittee on Surface Transportation, Senate Committee on Commerce, 92nd Congress, 1st Session, 1971.

improvement of rail service are the passenger rail ridership file for trains operated by Amtrak throughout the country and the rail network developed for the Midwest and Northeast region of the United States.

3.10.1 Passenger Rail Ridership Files

Railroad station codes for stations serving passenger traffic were developed for use in two information systems, one maintained by FRA and the other by Amtrak. Both systems are designed for the collection and analysis of ridership data on trains operated by Amtrak. The input data for these systems are passenger counts recorded by passenger train conductors and crewmen on a daily basis by route. This data is published in monthly aggregates by the FRA.

Since the data are entered directly onto input forms for machine processing, coded station names are necessary for data reduction. At the same time, however, meaningful codes are desirable to facilitate memorizing the codes and to reduce errors. Therefore, alphabetic codes, composed of letters selected from the station names themselves were deemed most practical (See Figure 3.31). Other existing coding systems were considered, but none were found appropriate. For example, although a Standard Point Location Code (SPLC, See Section 3.6) exists for each geographic location served by Amtrak, SPLC has two major disadvantages for this particular purpose in that it is numeric and thus difficult to remember, and it is larger (six digits) than necessary to provide unique identification for passenger railroad stations.

CITY ABBREVIATION LISTING			CITY ALPHA LISTING			STATE/CITY ALPHA LISTING		
City Abrv	City Name	State	City Abrv	City Name	State	City Abrv	City Name	State
ABQ	Albuquerque	NM	AKN	Akron	CO	ATN	Anniston	AL
ADM	Ardmore	OK	ALB	Albany	NY	BHM	Birmingham	AL
AGS	Augusta	GA	ABQ	Albuquerque	NM	BHS	Birmingham	AL
AKN	Akron	CO	ALX	Alexandria	VA	DCT	Decatur	AL
ALB	Albany	NY	ALH	Alhambra	CA	KHN	Dothan	AL
ALH	Alhambra	CA	ALP	Alpine	TX	EUT	Eutaw	AL
ALN	Alton	IL	ANA	Altavista	VA	LIN	Livingston	AL
ALP	Alpine	TX	ALN	Alton	IL	MGM	Montgomery	AL
ALT	Altoona	PA	ALT	Altoona	PA	TCL	Tuscaloosa	AL
ALX	Alexandria	VA	AMS	Amsterdam	NY	FLG	Flagstaff	AZ
AMS	Amsterdam	NY	ANG	Angora	PA	KNG	Kingman	AZ
ANG	Angora	PA	ARB	Ann Arbor	MI	PHX	Phoenix	AZ
ARA	Aurora	IL	ATN	Anniston	AL	SLG	Seligman	AZ
ARB	Ann Arbor	MI	ADM	Ardmore	OK	TUS	Tucson	AZ
ARD	Ardmore	PA	ARD	Ardmore	PA	WLO	Winslow	AZ
ARH	Ardsley-Hudson	NY	ARH	Ardsley-Hudson	NY	YUM	Yuma	AZ
ARK	Arkansas City	KA	ARK	Arkansas City	KS	LRK	Little Rock	AR
ARL	Arlington	TX	ARL	Arlington	TX	ALH	Alhambra	CA

Source: Federal Railroad Administration

Figure 3.31 Federal Railroad Administration: Amtrak Matrix System

The passenger station codes are currently in use on the ridership files of Amtrak and FRA including the fully automated metroliner segment of the information system. These codes have also been submitted to the U. S. Department of Transportation Data Standardization Group for approval as an official DOT data standard.

3.10.2 Regional Rail Network

In accordance with the Regional Rail Reorganization Act of 1973, the Federal Railroad Administration sponsored a study and report to the Secretary of Transportation on the railroad industry in the Northeast and Midwest United States. The Act specified that this report contain ". . . conclusions and recommendations with respect to the geographic zones within the region in and between which rail service should be provided . . ."[1] This zone approach was intended to facilitate the analysis of the region's rail requirements in four ways:

- It permits the analysis of an area which is small enough to allow the inclusion of a substantial amount of detail and the identification of each traffic generating point.

- It emphasizes local rail service requirements by excluding the need to consider traffic which now flows through the zone on mainlines but which may not continue to be routed through the zone in the future. If the points in a zone generate sufficient traffic to justify local rail service, that service will continue regardless of the ultimate configuration of the mainlines which carry traffic through the area.

---

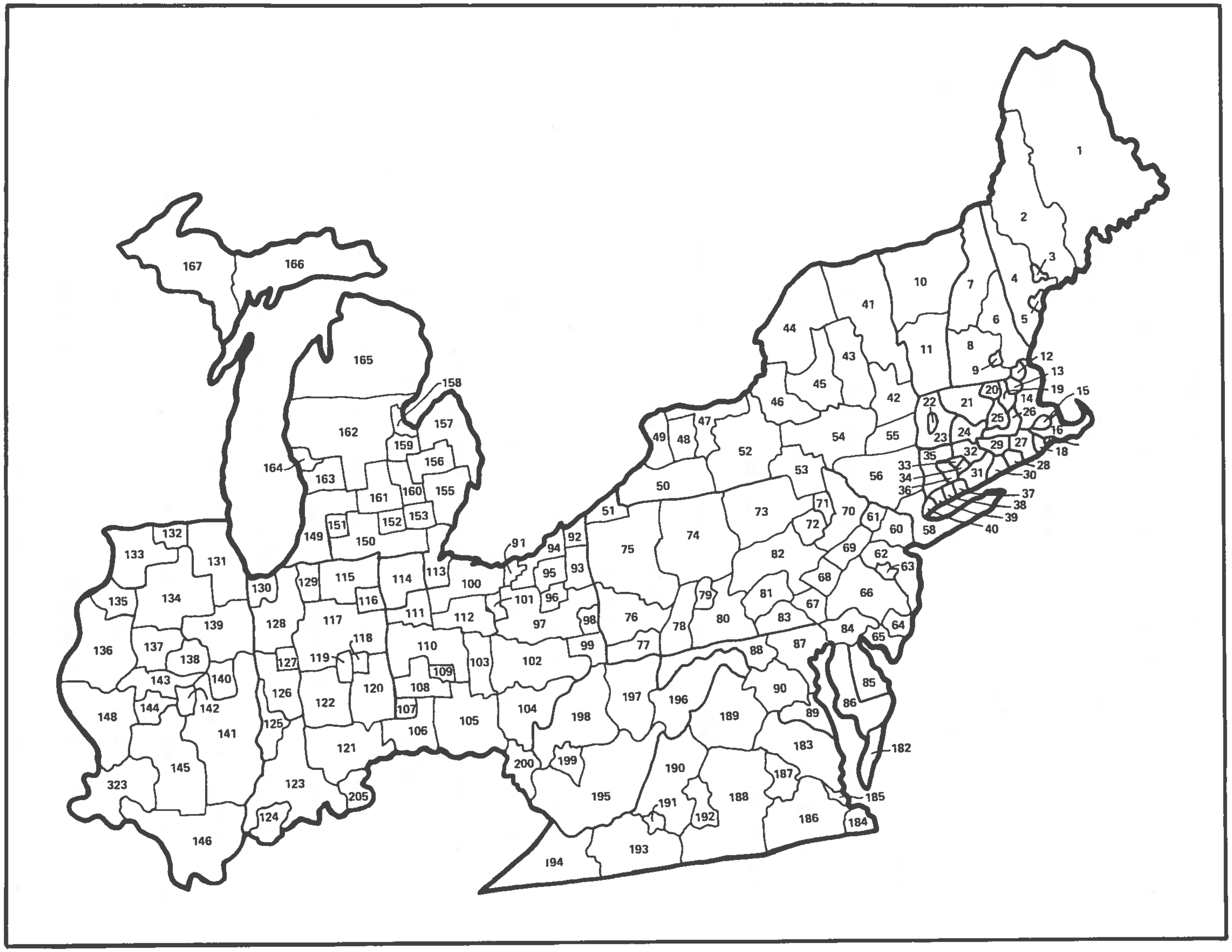
[1] U. S. Department of Transportation, Rail Service in the Midwest and Northeast Region, A Report by the Secretary of Transportation, Washington, D. C.: U. S. Government Printing Office, 1974, p. 17.

- It maintains a neutral position with respect to the present railroad corporate structures and their related track networks. The zone approach emphasizes the area's basic demand for rail service--where the traffic originates or terminates--and defines alternatives that allow the rail system (not specific companies) to meet that demand as efficiently as possible.

- It facilitates the grouping of traffic generated in the zone in order to evaluate the requirements of the mainline rail network.

The zones delineated by FRA for this study are based on the transportation zones established by the Office of Systems Analysis and Information (OSAI, See Section 3.9) and updated to 1970 Standard Metropolitan Statistical Area (SMSA) configurations. The region was divided into a total of 184 zones as illustrated in Figure 3.32. The zones were created in two steps. First, the SMSAs were used as the basis for delineating the initial set of zones. Second, those areas of the region outside of SMSAs were aggregated into groupings of contiguous counties (townships in the New England States) with similar socioeconomic characteristics based on the earlier OSAI zone definitions. The resulting 184 zones include St. Louis and Louisville, which are in states contiguous with the states specified in the Act as constituting the region. With these two exceptions, the 184 zones do not include any zones in states contiguous with those specified states.

Of the 184 zones, 21 do not contain facilities or operations of bankrupt carriers. These zones are on the periphery of the region and include two zones in the Upper Peninsula of Michigan, six zones in northern and western Illinois, one zone in West Virginia, nine zones across southern Virginia and the three most northerly zones in Maine.



Source: Federal Railroad Administration

Figure 3.32 Transportation Zones in the Northeast and Midwest

Following the identification of the zones, information was gathered for each zone with regard to major highways, rail lines and their characteristics (e.g. number of tracks, type of signal system, traffic density, etc.), and railroad traffic generating points. This information is needed for an understanding of each zone's overall transportation system, the capacity of its rail network and the sources of rail freight traffic.

Traffic data for 1972 originations and terminations were then obtained from Class I railroads operating in the region. After these data were sorted by zone, they represented a comprehensive information base as to traffic, station loadings and the commodities involved.

Using this traffic data and rail facility information for each zone, the analysis was conducted on two levels with a continuous iterative process between the two. First, adjacent zones were analyzed on a statewide basis to determine connecting traffic flows between adjacent zones and to points outside the region. Next, zone traffic was aggregated to give a first cut at determining the major traffic generating centers which would support high volume, mainline interstate traffic in the region. The possible mainline routes connecting these centers were then identified on the zone maps.

From this point, various criteria were applied on a zone-by-zone basis to determine which points in each zone should be recommended for local rail service. Similarly, other criteria were applied on a regional basis to determine which major traffic generating centers should be recommended for competitive rail service on an interstate network.

For the 163 zones served by bankrupt railroads, the analysis went one step further by identifying potentially excess rail lines which are not necessary to serve points recommended for service or are duplicate light density feeder lines. This additional analysis is intended as a potential planning tool for the U.S. Railway Association and for the states and local communities in determining where their own analysis should be concentrated during the remainder of the planning process.

Currently, the Federal Railroad Administration is expanding the regional model to a nationwide network based on the OSAI rail network. In addition to the standard items coded to each link in the network--distance, speed, signal system, type of track, company controlling the track, etc.,--there will be a station cross reference file containing SPLC point numbers, freight station accounting codes, transportation zone numbers, urban area codes and several other general geographic designations. Furthermore, the FRA is refining the model and developing a more detailed network configuration for railroads in the United States. This network is intended for use as a general rail systems planning tool.

References:

U.S. Department of Transportation. Rail Service in the Midwest and Northeast. A report by the Secretary of Transportation. Washington, D.C.: U.S. Government Printing Office, 1974.

U.S. Department of Transportation. Federal Railroad Administration. Memorandum to the Department of Transportation Data Standardization Group. March 1974. Washington, D.C.: Federal Railroad Administration, U.S. Department of Transportation, 1974.

4. SPECIALIZED AREAL UNIT REFERENCE CODES

4.1 UNITED STATES POSTAL SERVICE

During the late 1950's and early 1960's it became evident that the postal collection and distribution system in the United States was rapidly becoming inefficient, even obsolete. Fundamental changes in the nature of mail and the increasing volume of mail required more sophisticated and more mechanized methods of handling and moving the mail. For many years whole cities and states were used as grouping and sortation areas with geopolitical names as sorting keys. By 1960 an overwhelming number of cities constituted units that were simply not small enough for sorting the increasing volume of mail efficiently and the use of geopolitical names necessitated that individual sorters memorize spatial and distributional relationships often within unreasonably large and geopolitically complex areas.

As a result of this situation the U. S. Post Office Department developed and implemented the Zone Improvement Plan (ZIP) for postal service in the United States. According to this plan, all the possible postal origins and destinations were collapsed into a series of small areas for gathering and grouping mail. When mail from one area reached another area, it could then be regrouped in even smaller universes until the carrier route was achieved and delivery made.[1]

The hierarchically structured postal units are identified by the five digit ZIP code which was implemented in July, 1963. The first

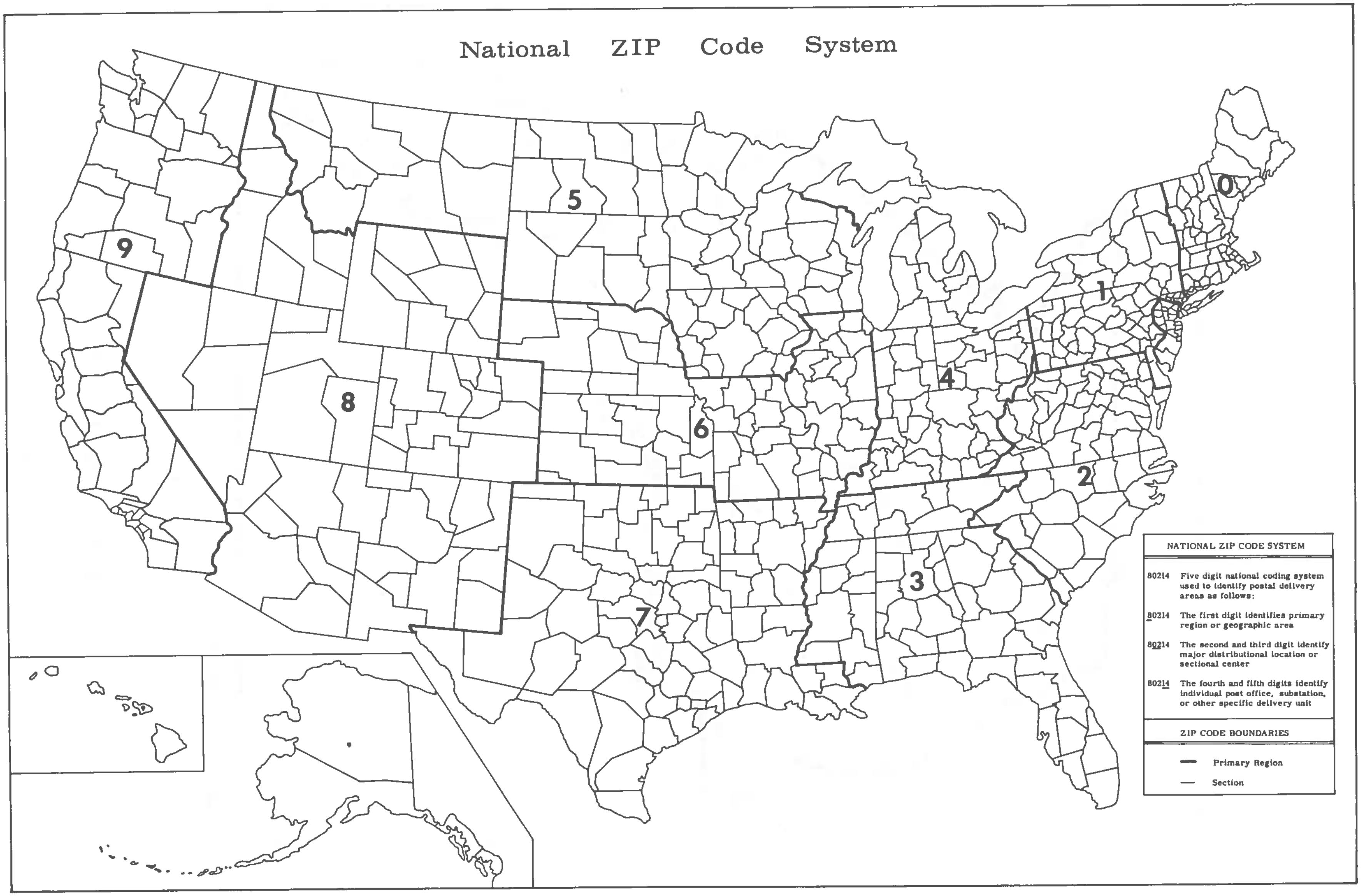
---

[1]Canada has also instituted the use of postal address codes. Refer to Appendix B for a description of the Canadian system.

digit of this code identifies one of ten major regions in the United States (See Figure 4.1). The second and third digits identify individual sections within these regions (See Figure 4.2). Sections are small apolitical areas defined in terms of local transportation patterns. Large cities constitute an entire section. One central post office gathers and redistributes mail for all other post offices in the section. In the case of the larger cities, the main office performs the same functions for stations and branches within the city. Jacksonville, Florida, for example, is the sectional center for section 320 and all mail coded to section 320 is forwarded to Jacksonville to be resorted for delivery to all post offices in that section. The city of Jacksonville is also a separate section coded 322. All mail destined for addresses within the city of Jacksonville is resorted at the main post office in the city for delivery to local branch offices. Originating mail is grouped at each sectional center for forwarding to the sectional center of destination.

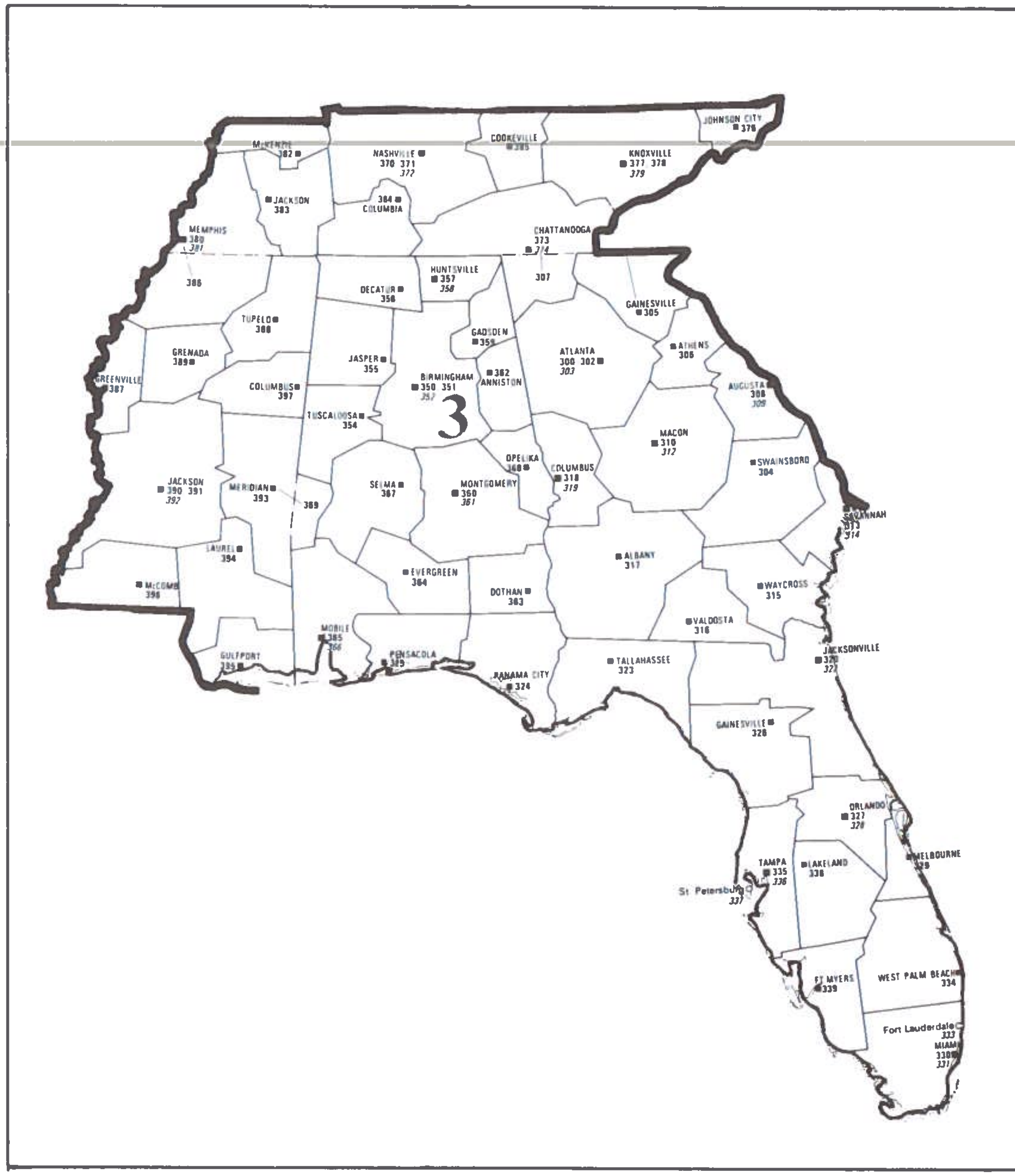
Although in most cases section boundaries conform to state lines there are exceptions as is evident in Figure 4.2. The size and shape of the section is determined by the local transportation pattern and as such tends to reflect the local economic pattern. Similarly, the multi-coded office is organized along the lines of the economic city, not the political city, and its stations and branches are consistent with population concentrations within the city, often with usable economic and social factors reflected as well.

Section numbers are serial and are assigned to the sections



Source: U.S. Postal Service

Figure 4.1 National ZIP Code System



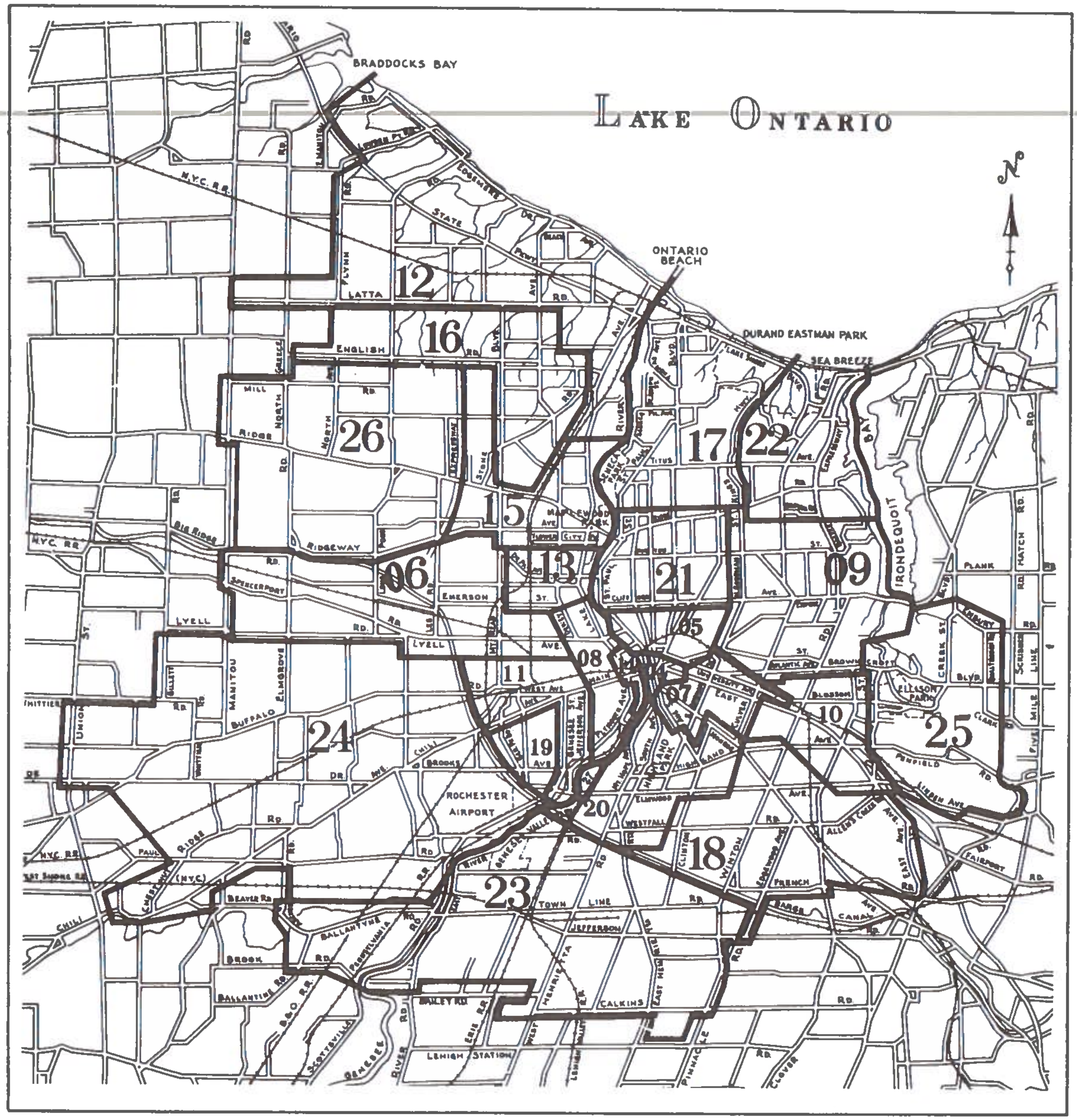
Source: U.S. Postal Service

Figure 4.2 ZIP Code Sections

grouped roughly by state. At the three digit, sectional level of the ZIP code there are potentially 1,000 numbers. Fifteen three digit numbers are assigned to military use (all APO and FPO addresses) and four are assigned to federal agencies in Washington, D.C. Of the remainder, 803 three digit numbers are assigned singularly; 32 sections have been assigned two three digit numbers; and 2 sections have been assigned three three digit numbers. 108 of the potential 1,000 are not used. Thus, there are 856 active three digit level ZIP code numbers.

The last two digits of the ZIP code represent the individual post offices within the section or the stations and branches of a city section post office. Usually the full five digit ZIP code zone covers a group of city blocks (See Figure 4.3) or a small rural area served by a single post office. However, separate ZIP codes may be assigned to single office buildings, institutions, or locations that generate and receive a large volume of mail. Thus, the size of the area represented by a five digit ZIP code varies considerably and reflects both concentrated, day time, business populations as well as residential population densities. At the five digit level, there are about 50,000 active ZIP codes, most of which (32,000) are assigned to individual post offices. The remaining five digit numbers are assigned to stations and branches of multi-coded post offices.

Although ZIP code was designed specifically for a mail distribution system and the ZIP code zones are unique to that system, the code is structured so that it may serve for a variety of other uses not



Source: U.S. Postal Service

Figure 4.3 Local Five Digit Level ZIP Code Zones Within Section 142, Rochester, New York envisioned when the system was conceived. ZIP marketing areas (See Section 4.6), for example, have been developed as a geographical model for sales analysis. Data from the 1970 Census of Housing and Population are coded to ZIP units on the summary tapes (by five digits for the country). The Internal Revenue Service is also using ZIP codes to tabulate some income data. The United States Postal Service[1] actively encourages such uses of the ZIP code and is developing a number of user services to facilitate the application of ZIP code for purposes other than collecting and sorting the mail.

References:

U. S. Postal Service. Customer Services Group. National ZIP Code Directory. Washington, D.C.: U. S. Government Printing Office, 1973.

Jones, Richard W. "New Developments in ZIP Code Marketing". A reprint of an article which appeared in Sales Management Magazine, June 10, 1969.

---

[1]The Postal Reform Act of 1970 created the U.S. Postal Service as an independent establishment of the executive branch, thereby replacing the old Post Office Department.

4.2 WATER RESOURCES COUNCIL

In response to the growing need for coordination among the various agencies concerned with water and related land resources development activities the Water Resources Council was created in 1965.[1] The primary function of this council is to encourage the conservation, development and use of water and related land resources in the United States on a comprehensive and coordinated basis through cooperation of all affected federal agencies, regional commissions, state agencies, local governments, individuals, corporations, and other concerned entities. More specifically the Council has a responsibility to:

- maintain a continuing study and assessment of the adequacy of water supplies necessary to meet the water requirements of the United States and the relation of regional plans and programs to the requirements of the nation

- appraise the adequacy of administrative and statutory means for coordination and implementation of policies and to make recommendations to the President with respect to federal policy

- establish principles and standards for federal participation in the preparation of regional plans and to coordinate the schedules and budgets involved in this planning process

---

[1]The Water Resource Council is an independent executive agency composed of the Secretaries of Agriculture; Army; Health, Education, and Welfare; Interior; Transportation; and the Chairman of the Federal Power Commission. Associate members of the Council include the Secretaries of Commerce; Housing and Urban Development; and the Administrator of the Environmental Protection Agency. The Attorney General, the Director of the Office of Management and Budget, the Chairman of the Council on Environmental Quality, and the Chairmen of the River Basin Commissions participate as observers.

- oversee the creation, operation and termination of federal/state river basin commissions and to review the plans prepared by these commissions
- assist the states financially in developing and participating in comprehensive water and related land resource plans or programs

In carrying out these duties the Water Resources Council has developed a system of regional framework studies and a major national assessment of water and related land resources which is reported every five years. Framework studies consist of an evaluation of regional water resource uses. They identify regions and river basins with complex problems which require more detailed investigation and analyses, and may recommend specific implementation programs in areas not requiring further study. The national assessment provides a general appraisal of the overall water supply and requirements including environmental quality aspects and the future needs for water related goods and services based upon correlated projections of population and economic activity. The national assessment is based upon and serves as a national constraint to regional framework studies.

This commitment on the part of the Water Resources Council to an ongoing data collection program and continuous national assessment necessitates the delineation of standard geographical subdivisions of the United States suitable for the organization of water resource related data and analyses. The Geographic system developed for the first national assessment in 1968 consisted of 20 water resource regions corresponding to the major drainage patterns in the United States and 110 subregions representing the river system components of large watersheds. In 1970 the North Atlantic water resource region was divided into the New England and the Middle Atlantic regions bringing the total number of regions to 21 (See Figure 4.4). The system of subregions was completely re-evaluated due to "the troubles encountered during the first assessment in aggregating availability requirement comparisons without taking account of internal return flows and flow requirements within individual channel reaches."[1]

A revised set of 205 water resource subregions were delineated according to the following criteria (See Figure 4.5):

- a subregion includes that area drained by reach of a river and its tributaries in that reach, a closed drainage basin, or a group of streams forming a coastal drainage area

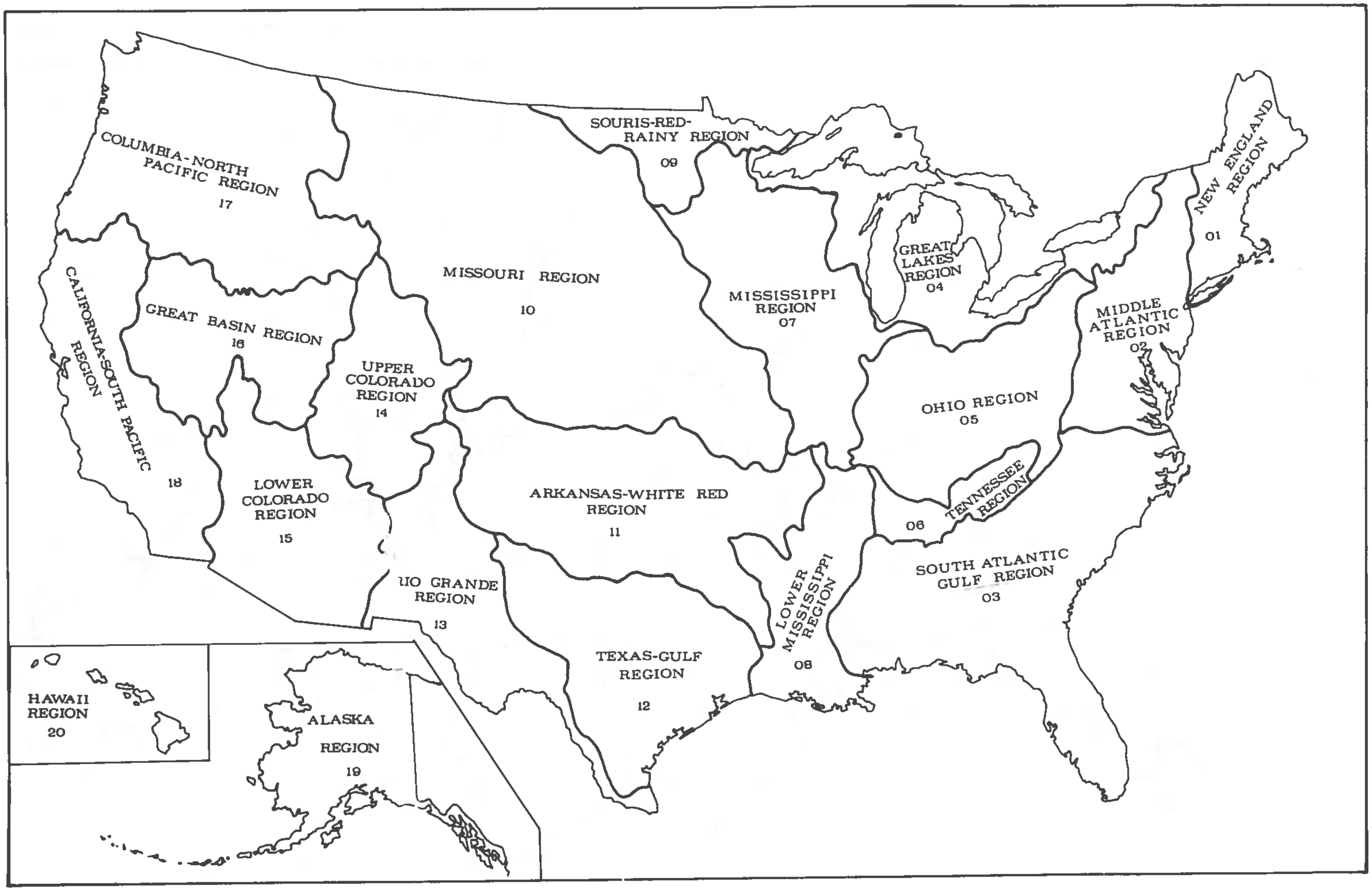
- a subregion is intended to represent a geographic area with common or unique water management problems relating to current or expected future aspects of water supply, use, quality, institutions, or organizations

- boundaries for subregions representing a reach of a river are defined at an established point such as a compact point or long term gauging station except that for selected reservoirs where a major use of the storage is for irrigation, the boundary is at the headwaters of the reservoir

This criteria reflects the current primary orientation of the national assessment toward surface hydrology. The geographic unit defined by surface hydrology is considered most appropriate for the analyses of in stream uses such as waste disposal, navigation, flood control and prevention, fish habitat, and most hydroelectric power

---

[1]Water Resources Council, Water Resources Regions and Subregions for the National Assessment of Water and Related Land Resources (Washington, D.C.: Water Resources Council, 1970), p. 2.



Source: Water Resources Council

Figure 4.4 Water Resource Regions

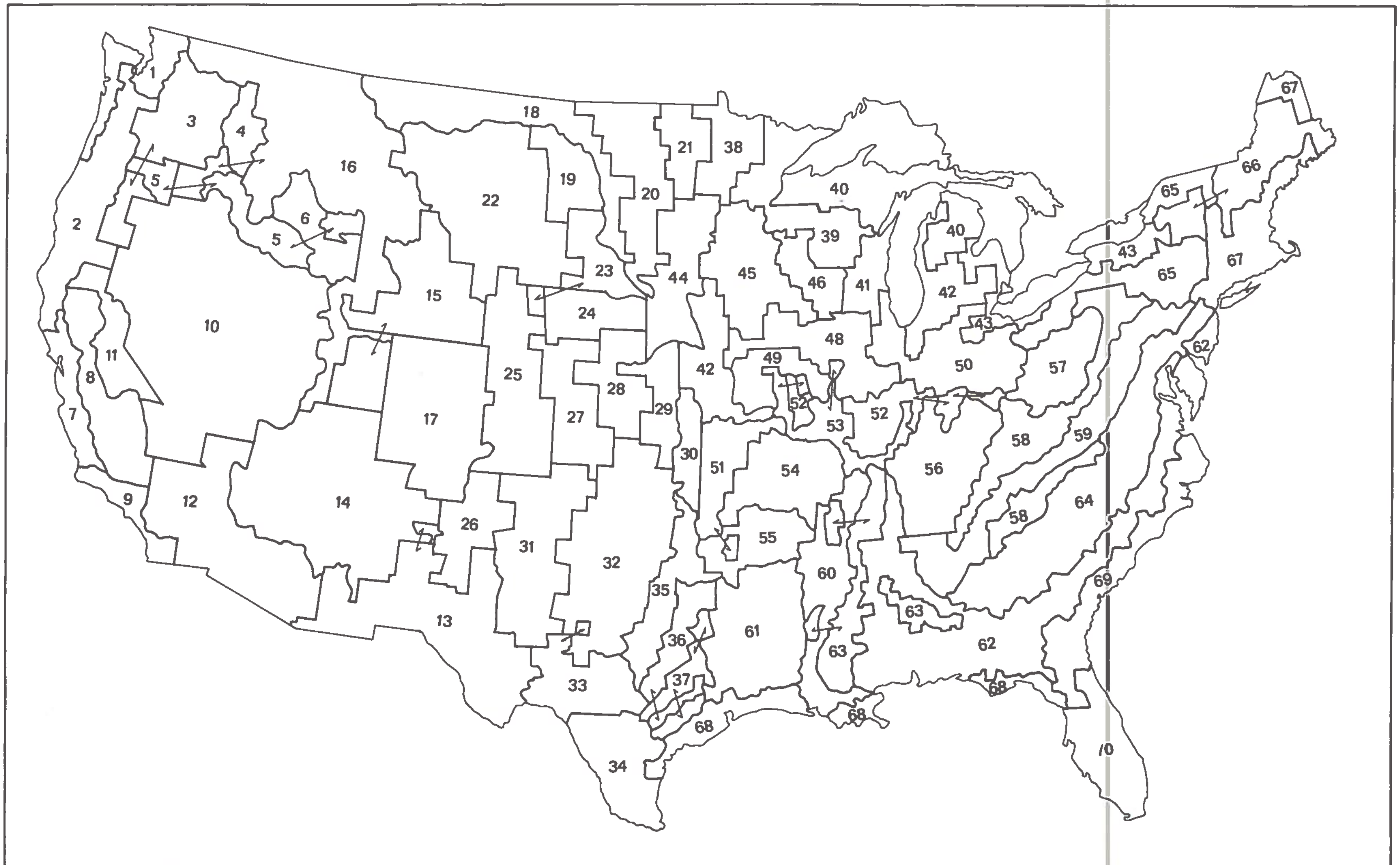
Source: Water Resources Council

Figure 4.5 Water Resource Subareas

generation. Groundwater units are less readily defined and less information is available for these systems at this time. However, as more data concerning groundwater resources is developed, separate groundwater subregions may be delineated.

In addition to the actual drainage area boundaries, each subregion is defined according to those county boundaries which most nearly approximate the drainage boundary. This delineation was necessary to develop approximate data series for subregions from sources which use counties as the lowest unit of geographic detail (See Section 4.5). The area delineated by county boundaries is called the subarea. The Canadian Subregion, for example, refers to the drainage area of the Canadian River and its tributaries, while the Canadian Subarea refers to the counties that approximate the area of the Canadian Subregion.

A second set of resource regions and regional subdivisions called land resource groups and land resource areas has been adopted by the Water Resources Council for use in the national assessment and the framework studies. The smaller geographic units, the land resource areas, are delineated by the Soil Conservation Service, U. S. Department of Agriculture as areas having similar patterns of soil (including slope and erosion), climate, water resources, land use, and type of farming. Like the water resource subareas, the boundaries of these units have been generalized to conform with county lines. Land resource groups are aggregates of land resource areas as defined and used by the Economic Research Service, U. S. Department of Agriculture (See Figure 4.6).



Source: Water Resources Council

Figure 4.6 Land Resource Groups and Areas

The geographic master file of the Water Resources Council contains a listing of the counties within each water resource subarea by water resource region (See Figure 4.7). For each of the counties listed in the file the following codes are recorded: the two digit water resource region code; the two digit water resource subarea code; the FIPS state, county and SMSA codes[1] (See Section 2.1); the three digit code signifying the BEA economic area (See Section 4.5) in which the county is located; and the appropriate five digit land resource codes.

Pending the approval of the Office of Management and Budget, the hierarchy of water resource regions and subareas maintained by the Water Resources Council will be expanded. The new system would be a five level hierarchy of nesting geographical units based on groups of whole counties and designed to provide several levels of areal aggregation for the information reporting process, the organization of data and the preparation of analyses. At the highest level of aggregation will be 21 regions as defined in 1970 (18 continental regions and one region each for Alaska, Hawaii, and the American territories in the Caribbean). A set of 106 aggregated subareas will comprise the next level of geographic subdivision. The 205 subareas which conform to county boundaries provide the third level. 350 county cluster entities called accounting units and 1,600 single county or county group combinations will make up the fourth and fifth levels of the expanded hierarchy of water resource geographic units.

---

[1]If a water resource subregion boundary passed through a multi-county SMSA, the subarea boundary was established along the SMSA boundary.

02. MIDDLE ATLANTIC REGION - The drainage within the United States that ultimately discharges into: a) the Atlantic Ocean, whose point of discharge is located within and between the States of New York and Virginia; and b) the Richelieu River, a tributary of the St. Lawrence River

Subregions

0201. Richelieu — The Richelieu River, which includes the Lake Champlain drainage

Region	Subarea	Name		FIPS State	FIPS County	Economic Area	Land Resources		FIPS SMSA
		County	State				Group	Area	
02	01	Clinton	New York	36	019	006	65	142	
		Essex	New York	36	031	006	66	143	
		Addison	Vermont	50	001	003	65	142	
		Chittenden	Vermont	50	007	003	65	142	
		Franklin	Vermont	50	011	003	65	142	
		Grand Isle	Vermont	50	013	003	65	142	
		Lamoille	Vermont	50	015	003	66	143	
		Rutland	Vermont	50	021	003	67	144	

Source: Water Resources Council

Figure 4.7 Water Resource and Subareas: Master File

References:

Water Resources Council. _Water Resources Regions and Subregions for the National Assessment of Water and Related Land Resources_. Washington, D.C.: Water Resources Council, 1970.

Water Resources Council. _The Nation's Water Resources_. Washington, D.C.: Water Resources Council, 1968.

4.3 OFFICE OF MANAGEMENT AND BUDGET

In addition to its responsibilities concerning the preparation of the national budget and the promotion of more efficient government through the coordination of interagency cooperation, the Office of Management and Budget (OMB)[1] is charged with the development of federal information systems and the coordination of all government statistical services. One important aspect of this latter responsibility has been the establishment of several standard geostatistical units for the tabulation and reporting of federal statistics. The Standard Federal Regions and the Standard Metropolitan Statistical Areas (SMSAs) are two types of geographic units established by OMB to provide such standards.

4.3.1 Standard Metropolitan Statistical Areas

Standard Metropolitan Statistical Areas were first defined in 1950 as "standard metropolitan areas". They were designed to replace the four different sets of geographical definitions then in use for various statistical series--metropolitan districts, metropolitan counties, industrial areas, and labor market areas. Data reported for these areas on population, industrial production and

---

[1]The Office of Management and Budget was established in the Executive Office of the President pursuant to a reorganization plan which became effective in 1970. Formerly OMB was the Bureau of the Budget (BOB).

labor force were nonrelatable because each set of geographic reporting units included slightly different territory. Thus, the primary objective in defining SMSAs was to establish a set of general purpose geostatistical units for agencies of the federal government that would facilitate the interfacing of various statistical series. The term standard metropolitan area was changed to standard metropolitan statistical area in order to describe more accurately the objective in defining these units.

An SMSA represents an integrated economic and social unit with a recognized large population nucleus. In order to serve the statistical purposes for which these areas are defined, their component parts must themselves be areas for which statistics are usually or often collected. Therefore, the county was chosen as the basic element in SMSA configurations. SMSAs consist of a central county or counties containing the population nucleus of the areas (central city or cities) and contiguous counties that are metropolitan in character and that are socially and economically integrated with the central city. In New England where townships are statistically more important than county units, the township is used as the basic component of an SMSA. Each SMSA must include one central city and there is no limit to the number of adjacent counties included in a single SMSA as long as the counties are integrated with the central city (See Figure 4.8). Many SMSAs are single county entities such as the Phoenix SMSA; the Atlanta SMSA includes 15 counties; and the Boston SMSA is composed of 92 townships.

With the advice of a committee composed of representatives from

1. Each SMSA must include at least one central city with 50,000 or more inhabitants; or twin cities having contiguous boundaries and constituting for general social and economic purposes, a single community with a combined population of at least 50,000 (the smaller of such twin cities must have a population of at least 15,000). The county or counties containing such a city or cities is the core area of the SMSA.

2. Adjacent non-core counties are regarded as integrated with core areas and thus part of an SMSA if:

   at least 75% of the labor force of the county is nonagricultural labor;

   and if:

   50% or more of its population is living in contiguous minor civil divisions (MCDs) of at least 150 persons per square mile in an unbroken chain of MCDs with such density radiating from the core;

   or

   the number of nonagricultural workers employed or living in the county equals at least 10,000 or 10% of the number of nonagricultural workers employed or living in the largest city of the core area;

   and if:

   15% of the workers living in the county work in the core county or counties;

   or

   25% of those working in the county live in the core county or counties.

3. If two or more adjacent counties each have a city or twin cities of 50,000 inhabitants or more and the cities are within 20 miles of each other, they will be included in the same SMSA unless there is definite evidence that the two cities are not economically and socially integrated.

4. In New England where the city and town are administratively more important than the county, and data are compiled for such minor civil divisions, towns and cities are the units used in defining SMSAs. In New England, because smaller units are used, and more restricted areas result, a population density criterion of at least 100 persons per square mile is used as the measure of metropolitan character.

5. The name of the largest city in an SMSA generally constitutes the SMSA title, however, up to two additional city names may be added to the SMSA title if the additional city or cities have a population equal to one-third or more of the population of the largest city in the area and a minimum population of 25,000. In several cases county names (e.g. the Nassau-Suffolk SMSA) or region of a state (e.g. the Northeastern Pennsylvania SMSA) may be substituted as the SMSA title.

Source: Office of Management and Budget

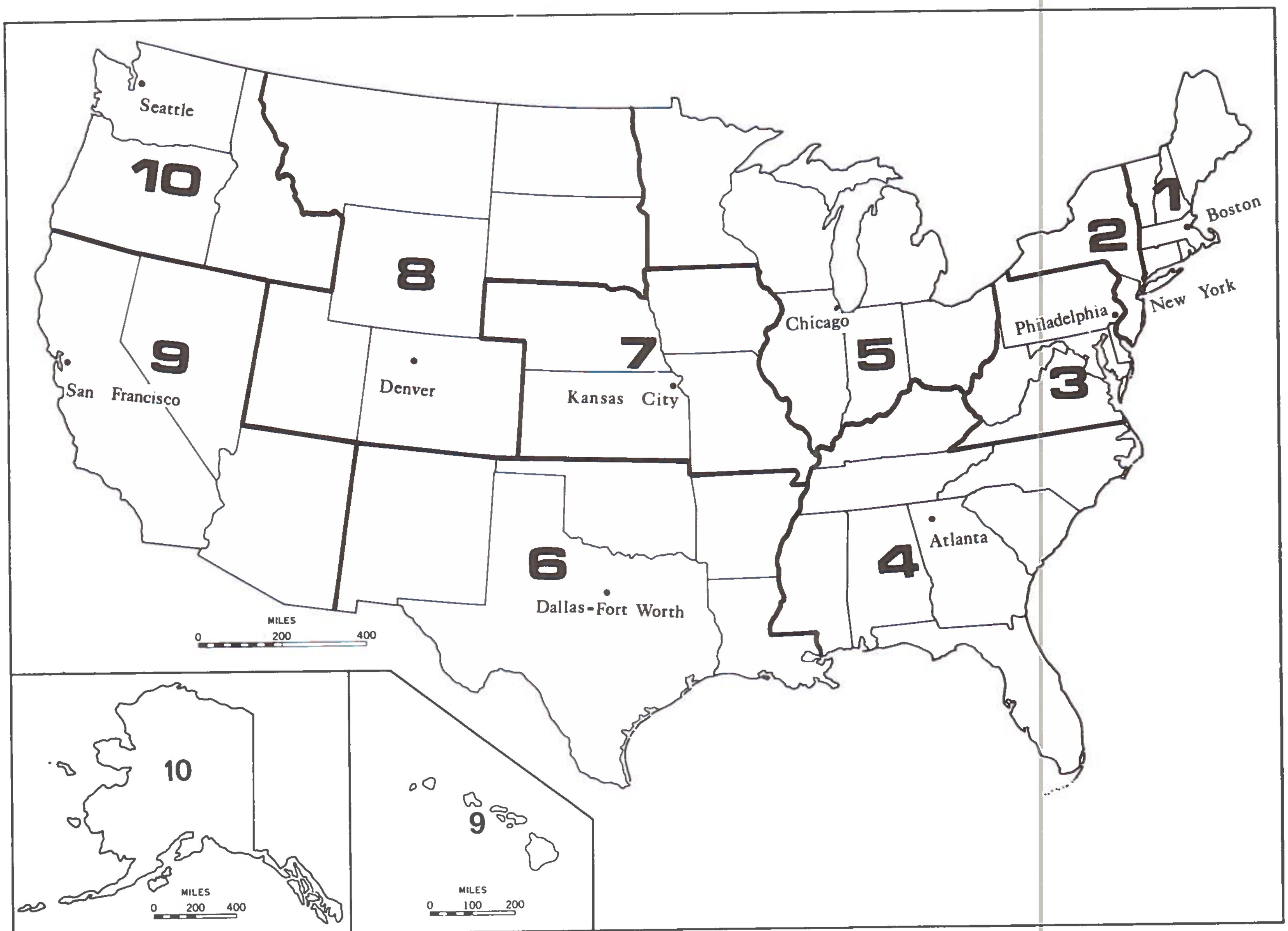
Figure 4.8 Criteria for Standard Metropolitan Statistical Areas the major federal statistical agencies, OMB periodically reviews and revises the designated SMSAs. New SMSAs may be created, the county configurations of existing SMSAs may be rearranged, older SMSAs may be split into two or more new SMSAs and sometimes the SMSA titles or code numbers are altered.[1] As of May 1973 there were 263 SMSAs in the United States with four additional SMSAs in Puerto Rico and each is assigned a four digit code number adopted by FIPS (See Section 2.1). The series of code numbers was originally gapped at intervals of forty and the last digit was always zero. However, as the universe of SMSAs expanded some gaps have been filled and there are three SMSA codes with non-zero final digits. Every one of the 50 United States except Wyoming and Vermont contain at least one SMSA and these units have gained widespread usage for the reporting and tabulating of data by private as well as public agencies (See Figure 3.6).

4.3.2 Standard Federal Regions

In 1969 the Office of Management and Budget established ten Standard Federal Regions of the United States. These regions divide the nation into groups of contiguous states with Alaska and Hawaii included in an appropriate continental sector (See Figure 4.9). Standard Federal Regions were delineated in order to establish uniform regional boundaries and regional office locations for federal agencies as a means of rationalizing the organization of service delivery

---

[1]The Bureau of Economic Analysis maintains a file on the historical composition of SMSAs (See Section 4.5).



Source: Office of Management and Budget

Figure 4.9 Standard Federal Regions

systems and to help assure that they are closely coordinated.

A regional council has been established in each one of the ten regions. Since it is important that these councils and the regional offices of federal agencies have access to easily relatable statistical information, OMB has directed that in all cases where federal agencies present data by region the ten Standard Federal Regions are to be the tabulating and reporting units. Agencies may for special or analytical purposes publish data for other regions in addition to the ten standard regions. However, only in exceptional cases where it can be demonstrated that the reliability and the validity of the data do not warrant the use of the ten Standard Federal Regions may other regional classifications be used. Thus, while the regions defined in the National Location Code, for example, were altered in order to conform to the Federal Standard Regions, other reporting units such as the economic areas defined by the Bureau of Economic Analysis were not changed because the Standard Federal Regions do not represent regional economic units of the United States.

References:

Executive Office of the President. Bureau of the Budget. Standard Metropolitan Statistical Areas. Washington, D.C.: U. S. Government Printing Office, 1967.

Executive Office of the President. Office of Management and Budget. "Publication of Statistical Data for Regions." OMB Circular A-46. Washington, D.C.: Office of Management and Budget, Executive Office of the President, 1971.

4.4 ECONOMIC DEVELOPMENT ADMINISTRATION

The primary function of the Economic Development Administration (EDA) is to promote the long range economic development of areas within the United States that have severe unemployment and low family income problems by aiding public and private enterprise to create new, permanent jobs in these areas. The EDA program includes public works grants and loans; business loans for industrial and commercial facilities; guarantees for private working capital loans; and technical, planning, and research assistance for economically distressed areas.

In carrying out these responsibilities the Economic Development Administration has established geographic subdivisions of the United States that serve as planning units, delineated areas according to the qualification for assistance criteria dictated by legislation, and defined growth centers.[1]

4.4.1 Regions and Economic Development Districts

There are six EDA regional commissions having jurisdiction over multi-state subdivisions of the United States (See Figure 4.10). These regions serve as large scale administrative, coordinating and planning units. Within these regions EDA has established a multi-county district program similar to that employed in France and the

---

[1]Economically distressed areas qualify for assistance under criteria established by the Public Works and Economic Development Act of 1965 (Public Law 89-136).

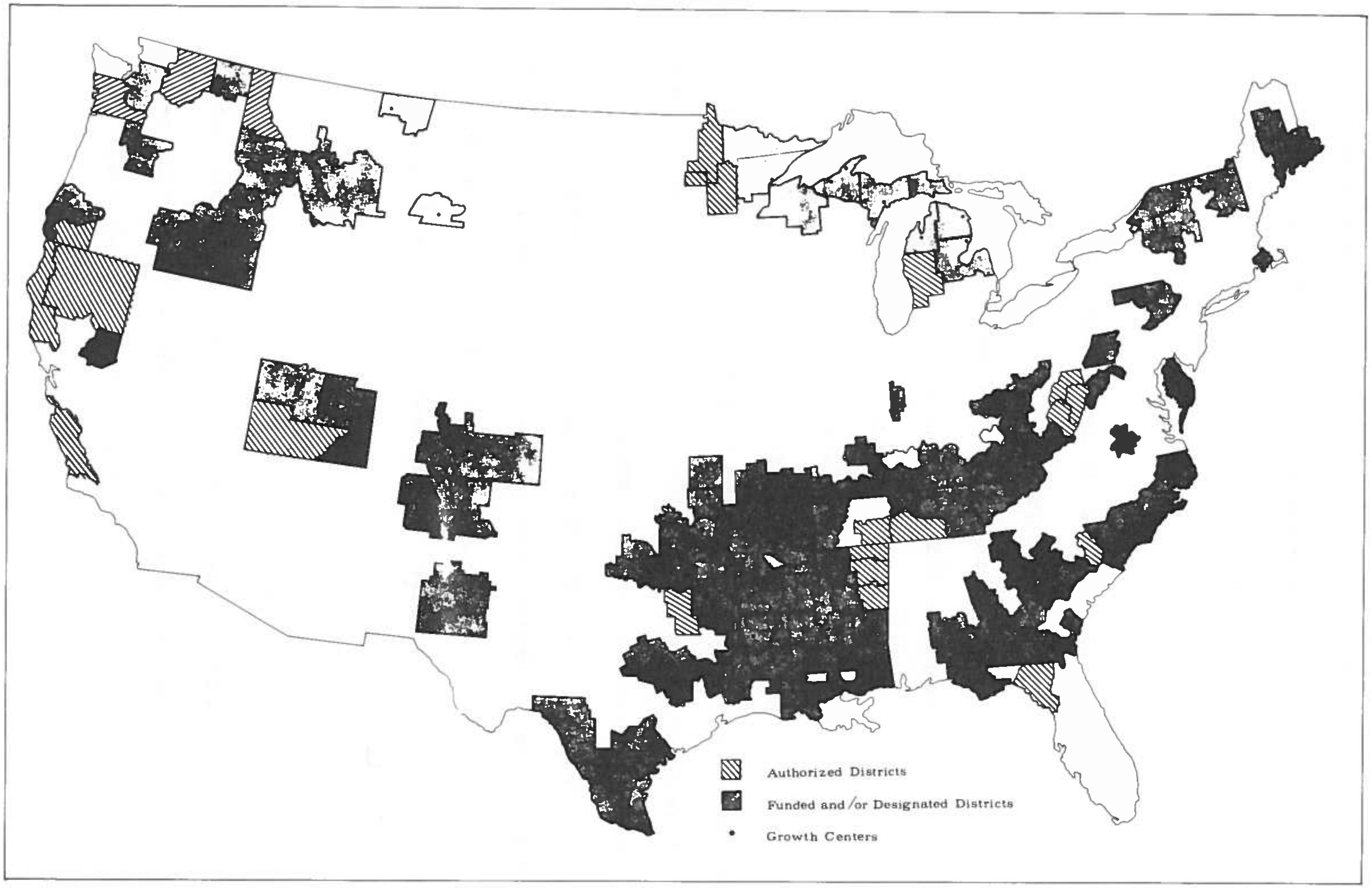
Source: Economic Development Administration

Figure 4.10 EDA Regions

United Kingdom. These districts are small scale planning and administrative units. Each of the economic development districts serves as the geographic framework in which the projects for qualifying individual redevelopment areas are coordinated and local resources pooled. There are approximately 120 authorized districts in the United States. Almost 100 of these have had initial overall economic development programs approved by EDA and have been officially designated as economic development districts (See Figure 4.11).

4.4.2 Redevelopment Areas

According to the legislation establishing the EDA program there are two types of redevelopment areas. Title I redevelopment areas are eligible only for public works grants and must have an annual average unemployment of 6 percent of more during the preceding calendar year in order to qualify for this type of assistance. There are only a few such designated redevelopment areas. Title IV redevelopment areas are eligible for the full range of assistance provided by EDA and must meet one of the following qualifying criteria: 1) unemployment 50% above the national average for three of the four preceding calendar years, or unemployment 75% above the national average for two of the three preceding calendar years, or unemployment 100% above the national average for one of the two preceding calendar years; or, 2) a population loss of 25% or more between the 1950 and 1960 Censuses; or, 3) a median family income of $2,264 or less. A redevelopment area may consist of an entire county, an Indian Reservation, a city, or



Source: Economic Development Administration

Figure 4.11 EDA Economic Development Districts combinations and segments of these units. There are approximately 450 designated Title IV redevelopment areas located within economic development districts in the United States.

4.4.3 Growth Centers

In addition to regions, economic development districts, and redevelopment areas, the Economic Development Administration has designated about 130 growth centers for economic development districts and over 50 growth centers for redevelopment areas. Several basic assumptions underlie the establishment of these centers. The first is that the provision of jobs, income, and local services in growth centers will benefit not only those centers, but residents of the surrounding areas as well. A second assumption is that accelerating and increasing the growth and prosperity of such centers will encourage residents of nearby depressed areas to seek work in a growth center instead of migrating to other areas. An additional assumption is that it would be less expensive to provide benefits to the target population (i.e., the unemployed and underemployed residents of depressed areas) through investments in growth centers rather than by funding projects in less developed areas.

4.4.4 Geographic Directory

The EDA geographic directory is a listing of all county and county equivalents by state in the United States with applicable EDA designations for each county and a list of redevelopment areas within each state (See Figure 4.12). The top of each printout page bears the name of the state and its Federal Information Processing Standard (FIPS) code, as well as the appropriate EDA regional office name and number. The following codes are included in the directory:

region: the region code beside a county name indicates that the county is included in the geographic area within the purview of a regional action planning commission (other than the EDA regional commission). Regional commissions are identified by name and code following the alphabetical listing of counties.

district: a district code on a line with a county name denotes that the county is part of the economic development district which is identified by name following the alphabetical listing of counties. A numeral nine appearing as the first digit in the three digit district code signals that the district boundary crosses state lines. Each numeric district code is followed by an alphabetic code of one or two characters signifying the current status of designation, funding or authorization of the district as follows:

A - district was authorized for establishment but does not receive any funding assistance from EDA.

F - district was authorized for establishment and receives funding assistance from EDA.

D - district is officially designated by EDA but does not receive funding assistance from EDA.

DF - district is officially designated by EDA and receives funding assistance from EDA.

MF - a planning unit which receives funding assistance from EDA but which is not eligible for designation as a district.

TF - a planning unit whose designation as a district has been terminated but which receives funding assistance from EDA.

T - a district whose designation has been terminated and which does not receive funding assistance from EDA.

STATE: 35 - New Mexico							REGIONAL OFFICE: 8 - Southwest
NAME	REGION	DISTRICT	COUNTY	AREA	SOD	GC	COMMENT
Bernalillo	04		001	0001	NE		
Catron	04		003	0003	PD		
Chaves	04	020 F	005	0005	NE		
Colfax	04	010 DF	007	0007	D		
Curry			009	0009	NE		
De Baca	04		011	0011	NE		
Dona Ana			013	0013	D		
Eddy		020 F	015	0015	D		
Grant	04		017	0017	NE		
Guadalupe	04		019	0019	D		
Harding	04		021	0021	NE		
Hidalgo			023	0023	NE		
Lea			025	0025	NE		
Lincoln	04	020 F	027	0027	E		
Los Alamos	04	010 DF	028	0028	NE		
Lina			029	0029	E		
McKinley	04		031	0031	D		
Mora	04	010 DF	033	0033	D		
Otero	04	020 F	035	0035	D		
Ouay			037	0037	NE		
Rio Arriba	04	010 DF	039	0039	D		
Roosevelt			041	0041	NE		
Sandoval	04	010 DF	043	0043	D		
San Juan	04		045	0045	E		
San Miguel	04	010 DF	047	0047	D		RDC Desig
Santa Fe	04	010 DF	049	0049	E		ELC Desig
Sierra			051	0051	NE		
Socorro	04		053	0053	D		
Taos	04	010 DF	055	0055	D		
Torrance	04		057	0057	D		
Union			059	0059	NE		
Valencia	04		061	0061	E		
Mescalero Res	04	020 F	035	0711	D		Indian Res
Zuni Res				0717	D		Indian Res
City of Las Vegas	04	010 DF	047	0047	D	002	RD Center
Santa Fe Dev Ctr	04	010 DF	049	0901	D	501	ED Center
North Central		010 DF					District
Southeastern		020 F					District
Four Corners	04						Region

Source: Economic Development Administration

Figure 4.12 EDA Geographic Directory

county: codes have been assigned to all counties or county equivalents. EDA's information system uses the Federal Information Processing Standard (FIPS) county codes.

area: a four digit numeric area code is assigned to identify the redevelopment areas listed at the end of the county sequence within a state. Each numeric area code has an alphabetic code in conjunction with it. The alphabetic code, which is displayed under the heading "SOD", describes the area status of designation as of the date the directory page was produced. The alphabetic codes used are as follows:

NE - an area which has never been eligible for designation as a redevelopment area.

E - an area which is currently eligible or qualified for consideration as a redevelopment area but which has either not submitted or has not received approval of an overall economic development plan.

D - an area which is currently designated as having met all conditions for receiving EDA assistance.

PE - an area which had an immediately prior status of eligible and from which it has been terminated in a previous fiscal year.

PD - an area which had an immediately prior status of designated and from which it had been terminated in a previous fiscal year.

TD - an area which had an immediately prior status of designated and from which it had been terminated in the current fiscal year.

growth centers: a three digit code in the column headed "GC" is used to signify a growth center. When the first digit is zero, the growth center is a redevelopment center; when the first digit is a five, the center is an economic development district center. The second and third digits identify the number of the center within the district.

comments: the comments used are primarily self-explanatory. When a county or area has a growth center located within it, a comment to that effect is displayed. Special names such as Indian Reservations, regions and districts are identified in this manner. If a center is situated in more than one county, this is explained in the comment.

The letters M, C, P, and U are used in this section to denote multi-county planning units. These counties are grouped for planning purposes and receive EDA planning assistance, but are not eligible for designation as districts.

References:

U. S. Department of Commerce. Economic Development Administration. "Geographic Directory". An unpublished, inhouse document. Washington, D.C.: Economic Development Administration, U. S. Department of Commerce, 1972.

U. S. Department of Commerce. Economic Development Administration. _Economic Development Districts: Composition and Status_. Washington, D.C.: U. S. Government Printing Office, 1970.

4.5 BUREAU OF ECONOMIC ANALYSIS

The bureau of Economic Analysis (BEA) which was formerly the Office of Business Economics (OBE) is engaged in an ambitious program of regional measurement, analysis and projection of economic activity. Initiated in 1965, this program represents the long term cooperative efforts of the U.S. Department of Commerce and the Economic Research Service (ERS), the U.S. Department of Agriculture with assistance from the Forest Service, the Water Resources Council (WRC) and the BEA. Several years ago the program acquired the acronym of OBERS.[1]

The objectives of the OBERS program are the development and maintenance of: 1) a regional economic information system with provisions for rapid and flexible data retrieval; 2) near term (1980-1990), mid term (2000), and far term (2020) projections of population, economic activity and land use for the United States and its geographic subdivisions; and 3) special analytical systems designed for use in water resource and other public investment planning.

With the publication of 1972 OBERS Projections: Regional Economic Activity in the United States (five volumes), the first two objectives have been accomplished. The OBERS report presents projections of economic activity for the United States as a whole, functional economic

---

[1] On January 1, 1972 the Office of Business Economics and the Bureau of the Census were merged to form the Social and Economic Statistics Administration (SESA). In the reorganization, the OBE was renamed the Bureau of Economic Analysis. Although the OBE has been renamed, widespread acceptance of the term OBERS has led to its continued use as a descriptive title of the projection project.

subdivisions of the United States, water resources regions and subareas, and the fifty states at decade intervals from 1980 to 2020. Included are projections of population, personal income, employment earnings of persons and output, with the last three items shown by industry. Also included are projections of land use by broad categories for the same period and the historical information from which the projections were derived in essentially the same geographic and industrial details as the projections. Further development will continue as economic measures and projection methodologies are improved, and as new data become available, updated and revised projections will be produced.

Analytical systems, the third objective, are in the process of development. These analytical systems are built upon the information system and the projections. They will provide systematic procedures for identifying the nature and estimating the magnitude of economic impacts resulting from alternative types and scales of resource development which may be considered.[1]

The geographic identification scheme employed in the OBERS information system is based on the county unit. County units serve as the building blocks from which a set of regional areas can be aggregated, then disaggregated and reaggregated into another set of regional areas delineated on the basis of different functional criteria.

---

[1]A more detailed description of the BEA analytical systems which is beyond the scope of this survey is contained in 1972 OBERS Projections: Economic Activity in the United States, Volume 1: Concepts, Methodology, and Summary Data (Washington, D.C.: U. S. Government Printing Office, 1972).

Also county units are well-established, fairly stable geopolitical areas for which the historical time series data necessary to OBERS projections are available. The three basic elements of the OBERS geocoding system are: BEA economic areas, WRC water resource regions and subareas, and a base file of geographic codes for counties.

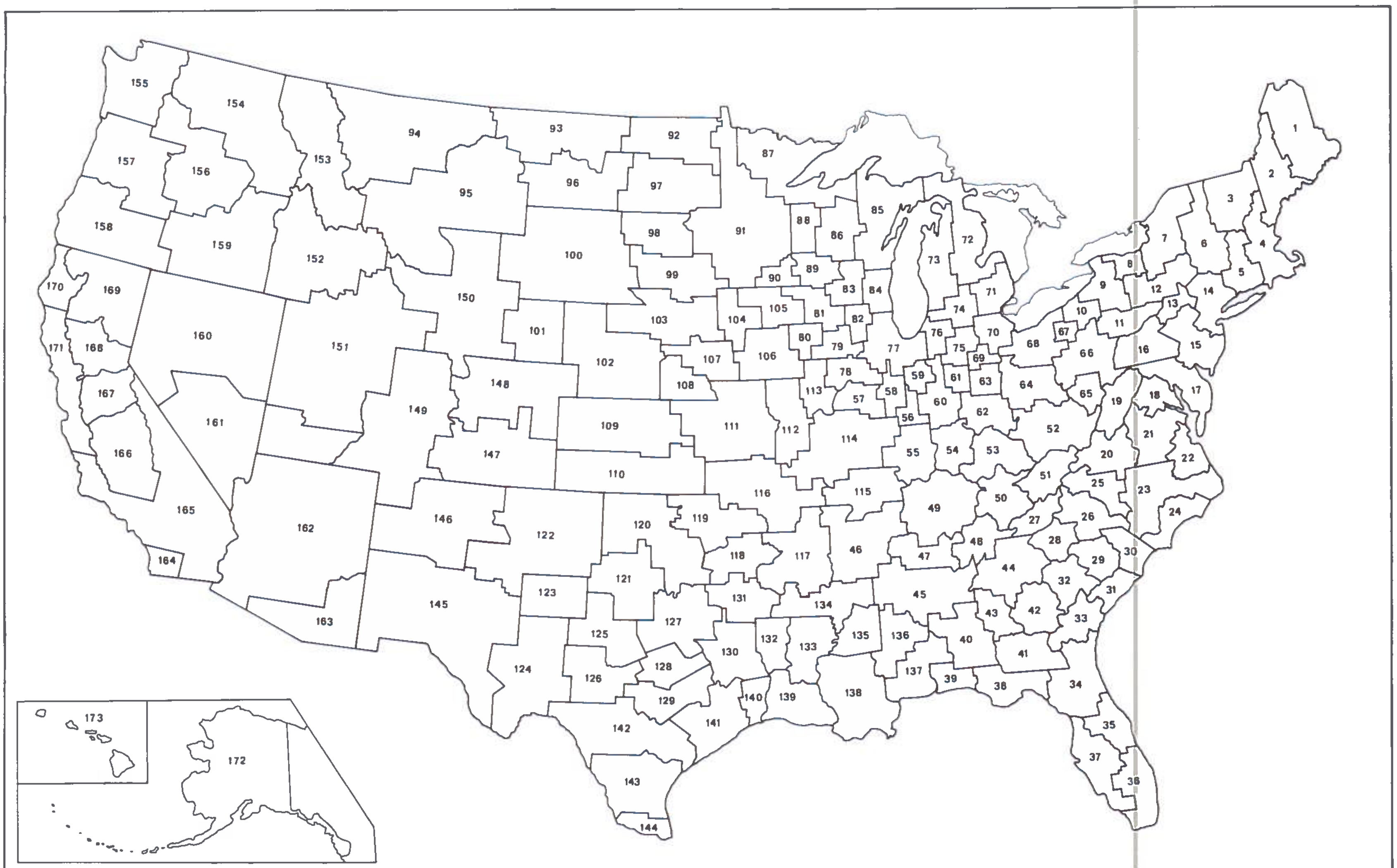
4.5.1 BEA Economic Areas

Central place theory with its emphasis on cities as the hubs around and within which integrated economic activity concentrates, provided the conceptual basis for the delineation of the BEA economic areas.[1] The application of this theory to the economic data relating to the counties of the nation resulted in 173 city-oriented areas, each with its hinterland in which there is a definite interaction of the various parts with the center and in which the establishments, both businesses and households, are functionally related. These nodal-functional economic areas have been designated BEA Economic Areas (See Figure 4.13).

One of the functional characteristics of these regions is that each combines its labor market and labor supply--the place of work and the place of residence of the labor force. There is, therefore, a minimum of commuting across economic area boundaries.

---

[1] A description of central place theory and further references on this subject are contained in: Brian J. L. Berry and William L. Garrison, "The Functional Bases of the Central Place Hierarchy," Economic Geography XXXIV (April 1958), pp. 145-154.



Source: Bureau of Economic Analysis

Figure 4.13 BEA Economic Areas

Each economic area has essentially two types of industries. One group constitutes the basic, or export, industries which produce goods and services most of which are exported to other areas, thus earning the means with which to purchase the specialized goods and services of other areas.

The production location of export types of goods and services is determined mainly by the costs associated with production processes requiring different input relationships and the comparative advantage of an area for the production of a commodity is determined by the area's relative endowment of the factors of production.

In addition to the basic, or export, industries, each area has another group of industries, termed residentiary, which are functionally related to the households and businesses of the area in that they produce most of the services and some of the goods required by the household sector and by other local business as intermediate products. Each of the areas approach self-sufficiency in regard to these residentiary industries which include general and convenience retail and wholesale trade activities, and those other services which are difficult or impossible to transport and which are most efficiently consumed in the vicinity of their production. Thus, the economic areas correspond to the closed trade areas of central place theory in which the number and type of residentiary establishments and their size and trade areas are bounded by the relative transportation costs from hinterland to competing centers.

Among economic areas the relationship of basic and residentiary industries and the composition of the latter vary according to factors such as type of basic industry, economic size of area, level of per capita income, economic maturity and type of surrounding areas. Despite these differentiating factors, interindustry relationships within each area exhibit general similarity and substantial stability although they do change as a result of secular trends and developmental threshholds (points at which local markets for intermediate or consumer products become large enough for local production to supplant all or a portion of imports). These characteristics of similarity and stability are what make the economic areas superior for projection purposes to other geographic areas.

The first step in the economic area delineation was the identification of the economic centers. Standard metropolitan statistical areas (SMSAs) which are general trade and labor market centers were chosen as the nodal centers. However, not all SMSAs are centers of economic areas because some are integral parts of larger metropolitan complexes. For example, the Jersey City, Newark, Paterson-Clifton-Passaic, Stamford, Norwalk, and Bridgeport SMSAs are all part of the New York City complex. In rural parts of the country where there are no SMSAs, cities with 25,000 to 50,000 population are designated as the economic centers.

After identifying the economic centers, the economic focus of the remaining counties was determined. The primary data used in this determination were the journey-to-work data from the 1960 Census of Population. Those data were summarized and posted on maps so as to show the gross commuting for each individual county to each adjacent county and to as many as 13 counties altogether if such commuting occurred. Counties were then associated with the economic centers in accordance with the commuting pattern.

In places where the commuting pattern of adjacent economic centers overlapped, counties were included in the economic area containing the center with which there was the greatest commuting connection. In the case of cities where the commuting pattern overlapped to a great degree, no attempt was made to separate them. Instead, both cities were included in the same economic area. Many counties were associated with an economic area not because of their commuting ties to the city itself, but because of their association with other counties which were tied to the economic center. Thus, for the first ring of counties around the central county, the criterion was commuting to the central county or to the first ring.

In the more rural parts of the country, the journey-to-work information was insufficient to establish boundaries of the economic areas. For these areas the road network and certain geographic features which would affect the possibility and time of travel to the economic centers, and the linkage of counties by other socioeconomic ties such as communications, cultural, recreational, and trade activities were major determinants.

4.5.2 Water Resource Regions and Subareas

In the publication entitled Water Resources Regions and Subregions for the National Assessment of Water and Related Land Resources, July 1970, the Water Resources Council presented a delineation of the nation into twenty-one water resources regions corresponding to major drainage patterns (See Section 4.2).

By application of a consistent set of criteria using hydrologic boundaries these regions have been further divided into tributary and main stem reaches entitled water resources subregions. These subregions cut across county lines where drainage conditions so indicate.

In the OBERS program counties and multi-county SMSAs form the geographic building blocks in any geographic classification system. It was necessary, therefore, to conform the water resources subregions to county and SMSA boundaries. The resulting multi-county delineations have been designated as water resources subareas. The water resources subareas, therefore, are county approximations of the hydrologically defined water resources subregions (See Figure 4.5).

4.5.3 County Base File

The basic file of geographic codes maintained by BEA in order to process and tabulate the information for the OBERS program and to aggregate statistics for economic regions, water resource areas and the states is a county file. This file contains a listing of county names along with the appropriate state name, state and county code, SMSA code, water resource region and subarea codes, BEA economic area code, and codes for land resource divisions (See Section 4.2).

Insofar as possible the Bureau of Economic Analysis has adopted the Federal Information Processing Standard for state and county designations (See Section 2.1). There are, however, some exceptions and alterations to the standard on the BEA file. In Virginia, for example, BEA does not recognize single independent cities as county equivalents. Independent cities are either listed in conjunction with the county to which it is most closely associated geographically and assigned the code of that county, or, in two cases, grouped to form a county equivalent unit and assigned a single code. Thus, while the current FIPSPUB 6 lists 95 counties and 39 independent cities in Virginia, each with a unique code, BEA recognizes only 98 coded county units in Virginia. There are other minor alterations of the FIPS standard in the BEA county file.

The Bureau of Economic Analysis also maintains a unique system of SMSA codes and a separate geographic file which indicates the historical county composition of the SMSAs. In addition to the SMSAs defined by the Office of Management and Budget (See Section 4.3), BEA has designated two other areas in the states where no official SMSA exists as SMSA equivalents. These are Burlington, Vermont (Chittendon County) and Cheyenne, Wyoming (Laramie County). In New England where the official SMSAs are made up of township units rather than whole counties, BEA has redefined the existing SMSAs so that they constitute aggregates of whole counties. Each of the SMSAs and BEA SMSA equivalents is assigned a three digit number. They are listed by title along with their county components and historical information on a special computer file which records the various SMSA changes since 1950 (See Figure 4.14).

SMSA	County	Title	State	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
338		Chicago, Ill.																										
	Lake		Ind.	*	*	*	*	*	*	*	*	*	06	(372)														
	Cook		Ill.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Du Page		Ill.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Kane		Ill.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Lake		Ill.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	McHenry		Ill.										06	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Will		Ill.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
339		Cincinnati, Ohio-Ky.-Ind.																										
	Clermont		Ohio														10	*	*	*	*	*	*	*	*	*	*	*
	Hamilton		Ohio	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Warren		Ohio														10	*	*	*	*	*	*	*	*	*	*	*
	Dearborn		Ind.														10	*	*	*	*	*	*	*	*	*	*	*
	Boone		Ky.														10	*	*	*	*	*	*	*	*	*	*	*
	Campbell		Ky.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Kenton		Ky.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
340		Cleveland, Ohio																										
	Cuyahoga		Ohio	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Geauga		Ohio														10	*	*	*	*	*	*	*	*	*	*	*
	Lake		Ohio	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	Medina		Ohio														10	*	*	*	*	*	*	*	*	*	*	*

Number followed by asterisks denotes month in which county joined SMSA
Number preceded by asterisks denotes month in which county left SMSA
Number in parentheses denotes new SMSA of departing county (991 = no SMSA)

Source: Bureau of Economic Analysis

Figure 4.14 Historical Composition of the SMSAs (1950-1974)

References:

U. S. Water Resources Council. 1972 OBERS Projections: Regional Economic Activity in the United States. Volume 1: Concepts, Methodology and Summary Data. Washington, D.C.: U.S. Government Printing Office, 1972.

4.6 ZIP MARKETING AREAS

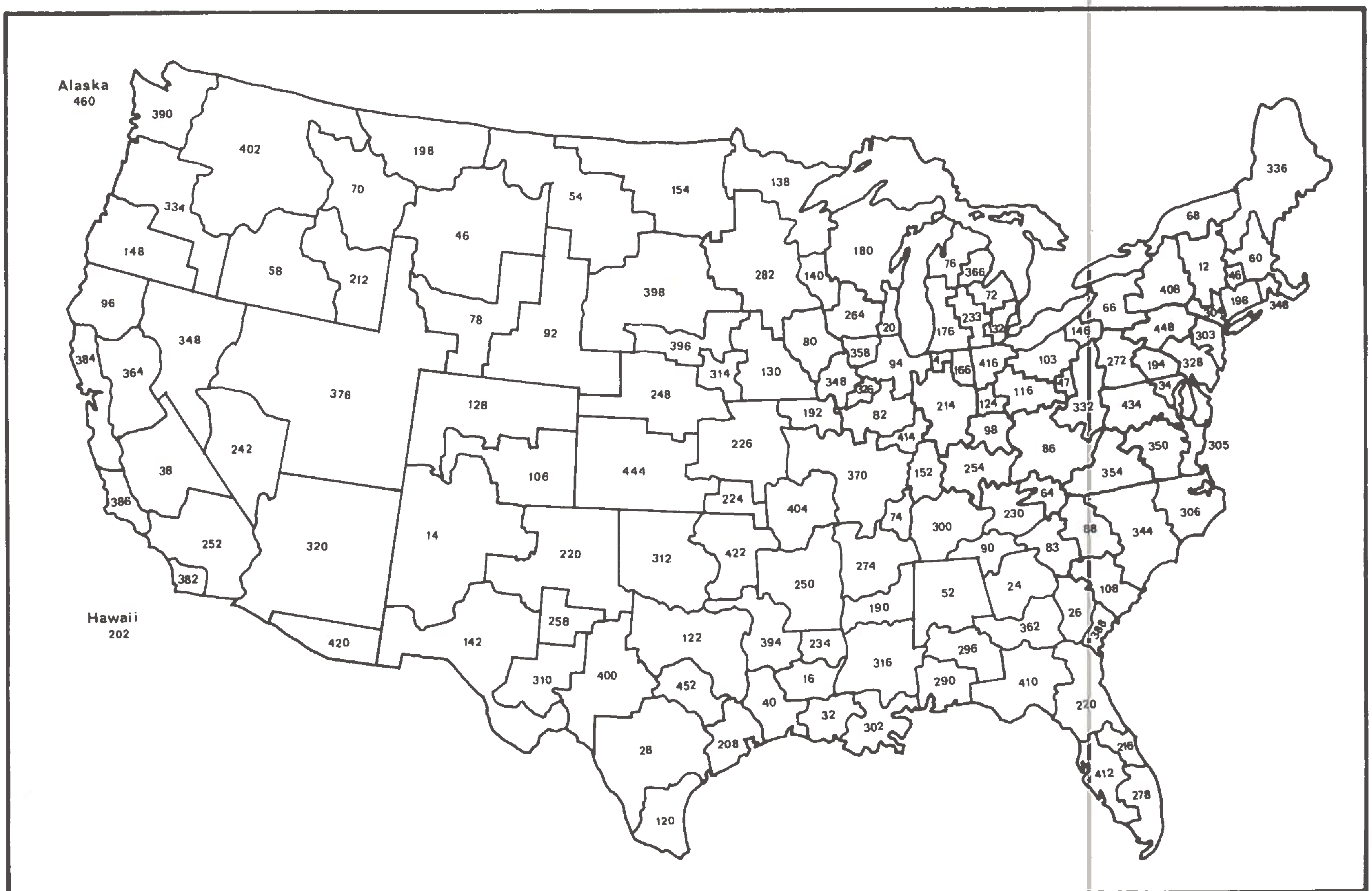
Since the ZIP code system was established in the United States by the U. S. Postal Service (See Section 4.1) many private companies and government agencies have found it necessary or advantageous to include ZIP codes among the geographic references maintained in their information systems. The development of mail-out/mail-back questionnaires for the 1970 Census, for example, required address coding guides with detailed ZIP code information. Some magazine publishers take advantage of the hierarchical structure of ZIP codes by assigning appropriate regional advertising and editorial pages on the basis of the first three digits of the ZIP code. In the private sector particularly, there are many other current applications of the ZIP code system to marketing analyses and experimentation. One such application was the development of a relatively stable geographic model for sales analysis based on three digit level ZIP code areas and television marketing areas called areas of dominant influence. These ZIP marketing areas were developed by Dik Twedt while he was working for Oscar Mayer and Company in 1967.

The first step in delineating ZIP marketing areas was defining the boundaries of the approximately 850 ZIP code sections reflected in the first three digits of the ZIP code. Since ZIP code sections do not follow county lines and much of the available sales data is tabulated by county, it was necessary to "force a fit" between counties and sections. ZIP sectional areas that split counties were extended to include all of each county in which section included the the majority of the population.

Next, ZIP code sectional areas were grouped to match the 205 television market areas of dominant influence as closely as possible. The area of dominant influence (ADI) is a geographic market design developed by the American Research Bureau which defines each market exclusive of another based on measurable television viewing patterns. An ADI consists of all counties in which the home market stations receive a preponderence of viewing. Viewing records obtained from more than a quarter million households serve as the source for the estimated total viewing hours in each county and the percent of these viewing hours shared by the combined stations in each market serving the county.

Finally the areas defined by matching county boundaries with ZIP code sectional areas and aggregated to areas of dominant influence are further aggregated into 136 ZIP marketing areas (See Figure 4.15) by overlaying known distribution patterns of major foods chains. Each marketing area is named for a major city within the area, they are listed alphabetically and then assigned a three digit code number ranging from 040 to 410 with sufficient interval gapping of numbers to allow for the designation of additional areas.

The principal advantage of a ZIP code based geographic model for sales analysis is that the ZIP code sectional areas are fairly stable units and the ZIP code number is widely recognized by all segments of business and general public, thus the users look-up problem is mini-



Source: Mr. Dik W. Twedt

Figure 4.15 ZIP Marketing Areas

mized. Also, the ZIP code areas are geared to the problem of delivering mail, and there is a relatively close parallel between highway networks, mail distribution, and the normal channels of distribution used by marketing organizations.

The disadvantages of this geographic system involve the inaccuracies that result from force matching various sets of boundaries. In the arbitrary allocation of split counties to a given ZIP marketing area, certain county sections are included in the wrong ZIP code sectional area. This was a problem in about 10 percent of the counties. However, in most cases the split was so disproportional that the allocation decision was an obvious one and only a minority of any given county areas was incorrectly assigned to a section. In every case where the problem existed, it was in fringe areas of a rural nature with low population density.

Once ZIP code sectional area boundaries were reconciled to county boundaries further compromises had to be made in reconciling these areas with ADIs. Some counties fell into one ZIP marketing area but were included in another area of dominant television influence. As in matching ZIP code sectional areas to county boundaries, this situation occurred only in fringe population areas. It was found that less than four percent of the population of the United States lived in counties which had to be switched from the correct area of dominant influence. In terms of the existing distribution pattern two out of the three switched counties actually seemed to belong more appropriately in the ZIP marketing area to which they were assigned rather than the actual

ADI county configuration simply because the ZIP code system was planned in accordance with actual distribution traffic patterns, whereas ADI data are subject to normal sampling error.

ZIP marketing areas have become a standard geographic model for sales and marketing analysis. In addition these areas were used in studies conducted for the Office of Civil Defense (See Section 4.9). It was found that although ZIP marketing areas do not follow topographic features as closely as the state economic areas and economic subregions delineated by D. J. Bogue and C. L. Beale (See Section 3.1), ZIP marketing areas reflect the more important correlation between population concentration and associated resources.

References:

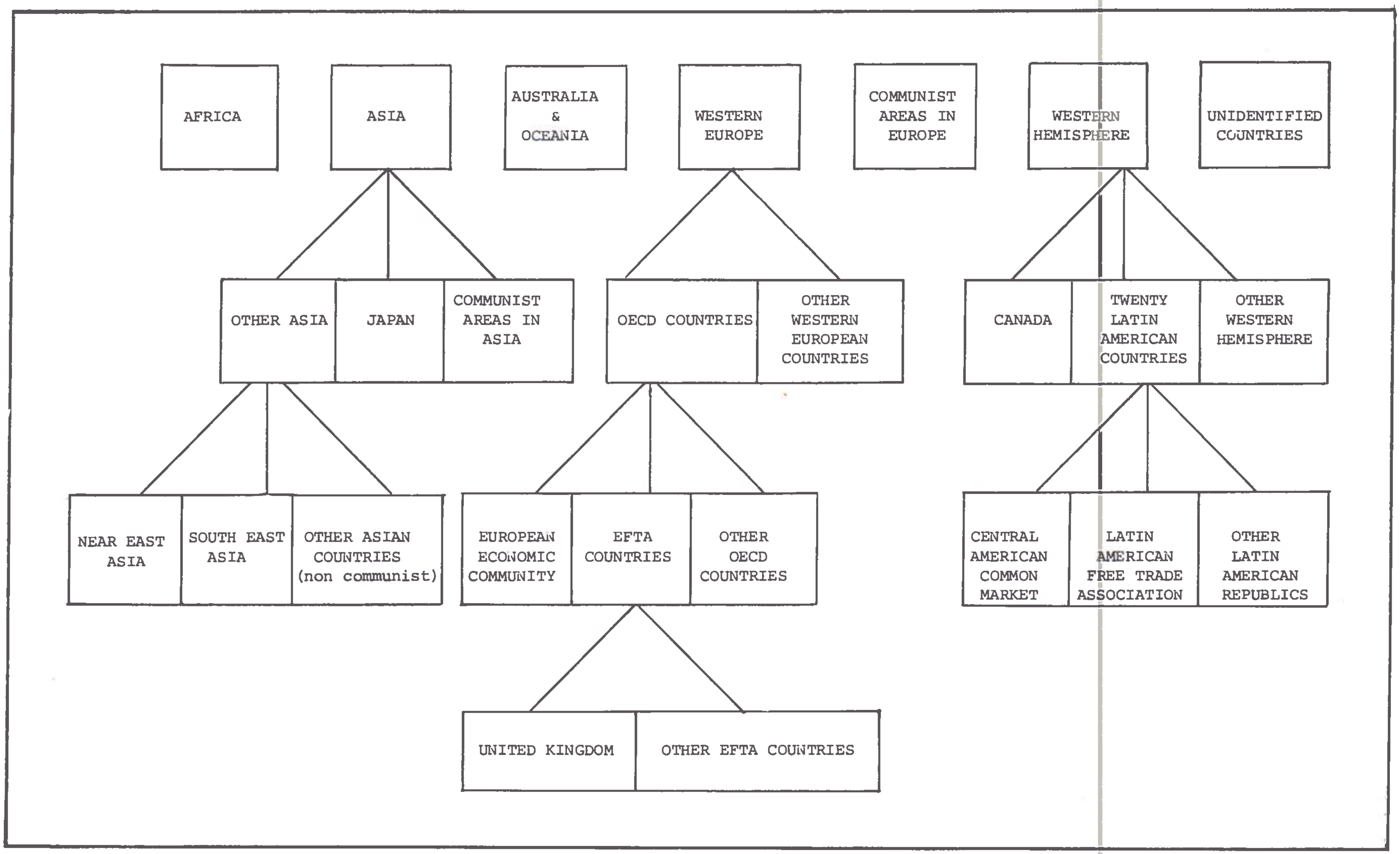
Dik Twedt. "Status of 'ZIP' as a Tool of Marketing Research." Proceedings of the Social Statistics Section, American Statistical Association. Washington, D.C.: American Statistical Association, 1967.

4.7 BUREAU OF THE CENSUS: FOREIGN TRADE DIVISION

The foreign trade statistics program, conducted by the Bureau of the Census, involves the compilation and dissemination of a large body of data relating to the imports and exports of the United States. These statistics are collected and tabulated to serve the needs of both government and nongovernment users who have a wide range of interests. The program, therefore, includes a variety of data presented in many different geographical and statistical arrangements and released in the form of reports or machine tabulation.

The unique geographic coding system used in compiling these statistics was devised by the Foreign Trade Division, U.S. Bureau of the Census and consists of three major code series: 1) coastal districts, subdivisions, and ports for foreign waterborne commerce; 2) country designations for all foreign commerce; and, 3) domestic customs districts and ports. In the tabulation of aggregate statistics the world is divided into the World Trade Areas represented in Figure 4.16.

Coastal districts, coastal subdivisions and foreign port designations are used in the tabulation of waterborne foreign trade statistics. The world is divided into 20 major coastal districts each of which is assigned a two digit number. These areas are further subdivided generally into coastal zones along a cluster of countries and coded uniquely by adding one digit to the code of the major district in which the subdivision is located. For example, west coast Central America and Mexico constitutes coastal district 04. Within this



Source: Bureau of the Census, Foreign Trade Division

Figure 4.16 Foreign Trade Statistics World Areas district the west coast of the Canal Zone (041) and the west coasts of Mexico, Guatemala, Honduras, Nicaragua, Costa Rica, and Panama (049) are the two subdivisions. Individual ports are assigned a five digit serial number within subdivisions. The port code, however, is independent of both the district and subdivision codes, providing unique identification in five digits.

The country designations for the foreign trade statistics is a three digit number code. Countries are sequenced by proximity in areas of continental land masses. Each country within the major divisions is assigned the same first digit as follows:

North America, northern area (1)	Asia (5)
North America, southern area (2)	Australia and Oceania (6)
South America (3)	Africa (7)
Europe (4)	Possessions of the United States (9)

There is no systematic assignment of the last two digits.

The country reported in the statistics as the country of origin is defined as the country where the merchandise was grown, mined or manufactured. In instances where the country of origin cannot be determined, the transactions are credited to the country of shipment. In foreign trade reports and tabulations, countries are presented by name and/or code. In some import reports, the country code is followed by the code 0 to indicate country of origin, or the code 1 to indicate country of shipment. The country designation UNIDENT is used to signify imports from countries which could not be identified because of illegible reporting on import entries or other shipment documents.

The country of destination is defined as the country of ultimate destination or the country where the goods are to be consumed, or further processed or manufactured. However, the statistics do not always reflect this definition since, at the time of exportation, the shipper may not know the ultimate destination of the shipment. In such cases, the shipments are credited to the last country to which the exporter knows that the merchandise will be shipped in the same form as when exported. It follows, therefore, that the statistics will tend to be overcounted for shipments to transshipment countries such as Hong Kong, Belgium, Netherlands, etc., and undercounted for other countries.

The 43 customs districts designated in the United States are coded with a two digit number ranging from 01 through 54. This irregular series of numbers is assigned to districts in a geographical sequence beginning with Portland, Maine, down the east coast, across the Gulf of Mexico, up the west coast and back to Chicago and St. Louis. Each of the ports within a district is designated by another two digit code. The number system for the 360 ports is irregular and unique only within each district. Therefore, the port codes must be prefixed with the district numbers in order to provide unique identification. Customs districts are grouped into nine regions, but these regions do not appear as a part of the location code.

Imported merchandise is coded to the district in which the merchandise is cleared through customs for entry into consumption channels or entered into customs warehouses. In either case, this may not be the district at which the merchandise arrived in the United

States. Imports by mail are credited to interior districts after having been transported in customs bond from the border point or seaport of entry into the United States. Imports of vessels moving under their own power or afloat are included under a single code (District 60) and are not reported by individual customs district of entry. Also, estimates for shipments valued at $250 and under are reported under a single code (District 70).

For exports transported by vessel or air, the shipments are coded to the customs district in which the merchandise is loaded on the vessel or aircraft which takes the merchandise out of the country. For exports by rail, truck, or pipeline, the shipments are coded to the customs district through which the merchandise crosses the United States border into foreign territory. Aircraft exported under their own power are credited to the customs districts from which they are flown out of the United States.

All mail exports are credited to one special district (District 80) and are not reported according to the geographical location of the post office from which the merchandise was mailed nor the customs district from which the mail leaves the United States.

Vessels exported under their own power or afloat are reported under District 60, and shipments individually valued under $100 are reported under District 70.

References:

U.S. Department of Commerce. Bureau of the Census. Foreign Trade Division. Guide to Foreign Trade Statistics: 1970. Washington, D.C.: U.S. Government Printing Office. 1970.

U.S. Department of Commerce. Bureau of the Census. Foreign Trade Division. Classification of Customs Districts and Ports: Schedule D and Classification of Country Destinations: Schedule C. Washington, D.C.: U.S. Government Printing Office, 1970.

4.8 INTERSTATE COMMERCE COMMISSION

When the Interstate Commerce Commission (ICC) was established by the Interstate Commerce Act of 1887, the railroads were the dominant mode of interstate transportation in the United States. The Act of 1887 was made applicable to all common carriers by railroad engaged in interstate or foreign commerce. It did not apply to common carriers wholly by water, but it did include common carriers partly by water and partly by rail, where they were under common control or arrangement for continuous carriage or shipment. Thus the primary responsibility of the ICC for many years was the regulation of railroads. As other modes of commodity transportation developed the jurisdiction of the ICC broadened to include trucking companies, bus lines, freight forwarders, oil pipelines, transportation brokers, and express agencies. Rail freight transportation, however, is still a major concern of the Interstate Commerce Commission and the source of one of its most important statistical series, the Rail Carload Waybill Statistics program which was initiated in 1946.[1]

---

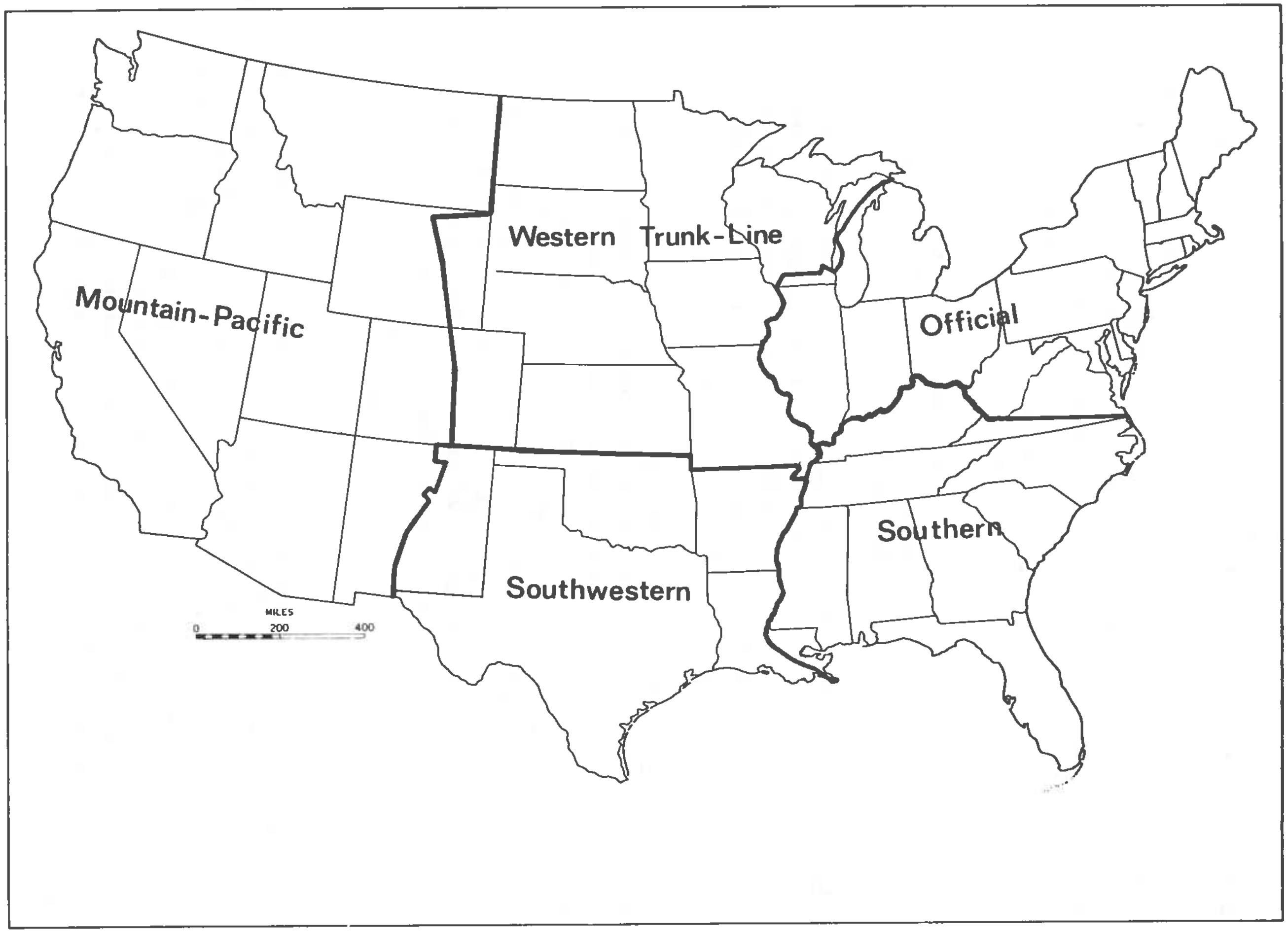
[1] In 1968 the Rail Carload Waybill Statistics program had to be discontinued as a result of budget restrictions. The U.S. Department of Transportation (DOT) recognizing the importance of these statistics began providing financial assistance to the Commission to maintain incoming waybills until funding and a processing plan could be re-established. In 1970, the Office of Systems Analysis and Information, DOT, began the development of a computer based system to process and tabulate waybill type statistics for any mode using the 1969 rail waybill statistics as a prototype. Future waybill tabulations will be under the direct supervision and management of the Federal Railroad Administration in the DOT.

The waybill statistical series is based on a one percent sample of audited revenue waybills for carloads terminated by line-haul operating railroads (as distinguished from a switching and terminal company) having three million dollars or more average operating revenues over a three year period and filed with the ICC.[1] The waybill sample includes import, export, transit, rebilled, and piggy back (TOFC: trailer-on-flatcar) traffic in the United States excluding all traffic originating in Mexico or Canada. It is a proportional sample and therefore it tends to yield a sample size for any group proportional to the number of waybills in the total frame. This means that commodities like bituminous coal and iron have a very large sample whereas commodities moved less frequently by rail have a lesser sample size. Hence, some of the bulk commodities are actually oversampled relative to hundreds of other commodities. Nevertheless, the waybill statistics are the most comprehensive, detailed and accurate reflection of the flow of rail freight by commodity in the United States.

Geographically the waybills of the one percent waybill sample are classified territorially according to the 5 major freight rate territories and 7 boundary or gateway territories established by the Interstate Commerce Commission. The five major freight rate territories are termed Official, Southern, Western Trunk-Line, Southwestern, and Mountain-Pacific. They are defined as follows (See Figure 4.17):

---

[1]An audited waybill is a document on which rate and freight charges have been verified in the railroad accounting offices.



Source: Interstate Commerce Commission

Figure 4.17 The Five Major ICC Freight Rate Territories

Official Territory: Commencing at the eastern terminus of the United States and Canadian boundary on the Atlantic Ocean and proceeding westwardly along the border to the Straits of Mackinac, thence southwesterly across Lake Michigan to Kewaunee, Wisconsin, thence southward along the shore of Lake Michigan to Manitowoc, Wisconsin, thence southward along the line of the Chicago and North Western Railway to Milwaukee, Wisconsin, thence northwest along the Milwaukee Railway to Rugby Junction, Wisconsin, thence south along the Soo Line to Duplainville, Wisconsin, thence west along the Milwaukee Railway through Watertown to Madison, Wisconsin, thence west and south along the North Western Railway through Montfort Junction, Wisconsin, to Benton, Wisconsin, thence southwest by airline to the intersection of the Wisconsin and Illinois boundary with the Mississippi River, thence south along the Mississippi River to the mouth of the Ohio River, thence eastward along the Ohio River to Cincinnati, Ohio, thence eastward along the Chesapeake and Ohio Railway to Kenova, Virginia, thence eastward along the Norfolk and Western Railway to its intersection with the Virginian Railway west of Roanoke, Virginia, thence east along the Virginian Railway to Suffolk, Virginia, thence northeast along the Norfolk and Western Railway to Norfolk, Virginia, and thence northeastward along the Atlantic Coast to the point of beginning.

Southern Territory: Commencing at Norfolk, Virginia, and proceeding westwardly along the southern border of Official Territory, as described above, to the mouth of the Ohio River, thence south along the Mississippi River to its mouth and thence east and north along the Gulf and Atlantic Coast to the point of beginning.

Western Trunk-Line Territory: Commencing at the Straits of Makinac and following the international boundary northeastward and thence westward to the western boundary of North Dakota, thence south along the Dakota and Montana line to the northern boundary of Wyoming, thence southwest by airline to Sheridan, Wyoming, thence southward by airline to Casper, Wyoming, thence southward along the line of the Burlington system to the Colorado and New Mexico line, thence eastward following the northern boundary of New Mexico, Oklahoma, and Arkansas to the Mississippi River, thence northward along the Mississippi River to the Wisconsin and Illinois line, and thence back to the point of beginning following the northwest boundary of Official Territory as described above.

Southwestern Territory: Commencing at the intersection of the Missouri and Arkansas boundary with the Mississippi River and proceeding westward along the southern boundary of Missouri, Kansas and Colorado to the point where the Santa Fe Railway crosses the Colorado and New Mexico line, thence southward along the Santa Fe Railway to El Paso, Texas, thence following the international boundary to the mouth of the Rio Grande River, thence along the Gulf Coast to the mouth of the Mississippi River and thence northward along the Mississippi River to the point of beginning.

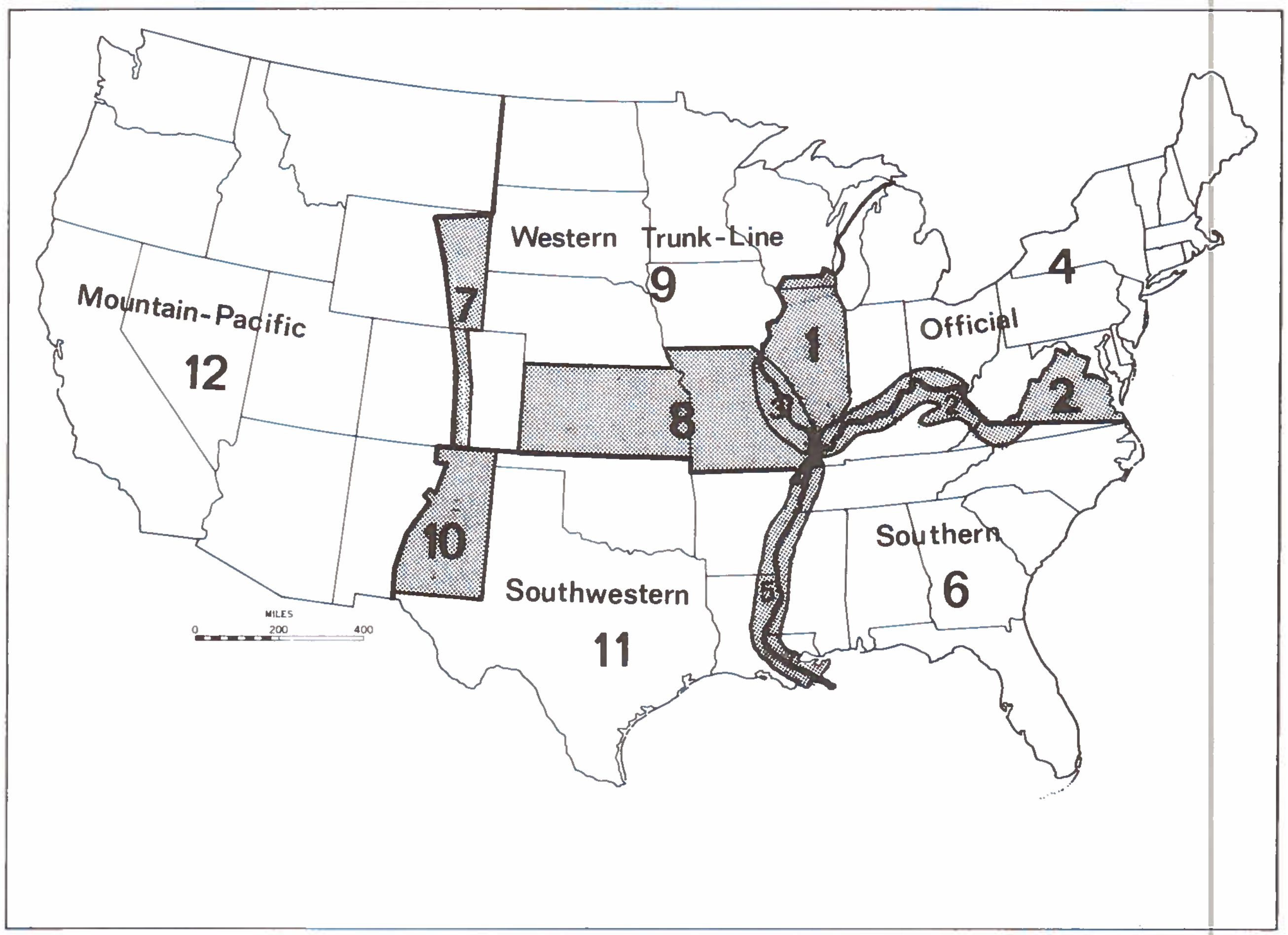
Mountain-Pacific Territory: That portion of the United States which lies west of the western boundaries of Western Trunk-Line and Southwestern Territories as described above.

Border zones and designated groups of gateway stations constitute another set of overlapping territories as shown in Figure 4.18. Traffic originating or terminating at border points or gateways is allocated to a territory according to the direction of movement. For example, traffic between Western Trunk-Line Territory and all points in Illinois, points in Indiana in the Chicago switching districts, and points in Wisconsin south and east of the northwest boundary of Official Territory is treated as Western Trunk-Line traffic.[1] Rate territory coding is assigned according to the schedule presented in Figure 4.19.

As part of the Rail Carload Waybill Statistics Program the straight-line or great circle distance between any two points recorded in the waybill sample was calculated using a network of unique Standard Point Location Code (See Section 3.6) coordinates. The calculation of straight-line miles is fully automatic and provides results of uniform high accuracy. Periodically, this file will be updated to accomodate additional points for existing modes and to incorporate new modes of transportation.

---

[1]For a detailed, official description of the border territories and gateway station groups shown in Figure 4.17, refer to "Interstate Commerce Commission: Freight Rate Territory Classification Used in Waybill Analysis", (Washington, D.C.: Interstate Commerce Commission, U.S. Department of Commerce, 1968) or Carload Waybill Statistics 1969, (Washington, D.C.: Office of Systems Analysis and Information, U.S. Department of Transportation, 1970).



Source: Interstate Commerce Commission

Figure 4.18 ICC Freight Rate Territories and Border Zones

BETWEEN	BETWEEN
1 and 1: Official	5 and 5: Southern
1 and 2: Official	5 and 6: Southern
1 and 3: Official	5 and 7: Southwestern - Western Trunk Line
1 and 4: Official	5 and 8: Southwestern
1 and 5: Official - Southern	5 and 9: Southwestern - Western Trunk Line
1 and 6: Official - Southern	5 and 10: Southwestern
1 and 7: Western Trunk Line	5 and 11: Southwestern
1 and 8: Western Trunk Line	5 and 12: Southwestern - Mountain Pacific
1 and 9: Western Trunk Line	
1 and 10: Western Trunk Line - Southwestern	6 and 6: Southern
1 and 11: Western Trunk Line - Southwestern	6 and 7: Southern - Western Trunk Line
1 and 12: Western Trunk Line - Mountain Pacific	6 and 8: Southern - Southwestern
	6 and 9: Southern - Western Trunk Line
2 and 2: Official	6 and 10: Southern - Southwestern
2 and 3: Official	6 and 11: Southern - Southwestern
2 and 4: Official	6 and 12: Southern - Mountain Pacific
2 and 5: Southern	
2 and 6: Southern	7 and 7: Western Trunk Line
2 and 7: Official - Western Trunk Line	7 and 8: Western Trunk Line
2 and 8: Official - Western Trunk Line	7 and 9: Western Trunk Line
2 and 9: Official - Western Trunk Line	7 and 10: Western Trunk Line - Southwestern
2 and 10: Official - Southwestern	7 and 11: Western Trunk Line - Southwestern
2 and 11: Official - Southwestern	7 and 12: Mountain Pacific
2 and 12: Official - Mountain Pacific	
	8 and 8: Western Trunk Line
3 and 3: Official	8 and 9: Western Trunk Line
3 and 4: Official	8 and 10: Southwestern
3 and 5: Official - Southern	8 and 11: Southwestern
3 and 6: Official - Southern	8 and 12: Western Trunk Line - Mountain Pacific
3 and 7: Western Trunk Line	
3 and 8: Western Trunk Line	9 and 9: Western Trunk Line
3 and 9: Western Trunk Line	9 and 10: Western Trunk Line - Southwestern
3 and 10: Southwestern	9 and 11: Western Trunk Line - Southwestern
3 and 11: Southwestern	9 and 12: Western Trunk Line - Mountain Pacific
3 and 12: Western Trunk Line - Mountain Pacific	
	10 and 10: Southwestern
4 and 4: Official	10 and 11: Southwestern
4 and 5: Official - Southern	10 and 12: Mountain Pacific
4 and 6: Official - Southern	
4 and 7: Official - Western Trunk Line	11 and 11: Southwestern
4 and 8: Official - Western Trunk Line	11 and 12: Southwestern - Mountain Pacific
4 and 9: Official - Western Trunk Line	
4 and 10: Official - Southwestern	
4 and 11: Official - Southwestern	
4 and 12: Official - Mountain Pacific	12 and 12: Mountain Pacific

ırce: Interstate Commerce Commission

Figure 4.19 ICC Freight Rate Territory Coding

References:

U.S. Department of Commerce. Interstate Commerce Commission. "Interstate Commerce Commission: Freight Rate Territory Classification Used in Waybill Analysis." Washington, D.C.: Interstate Commerce Commission, U.S. Department of Commerce, 1968.

U.S. Department of Transportation. Office of Systems Analysis and Information. Carload Waybill Statistics 1969. Washington, D.C.: Office of Systems Analysis and Information, U.S. Department of Transportation, 1970.

4.9 INSTITUTE FOR DEFENSE ANALYSES

Working for the Office of Civil Defense, the Institute for Defense Analyses (IDA) has conducted several studies concerning the distribution of population with respect to the concentration of other resources in the United States. The purpose of this research is to anticipate the spatial distribution and communication problems that might arise in recovery from nuclear attack and to determine the optimal geographic units for civil defense planning.

4.9.1 County Base Analyses

The population and economic resource data base compiled for these studies is a county data base. For each county a series of selected economic measures was recorded. These measures include among others: value added by manufacture, number of manufacturing establishments, iron and steel mills, petroleum refineries, electric generating stations, employment, retail trade sales, food sales, wholesale, trade sales, agricultural acreage, value of crops and livestock sold, and value of mineral shipments. Five separate sets of state and county codes were used in compiling the data base. In addition the designation for BEA economic areas (See Section 4.3) and ZIP marketing areas (See Section 4.6) were added to the file in order to gain a maximum of geographic flexibility.

The data, arrayed by county, were compared to population bases consisting of the counties containing the first 20 to 60 percent (at 10 percent intervals) of the total United States population. The counties included in the population percentages were designated and analyzed as core counties and the counties contiguous to them were also investigated. This core and contiguous county arrangement led to the adoption of ZIP marketing areas as a frame for civil defense recovery operations.

The core and contiguous county arrangement of data presentation was used because, in the event of a nuclear attack, the weapons effects on a target county would not be limited to the target (core) county, but would also affect those counties contiguous to it. Counties representing population levels from 20 to 60 percent of the national population were designated as the core counties and those counties geographically juxtaposed, the contiguous counties. Where core counties were found to be juxtaposed, the two or more counties were treated as a single core area. Each county code in the data base is listed with an alphabetic notation for a core or contiguous county and a group number was given to each cluster of counties. These clusters were then ordered according to the population ranking of the most populous county within the group. In no instance was a contiguous county listed more than once.

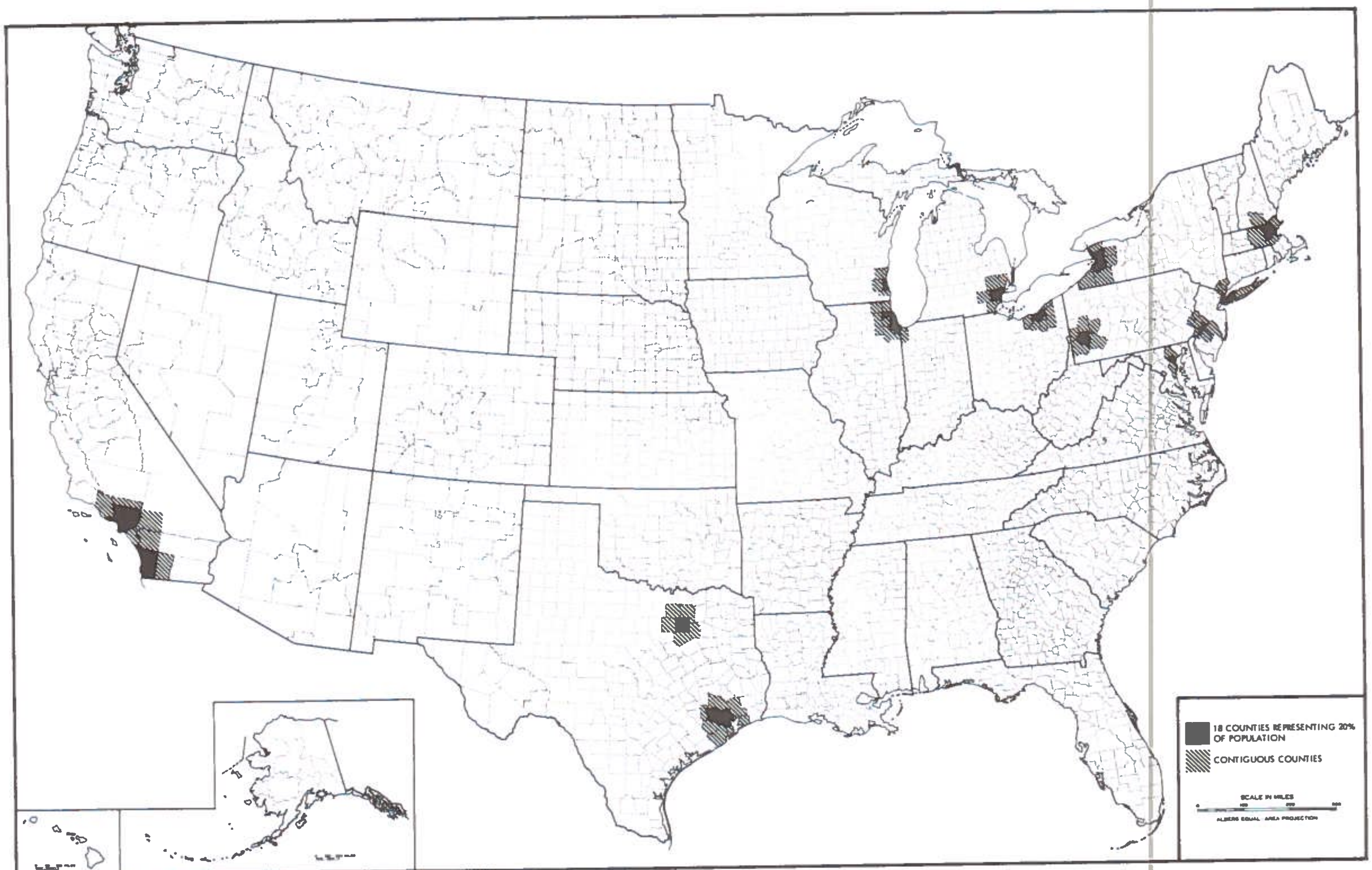
When data on contiguous counties were being assembled for this study, it was discovered that several outsized counties were causing analytical difficulties. Thus, the area of the 18 outsized counties was reduced to the average area of a county in the United States, i.e., approximately 1,100 square miles. Population and resources were assumed to be located in this 1,100 square mile area and the area was assumed to surround the main population center in the county.

Figures 4.20 through 4.22 illustrate the core counties at three of the cumulative population levels. Each cumulative 10 percent level was indicated separately to show the spread of population relative to land area as the base percentage increased. At the 20 percent level there were 18 core counties; at the 30 percent level there were 42 core counties; at the 40 percent level there were 76 core counties; and at the 50 and 60 percent levels there were 136 and 237 core counties respectively.

The concentration of each of the selected economic resources was presented in a manner similar to that used for the basic population data, all material for each resource was entered on tape by county. The data were then manipulated to show the relationship of the individual economic activity to population and other activities within the county. County totals were calculated and ranked as a percentage of the total national capacity. Manufacturing establishments, retail trade sales, population, wholesale trade sales and value added by manufacturing were also given as a percentage of the national capacity.

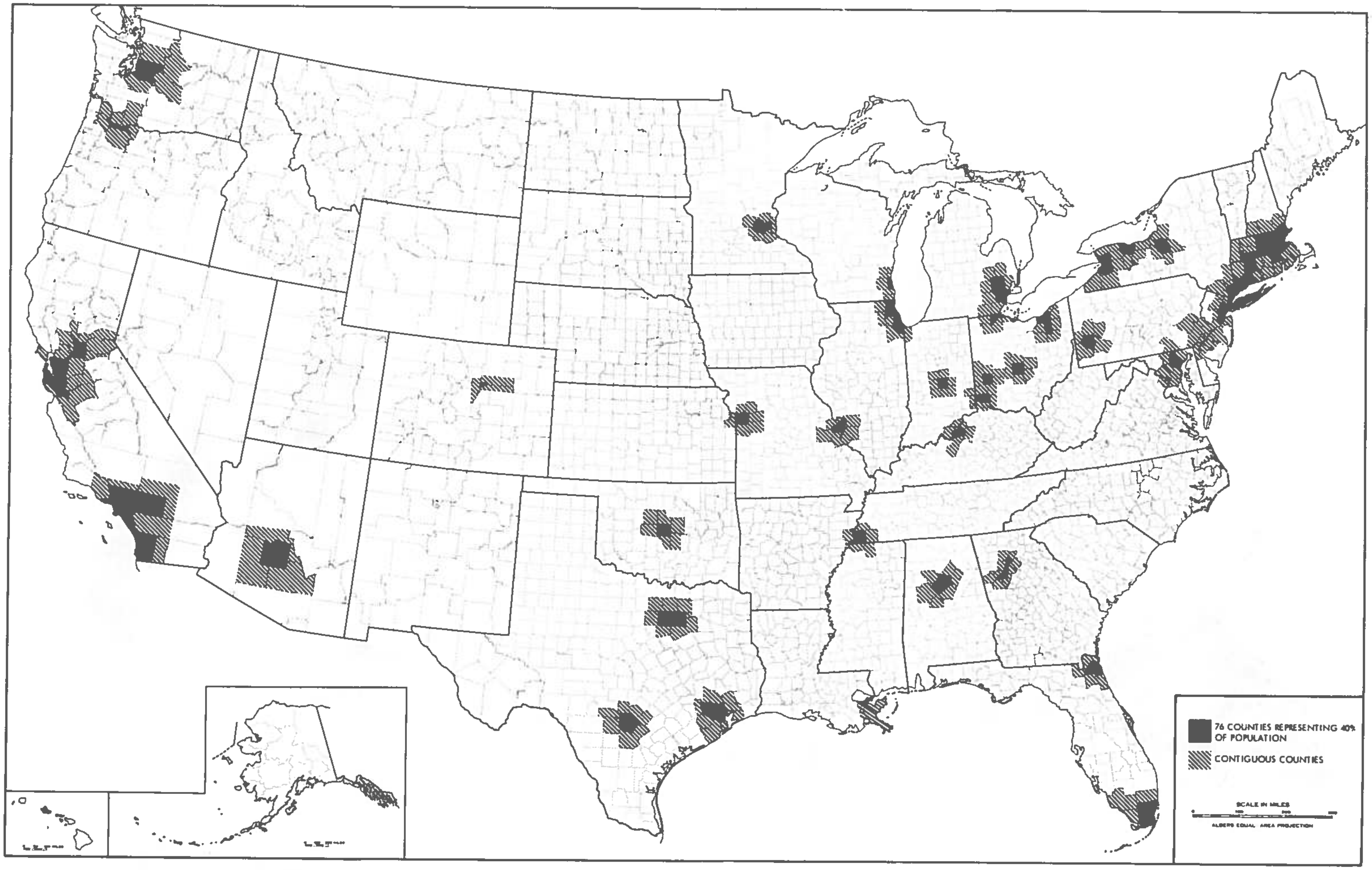
The results of the Institute of Defense Analyses study can be summarized as follows:

> 20% Population Base. At 20 percent population base the contiguous counties represent approximately 9.3 percent additional population yet occupy four times the land area of the core counties. Agricultural crop sales and mineral shipments are similarly quadrupled for contiguous counties compared to core counties, but the retail trade sectors are only half as large. Wholesale trade sales, however, are 7.3 percent for the contiguous counties as compared to 39.3 for the core counties.



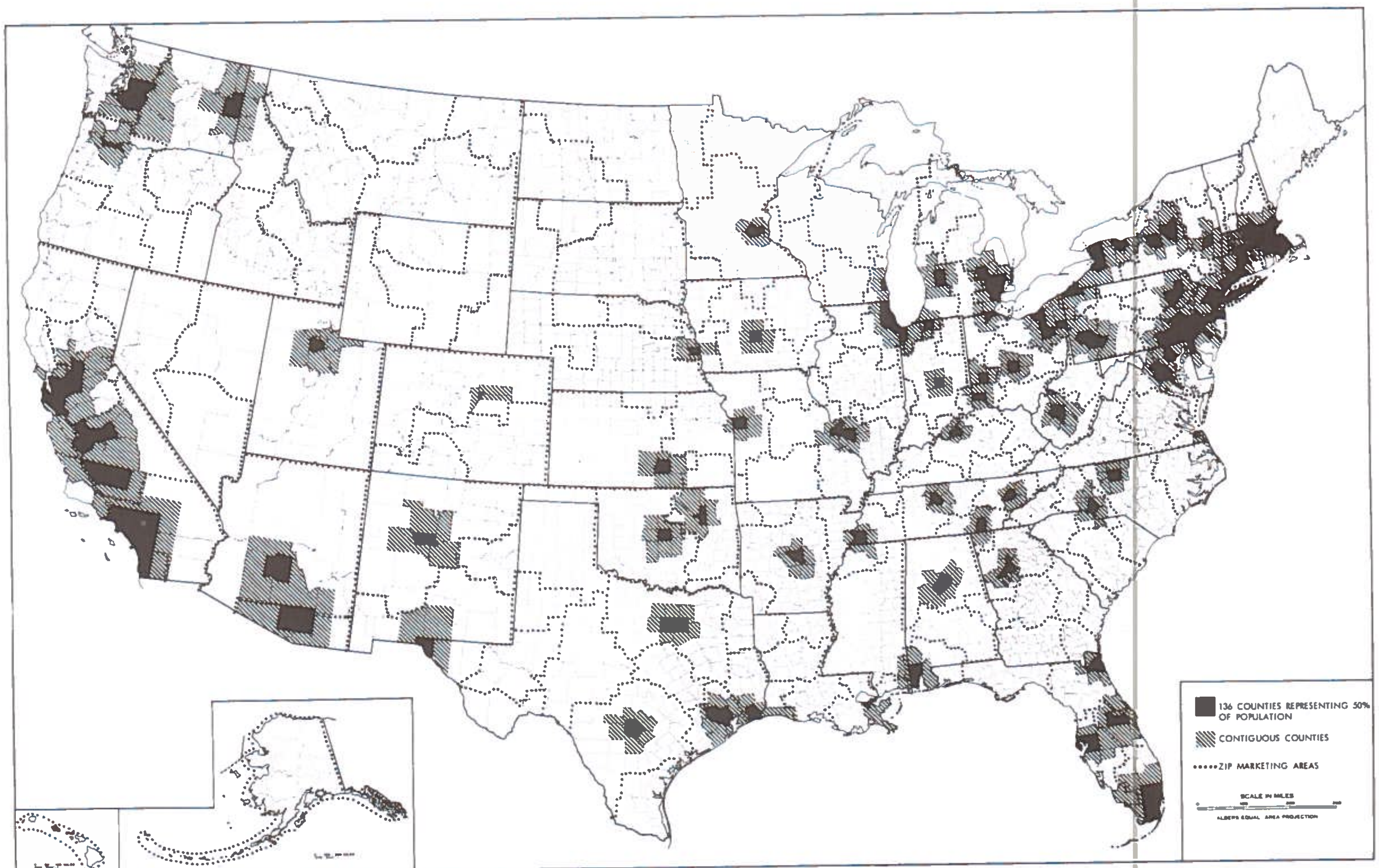
Source: Institute for Defense Analyses

Figure 4.20 IDA Core and Contiguous Counties Representing 20% Population Level



Source: Institute for Defense Analyses

Figure 4.21 IDA Core and Contiguous Counties Representing 40% Population Level



Source: Institute for Defense Analvses

Figure 4.22 IDA Core and Contiguous Counties Representing 50% Population Level

Energy and Steel Production. Approximately 50 percent of the iron and steel rolling mill capacity is located in only seven counties. Destruction of 50 percent of rolling mill capacity of the United States would result in little damage to collateral economic activity and would directly affect only about five percent of the population. Approximately 50 percent of the petroleum refining capacity is located in 20 counties, but these counties are more widely scattered than those representing the first 50 percent of iron and steel mill capacity. Removal of this capacity by attack would only have a small effect on other economic activity within the country and approximately 10 percent of the population would suffer damage. Fifty percent of the natural gas processing capacity is located in only 22 counties. These plants are located where there is very little economic activity or population concentration except for Los Angeles County. (In Los Angeles 1.2 percent of the total capacity is co-located with 3.4 percent of the population). Fifty percent of the operating electrical generating capacity is located within 59 counties; the remainder of the capacity is located in an additional 582 counties. Attack damage to these 59 counties would affect approximately 20 percent of the U. S. population and 40 percent of the total wholesale trade sales. Figure 4.23 illustrates those counties representing the first 50 percent capacity of these major resources within ZIP marketing areas.

Figure 4.22 includes the boundaries for ZIP marketing areas and illustrates the close relationship of the major core and contiguous county populations with the major marketing areas of the country. The 136 core counties of the 50 percent population base are located within 73 of the 136 ZIP marketing areas. Contiguous counties account for an additional four ZIP marketing areas. This close congruence between ZIP code marketing areas and the 50 percent population base county clusters prompted IDA to recommend that ZIP code marketing areas serve as the regional framework for civil defense recovery operations.

30% and 40% Population Bases. These two sets of core and contiguous counties have several similarities, and reflect certain trends as the core population base is increased. Though the core county population increases by 10 percent, the population in the contiguous counties remains at an almost constant 12.5 percent. The land area for each almost doubles with each increase in population. Manufacturing value added also increases with population, but different trends appear for retail trade and wholesale trade sales. In retail trade, sales for core counties continue to increase by only a moderate 10 percent and remain fairly stable. On the other hand, while wholesale trade sales in core counties increase by some 15 percent, an average 0.6 percent decrease is shown for contiguous counties. Thus, while land areas of the contiguous counties are still double that of the core counties (which have doubled as the population increased), the wholesale trade sales and general industry characteristics decrease simultaneously.

50% Population Base. As is easily seen in Figure 4.22, three major consolidated areas are clearly defined at the 50 percent population base: the California coast, the Great Lakes region, and the northeast corridor extending from Boston to Washington. At lower levels of population, these areas are not as distinct and at 60 percent they lose clarity from proliferation. The 60 percent break-out shows that one-half of the U. S. population resides in only 136 of the total 3,108 counties or 4.2 percent of the total land area. The contiguous counties represent an additional 14.2 percent of the population in 9.8 percent of the land area. Again, for the 10 percent increase in population, the land areas of both core and contiguous counties double and the number of contiguous counties also doubles. Agricultural crop sales and mineral shipments increase proportionately with land area, but only when the counties between the 40 and 50 percent population bases are included (the core area accounts for only approximately 11 percent of these sales). There is relatively little agricultural activity in these core counties, but a high percentage of industrial activity. The core counties account for 62.2 percent of the total manufacturing value added, 55.3 percent of retail trade sales, and 77.3 percent of wholesale trade sales. As seen in the 30 and 40 percent areas, the contiguous counties are responsible for a steady 13 percent of the total value added and retail trade sales, while there is a continued decrease in wholesale trade sales (only 5.6 percent for the contiguous counties of the 50 percent core counties).

Source: Institute for Defense Analyses

Figure 4.23 Major Resource Locations

4.9.2 Geometrical Cluster Analyses

Another study conducted by the Institute for Defense Analyses involves a hierarchical structuring of the national population based on aggregation of clusters of people into units called nodes. Within each county, clusters of urban populations are aggregated into nodes, with rural populations in the county aggregated into pseudo nodes. This "National Nodal Network" is delineated according to a set of rules that mathematically constrains each node into ellipsoidal or circular Gaussian configurations weighted by population. As part of the representation, which is the projected distribution of the national population in 1975, coordinates of the centers of gravity within each node, angles of rotation, and semi-major and semi-minor axes were calculated and included in the file. The hierarchical system of population clusters was designed to assess the relative problems of evacuation, involving minimum distances from target nodes and overcrowding ratios in outlying urban and rural nodes, in case of nuclear attack. It has also been used to supply national population and areal data for damage assessment models such as ANCET (Analytical Nuclear Casualty Estimating Techniques). Wider use of these constructs, however, requires that more data be added to the nodes to achieve more comprehensive nodal descriptions.

References:

Sachs, Abner and Timmermans, Judith A. Economic Structure of the United States Using the County as a Functional Base. Research Paper P-511 prepared for the Office of Civil Defense. Arlington, Va.: Institute for Defense Analyses, Program Analysis Division, 1969.

Timmermans, Judith A. User's Guide to Economic Data and Programs for IDA Civil Defense Studies. N-655 (R). Arlington, Va.: Institute for Defense Analyses, Programs Analysis Division, 1970.

Peterson, Dietrich L. and Schmidt, Leo A., Jr. Arrangements of U.S. Population by Urban and Rural Geometrical Clusters. Research Paper P-706 prepared for the Office of Civil Defense. Arlington, Va.: Institute for Defense Analyses, Program Analysis Division, 1970.

5. GRID AND COORDINATE CODES

5.1 GEODETIC LONGITUDE AND LATITUDE

Although the shape of the earth is actually that of an oblate ellipsoid, and its surface is highly irregular, the earth is treated as a perfect geometric sphere in regard to the network of intersecting lines inscribed upon the globe in order to define the location of a point upon the surface of the earth.[1] Thus, the geometric properties of the sphere, the circle and the arc are directly applicable to the earth and the geodetic grid which is inscribed upon it.

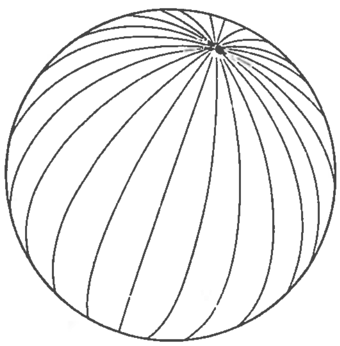
The geodetic grid is a system of "great" and "small" circles arranged in relation to the rotational axis of the earth (the poles).[2] It consists of a set of north-south lines called meridians and a set of east-west lines called parallels. All meridians are halves of great circles spreading farthest apart at the equator and converging to common points at the poles. An infinite number of meridians may be drawn on a globe. Thus, a meridian exists for any given point on a globe. All of the east-west lines called parallels are entire small circles except for the equator which is a complete great circle.

---

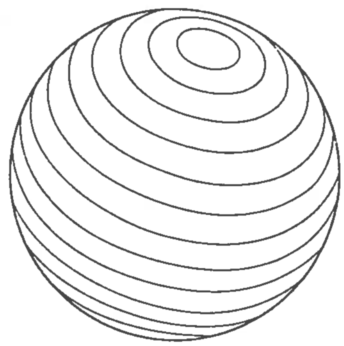
[1] An oblate ellipsoid is a spherical globe, compressed along the polar axis and bulging slightly around the equator. The earth, being about 26 miles less in diameter through the poles than in the plane of the equator is more accurately an oblate ellipsoid rather than a sphere. However, the spheroid deformation is minor, and the earth is generally treated as a perfect sphere with a smooth surface at sea level.

[2] If a perfect sphere is divided exactly in half by a plane passed through the center, the intersection of the plane with the sphere is the largest circle that can be drawn on the sphere and is known as a "great circle." Circles produced by planes passing through a sphere anywhere except through the center are smaller than great circles and are designated "small circles."

Parallels are always parallel to one another, and they always intersect meridians at right angles, except at the two poles. An infinite number of parallels may be drawn on the globe and therefore, every point on the globe except the north or south pole, lies on a specific parallel. These are the geometrical properties of "great" and "small" circles upon which the geodetic coordinate system is based.



Meridians



Parallels

The location of points on the earth's surface is defined by the length of arcs measured along these meridians and parallels. Parallels are measured in degrees north or south from the equator. Meridians are measured in degrees east or west from the prime meridian.[1] The longitude and latitude of a point read in degrees, minutes, and seconds, accurately and uniquely locate that point on the surface of of the earth in reference to the geodetic grid.[2]

---

[1]The prime meridian is almost universally accepted as that which passes through the Royal Observatory at Greenwich, near London, England, and is often referred to as the meridian of Greenwich. This meridian has the value of 0° longitude.

[2]There are 60 minutes to a degree, and 60 seconds to a minute.

The longitude of a place is the length of the arc of a parallel between that place and the prime meridian. The longitude of any given point on the globe is measured eastward or westward from this meridian, whichever is the shorter arc. Longitude may thus range from 0° to 180°. The actual length, in miles of a degree of longitude will depend upon where it is measured. At the equator one degree of longitude equals approximately 69 statute miles. At the 60th parallel the length of one degree longitude is about one-half that distance, 34 1/2 miles.

The latitude of a place is the length of the arc of a meridian between that place and the equator. Latitude may thus range from 0° at the equator to 90° north or south at the poles. The length of a degree of latitude is almost the same as the length of a degree of longitude at the equator, slightly over 69 statute miles. Assuming that the earth is a perfect sphere, the length of one degree of latitude never varies. However, to be more precise, and to take into account the oblateness of the earth, a degree of latitude actually changes slightly in length from the equator to pole (See Figure 5.1).

Along the 40th parallel, the area defined by one degree of latitude and one degree longitude covers approximately 3,557 square miles. Thus, if the longitude and latitude of a point in Harvard Square, Cambridge, Massachusetts is given to the nearest degree (42° latitude, 71° longitude), this reading would locate that point within a 3,557 square mile grid zone. If the longitude and latitude of this point is read to the nearest minute (42° 12' latitude, 71° 05' longitude)

Latitude (Degrees)	Length of 1° of Latitude		Length of 1° of Longitude	
	Statute Miles	Kilometers	Statute Miles	Kilometers
0	68.704	110.569	69.172	111.322
5	68.710	110.578	68.911	110.902
10	68.725	110.603	68.129	109.643
15	68.751	110.644	66.830	107.553
20	68.786	110.701	65.026	104.650
25	68.829	110.770	62.729	100.953
30	68.879	110.850	59.956	96.490
35	68.935	110.941	56.725	91.290
40	68.993	111.034	53.063	85.397
45	69.054	111.132	48.995	78.850
50	69.115	111.230	44.552	71.700
55	69.175	111.327	39.766	63.997
60	69.230	111.415	34.674	55.803
65	69.281	111.497	29.315	47.178
70	69.324	111.567	23.729	38.188
75	69.360	111.625	17.960	28.904
80	69.386	111.666	12.051	19.394
85	69.402	111.692	6.049	9.735
90	69.407	111.700	0.000	0.000

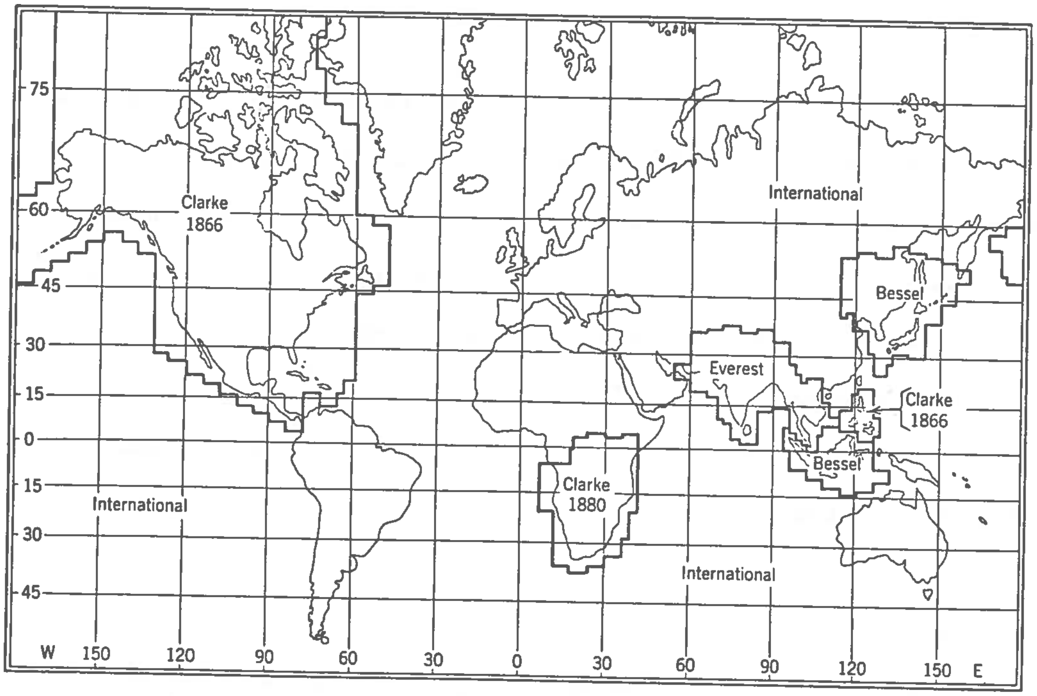
Source: Arthur N. Strahler, Physical Geography

Figure 5.1 Lengths of One Degree Longitude and One Degree Latitude

Harvard Square would be located within approximately 1.32 square miles. And, a reading of longitude and latitude which is carried to the nearest second would pinpoint this location within approximately one hundred square feet.

Plotting the geodetic network of meridians and parallels upon a plane surface involves two important transformations: determining the exact lengths of degrees of latitude and longitude based on the dimensions of the earth's ellipsoidal form; and projecting a representation of the earth's surface upon a plane.

Five sets of ellipsoid dimensions are currently employed to determine the lengths of degrees of latitude and longitude. These are: 1) the International Ellipsoid, whose values were computed by J. F. Hayford of the U. S. Coast and Geodetic Survey in 1909 and adopted by the International Geodetic and Geophysical Union in 1924; 2) the Clarke Ellipsoid of 1866 computed by A. R. Clarke, the head of the English Ordnance Survey; 3) the Clarke Ellipsoid of 1880, a recomputation by General Clarke; 4) the Bessel Ellipsoid computed in 1841 by a Prussion astronomer; and 5) the Everest Ellipsoid of 1830. Each ellipsoid is identified by specific values of its semi-major axis and semi-minor axis. For purposes of a unified system of international mapping, the world is divided up into areas, each assigned to one of the five ellipsoids.



The representation of the surface of the globe in an approximately true relationship on a flat surface is accomplished through some form of map projection at a reduced scale. The resulting maps always contain certain distortions in direction or distance between two points on the ground and the same two points as projected on the map. However, there are a number of ways in which to manipulate this deformation. Many types of map projection have been devised; each intended to achieve a certain objective such as true area, true shape, or true direction from a specific point. Of the many projections only two are commonly used for maps of areas lying between latitudes 80° North and 80° South. One is a projection onto a secant cone whose axis is identical with the axis of rotation of the earth and is called the Lambert Conformal Conic projection. The other is a projection onto a cylinder whose axis lies in the plane of the equator. It is called the Transverse Mercator projection in the United States but in Europe is usually called the Gauss-Kruger projection.

The curvature of the surface of the earth is so slight that for small areas the earth may be regarded as a plane without the introduction of serious distortions.

References:

Arthur N. Strahler. Physical Geography. 2nd Edition. New York: John Wiley and Sons, Inc., 1960.

Glenn T. Trewartha; A. H. Robinson; and E. H. Hammond. Fundamentals of Physical Geography. New York: McGraw Hill Book Company, 1961.

5.2 MILITARY GRID REFERENCE SYSTEM

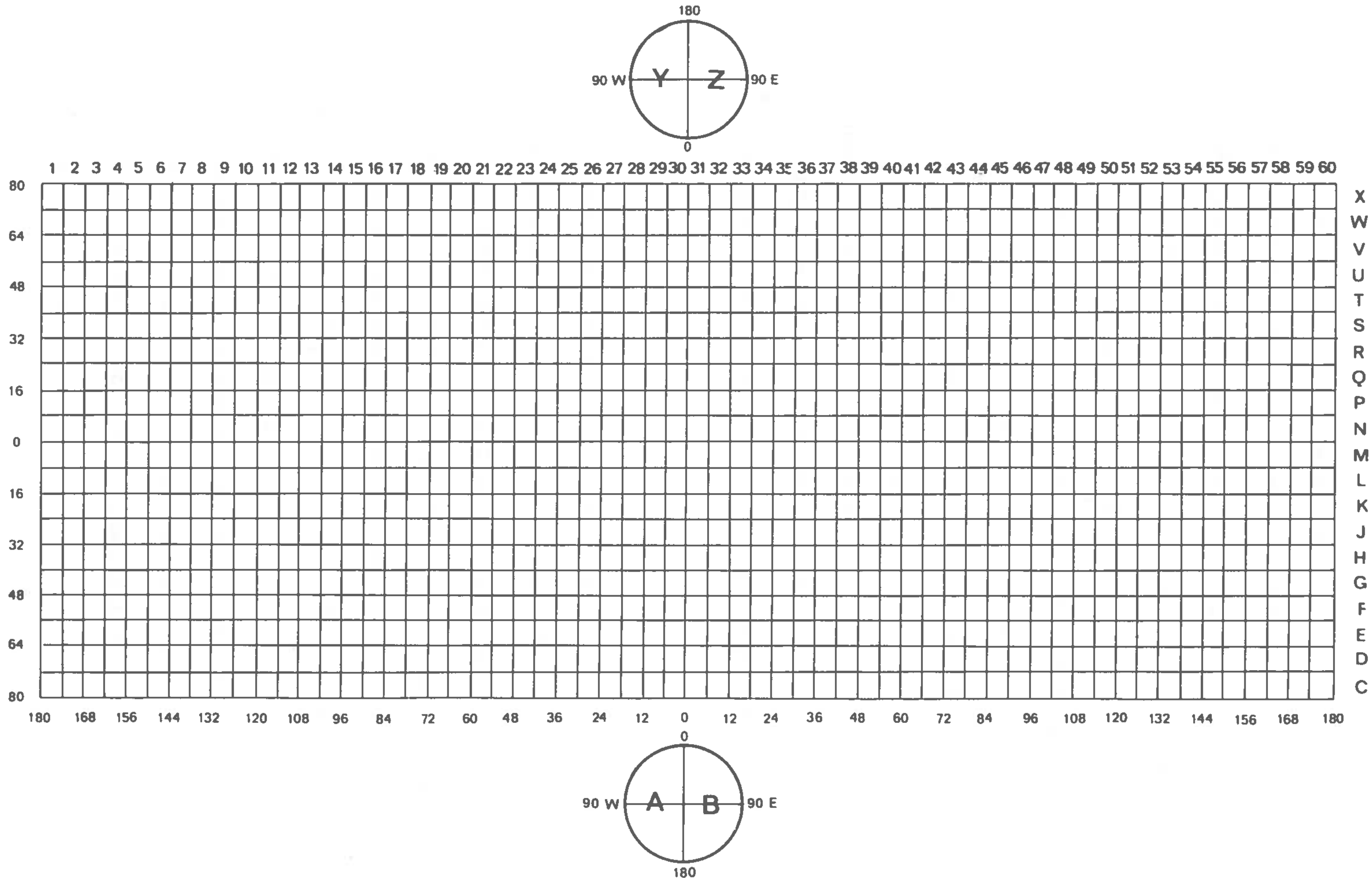
In 1951 the United States Army drafted a standardized worldwide geographic reference system known as the Military Grid Reference System which was developed in order to eliminate the irregularities of the various grid systems used previously by the Army, Navy and Air Force. The system has several major advantages in that only 60 zones are needed to extend around the entire world, all zones are exactly alike, every zone uses the equator and its own central meridian as the origins of coordinates, and the applicable mathematical equations are the same for all zones. It is global in scope, and yet it is precise enough to satisfy artillery requirements.

The Military Grid Reference System is designed for use on maps with Universal Transverse Mercator (UTM) and Universal Polar Stereographic (UPS) grids. The Universal Transverse Mercator is a conformal or orthomorphic projection which correctly represents the shape of smaller geographic features at the expense of changing scale over large tracts of land and bodies of water. This type of projection best meets the needs of military use between 80° north and 80° south latitudes. An azimuthal map projection has characteristics that are desirable for mapping polar areas and for construction of air distance charts; thus the Universal Polar Stereographic grid, which is one such azimuthal projection, is employed by the military in the polar regions.

Using these UTM and UPS projections the world is divided into large geographical areas extending 6° in longitude and 8° in latitude. They are identified by a unique two character grid zone designation. The grid zones are subdivided into 100,000 meter squares and each square is assigned another two character identification. Geographic references within these 100,000 meter squares are given to the desired accuracy in terms of the E (east) and N (north) grid coordinates of the point. Ordinarily, a reference is expressed only in numerical terms. When reporting beyond a 100,000 meter square, however, the numerical reference is prefixed with the 100,000 meter square identification consisting of two letters. When reporting beyond the 6° by 8° geographical area, the reference is prefixed with the grid zone designation consisting of one number and one letter.

The basic unit of measure in the Military Grid Reference System is the meter. Grid intervals of 1,000 meters are applied to large scale maps (1:25,000 and 1:50,000). Grid intervals of 10,000 meters are applied to medium scale maps (1:100,000 and 1:250,000).

Grid Zone Designations. Between 80° north and 80° south, the columns of the quadrilateral zones 6° wide are identified by UTM grid zone numbers (See Figure 5.2). Starting at the 180° meridian and proceeding easterly, the columns are numbered 1 through 60 consecutively. The rows 8° high are identified by letters. Starting at 80° south and proceeding northward to 80° north, the rows are lettered consecutively C through X (with the letters I and O omitted).

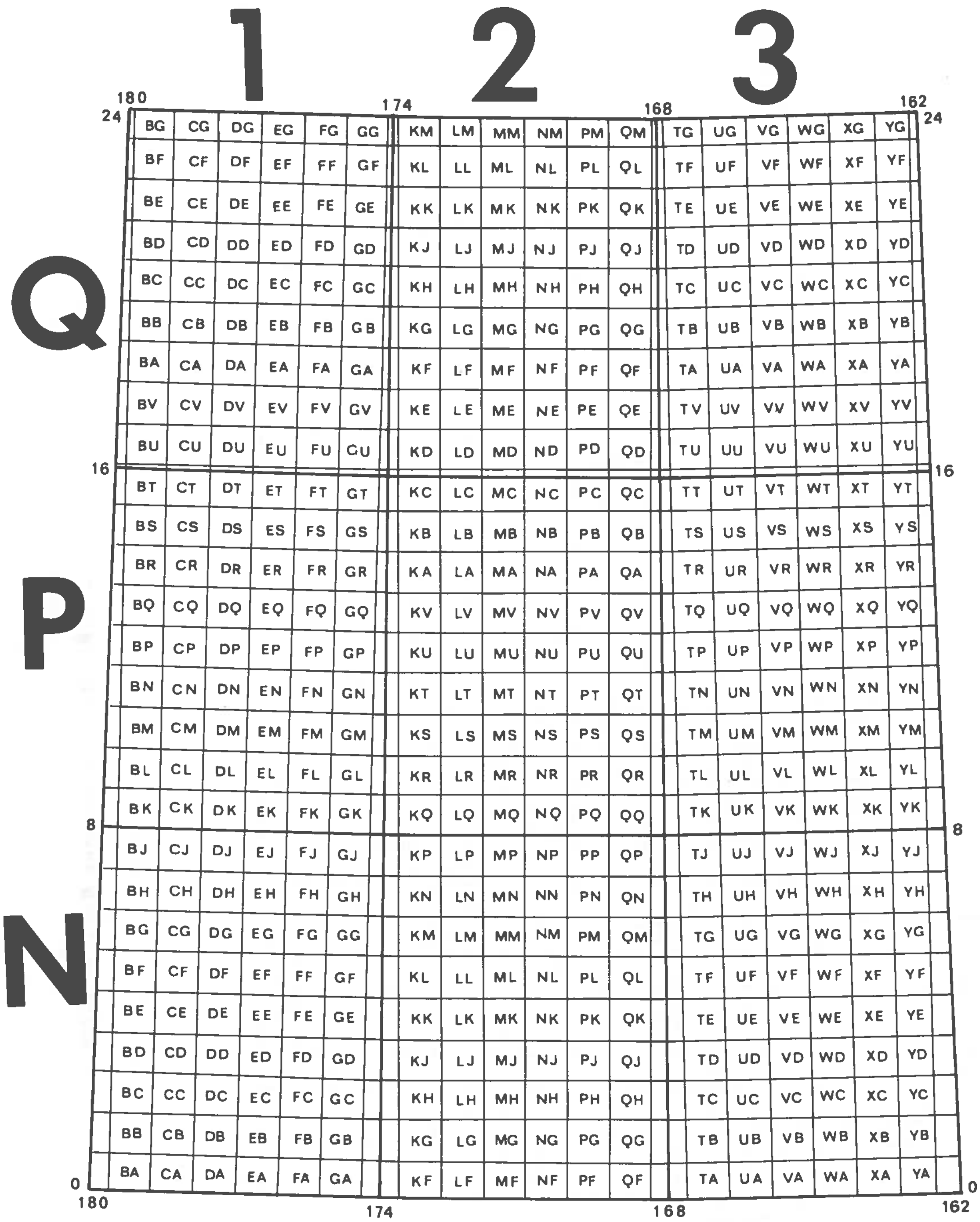


Source: U.S. Department of the Army

Figure 5.2 The Military Grid Reference System Grid Zones

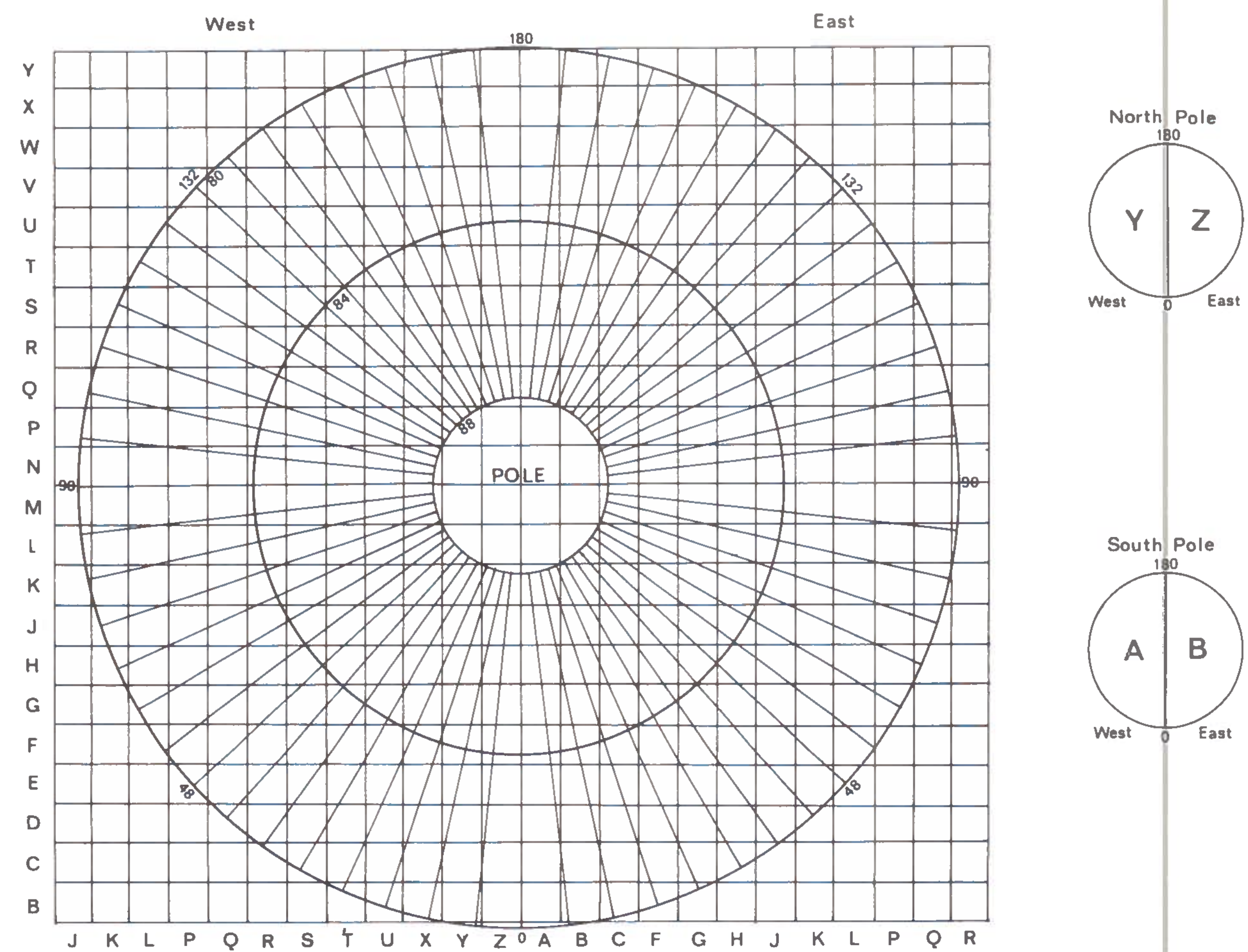
100,000 Meter Squares. Columns of squares are identified by letters. Rows of squares are also identified by letters (See Figure 5.3). Starting at the 180° meridian and proceeding easterly along the equator for 18°, the 100,000 meter columns (including partial columns along grid junctions) are consecutively lettered A through Z (I and O omitted) reading from south to north. This partial alphabet repeats every 3,000,000 meters. North of the equator, lettering starts with A at the equator in odd numbered UTM zones, and with F at the equator in the even numbered zones. The resultant staggering lengthens the distance between 100,000 meter squares of the same lettered identification. South of the equator, the 100,000 meter letters also read from south to north, complementing the letters above in any one particular zone.

Polar Designations. The north pole is divided into two parts by the 180° and 0° meridians. The west half is given the grid zone designation Y; the east half is given the grid zone designation Z. Similarly, the south polar area is divided into two parts by the 0° and 180° meridians. The west half is identified as A; the east half is identified as B (See Figure 5.4). The 180° - 0° meridians coincide with a 100,000 meter north and south grid line, and the 90° - 90° meridians coincide with a 100,000 meter east and west grid line. In the half of the area identified by the grid zone designation Y or A, the 100,000 meter columns are labeled J through Z, consecutively, from left to right. In the half identified by the grid zone designation



Source: U.S. Department of the Army

Figure 5.3 The Military Grid Reference System 100,000 Meter Square



Source: U.S. Department of the Army

Figure 5.4 The Military Grid Reference System Polar Identification Scheme

Z through B, 100,000 meter columns are labeled A through R, consecutively, from left to right. In both cases the letters I and O are omitted. To avoid duplicating the 100,000 meter squares in adjoining UTM zones, the letters D, E, M, N, V, and W are also omitted.

A full military grid reference is cited as a continuous series of numbers and letters without spaces, parentheses, dashes, or decimal points:

3PWN. . . . . locates a point within a 100,000 meter square

3PWN55. . . . locates a point within a 10,000 meter square

3PWN5354. . . locates a point within 1,000 meters

3PWN539544. . locates a point within 100 meters

The Military Grid Reference System (sometimes called the Universal Transverse Mercator System) is also used in the National Location Code (See Section 3.2), Statistics Canada's Geographically Referenced Data Storage and Retrieval System (See Appendix B) and approved by the American National Standards Institute (See Section 2.5).

References:

U.S. Department of the Army, and U.S. Department of the Air Force. The Universal Grid System. TMS-241 to 16-1-233. Washington, D.C.: U.S. Government Printing Office, 1951.

5.3 THE SMITHSONIAN INSTITUTION

The Smithsonian Institution's Information Systems Division in cooperation with the National Museum of Natural History has, over the past six years, developed a comprehensive computerized system for museum data processing. One of the basic problems involved in the development and implementation of this system was that of geographic notations or geocoding. Traditionally, scientists of the biological community have used latitude and longitude coordinates, Marsden Squares[1], or simply the name of a geographical or a geopolitical area for reporting the location at which a specimen is collected. Thus, the approach taken by the Smithsonian involved: 1) the creation of the Global Reference Code (GRC) to interrelate the various location reporting techniques; 2) the development of procedures for automatic updating of specimen records to include the appropriate GRC; and 3) the development of mathematical techniques to retrieve data on geographical areas of arbitrary shape or size.

5.3.1 Global Reference Code

The Global Reference Code is an internal computer code devised to interrelate the various location reporting techniques used by scientists in the musuem community. The GRC is based on a hierarchy

---

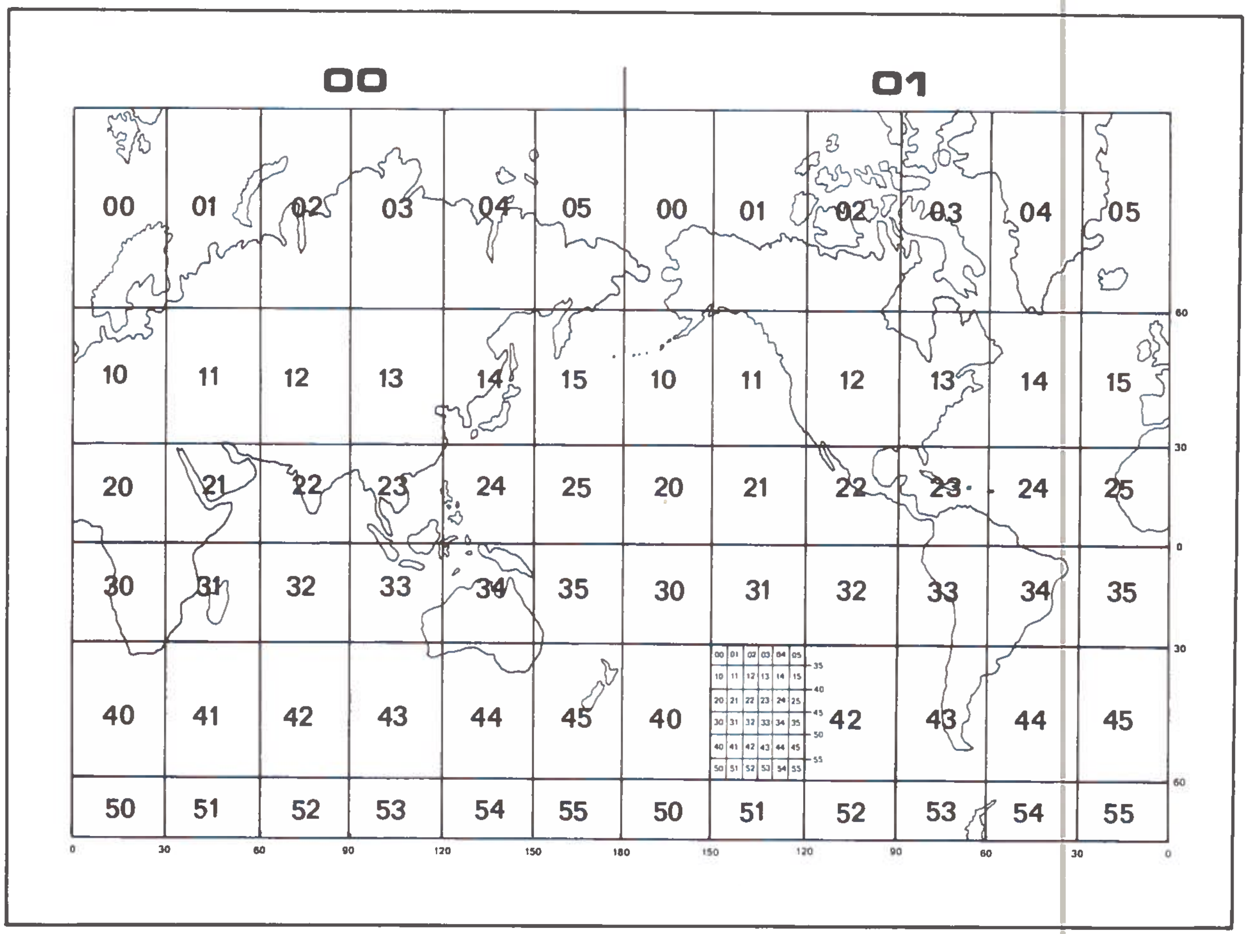
[1]A system of numbered areas each 10° latitude by 10° longitude, based on the Mercator projection, and used chiefly for identifying geographic positions and showing distribution of worldwide oceanographic and meteorologic data on a chart. Each square is subdivided into 100 one degree subsquares which are numbered from 00 to 99 starting with 00 nearest the intersection of the equator and the Greenwich meridian. The system was introduced in 1831 by William Marsden, an Irish orientalist.

of grids that can be mathematically related to other systems such as latitude and longitude or Marsden Squares. Each area referenced is designated as a subset of larger areas by the positional notation inherent in the code.

There are two first level grids labeled 00 and 01 which divide the earth's surface into cells measuring 180° by 180°. A mercator projection of the globe is used to establish the guidelines of longitude and latitude. At the second level, each 180° cell is divided into a 6 by 6 array of 30° cells (See Figure 5.5). At the third level, each 30° cell is divided into a 6 by 6 grid so that each cell will be exactly 5° on a side. The fourth level is a 5 by 5 array of 1° cells, the fifth level is a 6 by 6 array of 10' cells, and the sixth level is a 5 by 5 array of 2' cells.

Grid Level	Array Size	Cell Size
1	6 X 6	180° X 180°
2	6 X 6	30° X 30°
3	6 X 6	5° X 5°
4	5 X 5	1° X 1°
5	6 X 6	10' X 10'
6	5 X 5	2' X 2'

Each cell is referenced by its row and column designations. To reference a particular cell of the smaller set, a positional notation is imposed in which the label of the cell of the larger set is written first with the labels of the smaller set placed to the right. The division of cells may be repeated to any number of levels, and subscripts are used to distinguish the cells of each level. The first level would be referenced as $0_1\ 0_1$ and a cell at the second level would be $0_1\ 0_1\ 2_2\ 0_2$. The GRC coordinates represent the distance



Source: The Smithsonian Institution

Figure 5.5 Global Reference Code Grid

from the two axes, the 90° N latitude (north pole) and the 0° meridian.

5.3.2 Location Input Thesaurus

When coordinates of latitude and longitude or Marsden Square codes appear on a specimen record, mathematical functions transform the expression into a Global Reference Code. If a verbal expression of locality is all that appears on the specimen record, the computer system checks the Location Input Thesaurus of previously encountered expressions for a precedent to enable automatic processing of the record. If matching verbiage is found, the associated Global Reference Code is placed into the specimen record. However, if matching verbiage is not found in the thesaurus, the verbal expression on the specimen record is placed into the thesaurus as an "account receivable." The use of this term implies that support personnel will have to resolve the discrepancy of location manually.

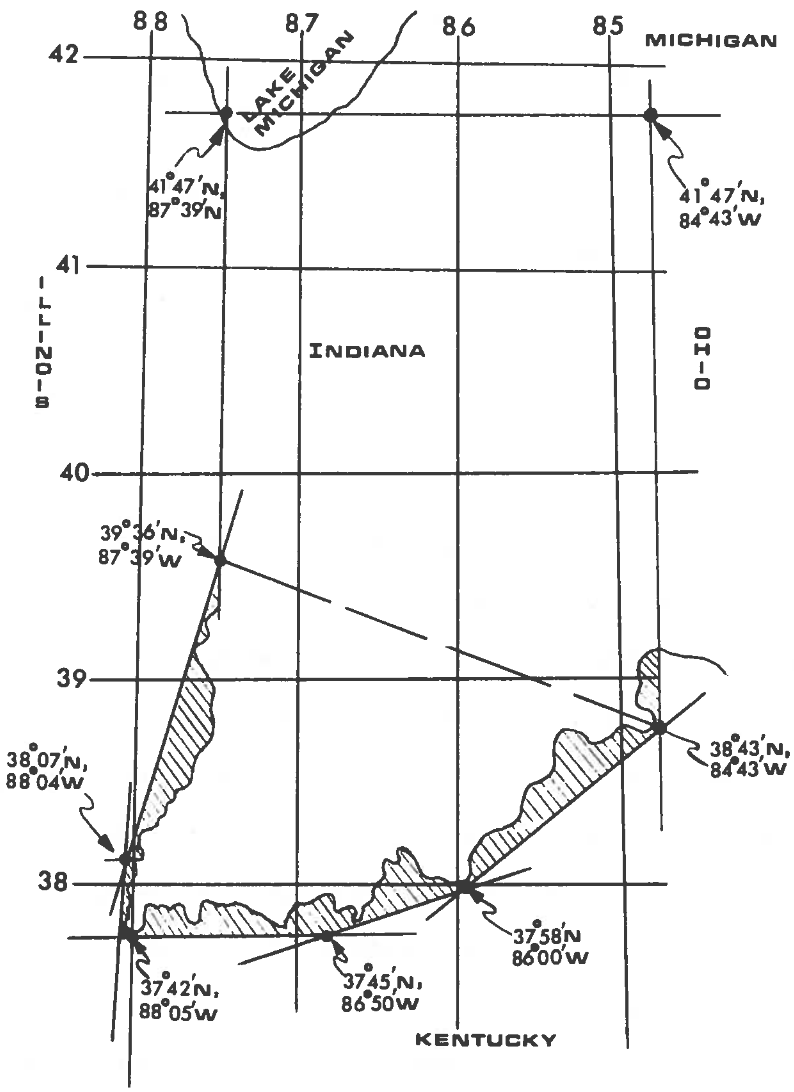
The system can be instructed to retain in the thesaurus those expressions known to be valid. Availability of direct access storage controls the number of items contained in the thesaurus. The choice of items is controlled by the popularity of the item as determined by a frequency count maintained by the system. Although the thesaurus obviously can never completely automate all location identification procedures, as the thesaurus builds up, fewer and fewer verbal expressions require human attention.

5.3.3 Selgem Data Retrieval

Selgem is an acronym derived from SELf-GEnerating Master. It is a generalized system for information storage, management, and retrieval especially suited for collection management in museums. It addresses the needs of the museum where collections grow rapidly, in that the system offers the ability to adjust a record's length, content, and format on a spontaneous basis and accomodates unanticipated data without reprogramming. Modularized components provide a choice according to budget and need among various levels of sophistication.

Selgem provides a special retrieval capability with respect to a geographic location which involves the retrieval of records having an expressed locality that falls within any polygon defined at retrieval time in terms of latitude and longitude, Marsden Squares or other reporting schema relatable in a mathematical sense to latitude and longitude coordinates or the Global Reference Code. An area of virtually any size or shape may be delimited by the user as a sector of interest, and records with coordinates within the sector will be retrieved. The user must supply the latitude and longitude points which, when connected by straight lines, bound the area of interest. The polygon created by connecting the points must consist of between three and eight sides and it must be convex (See Figure 5.6).

Convex polygons are used to describe planar areas in the Selgem retrieval system because it is possible to devise a simple algorithm for determining whether or not a point lies within such an area. By convention, the vertices of the convex polygon are numbered sequen-



The dashed line joining latitude and longitude points 39°36'N, 87°39'W, and 38°43'N, 84°43'W divides the general polygon into two convex polygons

Source: The Smithsonian Institution

Figure 5.6 Selgem Retrievel: A Polygon Approximation of Indiana tially in a clockwise direction the beginning point of which is arbitrary. The slopes are calculated once during the initial phase of the query process and are available thereafter. A linear function test is applied to each polygon side to determine if the point lies inside the polygon. If the test fails for any line, testing is terminated and the point is rejected. Conversely, if all tests are successful the point is accepted. The linear constraints used in this testing process apply only to the convex polygon and are not satisfactory constraints for the general polygon.

The retrieval can be made as accurate as desired by increasing the number of points that bound the area of interest. The lines drawn from point to point form a large general polygon which can be broken into smaller convex polygons. Retrieval of specimen data is effected when the location coordinates of the data lie within one of the convex polygons. The convex polygons need not be adjacent to one another. They may be completely separated (as in the case of polygons describing the Hawaiian Islands). Selgem geographic retrieval system allows each individual user to describe his own geographical area of interest which may range from a section of a state, a whole state, a river basin or a large regional area.

Due to the fact that short term payoff failed to justify costs in terms of computer time and personnel time, further development and use of the Input Location Thesaurus has been deferred by the Smithsonian Institution. Therefore, currently, geographical notations on specimen records are, when possible, converted directly to latitude and longitude coordinates rather than the Global Reference Code. The Selgem system for information management, however, has gained widespread acceptance in the museum community. At the Smithsonian, Selgem is used by the National Museum of Natural History, the Bicentennial Inventory of American Paintings, the Oceanographic Sorting Center, the National Collection of Fine Arts and the Office of Environmental Sciences. In addition some 100 persons at 36 museums, universities and institutions across the United States have adopted the Selgem system.

References:

Creighton, Reginald A. and Piacesi, Dante. "An Approach to the Geography Problems in Museums." Smithsonian Institution Information Systems Innovations. Vol. II, No. 1. Washington, D.C.: Smithsonian Institution, 1970.

Creighton, Reginald A. and Crockett, James J. "Selgem: A System for Collection Management." Smithsonian Institution Information Systems Innovations. Vol. II, No. 3. Washington, D.C.: Smithsonian Institution, 1971.

Creighton, Reginald A.; Packard, Penelope; and Linn, Holley. "Selgem Retrieval: A General Description." Smithsonian Institution Procedures in Computer Sciences. Vol. I, No. 1. Washington, D.C.: Smithsonian Institution, 1972.

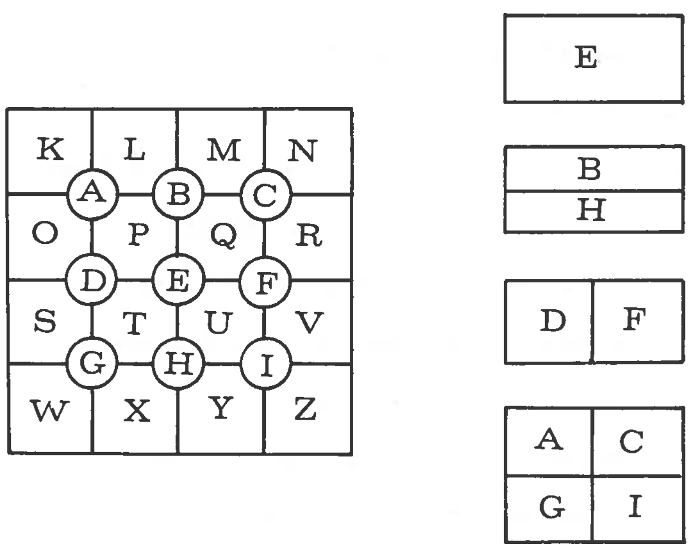
5.4 GEO-CODE

Geo-code is a geographic reference system designed by a biologist named Sydney W. Gould during the 1950's. The system was developed specifically as a machine readable shorthand to describe the location or range of plant and animal species in scientific monographs, manuals and indexes. The code is a four letter designation which hierarchically identifies a small geopolitical area within a succession of larger geopolitical or geographic divisions. In the United States, for example, a Geo-code identifies a county within a state, within a group of states, within the United States, within the North American Continent or Western Hemisphere.

Conceptually the Geo-code system is based on a set of four major grids each with an overlay grid. The basic grids divide the surface of the globe into successive blocks consisting of sixteen parts so that the fourth grid areas are equivalent to a small geopolitical subdivision such as a county. Since the surface of the globe contains about 197 million square miles, of which 140 million is water and 57 million is land, the area of the some 65,000 spaces in the fourth Geo-code grid equals about 1140 square miles.

Grid 1	16	16 cells
Grid 2	16 X 16	256 cells
Grid 3	16 X 16 X 16	4,096 cells
Grid 4	16 X 16 X 16 X 16	65,536 cells

In all four of the basic Geo-code grids each cell is represented by one of the last sixteen letters of the alphabet, K through Z. The overlay grid of each level consists of nine parts represented by the letters A through I. The letters of the overlay grid represent two or more cells of the base grid hierarchically. For example, in the first Geo-code grid the letter C includes the areas covered by both the N cell and the R cell of the base grid, thus C represents Asia; the letter B includes the area covered by overlay grid cells A and C, and thus represents the Northern Hemisphere; the letter F includes the area covered by overlay grid cells C and I, thus F represents the Eastern Hemisphere; the letter E represents all the other letters of the overlay and the base grid, hence the globe is coded EEEE.



E = A + C + G + I = B + H

B = A + C

H = G + I

D = A + G

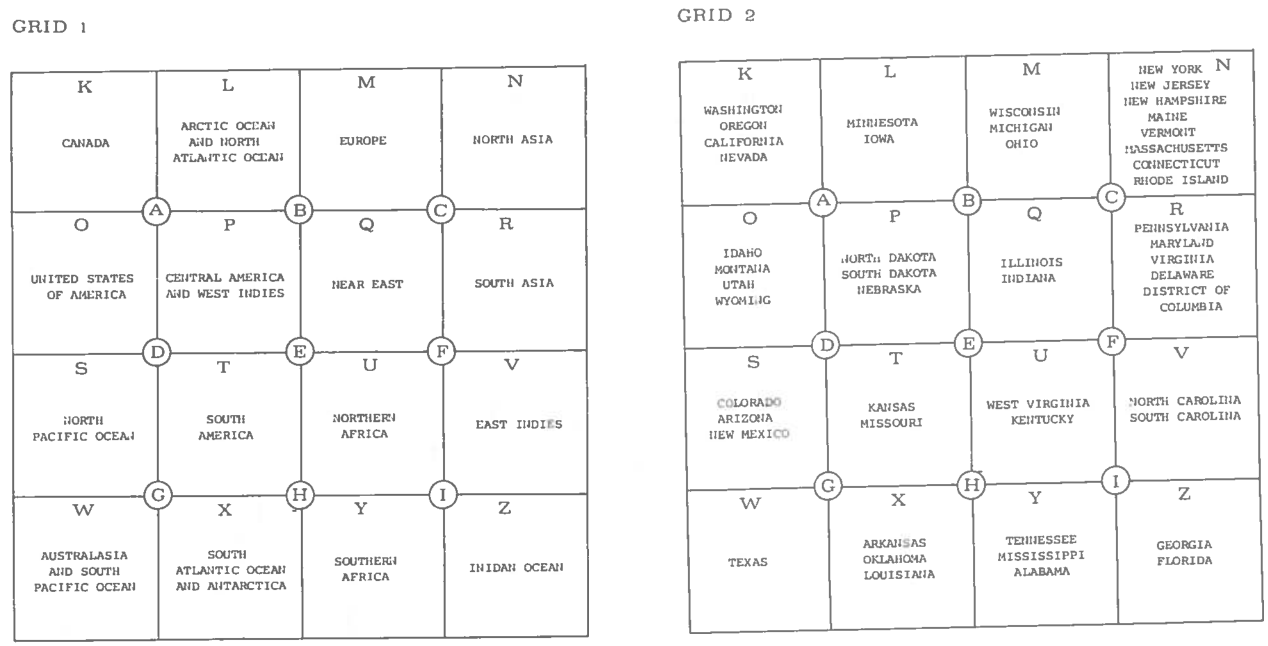
F = C + I

A = K + L + O + P
C = M + N + Q + R
G = S + T + W + X
I = U + V + Y + Z

The first Geo-code grid includes the entire surface of the globe divided into sixteen parts. Since the United States occupies one entire cell (0) of this grid and uses all of the subsequent subdivisions of that cell, it is coded OEEE. The second Geo-code grid subdivides one cell of the first grid into another sixteen parts. Thus the United States is subdivided into sixteen cells composed of two or more states (with the exception of one cell containing only one state, Texas). The third Geo-code grid subdivides the group of states contained in a cell of the second grid into single state blocks which generally occupy more than one of the sixteen available cells. Therefore the third letter of the Geo-code for states of the United States is usually a third level overlay grid designation which indicates the fact that one state occupies more than one cell of the third Geo-code base grid. The fourth Geo-code grid divides each cell of the third order base grid into sixteen additional subcells, each of which consists of a whole county (See Figures 5.7 and 5.8).

Geo-codes have been assigned to the second level grid subdivisions of all major oceans illustrated in Figure 5.9. The horizontal boundaries are parallels of latitude, and vertical boundaries are meridians of longitude based on the Universal Transverse Mercator Projection. All Geo-codes relating to the ocean end with the letter J.

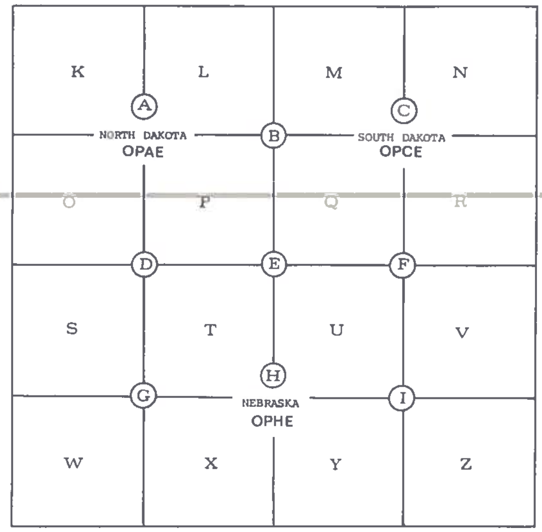
Although Geo-code boundaries are generally designed to represent geopolitical boundaries for practical reasons , the primary reference is to a grid cell area of land or sea. Reference to a geopolitical area such as county or state is really only secondary. A good example



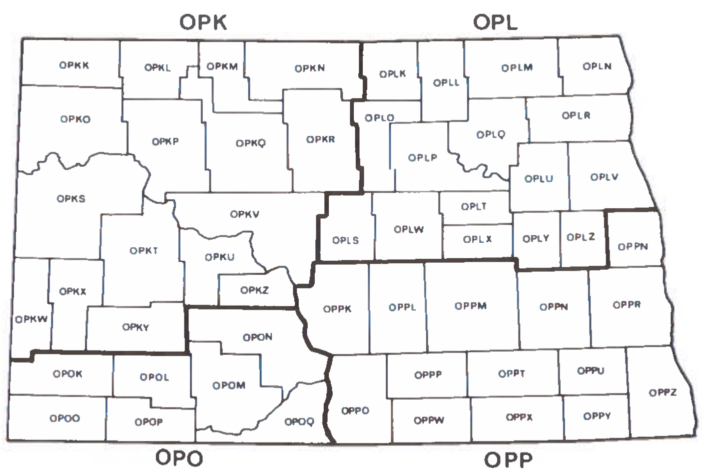
Source: Sydney Gould, Geo-Code

Figure 5.7 Geo-Code Grid One and Grid Two

GRID 3

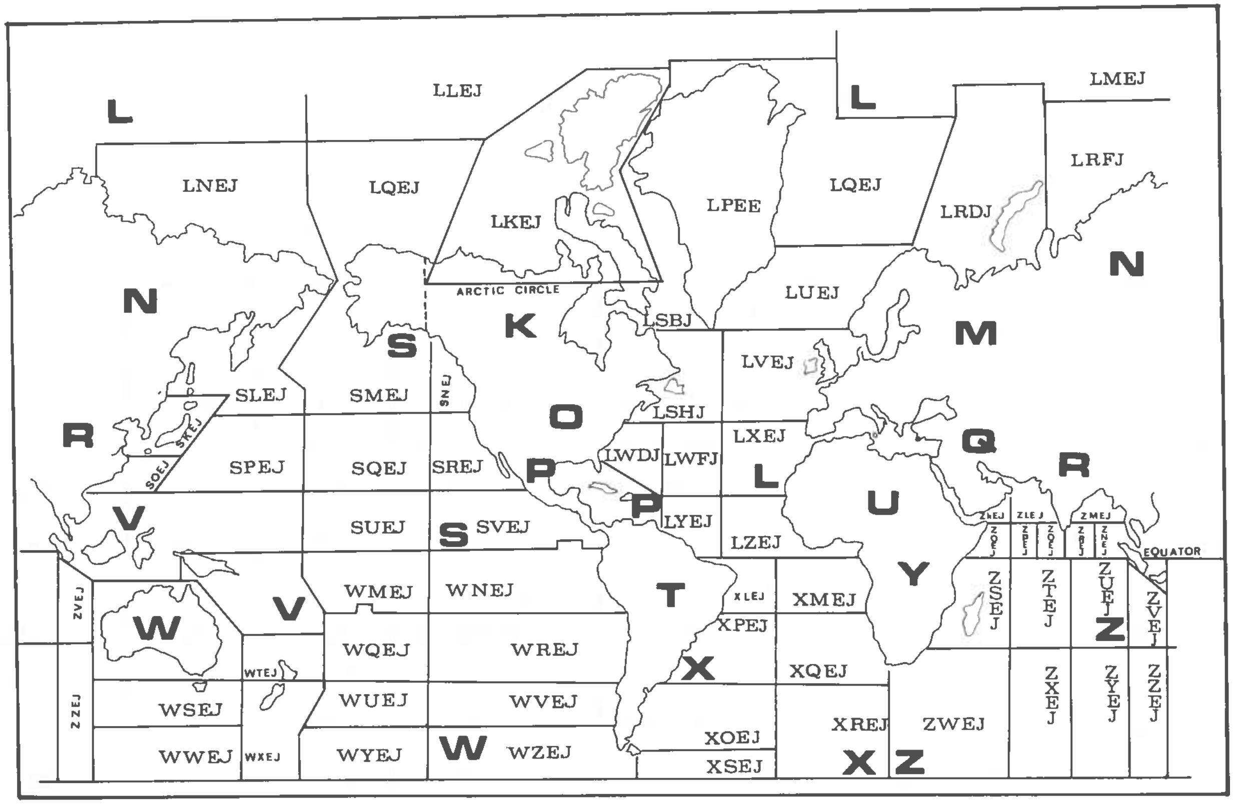


GRID 4



Source: Sydney Gould, Geo-Code

Figure 5.8 Geo-Code Grid Three and Grid Four



Source: Sydney Gould, Geo-Code

Figure 5.9 Geo-Codes of the Ocean

Country	State	City	City Name	State Name	Country	Year	Population
OEEE	OQFE	1OQRN	Fort Wayne	Indiana	USA	1960	161 776
OEEE	OWEE	1OWML	Fort Worth	Texas	USA	1960	365 268
OEEE	OKGE	1OKWU	Fresno	Californis	USA	1960	133 929
OEEE	OQFE	1OQMK	Gary	Indiana	USA	1960	178 320
OEEE	OKGE	3OKXM	Glendale	California	USA	1960	119 442
OEEE	OMCE	1OMGX	Grand Rapids	Michigan	USA	1960	177 313
OEEE	OVBE	1OVLP	Greensboro	North Carolina	USA	1960	119 574
OEEE	OQFE	2OQMK	Hammond	Indiana	USA	1960	111 698
OEEE	ONUE	1ONUP	Hartford	Connecticut	USA	1960	162 178
SQEJ	SQUJ	1SQTJ	Honolulu	Hawaii	USA	1960	294 194
OEEE	OWEE	1OWVK	Houston	Texas	USA	1960	938 219
OEEE	OQFE	1OQVO	Indianapolis	Indiana	USA	1960	476 258
OEEE	OYGE	1OYSQ	Jackson	Mississippii	USA	1960	144 422
OEEE	OZIE	1OZUR	Jacksonville	Florida	USA	1960	201 030
OEEE	ONGE	1ONT	Jersey City	New Jersey	USA	1960	276 101

Source: Sydney Gould, Geo-Code

Figure 5.10 Geo-Code City Codes

of this dichotomy is the Geo-codes for the State of Alaska, which is politically part of the United States (OEEE) but the land area is in Canada (KEEE) if Canada is considered not as a country but as the northern part of North America. Thus the Geo-code for Alaska is KKEE, not OEEE. This distinction is even more evident in the Aleutian Islands, which is a named political division of Alaska, one part of which is on the mainland, but the other part, the islands, stretches a long way westward in the Pacific Ocean, therefore this part of the Aleutians bears an oceanic geo-code, SMDJ.

By prefixing a numeral to the Geo-code of a county the major cities within it may be given a unique identification and such prefixes have been assigned to cities in the United States and Canada having a population over 100,000 (See Figure 5.10). Furthermore, the application of suffix numerals has been suggested for use in museums to locate the room, display, case or bin in which artifacts may be found or to express a third dimension below the earth's surface such as geological strata or ocean depths.

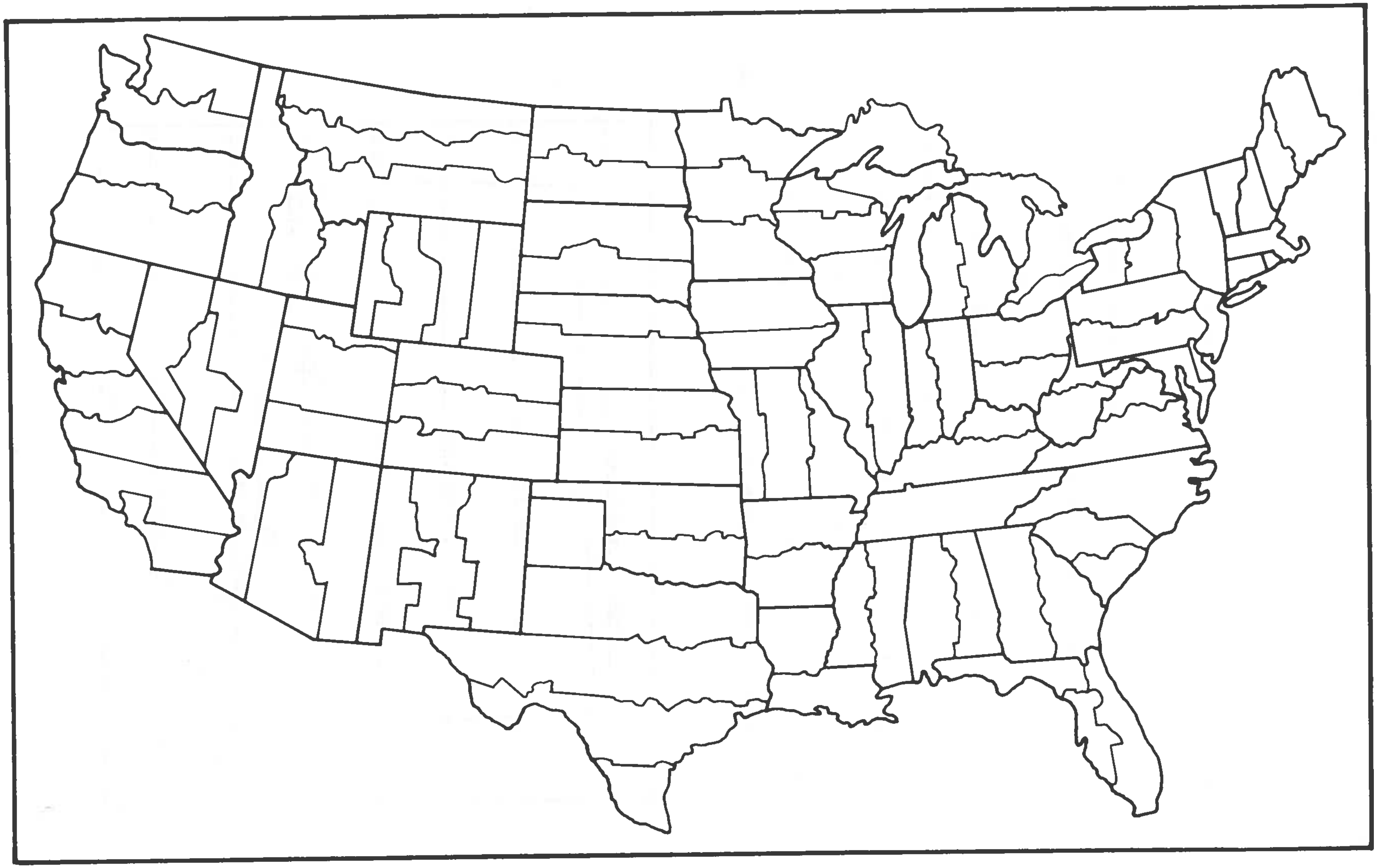
References:

Gould, Sydney W. Geo-code. Volume I, West Edition, New Haven, Connecticut: The Gould Fund, 1968.

5.5 STATE PLANE COORDINATE SYSTEMS

The State Plane Coordinate Systems devised by the United States Coast and Geodetic Survey in 1934 are a set of plane coordinate grids for states or state sections of the continental United States (See Figure 5.11). They are especially designed for use in local land surveys because the curvature of the surface of the earth is so slight that relatively small areas of land may be regarded as plane surfaces without the introduction of major distortions. Thus, a location within a local reference grid may be defined accurately by simple X Y coordinates measured in linear units from selected origins.

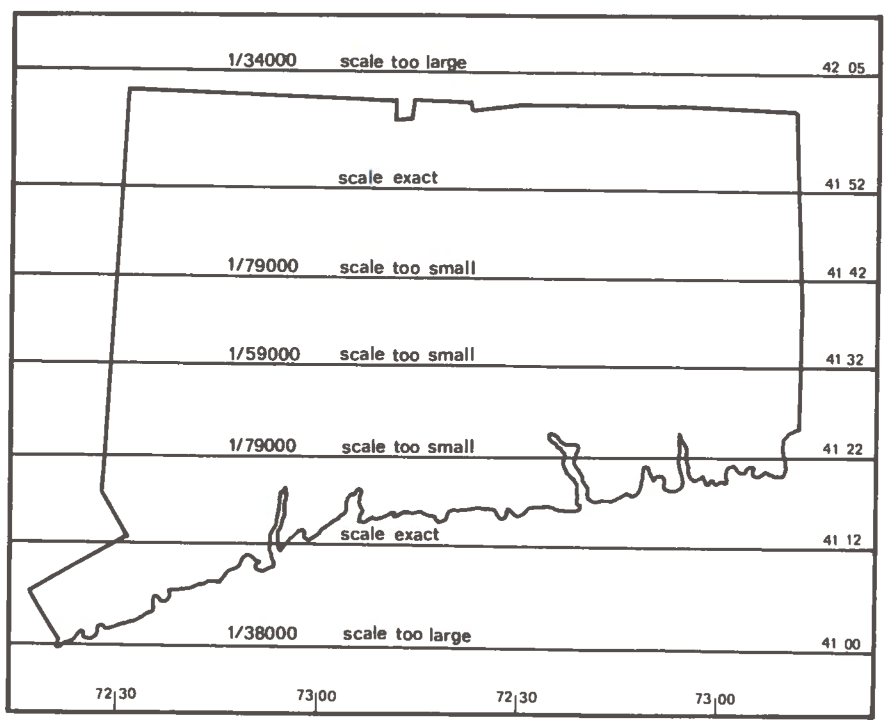
As with any plane rectangular coordinate system, a projection employed in establishing a state coordinate system may be represented by two sets of parallel straight lines, intersecting at right angles. One set of these lines is parallel to a meridian passing approximately through the center of the area. The grid line corresponding to that meridian is the axis of Y. Forming right angles to the axis of Y and to the south of the area shown on the grid is the axis of X. The position of a point represented on the grid is defined by stating the distance from each axis in feet. The X coordinate reading gives the position in an east-west direction and the Y coordinate reading gives the position in a north-south direction. All X coordinates in an area represented on a state grid are made positive by assigning the origin of the coordinates: X = 0 plus a large constant. For any given point, then, the X coordinate equals the value of X adopted



Source: U.S. Coast and Geodetic Survey

Figure 5.11 State Plane Coordinate Systems for the origin, plus or minus the distance (X^1) of the point east or west from the central meridian (axis of X).

The coordinate grids of the state plane systems are limited to a width of 158 miles, and they are always superimposed upon a conformal map projection. This insures that angles are preserved and that the deviation between grid and geodetic lengths will be maximum along the margins of the longer dimensions of the grid and midway between those margins; along the margins, the grid length of a line will be greater than its geodetic length; along the center line the geodetic length will be greater. Between these limits are two lines along which grid and geodetic lengths are equal. These are lines of exact



scale. When the width of an area covered by a single grid is 158 miles, the extreme difference between geodetic and grid lengths will be 1/10,000 of the length of the line which is the maximum tolerable limit of distortion error for property survey.

While a width of 158 statute miles was adopted as a standard in devising the state coordinate systems, departures from that width have been made where geographic conditions necessitate. If the width of a state is less than 158 miles, the width of the grid is decreased, and the effect of the scale factor thereby also decreases. Where a state is too wide to be covered by a single grid, it is divided into belts, called zones, for each of which a separate grid is adopted. The boundary lines between zones follow county lines.

As some states or zones have a width or smaller dimension extending in a north and south direction and others may have a smaller east and west dimension, two forms of geographic map projection serve as the foundation of the state coordinate system. For an area of small north and south dimension, the Lambert Conformal projection with two standard parallels is used. For an area of limited east and west dimension, another conformal projection, the Transverse Mercator projection is employed.

Plane coordinates present no serious computational problems as long as all points considered lie in the same local grid. However, points in different state plane grids must first be converted to latitude and longitude and the relationships to points in other state plane grids computed through the use of the applicable elliptical equations. In order to facilitate this procedure and help to unify the system of disparate local grids at a national level, the Coast and Geodetic Survey has calculated and compiled a series of conversion tables based upon the national triangulation network.[1] These tables provide a link between each local grid and any of the thousands of national triangulation monuments for which geodetic position data have been established. Thus, any point within a local system can be assigned longitude and latitude readings simply by calculating its relationship to the local triangulation monument.

Laws establishing the State Plane Coordinate Systems as official land survey standard have been adopted in the following 25 states:

Alabama	Maryland	Oregon
California	Massachusetts	Pennsylvania
Connecticut	Minnesota	Rhode Island
Delaware	Nevada	South Dakota
Georgia	New Jersey	Tennessee
Indiana	New York	Texas
Louisiana	North Carolina	Vermont
Maine	Ohio	Virginia
		Washington

[1]Triangulation is a method of surveying founded in the trigonometrical proposition that if three parts of a triangle are known, the other three parts may be obtained by computation.

References:

U.S. Department of Commerce. Coast and Geodetic Survey. The State Coordinate Systems: A Manual for Surveyors. Special Publication 235. Washington, D.C.: U.S. Government Printing Office, 1945.

5.6 UNITED STATES PUBLIC LAND SURVEY

The United States Public Land Survey, which was authorized by the Continental Congress in 1785, is a system of rectangular grids by which the land owned or controlled by the federal government was subdivided for administration, survey, and transfer of title under the public land laws of the United States. It included lands turned over to the federal government by the colonies or states, and those larger areas that were acquired later by purchase or treaty. The survey was required by the Ordinance of 1785 and began at the western boundary of Pennsylvania extending westward. It has been established at least partially in all states except the thirteen original states, Vermont, Maine, Kentucky, Tennessee, West Virginia and Texas.

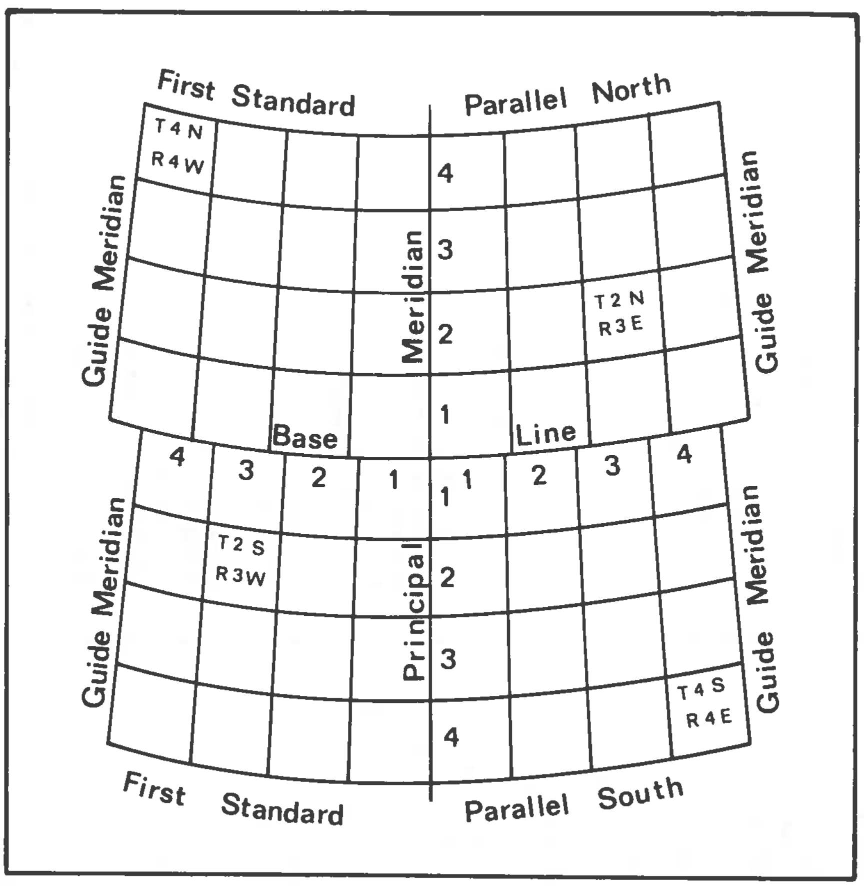
The public land survey system is a grid of north-south lines (meridians) and east-west lines (parallels) intersecting at right angles so as to form grid zones six miles square. These zones are known as townships. Each township is uniquely identified by a range and township number consisting of six characters,[1] (letter-number-letter-letter-number-letter). Range numbers are assigned to each grid zone east and west of a selected meridian called the principal meridian. Township numbers are assigned to each grid zone north and south of a selected parallel called the base line. There are 35

---

[1]In order to identify a township uniquely, this six character code must be prefixed by some form of principal meridian identification. Also in some cases town and range numbers may exceed 9, thus requiring more than 6 characters as in the Northwestern corner of North Dakota.

designated prinicpal meridians and 32 designated base lines in the survey system (See Figure 5.12).

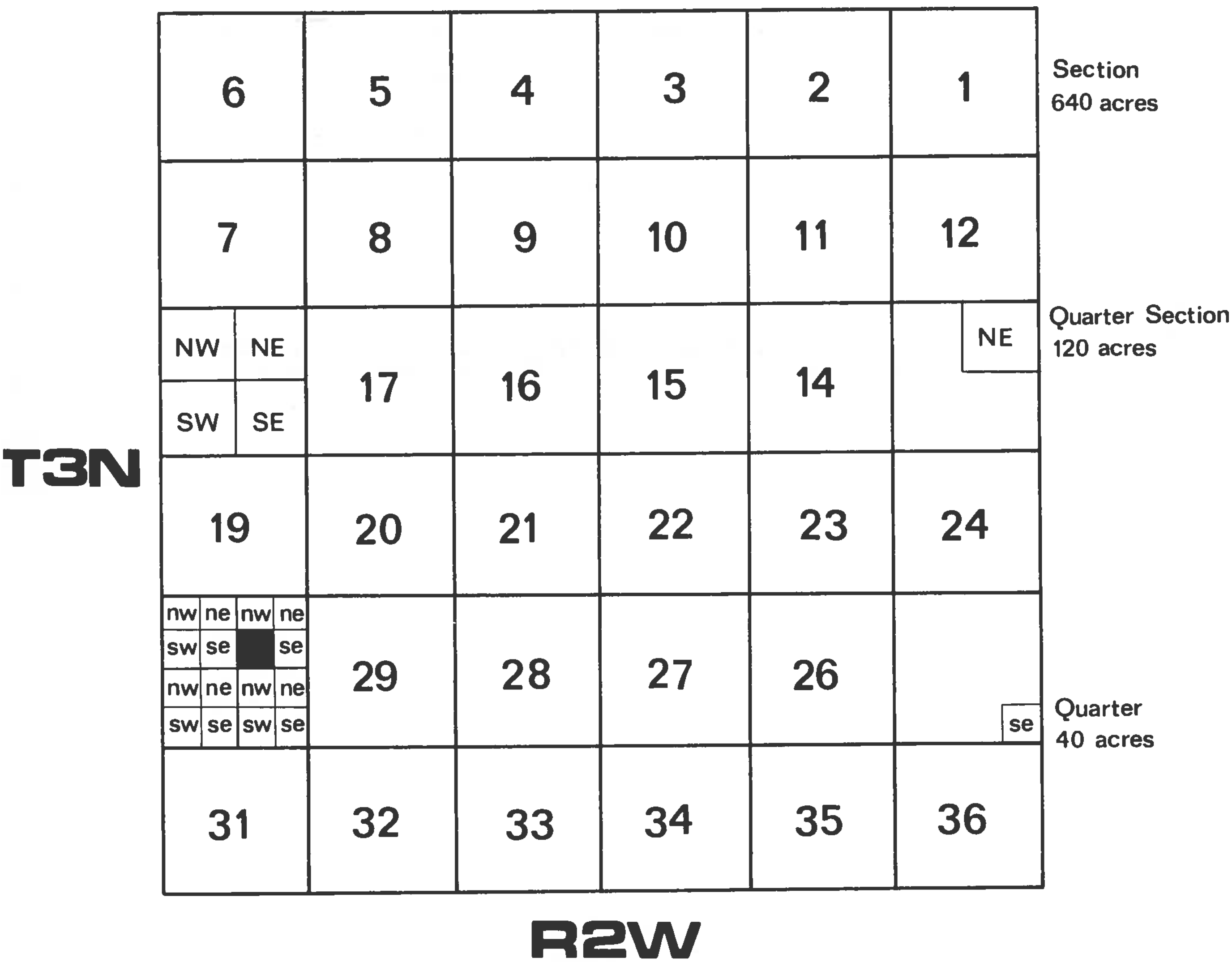
Governing lines, designated as guide meridians and standard or correction parallels, are established as reference lines for the extension of the survey. Guide meridians are established at 24 mile (four township) intervals to the east and west of the principal meridian. Standard parallels are established at 24 mile (four township) intervals to the north and south of the base line. The distortion caused by convergence of the meridians at the pole is taken into account by calculating the necessary adjustments at each standard parallel.



PRINCIPAL MERIDIANS	GOVERNING SURVEYS WHOLLY OR IN PART IN STATES OF	LONGITUDE OF MERIDIANS WEST FROM GREENWICH	LATITUDE OF BASE LINES NORTH FROM EQUATOR
Black Hills	South Dakota	104 03 00	44 00 00
Wind River	Wyoming	108 48 40	43 01 20
Boise	Idaho	116 24 15	43 22 31
Willamette	Oregon, Washington	122 44 20	45 31 00
Chickasaw	Mississippi	89 15 00	34 59 00
Washington	Mississippi	91 09 15	31 00 00
Ute	Colorado	108 33 20	39 06 40
Cimarron	Oklahoma	103 00 00	30 36 00
Choctaw	Mississippi	90 15 45	31 54 40
Utah	Utah	109 57 30	40 26 20
Copper River	Alaska	145 18 42	61 49 11
Third Principal	Illinois	89 10 15	38 28 20
Fairbanks	Alaska	147 38 33	64 51 49
Tallahassee	Florida	84 16 42	30 28 00
First Principal	Ohio	84 48 50	41 00 00
St. Stephens	Alabama, Mississippi	88 02 00	31 00 00
Fourth Principal	Illinois	90 28 45	40 00 30
St. Helena	Louisiana	91 09 15	31 00 00
Fourth Principal	Minnesota, Wisconsin	90 28 45	42 30 00
Seward	Alaska	149 21 53	60 07 26
Gila and Salt River	Arizona	112 17 25	33 22 40
Second Principal	Illinois, Indiana	86 28 00	38 28 20
Humbolt	California	124 08 00	40 25 12
San Bernardino	California	116 56 15	34 07 10
Huntsville	Alabama	86 34 45	35 00 00
Salt Lake	Utah	111 54 00	40 46 04
Indian	Oklahoma	97 14 30	34 30 00
Principal	Montana	111 38 50	45 46 48
Louisiana	Louisiana	92 24 15	31 00 00
New Mexico Principal	Colorado, New Mexico	106 53 40	34 15 25
Michigan	Michigan	84 22 24	42 26 30
Navajo	Arizona, New Mexico	108 32 45	35 45 00
Mount Diablo	California, Nevada	121 54 48	37 51 30
Sixth Principal	Colorado, Kansas, Nebraska, South Dakota, Wyoming	97 23 00	40 00 00
Fifth Principal	Arkansas, Iowa, Minnesota, Missouri, North Dakota, South Dakota	091 03 42	34 44 49

Source: Frank E. Clark, A Treatise on the Law of Surveying and Boundaries

Figure 5.12 Principal Meridians and Baselines of the United States Public Land Survey



The shaded area is identified as:
SW Quarter, NE Quarter Section, Section 30, R2W, T3N

Figure 5.13 United States Public Land Survey Townships

Townships are subdivided into 36 sections 1 mile square (640 acres) which are arranged in six tiers of six sections. They are numbered from 1 to 36 beginning at the northeast corner east to west, west to east, east to west, etc., and ending at the southeast corner. Sections are divided into quarter sections of 120 acres each, and quarter sections are then divided into quarters of 40 acres each (See Figure 5.13). Thus, any 40 acre lot covered by the United States Public Land Survey may be uniquely identified in the following manner: SW Quarter, NE Quarter Section of Section 30 in R2W, T3N.[1]

References:

Clark, Frank E. A Treatise on the Law of Surveying and Boundaries. 2nd edition. Indianapolis: The Bobbs-Merrill Company, 1939.

[1]There are exceptions to this numbering system, chiefly in Ohio and other areas that were surveyed in the early years of the U.S. Public Land Survey.

5.7 AMERICAN TELEPHONE AND TELEGRAPH

Since 1960 the Bell Telephone System has been using a method known as the Verticle-Horizontal Coordinate System to determine airline rate mileage for message and private line telecommunications services. Prior to this time, message rate mileage was determined from a rectangular grid overlaid on a map of the United States by measurement between the centers of 7 mile squares (blocks) or, for distances over 350 miles, between the centers of 35 mile squares (sections). The blocks and sections were identified by an arbitrary alphanumeric coding scheme and precomputed tables were used to determine distance between blocks or sections. Private line channel mileage was determined by direct scaling from a master map.

The V-H Coordinate System is a grid devised for the North American Continent based on the Donald Elliptic projection. This projection is essentially a modification of a two point equidistant projection. The two selected reference points are on a line running approximately through the center of the United States. The center is on the 97° W meridian and the reference points are 0.2 radian or 792 miles distant from the center. In this V-H coordinate system, the origin is at the east reference point and directions north and west of the origin are given a plus sign. These two point equidistant V-H coordinates were translated into final V-H coordinates by several operations involving a shift of origin, rotation of axes, inversion of the V axis, a unit conversion factor, and finally, an overall 0.3% shrinkage of scale.

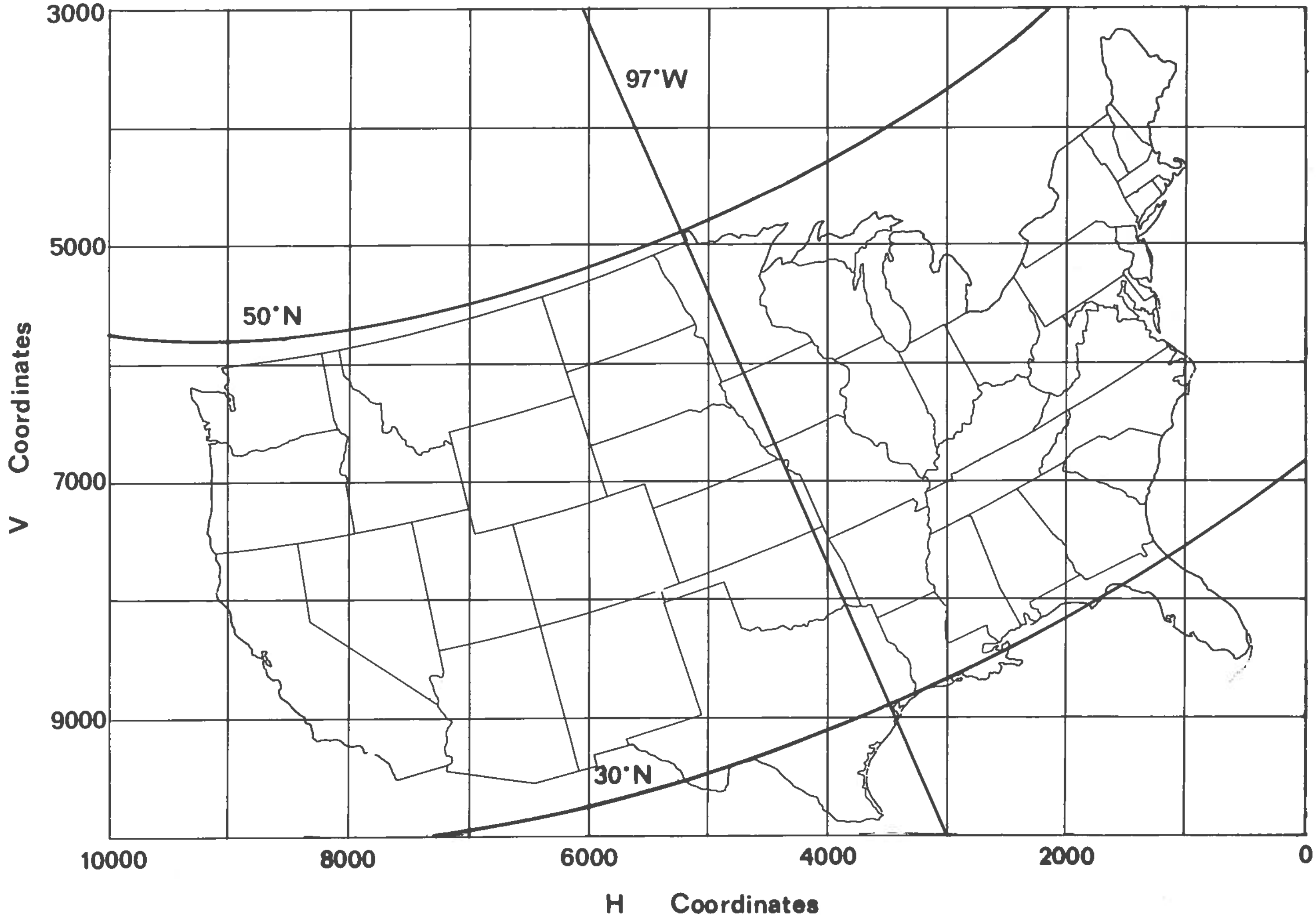
It is this final shrinkage of scale which changes the projection from a true two point equidistant projection to one which has been termed an elliptic projection. East-west along the Canadian and Mexican borders the scale error is only +0.3% scale error east-west across the center of the United States.

The final V-H Coordinate System is a square grid numbered 0 to 10,000 from north to south on the vertical axis and east to west on the horizontal axis (See Figure 5.14). One V or H unit is approximately 1,670 feet long and one V-H square covers an area of one tenth of a square mile. V and H coordinates are calculated from the latitude and longitude of the point selected as the measuring point for each telephone exchange area involved.

Precomputed tables of data for calculating the coordinates yield answers to three decimal places. For operational use the coordinates are rounded to the nearer integer.

Rate mileage for private line channels are determined by a look-up of the V-H coordinates of the two rate centers involved, calculation of the differences, $V_2 - V_1$ and $H_2 - H_1$ and then solving the resulting right triangle. The hypotenuse is the airline rate mileage between the two points.

Determination of message rate mileage is carried out in generally the same manner but with an additional mathematical adjustment which maintains a degree of parity with mileages determined from the former block and section system. Precomputed tables have been prepared to enable operating personnel to go directly from the coordinate difference to a rate, to mileage or to an index number.



Source: American Telephone and Telegraph Company

Figure 5.14 The V-H Coordinate System

References:

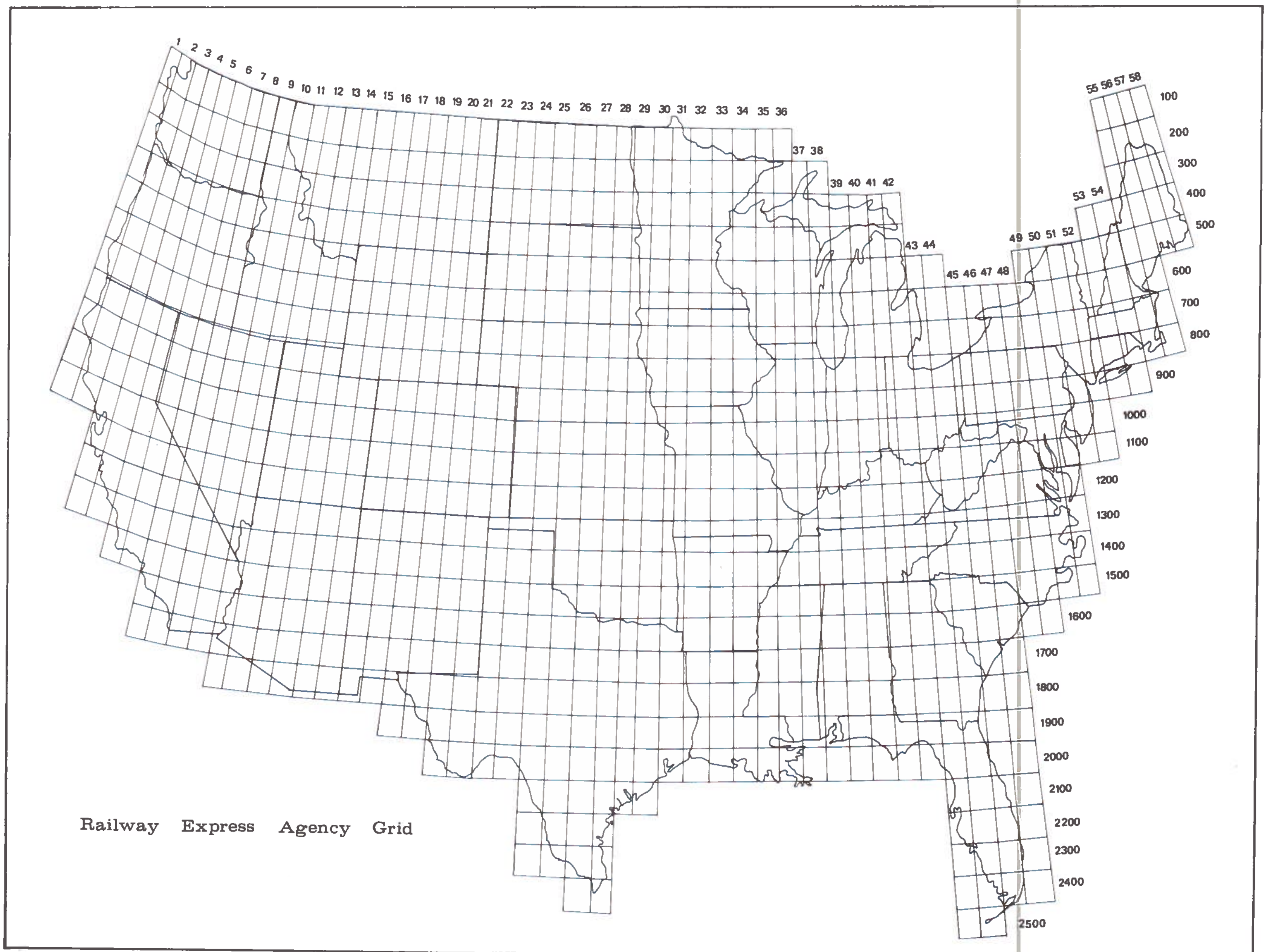
American Telephone and Telegraph Company, Long Lines Department, Administration of Rates and Tariffs. "Traffic F.C.C. Publication No. 255." New York: American Telephone and Telegraph Company, 1968.

5.8 RAILWAY EXPRESS AGENCY

The Railway Express Agency (REA) is a small shipment common carrier which maintains 2,386 offices and terminals in the United States and provides multimodal transportation service to both foreign and domestic locations. It has devised a system of geographic reference for locations in the United States which consists of an arbitrary grid. The grid is composed of two sets of parallel lines which intersect at right angles. One set of lines runs in a north-south direction, and the other set of lines runs in an east-west direction across the United States. This grid has no relation to the geodetic grid; it is simply a network of lines superimposed upon a map of the United States for the purpose of subdividing the areas into easily referenced units.

The base unit defined by the intersection of these grid lines is an area called a railblock. Each railblock is sixty-nine miles long and fifty miles wide. The entire grid system is composed of 25 units north to south and sixty units east to west. The code number assigned to each railblock consists of three or four digits. The first two positions are taken from the vertical scale (1 - 25) and the last one or two digits are taken from the horizontal scale (1 - 60). Out of the 1,500 possible railblock designations in this grid, there are actually 903 railblocks covering the United States (See Figure 5.15).

An air zone is composed of two contiguous railblocks. Each zone measures 138 miles north-south by 100 miles east-west. They are numbered serially from 1 to 44 beginning at the northwest corner



Source: Railway Express Agency

Figure 5.15 Railway Express Agency Grid

of the United States across the nation in tiers to the southeast corner. The code for air zones has no relation to either the vertical or horizontal railblock scales.

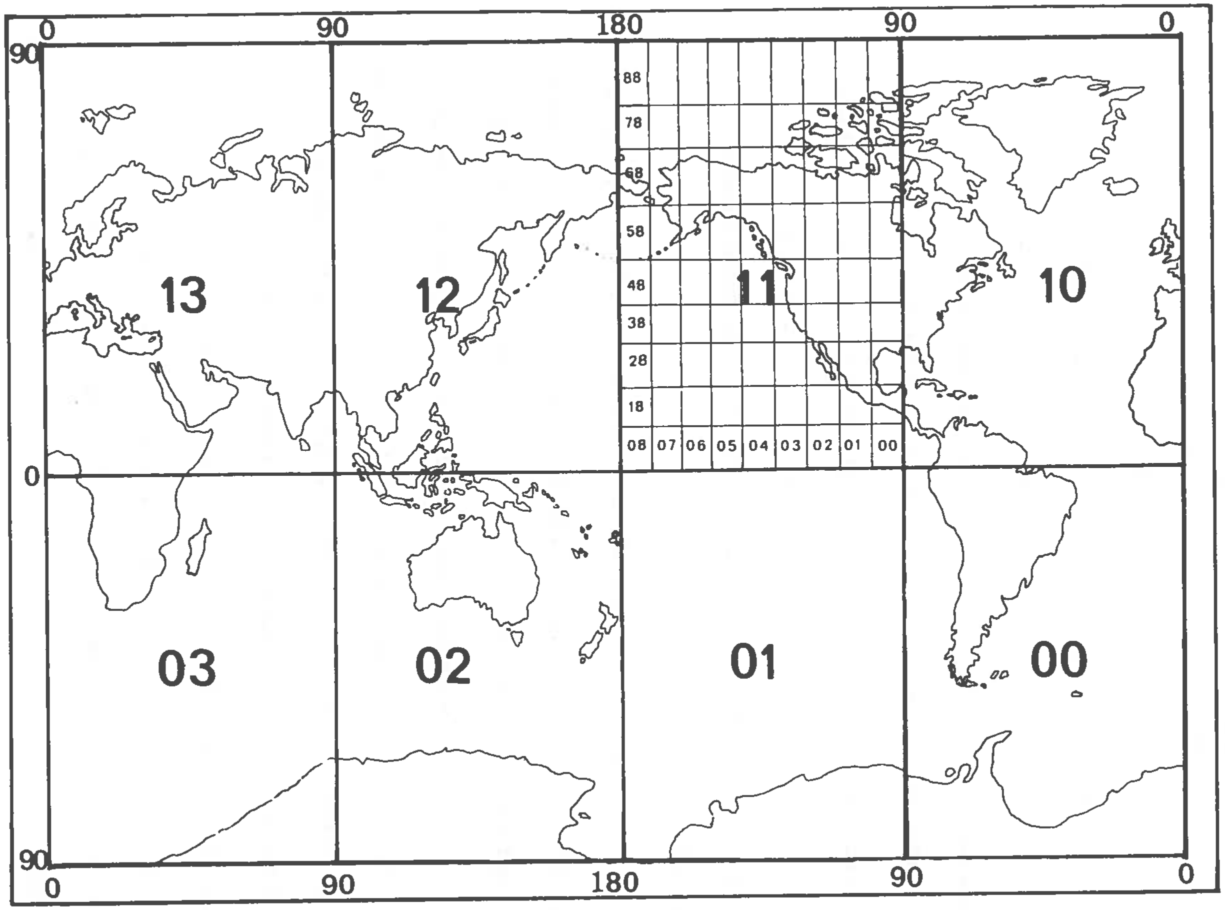
References:

Miano, G. T. Railway Express Agency, New York. Interview, 1971.

5.9 LINEAR GEOGRAPHICAL CODE

The linear geographical code is a grid based global geographic reference structure desiqned for the aggregation, processing and analysis of data in corporate information systems. Conceptually the linear geographic coding system is very similar to the Global Reference code of the Smithsonian Institution (See Section 5.3) and the Geo-Code system designed by Sydney Gould (See Section 5.4). It is a grid-in-grid structure focusing on the spaces defined at certain levels by lines of latitude and longitude.

In the linear geographical coding system the surface of the globe is subdivided by grids into masses, zones, spaces, areas and spots. The first order grid defining masses contains eight cells coded east to west around the world starting at the South Pole. Each cell is 90° by 90° (See Figure 5.16). Zones are defined in the second order grids which contain a 9 by 9 array of 10° cells. At the third grid level 1° spaces are defined by subdividing the second order 10° cells in a 10 by 10 array. The decimal grid is maintained at the fourth and fifth grid levels so that areas are defined as cells measuring 36" by 36". Each cell within these grids is identified by a two digit number. The first digit indicates location in the vertical or north-south direction and the second digit provides the east-west or horizontal reference.



Source: Jiri Kallab, "Linear Geographical Code for Management Information Systems"

Figure 5.16 Linear Geography: First and Second Order Grids

Grid Level	Area Name	Array	Cell Size
1	Masses	4 X 2	90° X 90°
2	Zones	9 X 9	10° X 10°
3	Spaces	10 X 10	1° X 1°
4	Areas	10 X 10	6' X 6'
5	Spots	10 X 10	36" X 36"

The grid-in-grid arrangement of the linear geographical coding system telescopes very rapidly from large 90° cells to very small 36" cells.[1] Application of this system on a worldwide or national scale to identify individual cities or place locations would not require the use of the fifth order grid because of the extremely small cell size at this level. However, in the metropolitan and urban application of linear geography the fifth order grid can be used to code single blocks or buildings.

Although the linear geographical coding system is based upon geographic subdivisions defined by regular grids and conforms with lines of latitude and longitude at various intervals certain applications of this code may require the realignment of the grid structure. As in the Geo-Code system, the grids can be altered to accommodate irregular areas such as states, counties or other geopolitically delineated entities. Furthermore, at a national or regional level very fine grain dimensions of the fifth level grid permits the use of each cell almost as a point reference. For example, the boundaries of a large city could be defined in terms of a set of fifth level grid cells outlining the area.

---

[1]In comparison the six level grid-in-grid system of the Global Reference Code begins with 180° cells and extends downward to cells measuring 2' X 2' (See Section 5.3).

The major advantage of linear geography and all other grid and coordinate coding structures is that the code not only uniquely identifies a given entity but reflects spatial relationships among the entities coded and can be manipulated mathematically to perform a number of distance and areal calculations.

References:

Kallab, Jiri. "Linear Geographical Code for Management Information Systems." New York, N. Y.: Artronic Information Systems, Inc., 1972.

CHART 1

GEOCODING SYSTEMS SUMMARY

Section	System	Code	Indication	Code Structure
2.1	Federal Information Processing Standards	25025 1120	The basic five digit FIPS code identifies a county or county equivalent within state; four additional digits qualify the county as part of a designated SMSA	state(2) county(3) SMSA(4)
2.2	General Services Administration	250250140	The basic nine digit GSA code identifies a place within a county within a state; code assignment is worldwide with an orientation towards military installations and other federal property	state(2) county(3) place(4)
2.3	International Business Machines Corporation	200250100	The basic nine digit IBM code identifies a place having a population of 2,500 or more as of 1960 within a county and state	state(2) county(3) place(4)
2.4	Dun and Bradstreet	410894863	The basic nine digit Dun code identifies a place within a county within a state; code assignment includes places in Canada and reflects population density and business activity; three additional digits qualify the place as part of a designated SMSA	state(2) county(3) place(4) SMSA(3)
2.5	American National Standards Institute	0110510	The basic seven digit ANSI code identifies a place within a state; code assignment covers eighteen possible place types including unpopulated locations	state(2) place(5)
3.1	Bureau of the Census	2612109523 4500221404	The basic six level Census code identifies an enumeration district within five other levels of geostatistical and geopolitical entities; in 1970 there were about 250,000 enumeration districts in the United States	state(2) county(3) MCD/CCD(3) place(4) tract(4) ED(4)
3.2	National Location Code	44200520022	The first seven digits of the NLC code identifies a county within a county cluster within a state and region; the last four digits of this code identify one of 43,000 designated standard location areas in the United States	region(1) state(1) area(4) county(1) SLA(4)
3.3	Geolocation Code	AALS	The four character Geoloc code identifies a place of significance to the Department of Defense; code assignment is worldwide and oriented to military use	place(4)
3.4	Waterborne Commerce Statistics	0716	The four digit U.S. Army Corps of Engineers Code identifies a division and waterway	division(2) waterway(2)
3.5	Federal Aviation Administration	BOS	The three character FAA code identifies major airports, heliports, or seaplane bases in the United States	airport(3)
3.6	Standard Point Location Code	142117	The six digit SPLC number identifies a location of significance to the trucking and railroad industries within a geopolitical hierarchy	region(1) state/state section(1) county group(1) county/county section(1) place(2)

CHART 1

GEOCODING SYSTEMS SUMMARY (CONTINUED)

Section	System	Code	Indication	Code Structure
3.7	PICADAD	4120533748	The ten digit PICADAD code of the Transportation Division, Bureau of the Census, identifies one of 56,000 designated key points within a county, state and region	region(1) state(2) county(3) key point(4)
3.8	Federal Highway Administration	1010	This four digit FHWA code identifies a county within a county cluster within a state	state(2) county cluster(1) county(1)
3.9	National Transportation Zones	014	This three digit code identifies an SMSA or other non-SMSA county cluster of relatively homogeneous socio-economic characteristics; code numbers 514 through 533 designate one of the ten highest U.S. ports ranked by foreign tonnage	zone(3)
3.0	Federal Railroad Administration	ADM	The three character Amtrak code identifies major passenger railroad stations in the United States	station(3)
4.1	ZIP Code	02120	The five digit ZIP code of the U.S. Postal Service identifies a post office or post office branch within a postal section and region of the United States	region(1) section(2) office(2)
4.2	Water Resources Council	0201 65142	The four digit water resources region and subarea code identifies a county cluster drained by reach of a river within a larger drainage system region; the five digit land resource group and area code identifies a county cluster area of homogeneous land patterns within larger land resource regions	water resource region(2) water resource subarea(2) land resource group(2) land resource area(3)
4.3	Office of Management and Budget	011120	The first two digits of this code identifies one of ten groups of states designated as standard federal regions; the last four digits identifies a county or county cluster designated as a standard metropolitan statistical area	region(1) SMSA(4)
4.4	Economic Development Administration	835010	This EDA code identifies a group of counties designated as an economic development district within a state and a region	region(1) state(2) district(3)
4.5	Bureau of Economic Analysis	151	This three digit BEA code identifies one of 173 designated economic areas in the United States conforming to county boundaries and delineated on the basis of central place theory	area(3)
4.6	ZIP Marketing Areas	040	This three digit code identifies one of 136 marketing areas which conform to county boundaries and which were delineated on the basis of three digit level ZIP code sections and areas of dominant influence	area(3)
4.7	Foreign Trade Statistics	0401	This four digit code established by the Foreign Trade Division, Bureau of the Census identifies a port within a U.S. customs district	district(2) port(2)
4.8	Interstate Commerce Commission	12	This two digit number identifies one of five major freight rate territories or seven border territories defined by ICC for the waybill statistical series	territory(2)

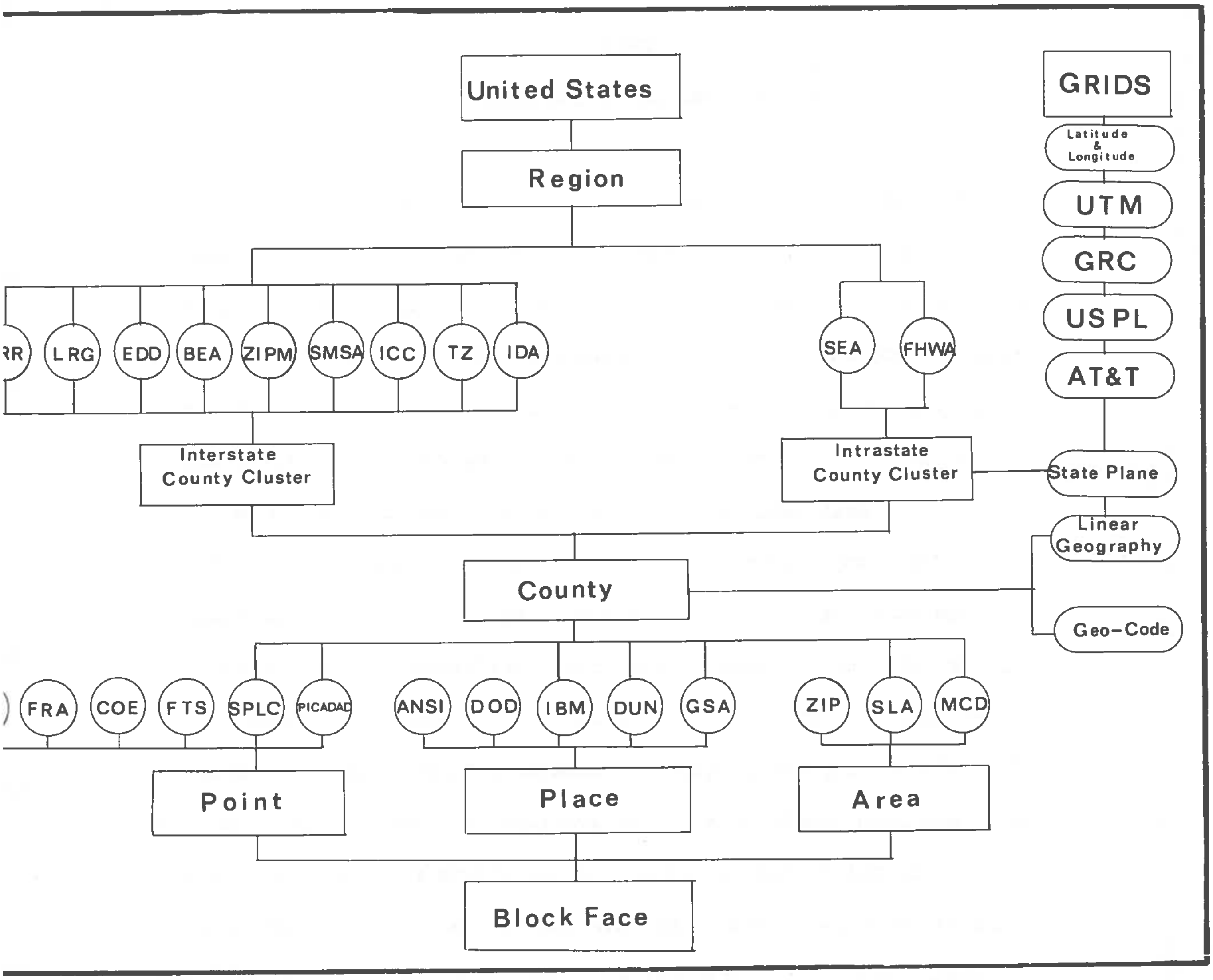
CHART 1

GEOCODING SYSTEMS SUMMARY (CONTINUED)

Section	System	Code	Indication	Code Structure
4.9	Institute for Defense Analyses	39	This two digit number identifies a cluster of core and contiguous counties designated by IDA to represent the distribution of percentage levels of the total U.S. population ranging from 20% to 60%	county cluster(2)
5.1	Geodetic Longitude and Latitude	42°12'30"N 71°05'45"E	Coordinates of longitude and latitude identify a point on the surface of the earth within degrees, minutes and seconds	degree(2) (2) minute(2) (2) second(2) (2) direction(1) (1)
5.2	Military Reference Grid	3PWN539445	The UTM code locates a point within 100 meters on a plane grid covering the entire globe	grid zone(2) 100,000 meter square(2) X coordinate(3) Y coordinate(3)
5.3	Global Reference Code	013122435143	This code identifies a location within a two minute square of longitude and latitude within six levels of grids covering the globe	grid 1 (2) grid 2 (2) grid 3 (2) grid 4 (2) grid 5 (2) grid 6 (2)
5.4	Geo-Code	OPKZ	This code identifies a geopolitical subdivision of the world or a subdivision of the oceans of the world within a series of four grids conforming generally to geopolitical boundaries	grid 1 (1) grid 2 (1) grid 3 (1) grid 4 (1)
5.5	State Plane Coordinate Systems	370.2N 671.4E	These coordinates locate a point on a plane grid covering a state or state section; error does not exceed one part in 10,000	X coordinate(5) Y coordinate(5)
5.6	United States Public Land Survey	T3NR2N17NESW	This code locates a 40 acre "quarter" within a "quarter section" (120 acres), a "section" (640 acres) and a "township" (6 square miles)	township(6) section(2) quarter section(2) quarter(2)
5.7	American Telephone and Telegraph V-H System	4422 1249	These coordinates locate a point within one-tenth of a square mile on a plane projection covering the United States and Canada	V coordinate(4) H coordinate(4)
5.8	Railway Express Agency	4750	This code locates an area 69 miles long and 50 miles wide on a random grid covering the United States	railblock row(2) railblock column(2)
5.9	Linear Geographical Code	1108213562	This code identifies a location within a 36 second square of longitude and latitude within a hierarchy of grids covering the globe	masses(2) zones(2) spaces(2) areas(2) spots(2)

NOTE: For the sake of brevity this summary includes only the basic or unique codes within each system. Numbers in parentheses in the code structure column indicate the number of digits required at each level of the code.

CHART 2
GENERALIZED HIERARCHICAL RELATIONSHIP OF VARIOUS GEOCODING SYSTEMS



KEY

WRR:	WATER RESOURCE REGIONS	SPLC:	STANDARD POINT LOCATION CODE
LRG:	LAND RESOURCES GROUP	PICADAD:	PLACE IDENTIFICATION/CHARACTERISTICS AND AREA/ DISTANCE AND DIRECTION
EDD:	ECONOMIC DEVELOPMENT DISTRICTS	ANSI:	AMERICAN NATIONAL STANDARDS INSTITUTE
BEA:	BUREAU OF ECONOMIC ANALYSIS	DOD:	DEPARTMENT OF DEFENSE
ZIPM:	ZIP MARKETING AREAS	IBM:	INTERNATIONAL BUSINESS MACHINES
SMSA:	STANDARD METROPOLITAN STATISTICAL AREA	DUN:	DUN AND BRADSTREET
ICC:	INTERSTATE COMMERCE COMMISSION	GSA:	GENERAL SERVICES ADMINISTRATION
TZ:	TRANSPORTATION ZONE	ZIP:	ZONE IMPROVEMENT PLAN
IDA:	INSTITUTE FOR DEFENSE ANALYSES	SLA:	STANDARD LOCATION AREAS
SEA:	STATE ECONOMIC AREA	MCD:	MINOR CIVIL DIVISION
FHWA:	FEDERAL HIGHWAY ADMINISTRATION	UTM:	UNIVERSAL TRANSVERSE MERCATOR
FAA:	FEDERAL AVIATION ADMINISTRATION	GRC:	GLOBAL REFERENCE CODE
FRA:	FEDERAL RAILROAD ADMINISTRATION	USPL:	UNITED STATES PUBLIC LAND SURVEY
COE:	U. S. ARMY CORPS OF ENGINEERS	AT&T:	AMERICAN TELEPHONE AND TELEGRAPH
FTS:	FOREIGN TRADE STATISTICS		

APPENDIX A

THE NATIONAL GEOCODING CONVERTER FILE

When the U.S. Department of Transportation became concerned with national geocoding systems, the success of the urban geocoding program conducted by the Bureau of the Census and supported to a large extent by DOT and other Federal agencies had already been established. The 1962 statutory requirement for a continuing metropolitan transportation planning process, establishment of the Census Advisory Committee on Small Area Data in the summer of 1965, and the beginning of the Census Use Study a year later, in 1966, were the prime stimuli for the forward thrust in urban geocoding. By 1971 a fairly standardized universe of urban geographic base files utilizing the DIME coding system had been implemented in almost all 267 SMSAs. A battery of DIME related computer programs, including UNIMATCH, GRIDS, CRAM and DACS, had been developed and were available for use (See Section 3.1).

In the United States comparable incentives for the development of national geographic base files have not existed. Indeed, national geographic coding is characterized by that fundamental chaos that results when a multitude of special purpose systems is brought into being by users needing to deal with the pressing urgencies of the moment. Unlike micro-level geocoding, which has exhibited a tendency towards standardization, uniformity and convergence of various approaches towards one major urban geocoding concept, macro-level geocoding tends towards divergence, specialization, and multiplicity of structures.

These fundamental differences in the development of urban and national geocoding dictated the approach taken by DOT in its efforts to achieve improvements in the utilization of national geocoding systems. Though a single, standard, DIME-like national geocoding system would seem desirable, the extent of investments in the many existing national systems and the lack of a geographic common denominator at the macro-level comparable to block face at the micro-level, make the comprehensive code approach infeasible. Therefore, DOT has focused its efforts on the development of a geographic code conversion capability. This code conversion capability is intended to facilitate the association of variously coded macro-level data files and to provide a flexible geographic cross referencing system for the benefit of users.

Drawing upon recommendations made at the National Geocoding Conference in 1971 and a systematic examination of existing systems, the DOT national geocoding program was designed to center on four potential converter files, as follows:

File 1 is a county file. It contains a record for every county and county equivalent in the United States, identifying for each county all major county codes and the associated county aggregate, state or regional codes.

File 2 is a county coordinate file. This is a DIME-structured file containing "link-node" records for all county segment boundaries in the United States. It is intended to provide mapping and graphic display capabilities at the county level.

File 3 is a county component file. It contains a record for each county and county equivalent in the United States, identifying for that county all major subcounty area, place, point, city or location codes.

File 4 is an international file containing the geographic codes used to designate countries and other national entities such as territories, zones and principalities.

File 1 and File 2 are completed and available from the Transportation Systems Center, Cambridge, Massachusetts. Feasibility studies for the development and structure of Files 3 and 4 are now under investigation.

The two county files were selected as the initial foci for converter design because a majority of the geocoding systems constructed and operated independently in government and industry contain this unit. The county unit is the only natural fit or linkage among these and, as indicated earlier, below the county level there is no geodefined areal unit common to any substantial portion of national information systems.

In addition to its value as a "most common denominator" among major national geographically referenced information systems, the county unit has several other characteristics which make it suitable to serve as the initial focus of geocoding converter design. Most importantly, counties have an exact, relatively stable, areal geometry with geodefinable boundaries that can be expressed as points, or nodes and lines, or links and curves. Although File 1 of the DOT National Geocoding Converter File is only an areal reference file, the combination of File 1 and File 2, will create the most flexible type of geographic reference system, a coordinate or point boundary system. The computational capabilities which this system will provide include simple but essential cross referencing by sort and print routines as well as more sophisticated point-in-polygon, polygon search and geo-manipulation routines.

The fact that the county unit occupies a position mid-way in the hierarchy of geographic units employed by macro and micro geographic identification systems is also significant. Having established geo-code convertibility and data manipulation capability at this level will greatly facilitate the expansion of converter capability above and below the county level. Eventually, a geocoding converter at this scale may provide the linkage between small area data and national data systems. In the meanwhile, the county converter file will be of great benefit to the coordination of data sets at the national, regional and state-wide levels.

On the basis of these advantages then, county and county equivalent were selected as the unit record for the National Geocoding Converter File 1. It became necessary, however, to incorporate a number of special county section units to account for certain anomalies. For example, although Standard Metropolitan Statistical Areas (SMSAs) generally consist of whole counties, the SMSAs in New England consist of townships that are segments of counties. In order to provide for the geographic definition of SMSAs, which do not consist of whole counties, and to provide for maximum conversion flexibility, SMSA county segments were uniquely identified by an arbitrarily assigned DOT County Section Identification Number and incorporated as separate records in the file. The same procedure was applied to incorporate SPLC county sections and the two sections of Yellowstone National Park located in Idaho and Wyoming.

The design of the DOT National Geocoding Converter File 1 includes a self documenting feature to explain codes with uncertain status, codes with which there is some special problem, fields for which no code is available and code changes over time. For example, although the list of counties that existed in 1950, 1960, and 1970 does not vary greatly, certain county equivalents such as independent cities are altered relatively frequently and these changes are documented on the file. Furthermore, in order to make it possible to reconstruct the exact composition of a given SMSA at any given point in time since 1950, National Geocoding Converter File 1 contains several fields to document the history of each SMSA. For each county, the year of entry into the current SMSA is recorded. If that county was once a member of another SMSA, the code identifying that former SMSA, the year of entry into that SMSA and the year of departure from that SMSA are also recorded. This information will enable the user to compare accurately county and SMSA units over time.

In addition to the Federal Information Processing Standard (FIPS) for state name, state abbreviation, state code, county name and county code which are the key fields in the file, each record of the National Geocoding Converter File 1 contains 61 fields, including three variations of state abbreviations, twenty-nine county and county aggregate codes, a record serial number, four fields documenting SMSA history, a DOT county section identifier and a DOT county type indicator. DOT now has the software programs for updating and maintaining the file as well as programs to obtain printouts of the file in either of two formats and in any sequence desired. When File 2, the county segment file, is coupled with File 1 a very powerful computer mapping capability is added, making it possible to associate, at the county level, data coded to different systems, and to graphically display the distribution of the data for counties and county aggregates.

The Office of Systems Analysis and Information at DOT views the construction and application of the National Geocoding Converter File 1 as a first step toward solving the problems of diversely coded data bases and meeting the needs of the U.S. Department of Transportation and other users in government and industry for a flexible national geographic base file.

References:

G. Boetje; C. Davis; and D. DeVorkin. United States Department of Transportation National Geocoding System: County Boundary Coordinate File for Geographic Display of Data. P-058. Cambridge, Mass.: The Charles Stark Draper Laboratory, Inc., 1974.

H.W. Bruck, and P.A. Werner. "National Geocoding Systems: Design of a Converter." A paper prepared for the Urban and Regional Information Systems Association Annual Meeting, San Francisco.

H.W. Bruck, and P.A. Werner. "National Geographic Base Files: An Initial Step." Geographic Base File System: Establishing a Continuing Program. Conference Proceedings, January 18-19, Seattle, Washington. Washington, D.C.: U.S. Government Printing Office, 1973.

H.W. Bruck, and P.A. Werner. "National Geographic Base Files: Future Developments." Geographic Base File Systems: A Forward Look. Conference Proceedings, April 16-17, Boston, Massachusetts. Washington, D.C.: U.S. Government Printing Office, 1974.

U.S. Department of Transportation. Transportation Systems Center. National Geocoding Converter File 1: Structure and Content. ed. Santo La Tores. Cambridge, Massachusetts: Transportation Systems Center, U.S. Department of Transportation, 1974.

P.A. Werner. "National Geocoding." Annals of the Association of American Geographers. Volume 64, Number 2. (June 1974).

P.A. Werner. The DOT National County Component Converter File: Propects, Problems, Feasibility. A report prepared for the Transportation Systems Center, U.S. Department of Transportation, Cambridge, Massachusetts, 1974.

APPENDIX B

GEOCODING SYSTEMS OF CANADA

B.1 Standard Geographical Classification System

The Standard Geographical Classification (SGC) is a system that was developed and is maintained by Statistics Canada for the identification and coding of geographical areas.[1] It consists of a heirachical set of geopolitical units, each identified by a unique code, and a list of place names related to these units. The objective of the system is to make available a standard set of geographical units which will facilitate the comparison of statistics for particular areas.

The first draft edition of the Standard Geographical Classification manual, based on the 1961 Canadian Census, appeared in the Fall of 1964. It was circulated as the official standard for coding, tabulating, and publishing geographical data in Statistics Canada. An interim edition appeared in 1968 with codes up to date as of June 1966. A second edition in 1970 provided revised codes to January 1970. A third edition was published in August 1972 with codes updated to January of that year. All three of these early editions were distributed on a limited basis and the codes were intended primarily for use within Statistics Canada. However, as a result of a

---

[1]Statistics Canada is roughly the Canadian counterpart of the Social and Economic Statistics Administration (SESA) in the United States.

growing number of outside users, the new SGC manual will be made available to the general public late in 1974.

The Standard Geographical Classification System is a three level hierarchy. Each level is represented by a two digit numeric code. The highest level of the hierarchy designated by the first two digits of the code identifies a province or a territory which are the first order political subdivisions of Canada. The third and fourth digits of the code identify a census division, and the fifth and sixth digits identify a census subdivision, which is the basic building block of the system. These six digits form the root code of the classification, and all other codes in the system are related to it. The other codes include designations for census metropolitan areas and census agglomerations, a municipality type code, a reference area code and place names.

<u>Provinces and Territories</u>. Provinces and territories are coded in the following manner:

(1) Atlantic Provinces
- 10 Newfoundland
- 11 Prince Edward Island
- 12 Nova Scotia
- 13 New Brunswick

(2) Quebec
- 24 Quebec

(3) Ontario
- 35 Ontario

(4) Prairie Provinces
- 46 Manitoba
- 47 Saskatchewan
- 48 Alberta

(5) British Columbia
- 59 British Columbia

(6) Territories
- 60 Yukon
- 61 Northwest Territories

If regional codes and codes for the territories are not required the provinces can be numbered by only the second of the two digits.

Census Divisions. Census divisions comprise the middle level of the hierarchy. They are intended to be stable over long periods of time (30-40 years) and, therefore, useful for historical analysis of statistical series. At the present time there are six types of census divisions as shown below:

Province or Territory	County	Census Division	Territory	Regional Municipality	District	Regional District
Newfoundland		X				
Prince Edward Island	X					
Nova Scotia	X					
New Brunswick	X					
Quebec	X		X			
Ontario	X			X	X	
Manitoba		X				
Saskatchewan		X				
Alberta		X				
British Columbia						X
Northwest Territories					X	

In the eastern part of Canada the county is the main type of census division, supplemented by districts and a territory in sparsely inhabited areas and regional municipalities in some densely populated areas. The first regional municipality was created in 1969, and since then five more have been established, replacing counties and districts in several areas of Ontario.

In Newfoundland, Manitoba, Saskatchewan and Alberta, provincial law does not provide for areal units equivalent to counties. Therefore census divisions have been established for statistical purposes in cooperation with provincial authorities. In 1970 the ten census divisions in British Columbia were replaced by a new structure of 29 regional districts designated by the province. These units were first used for data collection in the 1971 Census and are now being used in Census publications. The Yukon Territory forms one division by itself, and the Northwest Territories are divided into three districts.

Census Subdivisions. Census subdivisions are the smallest geographical units of the classification and constitute its basic building blocks. There are two categories of census subdivisions:

- organized (incorporated) municipalities with legally defined boundaries, the authority to levy taxes and some form of elected government; and
- non municipally organized subdivisions

Incorporated municipalities are created by provincial legislation, and usually come under the authority of a provincial department of municipal affairs. All incorporated municipalities are recognized as census subdivisions. There are about 4,500 such incorporated municipalities.

There are a number of different types of municipalities across Canada. Since each province has its own municipal legislation it is difficult to compare types of municipal organization among provinces except in the broadest terms. For example, the township in Ontario, the canton in Quebec, and the district municipality in British Columbia are all rural municipalities, but their powers and responsibilities for local administration differ. Nor should it be assumed that municipalities designated in the same manner in several provinces are organized and administered in the same fashion. For example, the criteria of minimum population required for the incorporation of cities varies from province to province. Nevertheless, broad comparisons are possible; cities are large units rarely having a population under 10,000, while most towns are in the 1,000 - 10,000 range.

Under certain conditions a municipality may be subdivided into two or more geographical units. In twelve counties of Nova Scotia, the rural part of each county comprises one municipality. In these cases the municipality has been split up into a number of census subdivisions by mutual agreement between the province and Statistics Canada. The purpose of these units is to provide a finer areal breakdown for statistics published by the Census of Canada. A similar situation occurs in New Brunswick and Prince Edward Island but in these cases lots and parishes respectively take the place of census subdivisions. For the rural areas of British Columbia, electoral areas were combined to form census subdivisions. In Saskatchewan a number of rural municipalities are located in more than one census division and are subdivided so the part in each division comprises a separate geographical unit. The cities of Flin Flon and Lloydminster are both located in two provinces and are both split into two geographical units, one in each province.

The sparsely populated areas of Canada, both in the provinces and the territories, are not municipally organized, but are administered directly by the province and territories. In many cases non-municipal subdivisions are those parts of census divisions (counties and districts or equivalent) which are left over after cities, villages, various types of rural municipalities and Indian reserves have been accounted for. In the northern areas of Quebec, Ontario, Manitoba, and Saskatchewan, there is one such subdivision per census division. In Newfoundland and the Territories, the unorganized part of each census division was split up into a number of subdivisions. These units are called subdivisions in Newfoundland and health districts in the Territories.

The other principal type of non municipal subdivision is the Indian reserve. These units are designated and administered by the Federal Department of Indian Affairs and Northern Resources. Originally, Indian reserves (approximately 3,200 in number) were combined geographically into 480 groups on the basis of a maximum of ten groups per census division. However, in the 1974 Standard Geographical

Classification, Indian reserve coding has been completely revised. Reserves are grouped according to the municipalities and unorganized areas within which they are located. There are now 280 groups. Both in the former and the revised systems, some groups contain one reserve. Indian settlements are grouped according to the same principles as Indian reserves.

Census Metropolitan Areas (CMA) and Census Agglomerations (CA). Census metropolitan areas are census geostatistical areas which have been established for the purpose of providing statistics on a comparable basis for all urban areas with a population in excess of 100,000. The 1971 Census of Canada defined 22 such areas. Generally, a census metropolitan area is made up of an urban core with a partly urban and partly rural fringe surrounding it. The core is a continuously built up area with a population over 100,000. It may contain one or more municipalities. Fringe municipalities are attached to the core on the basis of a number of criteria which measure the strength of linkages between the fringe and the core. The most important of these criteria is the volume of commuting; if there are substantial numbers of commuters from a municipality outside the core to the core, that municipality is considered to be part of the fringe. Another criterion for inclusion in a CMA is the rate of population increase. Rapid growth is characteristic of suburban municipalities which are closely linked to large urban centers.

The Calgary and Saskatoon CMAs do not have a fringe but consist of the City of Calgary and the City of Saskatoon alone. The number of municipalities in the CMA vary from one (Calgary and Saskatoon) to over one hundred (Montreal), but in all cases the components are complete census subdivisions.

Census agglomerations are census geostatistical areas which have been established for the same purpose as CMAs, but are considerably smaller, with a population in the 5,000 to 100,000 range. A CA must have an urban core of at least 1,000 persons plus an adjacent fringe area of at least 1,000 persons at a density in excess of 1,000 persons per square mile in a different municipality. CAs consist of complete municipalities as well, with a few exceptions.

Census agglomerations can be subdivided into a number of size classes. They are identified by a range of numbers in the CMA field.

Population Size Classes	CA Code Ranges	Number of CAs
50,000 - 99,000	040 - 099	12
25,000 - 49,000	100 - 139	26
10,000 - 24,999	140 - 199	31
5,000 - 9,999	200 - 249	10

It is worth noting that not all cities and towns in these population size categories are CAs. For example, St. Jean, Quebec, with a 1971 population of 32,863 is the core of a CA of the same name since two adjacent towns meet the fringe criterion mentioned above, while Chatham, Ontario, with a 1971 population of 35,317 is not the core of a CA since there are no adjacent municipalities which meet the fringe criterion. In the smaller CA categories, the number of CAs is only a fraction of the total number of urban centers in these size categories.

Place Names. "Place" is a general term for hamlets, settlements, localities, urban neighborhoods or subdivisions, post offices, railway stops, airports, and geographical townships. Places are generally vaguely defined small populated areas, or point locations in the case of railway stops and post offices. The term village, when used as a place name, usually refers to an unincorporated center in a rural area with a population in excess of 100 persons, occasionally reaching as many as 1,000 persons. Hamlets and settlements are small unincorporated centers with a population of under 100 persons. Localities contain scattered population, usually less than 100 people, but in some cases are uninhabited. Rural post offices are often located in hamlets or localities. The same is true of scheduled railway stops although in some cases there is no resident population at railway stops.

The primary source of information on rural place names are the provincial gazetteers published by the Canadian Permanent Committee on Geographical Names, Department of Energy, Mines and Resources. Each gazetteer (one per province) contains a complete list of official names for natural features and places in that province. The Standard Geographical Classification, with some exceptions, includes the names of all places in Canada. The names of natural features do not appear in the SGC except where some activity by man is associated with a feature. In northern Quebec and Ontario many lumber camps and some mining operations are located on or near the shores of lakes, and such activities often give the name of a lake as their location address.

The bulk of names within incorporated cities and towns represent urban neighborhood or subdivision names. The former is a homogeneous area with recognized social or economic characteristics or historical associations. Urban subdivisions are similar but are located in suburban areas. In many cases a subdivision is named by the developer or builder. In some cases other major features of an urban area such as larger parks are included in the SGC. Airports, which are located in an urban area or nearby in a rural township are included as well.

All the places named in the Standard Geographical Classification System are listed by census subdivisions within census divisions (See Figure B.1).

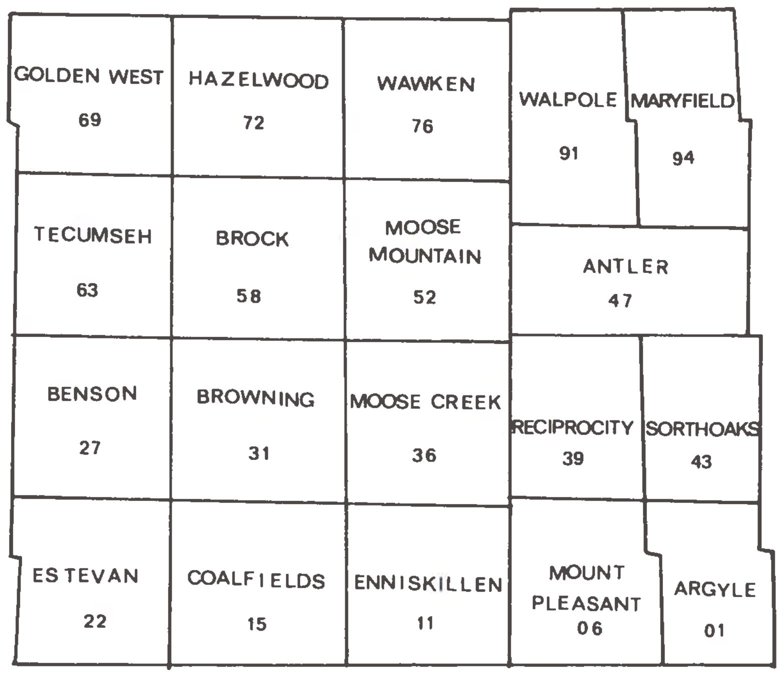
The Coding System. In the Standard Geographical Classification System codes are assigned to census divisions and subdivisions in geographical order. The first code is assigned to the subdivision or municipality located in the southeast or lower right corner of the division. The coding proceeds westward along the southern boundary of the division to the east and proceeds to the eastern boundary where it turns north again,and proceeds in the same snake like fashion. Numbers are not assigned consecutively, care being taken to leave digits free for new municipalities which may be created. By this system the code number of a municipality gives some indication of its geographical location and new municipalities are easily accomodated. Codes which have been deleted are normally not

PR	CD	CS	CMA	RE	NAME	TYPE
35	29				Brant County	
35	29	01		01	Onondaga	Twp.
35	29	03		01	Onondaga Indian Reserves	I.R.
35	29	04	073	04	Brantford	Twp.
35	29	06	073	06	Brantford	City
35	29	09		09	Oakland	Twp.
35	29	11		11	Burford	Twp.
35	29	16		16	South Dumfries	Twp.
35	29	18	073	16	Paris	Town
35	30				Waterloo Regional Municipality	
35	30	04	019	04	North Dumfries	Twp.
35	30	10	019	10	Cambridge	City
35	30	13	019	13	Kitchener	City
35	30	16	019	16	Waterloo	City
35	30	20	171	20	Wilmot	Twp.
35	30	27		27	Wellesley	Twp.
35	30	35		35	Woolwich	Twp.

Source: Statistics Canada

Figure B.1 Canadian Standard Geographical Classification Province of Ontario

reused. The following diagram provides an example of this concept.



Census divisions are numbered within provinces according to the same principle, the only difference being that numbers are assigned consecutively since census divisions are more stable and, when they do change, tend to change all at once for a province, requiring complete revision.

Some variations exist in the numbering system. In Manitoba, Saskatchewan and Alberta where census divisions have official numbers, but no official names, the Standard Geographical Classification codes for the divisions coincide with the official numbers although the numbers are not strictly in geographical order. In a few cases municipalities are not coded in order because the recent incorporation of a large number of units has strained the system--the most notable example being division

1 in Newfoundland where 95 of the 99 available codes have been utilized.

In addition to unique codes for each census subdivision the Standard Geographical Classification System provides for a municipality type code and a reference area designation. The municipality type code categorizes each census subdivision as one of eight types or organization such as Indian reserve, township, rural municipality, etc. The concept of the reference area permits small census subdivision areas--cities, villages, townships, etc.--to be grouped with the surrounding larger census subdivision (generally a rural municipality or unorganized area). Larger cities of over 25,000 inhabitants which are likely to have a larger population than adjacent rural or unorganized areas, and for which a reference area is not obvious, are not reference to a rural municipality or unorganized areas but to themselves. In some cases large cities can include villages and towns within their boundaries, and in those cases the city acts as a reference area for these villages and towns. In other cases towns or villages have amalgamated with adjacent townships, the result being a town or village referenced to itself.

Sometimes villages and towns are located on the boundary between two or three reference areas rather than within one area. In these cases the village or town is referenced to the area within which the larger part of the population is located.

The three volume SGC manual is prepared from the Standard Geographical Classification Master File. The master file contains additional geographic codes for special purpose districts and regions. These include postal zones, Statistics Canada economic regions, and the Canada manpower center areas.

The Standard Geographical Classification Master File

Field	Size	Title
1	1	Region
1	1	Province
3	2	Census Division
4	2	Census Subdivision
5	3	Record Number
6	30	Subdivision Name
7	6	All Canada Alpha Sequence Number
8	1	Type of Municipality Code
9	1	Alternative Name Indicator
10	1	Blank
11	1	Post Office Indicator
12	2	Reference Area Code
13	3	Census Metropolitan Area/Census Agglomeration Number
14	2	Federal Electoral District Number
15	2	Department of Defense Production Economic Region
16	2	Statistics Canada Economic Region
17	2	Type of Municipality Name
18	10	Blank
19	30	Place Name
20	30	Blank
21	8	Population
22	16	Universal Transverse Mercator Grid Coordinates
23	52	Blank
24	4	Canada Manpower Center Number

There is also a Canadian Standard Geographical Classification for foreign countries and states of the United States. Units of area defined as political entities (including territories, trust territories, overseas departments, dependencies, colonies, and protectorates) are grouped into six continents or large parts thereof. In general, the continents follow the commonly recognized boundaries except for Europe which includes all of the Soviet Union.

The continents are ordered so as to link regions which are closely related geographically and politically. They appear as follows: Europe, Africa, Asia, Oceania and Antarctica, South America, North America, and Central America and the Carribean.

Within each continental group of units of area are arranged contiguously, so far as possible (the contiguous order being broken sometimes to work islands into the system). The ordering is influenced by such factors as trade routes, political affiliations, language groups, etc., and therefore, the country chosen as a starting point in each group is the one which appears to provide the best chance of a useful ordering. The states of the United States are listed contiguously within nine regions of the country as shown in the publications of the U. S. Bureau of the Census (See Section 3.1).

The structure of the code is similar to that used for areas in Canada. It is limited to four digits, however, because for most purposes identification of the country or state of the United States is likely to be sufficient. If more specific foreign locations are required for a special purpose, two digits are available for such special codes---e.g., to designate provinces, cities or ports within countries. The first digit of the code is 9 in all cases indicating areas outside of Canada, first digits 1, 2, 3, 4, 5, and 6 having been used to designate regions within Canada. The second digit of the code designates continental subgroups, the second digits 0, 2, and 4 being reserved for such alternate groups as European Economic Community, Communist Block, or Commonwealth Countries. The third and fourth digits of the code designate the basic geographical units of area, the third and fourth digit numbers 01 to 19 being reserved throughout the classification for users who wish to follow the coding system but who require a few extra codes for special purposes. The code for the United States of America is the only code in the system to end in 0.

B.2 The Geographically Referenced Data Storage and Retrieval System (GRDSR)

The Geographically Referenced Data Storage and Retrieval System is the outcome of five years' research by Statistics Canada into solving the many problems associated with the storage and retrieval of statistics about small areas. In order to respond to the small area data requests Statistics Canada requires a flexible, efficient system to repeatedly assemble and tabulate information according to arbitrary special interest areas such as school districts, police precincts, and fire zones. Very much like the U. S. Bureau of the Census DIME system (See Section 3.1), GRDSR provides the necessary flexibility, the aggregation, disaggregation, and reaggregation capabilities by utilizing the block face as the basic building block in the geocoding of urban areas.

In the GRDSR file points at which streets intersect or curve sharply in the city pattern are referred to as nodes, each of which is assigned Universal Transverse Mercator coordinates (See Section 5.2). Every street is represented by a series of nodes connected by straight line segments.

A block face is defined as one side of a city street between consecutive intersections with other streets. Thus, up to two block faces can be formed by a pair of adjacent nodes, each located at a four way street intersection. However, a block face can also encompass several nodes. For example, a block face may contain one intermediate node marking a change in direction and another node representing an intersection on the opposite side of the street only.

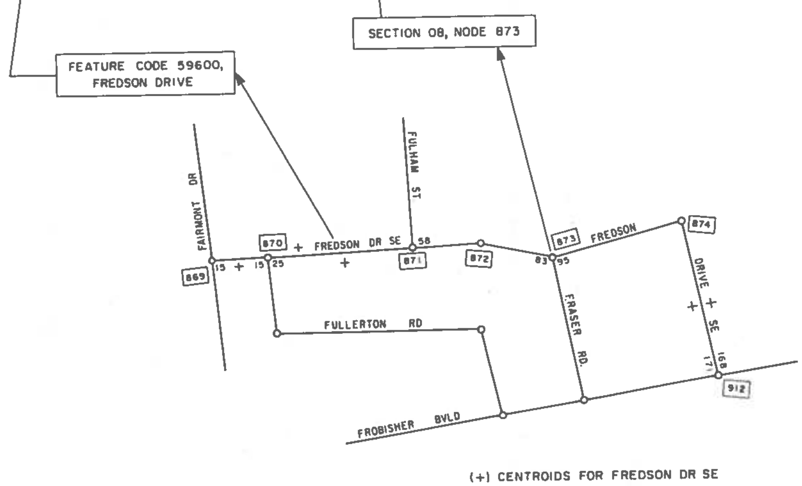
Whenever a block face is to be formed by a pair of nodes, these nodes must consitute the beginning and end of a valid civic address range.

Unlike DIME, however, all street addresses along a block face in the GRDSR file are assigned, and share, the coordinates of the block face centroid, which is simply a reference point offset from the street midway between the two nodes forming the block face. During the conversion operation, the address of each record or data observation is matched to a block face (using a list of valid street names and address ranges). From there, the correct centroid is known and its coordinates can be added to the record.

The actual geocoding operation (or assignment of coordinates to data) is carried out using GRDSR components known as Area Master Files (AMF). Area Master Files contain a logical representation of all city streets, plus some other features, in computer readable form. An AMF references every street, address range, block face and centroid coordinate in the covered area (See Figure B.2). Also itemized are other features (such as railroad tracks, rivers, and municipal boundaries) which help users to choose query areas.

Area Master Files have been created for major portions of 14

MUNIC CODE	FEATURE CODE	SEQ. No.	STREET NAME	TYPE	DIR	NODE NUMBER	NODE COORD		ADDRESS BEFORE		ADDRESS AFTER		CENTROID LEFT		CENTROID RIGHT		INTERSECTING FEATURES
							X	Y	L	R	L	R	X	Y	X	Y	
4835	59600	040E	FREDSON	DR	SE	08912	706954	5651540	168	171	—	—	706868	5651755	706861	5651673	FROBISHER BV
		035				08874	706844	5651755	—	—	—	—	—	—	—	—	
		030				08873	706771	5651722	—	83	—	95	—	—	706635	5651700	FRASER RD
		025				08872	706733	5651729	—	—	—	—	—	—	—	—	
		020				08871	706631	5651722	—	—	58	—	706516	5651734	—	—	FULHAM ST
		015				08870	706497	5651711	—	15	—	25	—	—	706452	5651687	FULLERTON RD
		010B				08869	706406	5651707	—	—	—	15	—	—	—	—	FAIRMONT DR



Source: Statistics Canada

Figure B.2 The Geographically Referenced Data Storage and Retrieval System

Canadian urban centers, which include a total of 16 cities. These files reference more than 225,000 block faces, corresponding to a population figure of approximately seven million.

B.3 The Canadian Postal Code

In 1969 the Canadian Post Office Department, virtually the last major mail service to mechanize sorting operations, engaged the consulting firm of Samson, Belair, Riddell, Stead, Inc., to identify the department's operating needs for a postal coding system and to recommend a code which would take full advantage of automated sorting machines and ensure the highest degree of utilization by both the department and the mailing public, especially large volume mailers. The outcome of the consultants' study was a unique six character alpha numeric postal code which identifies both the forward and local elements of the sortation system.

The first three characters of the Canadian postal code identify a forward sortation area in a city or in the country. The last three characters which are separated from the first half of the code by a single space, identify a local delivery unit in the forward sortation area. The area codes are always in an alpha numeric alpha format and the local codes are always in a numeric alpha numeric format. The code utilizes 20 alpha characters. Colons, dashes, oblique strokes and the letters D, F, I, O, Q, and U have not been used because of difficulties with optical character recognition equipment.

One of eighteen major geographic areas are identified by the first alpha character of the Canadian postal codes. The letters have been assigned from east to west and they represent a province or a large portion of a province. The significance of the remaining components of the code vary according to the urban or rural nature of the sortation area coded. Rural forward sortation areas are identified by a zero in the second position of the area code. This means that all Canadians having a code with the second digit "0" receive their mail by means other than letter carrier. Urban forward sortation areas served by a letter carrier are identified by the numbers 1 through 9 in the second position of the area code. Initially the area represented corresponded to an area served by a maximum of 25 letter carriers, however as the system develops, this areal criteria is expected to cease to be significant.

The local portion of the postal code identifies a local delivery unit. In urban areas this may be:

- a city block face
- a large apartment building (50 suites or more)
- a large business building (10 individual firms or more)
- a large volume receiver

The local area code also identifies a type of delivery service as either general, lock box, rural route, suburban, or mobile delivery service.

A letter addressed to 608 Portage Avenue in Winnipeg, for example, would be coded R3C OG5. The first letter, R, stands for all of the province of Manitoba; 3C represents a 120 block section of Winnipeg where 25,000 people are served by roughly 25 mailmen. The second half of the code, OG5, designates one of six addresses on an even numbered block face. Most urban family homes will share the same code with up to 50 other houses on the street.

The six digit alpha numeric code employed in Canada is more flexible than the five digit numeric ZIP code of the United States, because the use of numbers restricts the ZIP code to ten characters in each position. The first, third and fifth positions of the Canadian code, on the other hand, are letters. Although several letters have been rejected because of their similarity with numbers or other letters, there are still twenty letters to choose from in these positions. Generally speaking, the entire ZIP code gives the same information as the area code portion of the Canadian postal code. In urban areas, for example, the ZIP code only describes an address down to a relatively large section of a city, while the combination of area and local codes in the Canadian postal code defines a specific building or a block face. The Canadian postal code more closely resembles the British post code except that the British post code may vary from 5 to 7 characters and the alpha numeric order is not always the same.

The Canadian postal code was introduced April 1, 1971 in the city of Ottawa, Ontario. By November of 1973 all the provinces and territories had been coded.

References:

Canadian Department of Industry, Trade, and Commerce. Statistics Canada. GRDSR: Facts by Small Areas. Ottawa, Canada: Statistics Canada, Canadian Department of Industry, Trade, and Commerce, 1972.

Canadian Department of Industry, Trade, and Commerce. Statistics Canada. "Standard Geographical Classification Manual." Draft text of introduction. Ottawa, Canada: Statistics Canada, Canadian Department of Industry, Trade, and Commerce, 1974.

Canadian Post Office Department. Canada Post. "Canadian Postal Code." Ottawa, Canada: Canada Post, Canadian Post Office Department, 1973.

Samson, Belair, Riddell, Stead, Inc. *A Canadian Public Address Postal Coding System*. A consultant's report prepared for Canada Post. Ottawa, Canada: Canada Post. Canadian Post Office Department, 1969.

APPENDIX C

INTERNATIONAL GEOPOLITICAL CODING SYSTEMS

Over the past few years there has been a great deal of interest on the part of many organizations in establishing a standard set of international geocodes for geopolitical subdivisions of the world. Organizations such as the Universal Postal Union, the International Telecommunications Union, the International Organization for Standardization, the Defense Intelligence Agency and several others are involved in this standardization and geocoding activity. The geocodes described in this appendix include both those international geocoding systems currently operational and those which have been proposed for use as standards on an intra or inter agency basis.

C.1 Defense Intelligence Agency (DIA)

As the Assigned Responsible Agency (ARA) for the standardization of geopolitical data elements and related features in the Department of Defense Standardization Program (See Section 3.3) the Defense Intelligence Agency has developed an extensive set of international geocodes. There are standard DIA coding systems for:

> <u>Celestial bodies of the solar system</u> which includes any of the planets, satellites, asteroids, captured comets, meteor swarms, and other natural physical bodies held by the gravitational system of the sun and revolving around it
>
> <u>Terrestrial hemispheres</u> defined as that one-half of the globe divided by the equator into north and south latitudes, and by the prime meridian into east and west longitudes

Divisions of the world which include the seven major land masses (continents) and nine special purpose groupings of water and islands

Water bodies of the world which are continuous water areas representing separate or self-contained bays, seas, and straits

Countries of the world representing first order global geo-political subdivisions and their dependencies, areas of quasi-independence, and areas with special sovereignty associations; unrecognized political regimes and administrative divisions without sovereignty; and outlying areas of the United States, including islands in dispute

States and provinces of the world which represent first order geopolitical subdivisions of each entity coded as a country of the world

These standards were developed in order to resolve the differences between the various levels of aggregation, the various definitions, the various terminology and various data codes previously employed in Department of Defense information systems.

The two character, partially mnemonic alpha designations established by DIA for countries of the world has been published as a Federal Information Processing Standard (FIPS PUB 10), however as stated in the introduction to that publication

> The code scheme used in this standard may not be permanent. The work yet to be done at the American National Standards and International Organization for Standards levels may make it advisable to modify this federal standard. The use of this code in the interim is advised to facilitate interchange for an unknown period of time, and to allow an orderly conversion to a new code scheme if necessary.

The DIA codes for states and provinces of countries of the world are currently being considered for approval as a FIPS standard. In order

---

*U.S Department of Commerce, National Bureau of Standards, Federal Information Processing Standards Publications, FIPS PUB 10 (Washington, D.C.: U.S. Government Printing Office, 1970), p.2.

to provide unique identification with a minimum length code, these state and province codes are dependent upon the DIA country codes. They combine the two character alpha country designations with two digit numbers assigned serially to geopolitical subdivisions sequenced alphabetically by country.[1]

C.2 American National Standards Institute (ANSI)

The coding system proposed by ANSI in 1971 for the identification of countries, dependencies, and areas of special sovereignty of the world is actually a double system--one set of three character alpha-mnemonic designations and one set of three digit numeric designations.

The list of entities generated by ANSI for this standard includes all entities which appear in at least two of the following three sources: FIPS PUB 10, the United Nations Standard Country Codes, and the Universal Postal Union's International List of Post Offices. Alpha codes assigned to these entities are generated according to a set of rules designed to maximize the mnemonic quality of the code in relation to the full name of the entity. The three digit numbers assigned to countries are randomly generated. These codes have not yet become an official ANSI standard. Their acceptance at ANSI may well depend upon the action taken by the International Organization for Standardization to which this standard will also be submitted for approval.

---

[1] In New Zealand where there are over 99 geopolitical subdivisions, the DIA code goes from NZ99 to NZZ2 through NZZ9 and then NZB1 through NZB3.

C.3 General Services Administration (GSA)

In the 1972 edition of the General Services Administration's *Worldwide Geographical Location Codes* outlying areas of the United States, and foreign countries are listed in alphabetical sequence and are identified by a two character alpha code and a three digit numeric code. The two character alpha code which is derived from FIPS PUB 10 in combination with the three digit numeric code is used to comply with GSA Real Property Owned and Leased Inventory reporting requirements. All other interchange of data between agencies of the General Services Administration requires only the two character alpha code for identifying a country or an outlying area of the United States.

The two digit numerical prefix indicating continent which was formerly used in conjunction with the GSA country code has been abandoned. However, the GSA four digit numeric code for cities is still maintained and, in combination with the country codes, provides unique identification for approximately 2,000 cities outside the United States.

C.4 International Telecommunications Union (ITU)

The International Telecommunications Union is a specialized agency of the United Nations. It's major function is to "maintain and extend international co-operation for the improvement and rational use of telecommunications of all kinds."[1] Most of the work accomplished by ITU is carried out by its two international consultative committees, one of

---

[1] Gerd D. Wallenstein, "International Telecommunications: Where Co-Operation is the Message," *Telecommunication Journal* (Volume 39, 1972), p. 366.

which is concerned with radio communications, the other of which deals with non-radio aspects of telephone and telegraph communications.

One set of international geocodes employed by ITU and its members is a variable one to three character alpha code which is assigned to all countries and a certain number of special locations such as space research stations and stations of the United Nations Truce Supervision Organization in areas between the Armistice Demarcation Lines at Jerusalem. This code is used primarily for administrative purposes and it appears on all ITU service documents.

A second set of international geocodes sponsored by ITU and well-established in the world community with only one exception is the Telex Destination Codes.[1] This is a set of variable two or three digit numbers assigned to countries of the world for the purpose of telegraph communications. The nations are listed by geographic proximity and thus the first digit of this code indicates a world area.

A third series of international codes sponsored by ITU is the variable one to three digit number known as the international telephone code

---

[1]Only one of the major international telegraph companies does not subscribe to the Telex Destination coding scheme. When ITU was considering draft standards for their telex designation codes, International Telephone and Telegraph (ITT) submitted the list of codes it had developed independently and was currently using. This set of ITT codes was rejected in favor of the standard eventually adopted by ITU. However, ITT did not change their codes in accordance with the new standard. Thus ITT maintains a set of international telex destination codes for transmission within the United States. At the point where a message is transmitted to an overseas operator the code is translated into the ITU destination indicator and then forwarded.

which is assigned to individual countries grouped into nine world zones. This code is currently accepted by all members of the telephone industry with the exception of the American Telephone and Telegraph Company and the telephone company of Sweden.

C.5 Universal Postal Union (UPU)

The United States Postal Service has developed a set of international geocodes which would facilitate the sorting and dispatch of international mail and submitted this system to the Universal Postal Union for approval.

The draft standard which the U.S. Postal Service has proposed is a three digit number. The first digit identifies the continent or ocean area and is assigned in a north-south pattern beginning at the International Date Line and moving west. The last two digits indicate a country's or territory's position within the continent or ocean area. Numbering begins at the northern corner of the continent or ocean areas and individual countries are assigned codes in an east-west sequence. The countries identified by this code include all member countries of the Universal Postal Union and the entire list of entities included in the United Nations Country Code.

In the expository remarks appended to this draft standard the rationale for such a coding scheme is explained in the following statement:

> One important purpose of the code is the facilitation of the sorting or international mail. Since it appears at this time impossible to construct a code based on

> clearly defined, mail related variables (such as volume flow between the various member countries of the Union) due primarily to the lack of data, the best possibility for realizing this purpose seems to lie with a code that indicates to postal employees the distances and direction of countries of destination. That is, that the code be 'geographical'.[1]

C.6 International Standard Point Location Code

At a meeting sponsored by the Transportation Data Coordinating Committee (TDCC), a representative group of Standard Point Location Code (SPLC) users discussed the expansion of SPLC to include international locations. It was decided that "the goal of an international SPLC, accepted and usable by all major shipping countries, will be the long term goal of the committee . . . however, an interim code should be established which would be beneficial to United States carriers and TDCC membership as an immediate goal in lieu of attempting to establish an internationally accepted code at the outset which might take three to five years."[2]

This proposed interim code would be structured in the following manner: a three digit numeric code prefix containing the country designations derived from Schedule C of the U.S. Department of Commerce's foreign trade program (See Section 4.7) would be appended to the six digit SPLC. The first of the six digit SPLC positions would contain

---

[1]U.S. Postal Service, "Proposed Universal Postal Union Numerical Country Identification Code," Second Draft, 1972, p. 15.

[2]Transportation Data Coordinating Committee, "Summary Minutes of the First International Geographic Subcommittee meeting," 1972, p. 1.

the number 9 (the numbers 1 through 8 in this position indicate a region of the United States) and the remaining five digits would contain a city code.

The proposed long term international code would be established by adopting the coding method of each country for locations in that country in place of the interim SPLC. The three digit prefix would contain either the Schedule C country designations or the codes sponsored by the International Organization for Standardization.

C.7 The United Nations

The basic set of international geocodes developed and maintained by the Statistical Office of the United Nations is a series of three digit numbers gapped by intervals of four and assigned to countries and other first order geopolitical subdivisions of the world sequenced alphabetically by English language names. The list of national entities compiled by the United Nations for this geocoding system is one of the most comprehensive and frequently updated of any international geocoding system. Many other agencies such as ANSI and FIPS use this list as a standard reference.

In addition to the basic three digit country code, the United Nations has developed various code prefixes and suffixes to provide for certain levels of data aggregation or disaggregation. For example, a two digit prefix is used to classify world trade areas in which the first digit identifies a continent or global region and the second digit identifies a subdivision of the region which constitutes a multi-country trade area. Furthermore, basic country codes ranging from 900 to 999 have been reserved to provide for special purposes depending upon the user's needs. For many purposes, geocodes for nation geopolitical subdivisions are required. In such cases a two digit suffix can be added to the right of the three digit country code to identify, for example, the Channel Islands (826 01) as part of the United Kingdom (826).[1]

The three digit numerical country codes developed by the United Nations are employed for data compilation as well as clerical purposes. Data compilation activity at the United Nations includes numerous field studies and special topical studies and the preparation of a general Statistical Yearbook. This yearbook contains a prodigious amount of socio-economic data by country. National statistics for each country are submitted directly by individual countries or territories in a questionnaire supplied by the United Nations or gathered by means of published documents. The information gathered in this fashion includes data on: population, housing, medical care, social security, manpower, agriculture, forestry, fishing, mining, manufacturing, construction, fuel comsumption, produce consumption, transportation, communications, trade, economic aid, wages and prices, national income, education and culture.

---

[1]For designated geopolitical subdivisions of New Zealand (the only country with more than 99 first order subdivisions) a three digit numerical suffix would be required.

The United Nations also maintains, updates and publishes six sets of country name abbreviations composed of 4, 8, and 12 character representations in both the English and French languages.

C.8 The Agency for International Development (AID)

The Agency for International Development maintains a standard set of international geocodes for use on commodity procurement instructions and other AID forms to indicate the country of destination or source for commodity purchases. The first digit of the country code is used to identify the global region in which the country is located. The groupings were delineated for administrative and statistical purposes, and any political associations or affiliations are neither intended nor implied. The regional codes are as follows:

0 - United States, and areas of U.S. associated sovereignty	5 - Latin America
1 - Europe	6 - Africa
2 - Near East	7
3 - South Asia	8 - Viet Nam and Other
4 - East Asia	9

The second and third digits in this three digit AID country code identify the specific geographic entities, countries, or territories within each region. Important islands or island groups, at considerable distance from, but forming an integral part of some mainland entity, and mainland portions of an entity are listed after the name of the entity. Minor islands or island groups are not listed unless they form separate entities.

Additional codes are assigned for administrative purposes which do not refer to either a single country or region. For example, a separate code has been assigned to middle east oil areas for limited use when reference to petroleum purchases in that area is required.

C.9 Other International Geopolitical Coding Systems

The Federal Reserve Bank and the U.S. Treasury Department employs a four digit numeric code assigned to countries sequence alphabetically with an additional check digit to identify nations of the world. This code is applied for the collection and tabulation of balance of payment statistics.

The Bank of America employs a three digit numeric code to identify continents and countries of the world. This code is applied for record keeping purposes.

The Export Import Bank employs a three digit numeric code assigned to countries sequenced alphabetically for record keeping purposes.

Rand McNally and Company employs a three digit numeric code assigned to countries of the world. An additional fourth numeric digit identifies continent. Rand McNally also maintains an international city code. These codes are used in the preparation of the atlases and maps produced by Rand McNally.

The International Standards Organization (ISO) has adopted an international automobile registry code which consists of a variable one to three character alphabetic designation used to identify countries in which automobiles are registered.

The International Civil Aviation Organization (ICAO) has established a set of four character airport designations (including navigational aids and charted waypoints) which are used in all international flight telecommunications.

In addition to the various sets of geographic references for international geopolitical entities there are a number of global mapping systems. The most extensive and detailed of these is compiled by the Central Intelligence Agency (CIA). This automated CIA mapping system is being refined at the Laboratory for Computer Graphics and Spatial Analysis, Harvard University where an assortment of programs are being developed in order to enhance the applications of the basic system and provide several levels of geographic detail on a global scale.

References:

Agency for International Development. Bureau for Program and Policy Coordination. Office of Statistics and Reports. A.I.D. Geographic Code Book. M. O. 302.1 Revised. Washington, D.C.: Agency for International Development, 1970.

Defense Intelligence Agency. Geopolitical Data Elements and Related Features. Manual number 65-18A. Washington, D.C.: Defense Intelligence Agency, 1973.

General Services Administration. Office of Finance. Worldwide Geographical Location Codes. Washington, D.C.: General Services Administration, .972.

Union Internationale Des Telecommunications. "Plan et Methods de Numerotage Nationaux et le Plan de Numerotage Internationale." Avis E. 160. Genève, Suisse: Union Internationale Des Telecommunications, 1972.

United Nations. Department of Economic and Social Affairs. Statistical Office of the United Nations. United Nations Standard Country Code. Statistical Papers Series M, Number 49. New York, N.Y.: United Nations, 1970.

U.S. Department of Commerce. National Bureau of Standards. Federal Information Processing Standards Publications, FIPS PUB 10. Washington, D.C.: U.S. Government Printing Office, 1970.

Universal Postal Union. "Proposed Universal Postal Union Numeric Country Identification Code." Second Draft. Washington, D.C.: Universal Postal Union, 1971.

APPENDIX D

BIBLIOGRAPHY

Agency for International Development. Bureau for Program and Policy Coordination. Office of Statistics and Reports. A.I.D. Geographic Code Book. M. O. 302.1 Revised. Washington, D.C.: Agency for International Development, 1970.

American National Standards Institute. Task Group X3L84. "Meeting Minutes," 1970-1974.

American Telephone and Telegraph Company. Long Lines Department. Administration of Rates and Tariffs. "Traffic F.C.C. Publication No. 255." New York: American Telephone and Telegraph Company, 1968.

Brokke, Glenn E. "Nationwide Highway Travel." A paper prepared for the Western Association of State Highway Officials Highway Planning Conference, June 9, 1966, Santa Fe, New Mexico. Washington, D.C.: Bureau of Public Roads, U.S. Department of Commerce, 1966.

Canadian Department of Industry, Trade, and Commerce. Statistics Canada. GRDSR: Facts by Small Areas. Ottawa, Canada: Statistics Canada, Canadian Department of Industry, Trade, and Commerce, 1972.

Canadian Department of Industry, Trade and Commerce. Statistics Canada. "Standard Geographical Classification Manual." Draft text of introduction. Ottawa, Canada: Statistics Canada, Canadian Department of Industry, Trade, and Commerce, 1974.

Canadian Post Office Department. Canada Post. "Canadian Postal Code." Ottawa, Canada: Canada Post, Canadian Post Office Department, 1973.

Church, Donald E. "PICADAD: A System for Machine Processing of Geographic and Distance Factors in Transportation and Marketing Data." Washington, D.C.: Transportation Division, U.S. Bureau of the Census, August 1970 (Mr. Church was formerly Chief of the Transportation Division, U.S. Bureau of the Census).

Clark, Frank E. A Treatise on the Law of Surveying and Boundaries, 2nd ed. Indianapolis: The Bobbs-Merrill Company, 1939.

Creighton, Reginald A. and Piacesi, Dante. "An Approach to the Geography Problems in Museums." Smithsonian Institution Information Systems Innovations. Vol. II, No. 1. Washington, D.C.: Smithsonian Institution, 1970.

Creighton, Reginald A. and Crockett, James J. "Selgem: A System for Collection Management." Smithsonian Institution Information Systems Innovations. Vol. II, No. 3. Washington, D.C.: Smithsonian Institution, 1971.

Creighton, Reginald A.; Packard, Penelope; and Linn, Holley. "Selgem Retrieval: A General Description." Smithsonian Institution Procedures in Computer Sciences. Vol. I, No. 1. Washington, D.C.: Smithsonian Institution, 1972.

Defense Intelligence Agency. Geopolitical Data Elements and Related Features. Manual number 65-18A. Washington, D.C.: Defense Intelligence Agency, 1973.

Dun and Bradstreet, Inc. Marketing Services Division. County Code Book. Standard Metropolitan Statistical Area Code Book. City Code Book. Magnetic Tape Description. New York: Dun and Bradstreet, Inc., 1971.

Eckler, A. Ross. The Bureau of the Census. New York: Praeger Publishers, 1972 (Mr. Eckler was Director of the Bureau of the Census, 1965-1969).

Executive Office of the President. Bureau of the Budget. Standard Metropolitan Statistical Areas. Washington, D.C.: U.S. Government Printing Office, 1967.

Executive Office of the President. Office of Management and Budget. "Publication of Statistical Data for Regions." OMB Circular A-46. Washington, D.C.: Office of Management and Budget, Executive Office of the President, 1971.

Fajnor, John P. International Business Machines Corporation, Gaithersburg, Maryland. Conversations and correspondance, 1974.

Forstall, Richard L. "The American National Standard Code for Named Populated Places and Related Entities," The National Geocoding Conference Proceedings. Washington, D.C.: U.S. Department of Transportation, 1972.

General Services Administration. Office of Finance. Worldwide Geographical Location Codes. Washington, D.C.: General Services Administration, 1972.

General Services Administration. National Archives and Records Service. Office of the Federal Register. United States Government Organization Manual 1972/1973. Washington, D.C.: U.S. Government Printing Office, 1973.

Gould, Sydney W. Geo-Code. Volume I, West Edition. New Haven, Conn.: The Gould Fund, 1968.

International Business Machines Corporation. Data Processing Division. Numerical Codes for States, Counties and Cities of the United States. GC 20-8073-0. White Plains, N.Y.: IBM Technical Publications Department, 1961.

Jones, Richard W. "New Developments in ZIP Code Marketing." A reprint of an article which appeared in Sales Management Magazine, June 10, 1969.

Kallab, Jiri. "Linear Geographical Code for Management Information Systems." New York, N.Y.: Artronic Information Systems, Inc., 1972.

McIntyre, Harris B. "What is this New V-H System?" Background for Management No. 33 (December 1959).

National Motor Freight Traffic Association, Inc. "Statement of Work for the Creation and Maintenance of a Standard Point Location Code Master File." Washington, D.C.: National Motor Freight Traffic Association, Inc., 1973.

Office of Emergency Preparedness. Geographic Codes for Region, State, and County. ISG-111. Washington, D.C.: Office of Emergency Preparedness, 1971.

Petersen, Dietrich L. and Schmidt, Leo A. Jr. Arrangements of U.S. Population by Urban and Rural Geometrical Clusters. Research Paper P-706 prepared for the Office of Civil Defense. Arlington, Va.: Institute for Defense Analyses, Program Analysis Division, 1970.

Sachs, Abner and Timmermans, Judith A. Economic Structure of the United States Using the County as a Functional Base. Research Paper P-511 prepared for the Office of Civil Defense. Arlington, Va.: Institute for Defense Analyses, Program Analysis Division, 1969.

Samson, Belair, Riddell, Stead, Inc. A Canadian Public Address Postal Coding System. A consultant's report prepared for Canada Post. Ottawa, Canada: Canada Post, Canadian Post Office Department, November 1969.

Strahler, Arthur N. Physical Geography. 2nd ed. New York: John Wiley and Sons, Inc., 1962.

Timmermans, Judith A. User's Guide to Economic Data and Programs for IDA Civil Defense Studies. N-655(R). Arlington, Va.: Institute for Defense Analyses, Programs Analysis Division, 1970.

Transportation Data Coordinating Committee. _Standard Point Location Code Congress: Addresses and Panel Discussion_. Washington, D.C.: Transportation Data Coordinating Committee, 1970.

Trewartha, G.T.; Robinson, A.H.; and Hammond, E.H. _Fundamentals of Physical Geography_. New York: McGraw Hill Book Company, 1961.

Twedt, Dik. "Status of 'ZIP' as a Tool of Marketing Research." _Proceedings_ of the Social Statistics Section, American Statistical Association. Washington, D.C.: American Statistical Association, 1967.

Union Internationale Des Telecommunications. "Plan et Methods de Numerotage Nationaux et le Plan de Numérotage Internationale." Avis E.160. Genève, Suisse: Union Internationale Des Telecommunications, 1972.

U.S. Department of the Army. Office of the Chief of Engineers. "Reports and Statistics: Waterborne Commerce Statistics." Regulation No. 335-2-1, ENGCW-OM. New Orleans, Louisiana: U.S. Army Corps of Engineers, 1970.

United Nations Department of Economic and Social Affairs. Statistical Office of the United Nations. _United Nations Standard Country Code_. Statistical Papers Series M, Number 49. New York, N.Y.: United Nations, 1970.

U.S. Department of the Army, and U.S. Department of the Air Force. _The Universal Grid System_. TM5-241 to 16-1-233. Washington, D.C.: U.S. Government Printing Office, 1951.

U.S. Department of Commerce. Bureau of the Census. _National Location Code: OCP-OEP Region 2_. FG-D-3 1/2. Prepared for the Office of Civil Defense, Department of Defense, and the National Resource Evaluation Center, Office of Emergency Planning. Washington, D.C.: U.S. Government Printing Office, 1964.

________. 1970 Census Users' Guide, Part 1. Washington, D.C.: U.S. Government Printing Office, 1970.

________. "1970 Census Summary Tape User Memorandums." Washington, D.C.: U.S. Bureau of the Census, 1969-1972.

U.S. Department of Commerce. Bureau of the Census. Census Use Study. _The DIME Geocoding System: Census Use Study Report No. 4_. Washington, D.C.: U.S. Government Printing Office, 1970.

________. _DIME Workshops: An Interim Report_. Washington, D.C.: U.S. Government Printing Office, 1973.

U.S. Department of Commerce. Bureau of the Census. Foreign Trade Division. Guide to Foreign Trade Statistics: 1970. Washington, D.C.: U.S. Government Printing Office, 1970.

________. Classification of Customs Districts and Ports: Schedule D and Classification of Country Designations: Schedule C. Washington, D.C.: U.S. Government Printing Office, 1970.

U.S. Department of Commerce. Bureau of the Census. Geography Division. Use of Address Coding Guides in Geographic Coding: Case Studies. Conference Proceedings, November 19-20, Wichita, Kansas. Washington, D.C.: U.S. Government Printing Office, 1970.

________. Geographic Base Files: Plans, Progress and Prospects. Conference Proceedings, April 1-2, Jacksonville, Florida. Washington, D.C.: U.S. Government Printing Office, 1971.

________. Geographic Base File System: Uses, Maintenance, Problem Solving. Conference Proceedings, November 16-17, Arlington, Texas. Washington, D.C.: U.S. Government Printing Office, 1971.

________. Geographic Base File System: Establishing a Continuing Program. Conference Proceedings, January 18-19, Seattle, Washington. Washington, D.C.: U.S. Government Printing Office, 1973.

________. Geographic Base File Systems: A Forward Look. Conference Proceedings, April 16-17, Boston, Massachusetts. Washington, D.C.: U.S. Government Printing Office, 1974.

________. "Introduction to Small Area Geographic Subdivisions for Which the U.S. Bureau of the Census Collects and Tabulates Data." Prepared by Valerie McFarland and Ann D. Casey of the Data User Services Office. Washington, D.C.: Geography Division, U.S. Bureau of the Census, 1972.

U.S. Department of Commerce. Bureau of the Census. Transportation Division. PICADAD: A Computer File of Places with Geographic Characteristics and A Straight-Line Distance Computation Method. Public Use Tape Users' Manual. Washington, D.C.: Transportation Division, U.S. Bureau of the Census, March 1974.

U.S. Department of Commerce. Coast and Geodetic Survey. The State Coordinate Systems: A Manual for Surveyors. Special Publication 235. Washington, D.C.: U.S. Government Printing Office, 1945.

U.S. Department of Commerce. Economic Development Administration. "Geographic Directory." An unpublished, inhouse document. Washington, D.C.: Economic Development Administration, U.S. Department of Commerce, 1972.

U.S. Department of Commerce. Economic Development Administration. Economic Development Districts: Composition and Status. Washington, D.C.: U.S. Government Printing Office, 1970.

U.S. Department of Commerce. Interstate Commerce Commission. "Interstate Commerce Commission: Freight Rate Territory Classification Used in Waybill Analysis." Washington, D.C.: Interstate Commerce Commission, U.S. Department of Commerce, 1968.

U.S. Department of Commerce. National Bureau of Standards. Federal Information Processing Standards Publications. FIPS PUB Nos. 0, 5, 6, 8, 9, 10, and 12. Washington, D.C.: U.S. Government Printing Office, 1968-1974.

U.S. Department of Defense. Office of the Assistant Secretary of Defense (Comptroller). Department of Defense Manual for Standard Data Elements. 5000.12M. Washington, D.C.: Assistant Secretary of Defense (Comptroller), U.S. Department of Defense, 1970.

U.S. Department of Transportation. Rail Service in the Midwest and Northeast Region. A report by the Secretary of Transportation. Washington, D.C.: U.S. Government Printing Office, 1974.

U.S. Department of Transportation. Federal Aviation Administration. Air Traffic Service. "Air Traffic Operational Coding System." Order 7350.2B. Washington, D.C.: Federal Aviation Administration, U.S. Department of Transportation, January 1974.

________. Location Identifiers. Order 7350.4. Washington, D.C.: U.S. Government Printing Office, January 1974.

U.S. Department of Transportation. Federal Highway Administration. Urban Transportation Planning: General Information and Introduction to System/360. Washington, D.C.: Federal Highway Administration, U.S. Department of Transportation, June 1970.

U.S. Department of Transportation. Federal Railroad Administration. Memorandum to the Department of Transportation Data Standardization Group. March 1974. Washington, D.C.: Federal Railroad Administration, U.S. Department of Transportation, 1974.

U.S. Department of Transportation. Office of Systems Analysis and Information. Carload Waybill Statistics 1969. Washington, D.C.: Office of Systems Analysis and Information, U.S. Department of Transportation, 1970.

U.S. Postal Service, Customer Services Group. National ZIP Code Directory. Washington, D.C.: U.S. Government Printing Office, 1973.

Universal Postal Union. "Proposed Universal Postal Union Numeric Country Identification Code." Second Draft. Washington, D.C.: Universal Postal Union, 1971.

Walker, Henry D. "Report on Work of Subcommittee for Codes for Point Locations in the United States." A draft report presented to ANSI Task Group X3L84. Washington, D.C.: American National Standards Institute, 1971.

Water Resources Council. 1972 OBERS Projections: _Regional Economic Activity in the United States_. Volume 1: _Concepts, Methodology, and Summary Data_. Washington, D.C.: U.S. Government Printing Office, 1972.

Water Resources Council. _Water Resources Regions and Subregions for the National Assessment of Water and Related Land Resources_. Washington, D.C.: Water Resources Council, 1970.

________. _The Nation's Water Resources_. Washington, D.C.: Water Resources Council, 1968.

APPENDIX E

REPORT OF INVENTIONS

A diligent review of the work performed under this contract has revealed no new innovation, discovery, improvement or invention.

☆U.S. GOVERNMENT PRINTING OFFICE: 1975—601-399