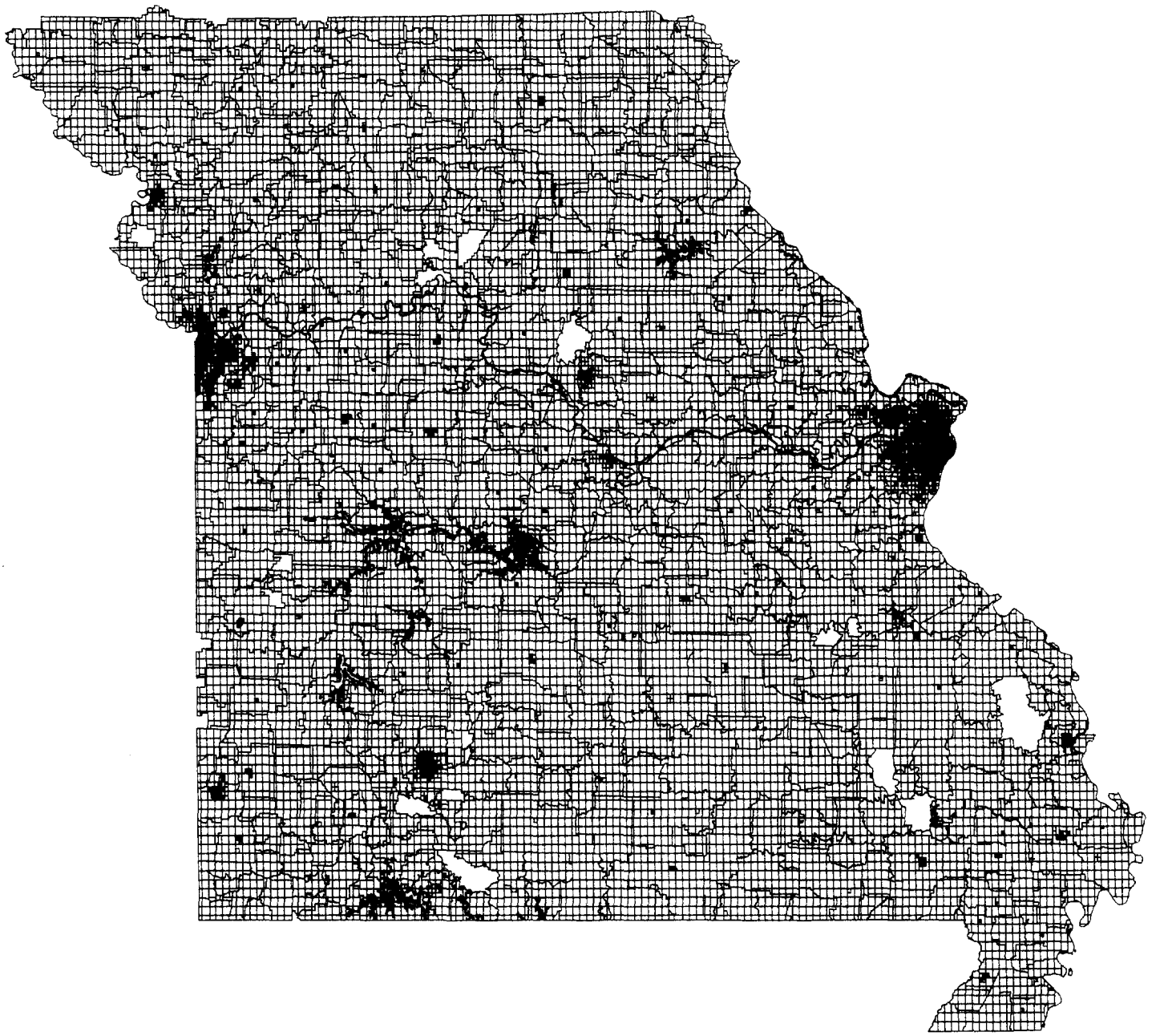
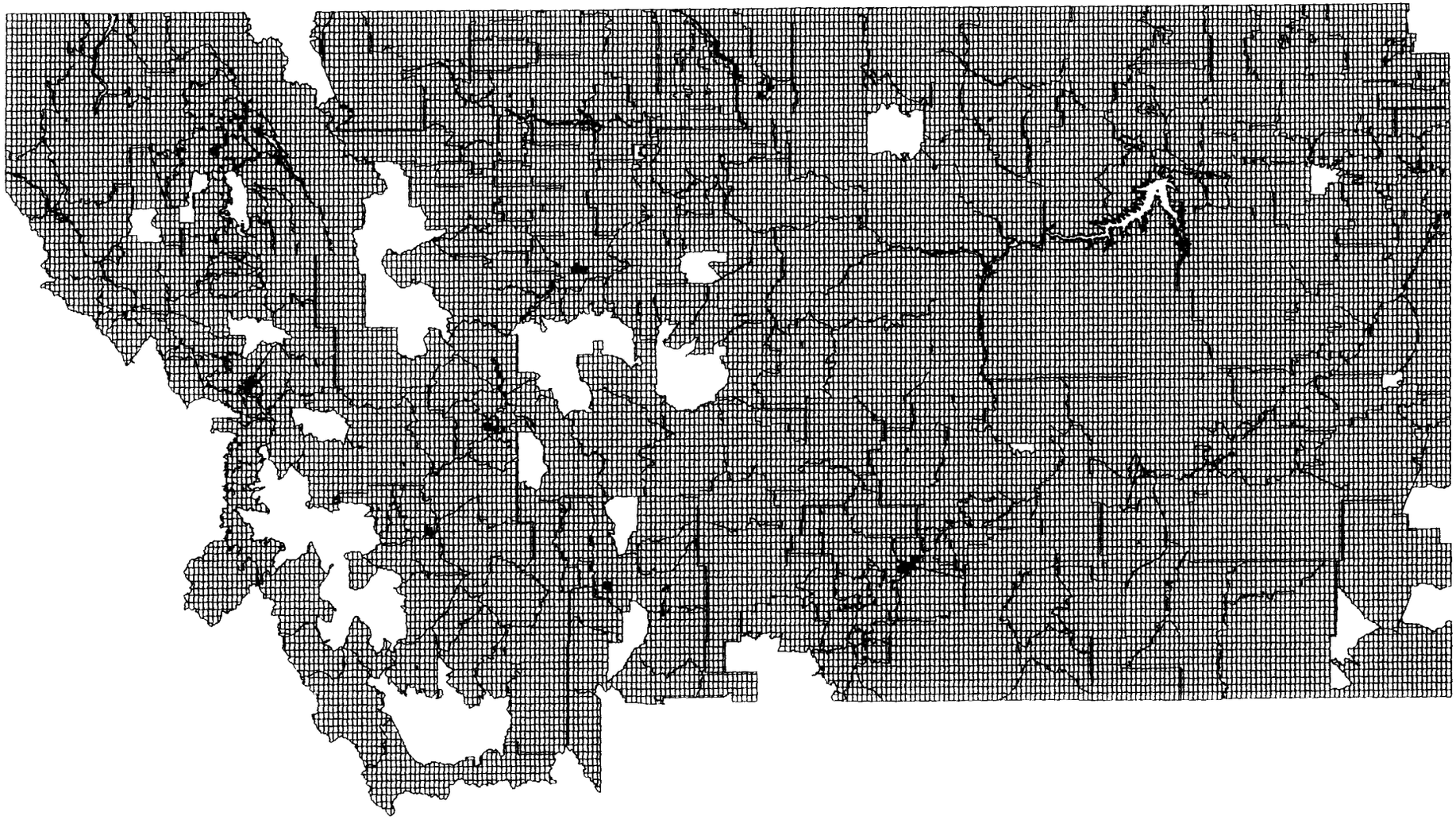
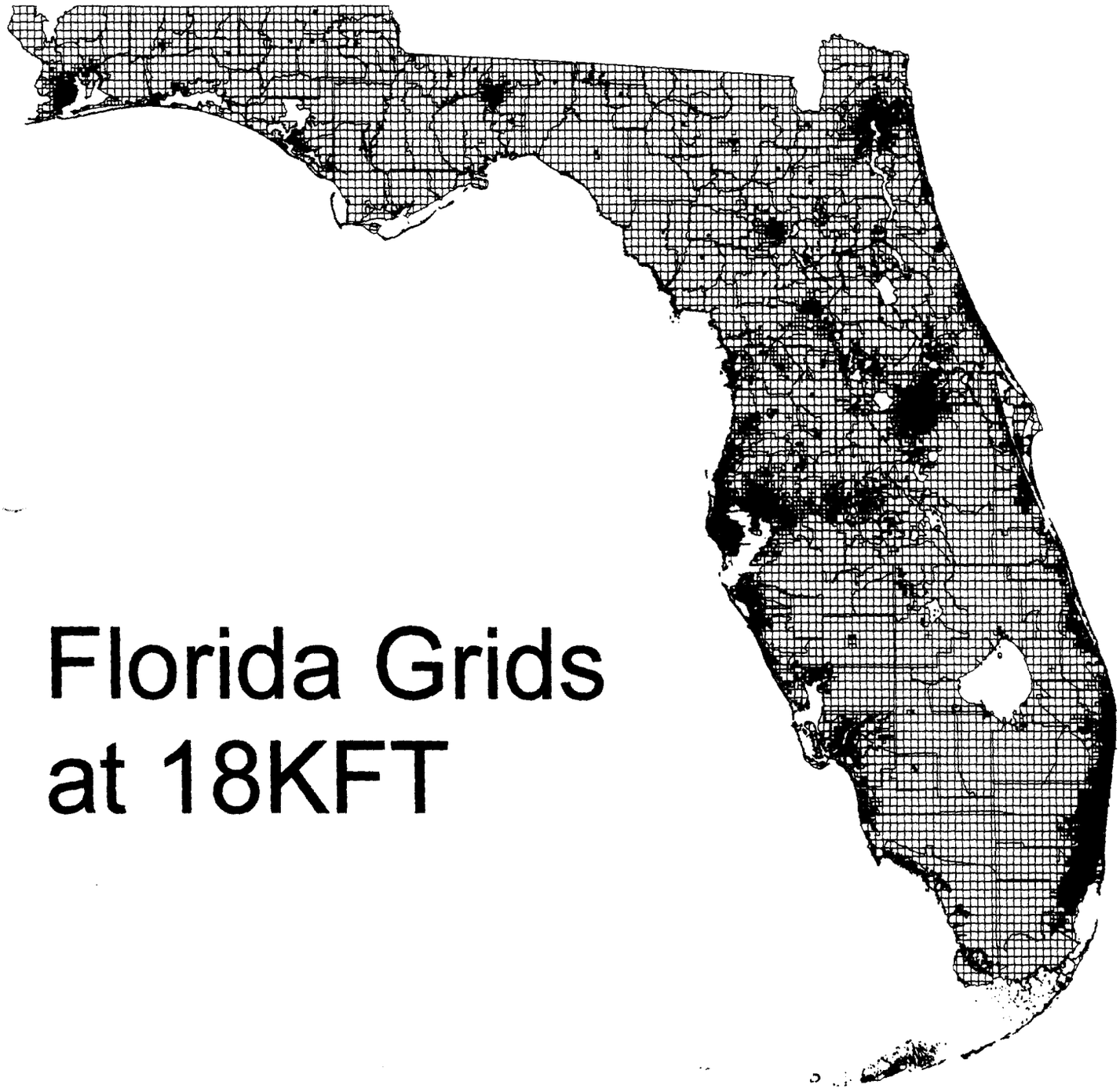
Maryland Grids at 12Kft

Missouri Grids at 12Kft



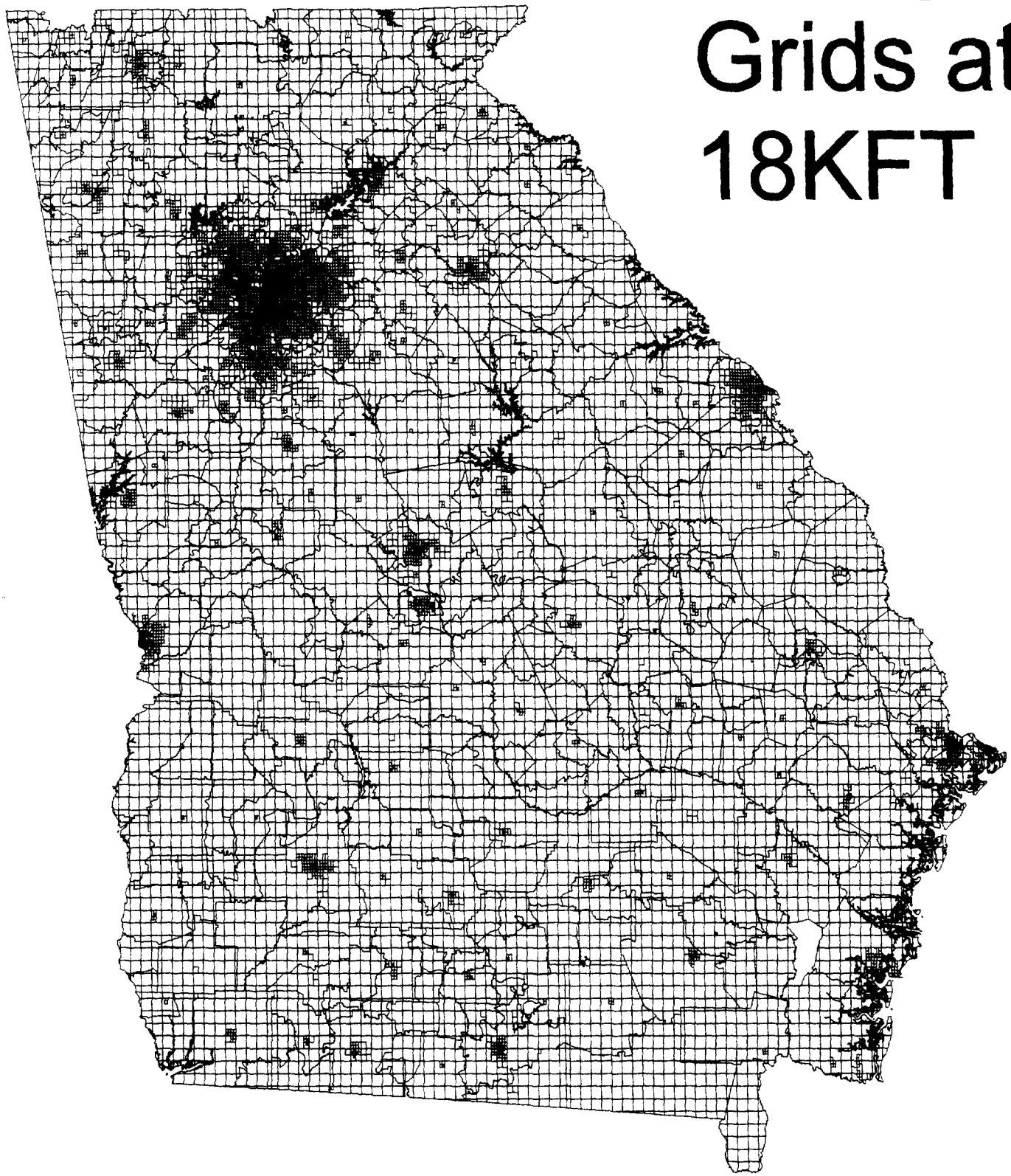
Montana Grids at 12Kft

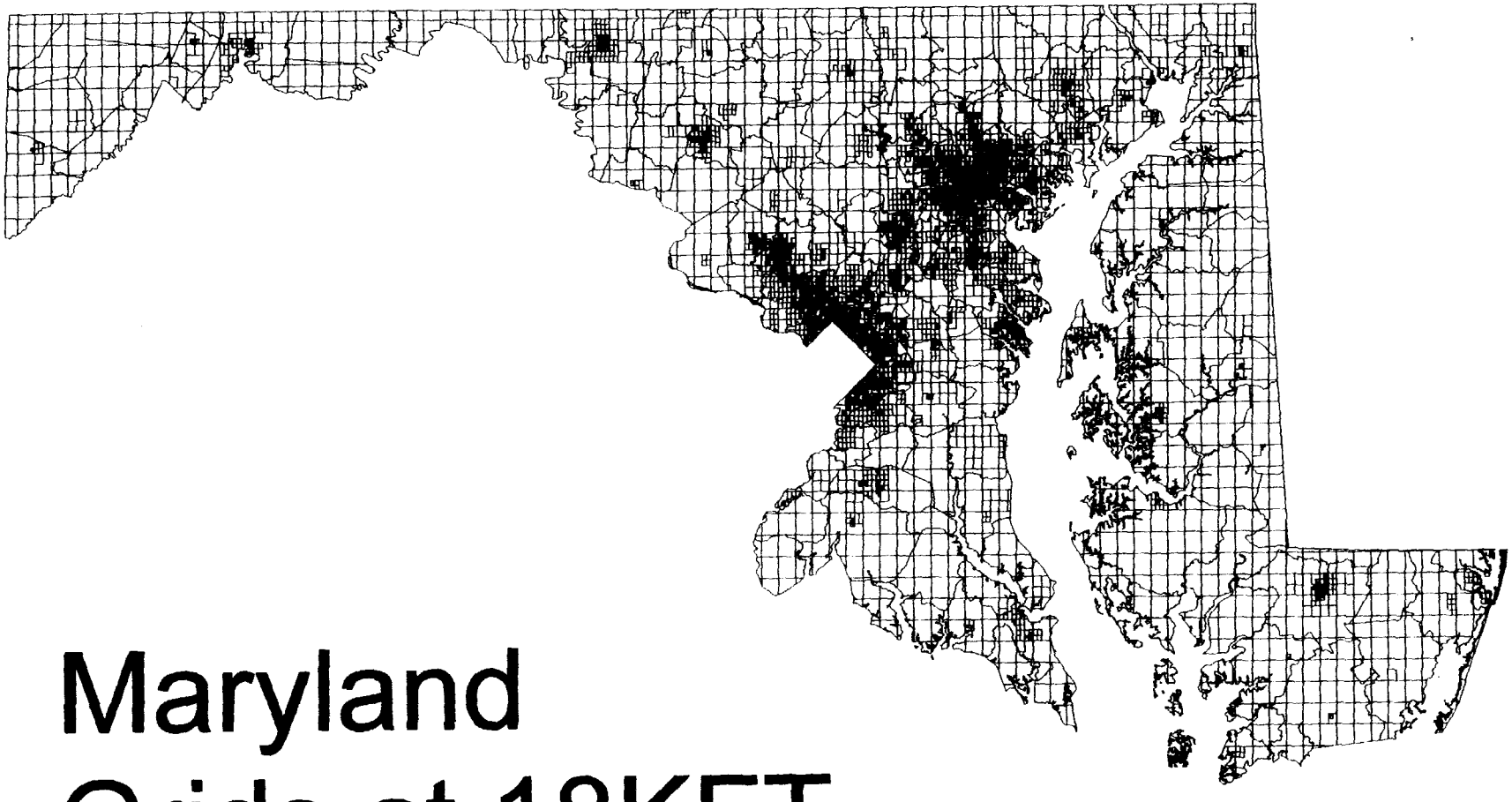




Florida Grids at 18KFT

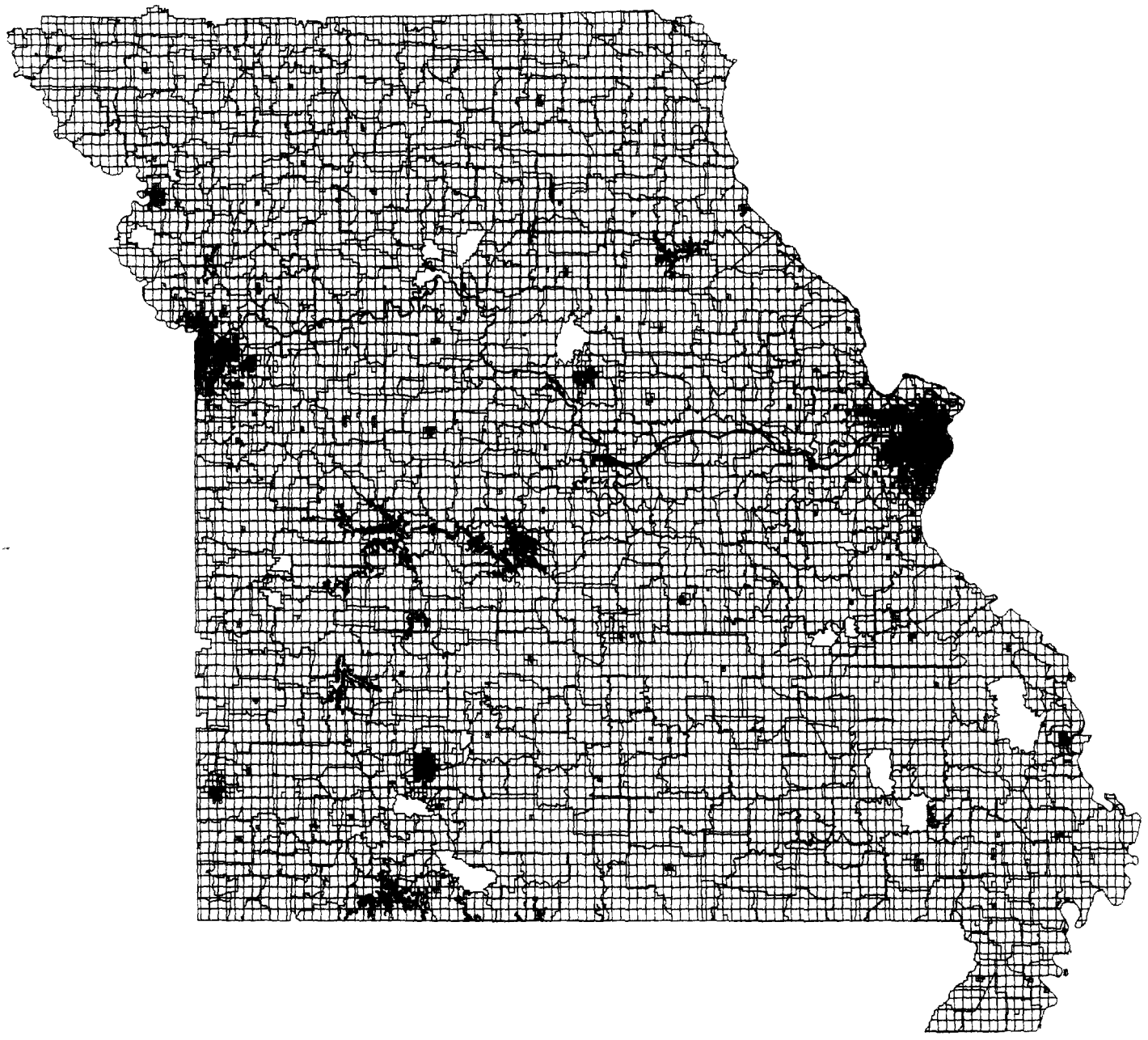
Georgia Grids at 18KFT

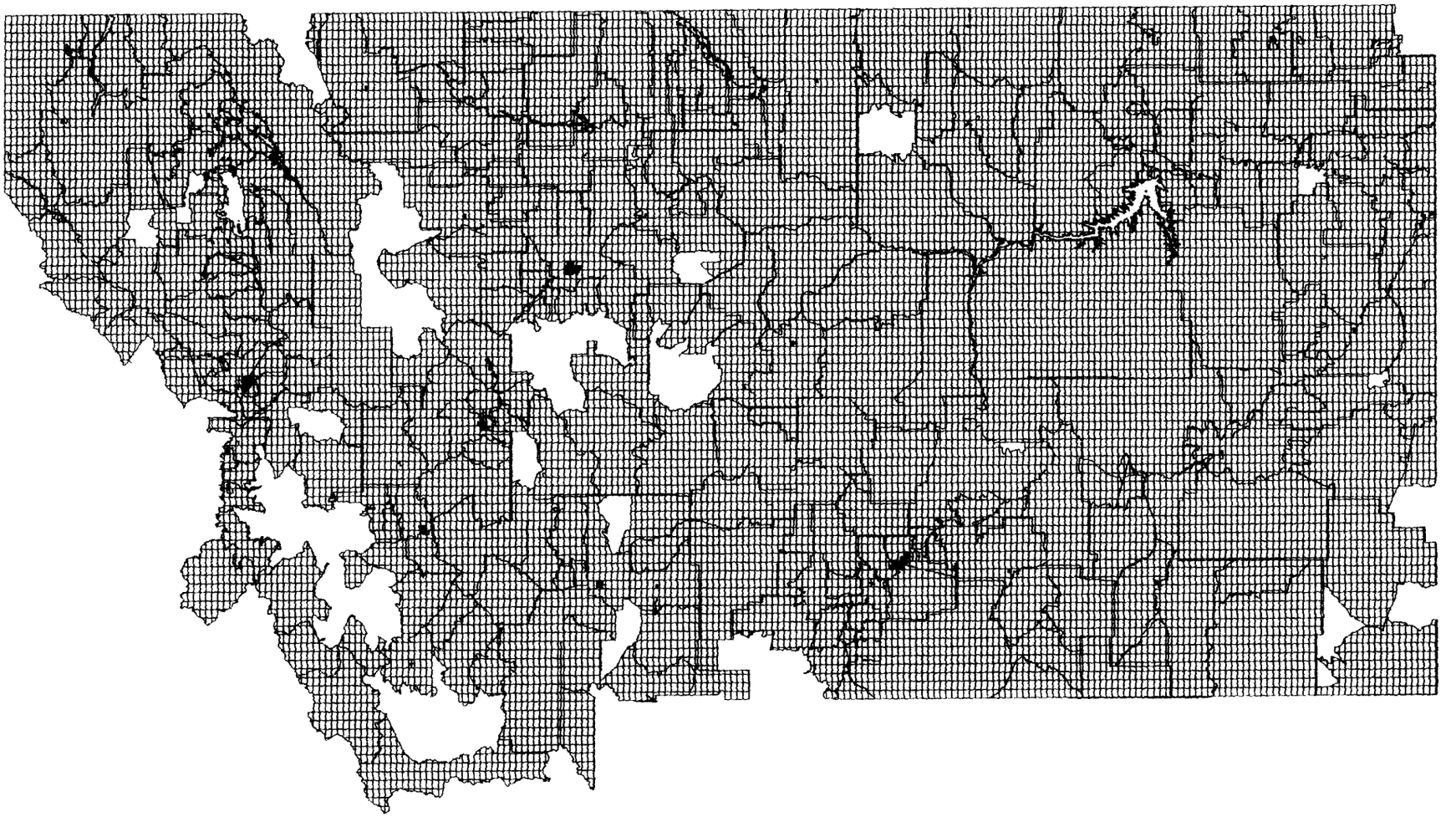




Maryland
Grids at 18KFT

Missouri Grids at 18Kft





Montana Grids at 18KFT

Cardinal®

Geocoding and Hatfield 5.0

The developers and proponents of the Hatfield Model 5.0 tout its use of geocoded customer locations. Indeed, the Joint Sponsors acknowledge the informational benefits of this type of information. However, the current state of geocoding does not allow for the use of geocoded data in cost proxy models that attempt to estimate the cost of universal service across a wide range of geographic densities. In particular, geocoding is only feasible for the dense, urban areas. In the sparsely populated, rural areas, the percentage of customer locations that can be geocoded is simply too low to be of use to a cost proxy model.

The above conclusion is supported by an analysis of geocoding undertaken by the Joint Sponsors. In particular, the Joint Sponsors acquired the customer addresses for the Albany and Vernon wirecenters in Texas and for 8 rural counties in the states of Montana, Utah, North Dakota, and North Carolina. The customer addresses were acquired from Metromail, the source of the customer addresses used by the developers of Hatfield 5.0. In addition, the Joint Sponsors geocoded these locations using the same geocoding software (Centrus) and the same road network data (GDT) used by the Hatfield Model developers. The intent of the analysis was to follow the Hatfield geocoding methodology as closely as possible, given the scant information available on the Hatfield methodology, to determine the share of actual customer locations that could be spatially located.

The primary findings of the Joint Sponsors analysis, described in more detail below, are:

1. The Metromail database of household addresses does <u>not</u> include the address of every housing unit in the country. In fact, the share of total U.S. housing units for which Metromail has an address is <u>smaller</u> that the figure touted by the proponents of the Hatfield Model.
2. An "address" in the Metromail database can be a P.O. Box or Rural Route. These addresses cannot be accurately geocoded. Hence, the number of <u>geocodable</u> Metromail addresses is <u>smaller</u> than the number of addresses in the Metromail database.
3. The Metromail address database contains both urban and rural addresses. The share of housing unit addresses can be substantially smaller in the rural areas than in the urban areas. Moreover, the geocodable share of housing unit addresses in rural areas can be even smaller given the preponderance of P.O Box and Rural Route addresses in rural areas.
4. Since it is more likely that a large share of Census housing units cannot be spatially located in rural areas, the Hatfield 5.0 customer location algorithm is reduced to an arbitrary algorithm, one that simply allocates locations that cannot be geocoded to the perimeter of the Census Block. This arbitrary algorithm is no different from the arbitrary algorithm in version 4.0 that allocated most of the low-density, Census Block Group housing units to 2 to 4 densely packed towns.
5. Since Census Blocks are very small in the urban areas, geocoding of customer locations in urban areas does not add much insight into the cost modeling process.

In the rural areas, because of the low percentage of housing unit locations that can be accurately geocoded, geocoding is also not very useful. Hence, the only value added by the Hatfield Model 5.0, over version 4.0, is its use of Census Block data.

6. Through the use of the existing road network in Census Blocks, BCPM3.0 does a much better job than Hatfield 5.0 of locating customers in the rural areas. Hence, BCPM 3.0 should more accurately estimate the cost of serving rural telephone customers.

Metromail Customer Database

The Metromail National Customer Database is primarily used for direct marketing. It is compiled from the white pages, drivers license data, tax roll information from county courthouse records, and U.S. Census data. The database is updated 65 times per year. As of December 5, 1997, the Metromail database contained 74.4 million named and unnamed address records for the 50 states.[1] The U.S. Census, on the other hand, estimates that in 1995 there were 107.9 housing units. Hence, the Metromail database contains only 69 % of the potential addresses necessary to spatially locate all of the existing housing units in the 50 states. The Hatfield documentation for Preliminary Release 5.0 claims that the Metromail database includes 90 % of the 1995 Census count.[2]

Table 1 shows the 1995 Census housing unit count and the Metromail address count for each of the 50 states. The percentage of housing units accounted for ranges from a high of 85.4 % in Minnesota to a low of 37.4 % in Alaska. The Joint Sponsors stress that not all of the addresses in the Metromail database can be accurately geocoded. Many addresses, particularly in the rural areas, are P.O. Boxes and Rural Routes. P.O. Boxes, for example, can only be geocoded to the Zip Code centroid. In Rural areas, the Zip Code range can be larger than a Census Block Group, which contains many Census Blocks. An indication as to the share of Metromail addresses in the rural areas that are geocodable is given below.

---

[1] Metromail distributes two types of address records. Approximately 90 % of their address records consist of names as well as mailing addresses. Those remaining are known as Ghost records. A Ghost record consists of a mailing address and at times, a phone number, but no customer name. For the purposes of accounting for Hatfield's customer data input set, the Joint Sponsors assumed that Ghost records had been included. The rationale is that customer location is the concern, not customer contact.

[2] Page 6 of the Hatfield Model Preliminary Release 5.0 documentation states, "The file consists of over 100 million households – which constitute over 90 % of the 110 million households that the U.S. Bureau of the Census reported for 1995." For the 50 states, the U.S. Census reports 96.9 million households. The Metromail database has 74.4 million address records. So even using the Census *household* count as the base, Metromail captures 76.8 %, not the 90 % asserted in the Hatfield documentation.

Table 1. Metromail Address Counts for the 50 States.

State	1995 Census Housing Units	Metromail Households	Metromail Percent of Actual
AL	1,751,352	1,136,471	64.89%
AK	254,224	94,984	37.36%
AZ	1,909,847	1,113,805	58.32%
AR	1,056,249	702,468	66.51%
CA	11,873,368	6,502,423	54.76%
CO	1,679,193	1,158,787	69.01%
CT	1,316,253	1,055,291	80.17%
DE	313,506	184,058	58.71%
DC	254,322	207,843	81.72%
FL	6,681,270	4,735,040	70.87%
GA	2,920,754	1,855,926	63.54%
HI	418,487	257,517	61.54%
ID	478,233	335,523	70.16%
IL	4,640,118	3,441,769	74.17%
IN	2,339,232	1,713,996	73.27%
IA	1,167,635	997,394	85.42%
KS	1,078,136	814,087	75.51%
KY	1,571,696	1,105,204	70.32%
LA	1,761,092	1,198,541	68.06%
ME	592,915	461,795	77.89%
MD	1,992,468	1,547,206	77.65%
MA	2,494,157	1,986,744	79.66%
MI	3,951,633	2,816,709	71.28%
MN	1,939,185	1,654,119	85.30%
MS	1,055,318	665,505	63.06%
MO	2,279,194	1,676,534	73.56%
MT	391,798	284,994	72.74%
NB	682,676	557,127	81.61%
NV	659,623	320,856	48.64%
NH	519,916	392,564	75.51%
NJ	3,159,530	2,282,496	72.24%
NM	700,919	396,066	56.51%
NY	7,275,606	5,221,639	71.77%
NC	3,055,821	1,980,185	64.80%
ND	278,278	237,393	85.31%
OH	4,482,349	3,437,918	76.70%
OK	1,460,936	939,910	64.34%
OR	1,315,737	911,509	69.28%
PA	5,010,275	3,767,022	75.19%
RI	409,467	319,129	77.94%
SC	1,497,543	1,011,833	67.57%
SD	304,878	240,613	78.92%
TN	2,173,864	1,476,805	67.93%
TX	7,692,280	4,837,260	62.88%
UT	679,703	493,984	72.68%
VT	280,064	189,221	67.56%
VA	2,658,035	1,815,120	68.29%
WA	2,263,852	1,476,309	65.21%
WV	794,656	516,222	64.96%
WI	2,147,298	1,764,155	82.16%
WY	214,565	149,279	69.57%
Total	107,879,506	74,439,348	69.00%
Maximum			85.42%
Minimum			37.36%

An indication of the shortcoming of the Metromail database in the rural areas is given by Table 2. Table 2 shows the Census housing unit counts and Metromail household address counts for eight rural counties, two from each state of Montana, North Carolina, North Dakota, and Utah. The wide variation in the Metromail "hit rate" is striking, both inter- and intrastate. In Montana, for example, Metromail has potential addresses for only 7.8 % of the housing units in Meagher County. In Rosebud County, the hit rate is higher but is still only 30.9 %.

Again, these addresses include many that cannot be accurately geocoded (i.e., P.O. Boxes and Rural Routes). Hence, the geocodable hit rate can be much smaller than that shown in Table 2. Of the 1,348 Metromail addresses for Rosebud County, for example, 567 or 42 % are P.O. Boxes or Rural Routes. ***Hence, the locations of only 18 % of the Census housing units in Rosebud County, MT can be geocoded to the level of accuracy used by the Hatfield Model developers. In other words, 82 % of the housing units in Rosebud County are subject to the Hatfield Model's "surrogate" algorithm, which simply allocates uncodable housing units to the perimeter of the Census Block.***

Table 2. Metromail Address Counts for Eight Rural Counties.

State	County	1995 Census Housing Units	Metromail Households	Metromail Percent of Actual
MT	Meagher	1,259	98	7.78%
MT	Rosebud	4,358	1,348	30.93%
NC	Halifax	23,136	11,597	50.13%
NC	Wilkes	23,439	19,372	82.65%
ND	Ransom	2,569	786	30.60%
ND	Stark	9,523	8,550	89.78%
UT	Sanpete	7,810	3,776	48.35%
UT	Wasatch	5,161	3,786	73.36%
Total		77,255	49,313	63.83%

Geocoding

To determine the level of accuracy of the Hatfield geocoding the Joint Sponsors acquired from Metromail their addresses for the Albany and Vernon wirecenters in Texas. The Joint Sponsors also acquired from GDT the necessary street data for the geocoding process and from Qualitative Marketing Software, the Centrus Desktop geocoding software. Both GDT and Centrus were used in the Hatfield geocoding methodology. The Joint Sponsors followed, as closely as the limited Hatfield documentation permits, the Hatfield geocoding methodology.

Table 3 shows the Census and Metromail statistics for these two wirecenters. Also shown are the shares of Census housing units that can be geocoded. For the Albany wircenter, the Metromail address count is only 53.5 % of the Census housing unit count. When the addresses that cannot be geocoded are removed, the Metromail address count ***falls to 16.3 % of the Census housing unit count***. The geocode hit rate is higher in the

Vernon wirecenter, although still only two-thirds of the Census housing units can be spatially located.

Table 3. Metromail Address Counts for the Albany and Vernon Texas Wirecenters.

Wirecenter	1995 Census Housing Units	Metromail Households	Metromail Percent of Actual	Geocodable Metromail Households	Geocode Percent of Actual
Albany	1,740	928	53.33%	283	16.26%
Vernon	6,547	4,935	75.38%	4,366	66.69%

The aggregate statistics shown in Table 3 hide the substantial differences in the geocode hit rate that exist between the urban and rural portions of these wirecenters. An appreciation of these differences can be obtained by inspecting maps of the wirecenters. These maps show the customer locations that can be geocoded using the Metromail and GDT data. They also show the actual locations of housing units. In the case of the Albany wirecenter, actual housing unit locations were obtained through the use of satellite photographs. These observations were augmented by Southwestern Bell data so that 100 % of housing unit locations are shown.[3] Southwestern Bell does not have GPS data for the Vernon wirecenter so only satellite observations are shown. Satellite observations are limited to the non-urban areas of the Vernon wirecenter where a high percentage of the existing housing units could be identified in satellite photographs.

Figure 1 shows the Albany wirecenter and Figure 2 shows the Vernon wirecenter. Geocoded housing units are indicated by a dark diamond while actual housing units are shown by a lightly shaded circle. What is immediately apparent is ***that most of the geocoded housing units are in the cities of Albany and Vernon, <u>not</u> in the rural areas of the wirecenter.*** Even in the Vernon wirecenter, where 66.7 % of the Census housing units can be spatially located, ***the bulk of the geocoded locations are in the urban area.***

The Albany and Vernon wirecenters clearly highlight the fact that customer locations in the rural areas cannot be accurately located using geocoding. Substantial enhancements to the household address database need to occur before geocoding is a viable option for locating telephone customers in rural areas.

The Hatfield 5.0 Customer Location Algorithm

The fact that the geocode rates in rural areas are low has a profound implication for the Hatfield 5.0 customer location algorithm. It essentially reduces the algorithm to an arbitrary determination of customer location, much like that of version 4.0.

According to the Hatfield Model Preliminary Relase 5.0 documentation and presentations made at NARUC in November of 1997, the Hatfield model arbitrarily places housing units that cannot be geocoded on the perimeter of the Census Block in

[3] Southwestern Bell determined actual customer locations through the use of Global Positioning Satellite (GPS) technology.

which they are located. Specifically, the difference between the predicted housing unit count in a Census Block and the number of units that can be geocoded are uniformly placed along the perimeter of the Census Block.[4] An undisclosed clustering algorithm is then used to group these actual and arbitrary ("surrogate") customer locations into "main" and "outlier" clusters.

In Rosebud County Montana, for example, only 18 % of the Census housing units can be accurately geocoded. The remaining 3,574 housing units are uniformly distributed along the perimeters of the Census Blocks in the county. In other words, in cases where the geocode rate is low, as it will be the case is the majority of rural areas, the Hatfield algorithm is reduced to an *arbitrary* allocation of housing units within a Census Block. This arbitrary allocation is no more defensible than the 4.0 arbitrary allocation of customers to 2 to 4 towns by a "town factor" (which had little relationship with customer clustering).

The cluster input database for the Hatfield Model 5.0 indicates that there are 2,209 "households" in Rosebud County. Of these 2,209 households, 91% are identified to reside in 4 "main clusters." However, these clusters have little meaning since the overwhelming majority of locations in the clusters are arbitrary. By collapsing most of locations into these main clusters, the Hatfield Model developers can minimize the amount of distribution and connecting cable "built" by the Model.

Accurate geocoding appears to be much more likely in the urban areas. However, since Census Blocks tend be very small in the urban areas, there is likely little added to our understanding of cost determination by the use of geocoding in the urban areas. In the rural areas, accurate geocoding of most customer locations is not possible given the currently available data. Hence, the only value added by the Hatfield 5.0, relative to version 4.0, is the use of Census Block data. This is true despite the considerable expense and resources devoted to customer location geocoding by the Hatfield Model developers.

The BCPM 3.0 Customer Location Algorithm

The BCPM 3.0 customer location algorithm is <u>not</u> reduced to an arbitrary allocation of customer locations in the rural areas. The algorithm is based on the road network, the location of which is known in every Census Block. Housing units are allocated to grids based on the share of the Census Block's road mileage that occurs in a given grid.[5] The underlying assumption is that housing units tend to be located near roads.

---

[4] The only justification given for this allocation is that there might be a road on the perimeter of the Census Block. However, it is unclear as to whether a strong relationship exists between Census Block boundaries and roads. It is clear that there is substantial variation in this relationship across the states. In the western states, for example, census blocks tend to be large and irregularly shaped and roads tend to traverse the interior of the Census Block.

[5] There are actually two methodologies for allocating housing units to grids used in BCPM 3.0. For large Census Blocks, those greater than 0.25 square miles in area, relative road lengths are used. For small Census Blocks, housing units are apportioned based on the land area of the grid relative to the Census

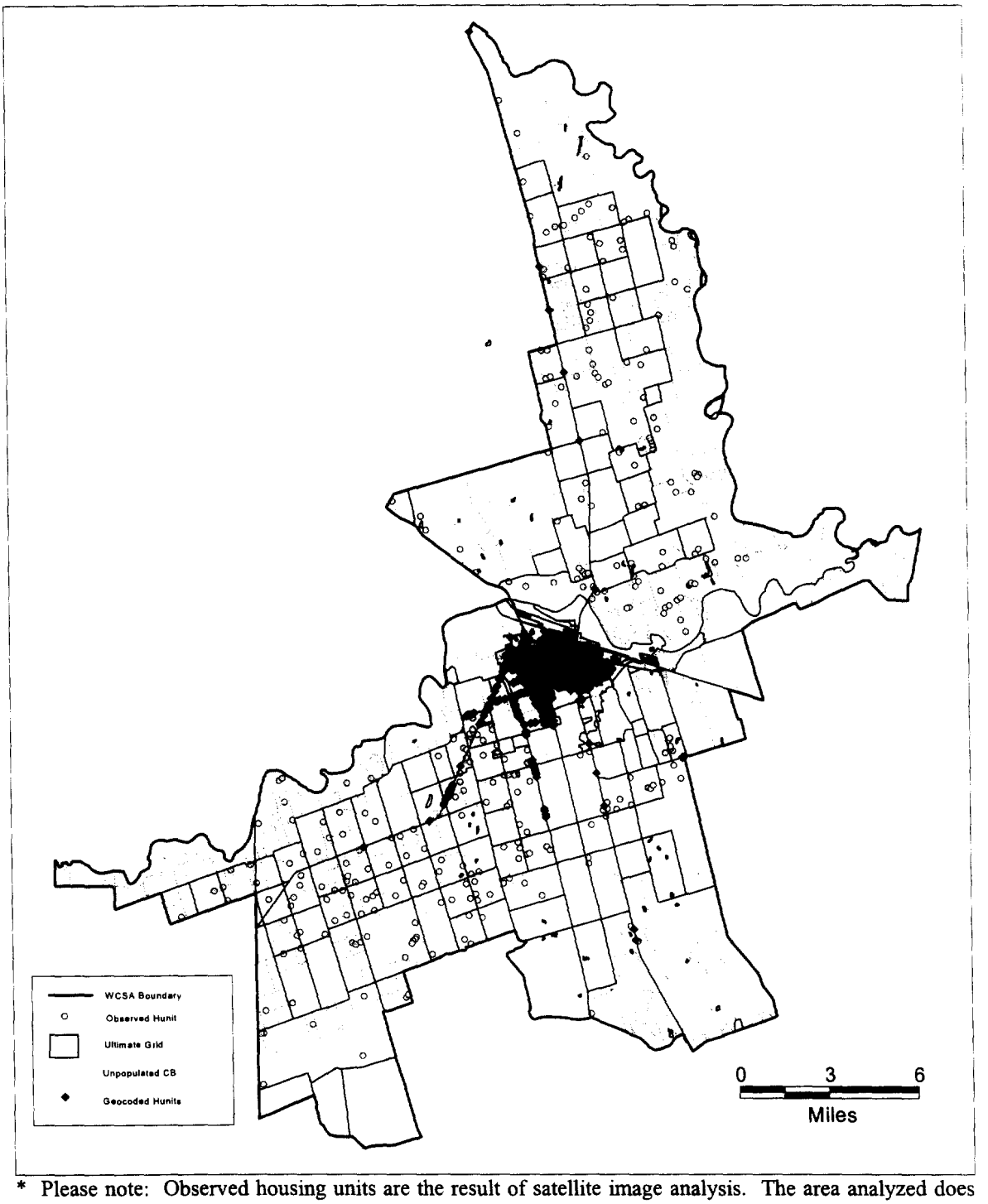
Note the difference between the BCPM and Hatfield methodology. The Hatfield methodology essentially says, "Since we don't know where in a Census Block customers are located and we don't know where the roads are located, let's place all of the customers we cannot geocode uniformly on the Census Block perimeter because we think that's where a road is located." The BCPM methodology says, "Since we don't know where in a Census Block customers are located but we do know where the roads are located, lets allocate the customers around a Census Block based on the road mileage in each section (grid) of the Census Block." Both Models acknowledge a relationship between customer location and the road mileage. However, it is the BCPM that makes use of this relationship and ends up more accurately locating customers within a Census Block than does Hatfield 5.0.

Block's total area. Since large Census Blocks characterize rural areas, the road methodology applies to rural areas.

Vernon, TX Wirecenter

Customer Location

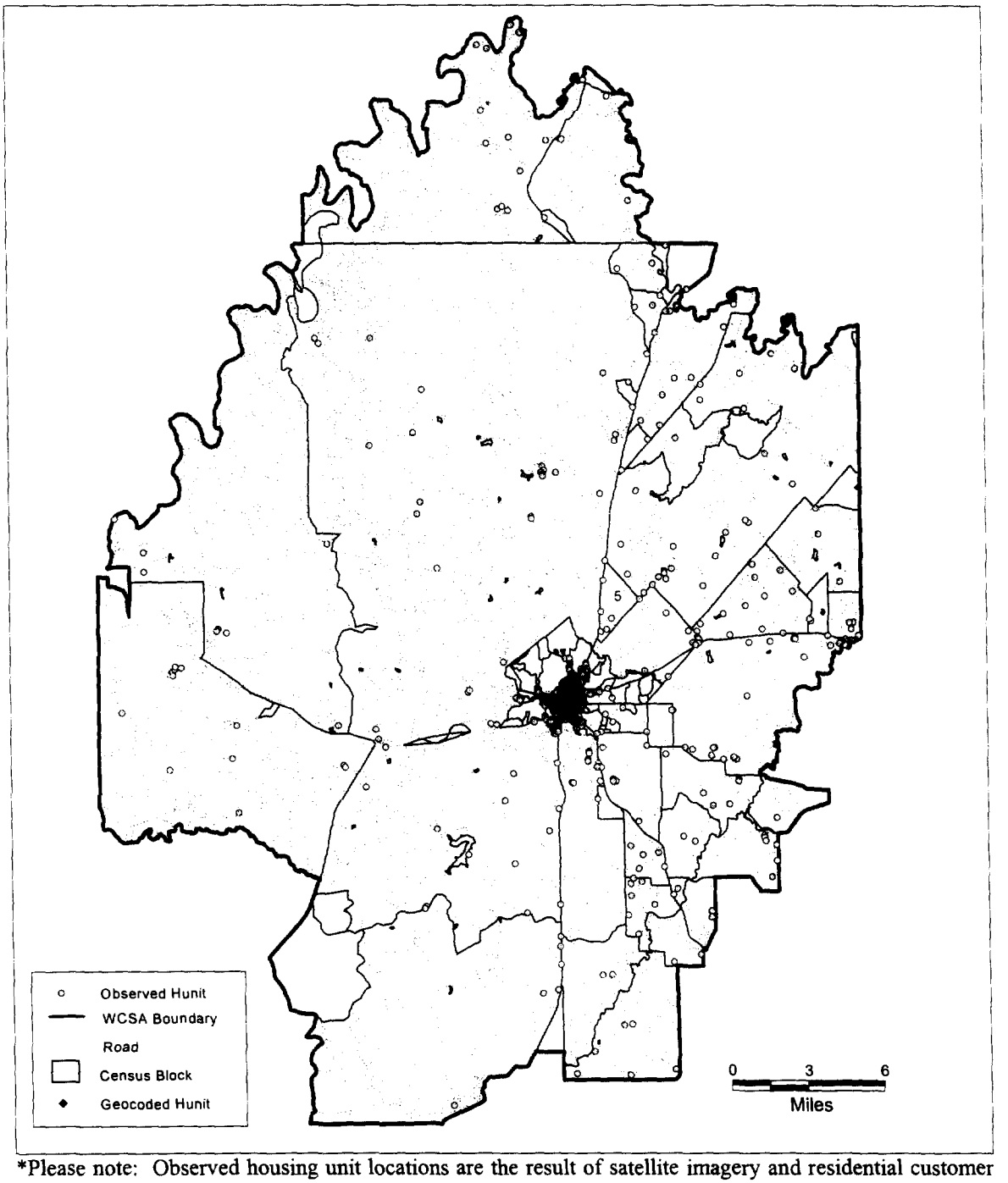
Observed and Geocoded Housing Units



* Please note: Observed housing units are the result of satellite image analysis. The area analyzed does not include the densely populated census blocks in and surrounding the city of Vernon but rather, the northern and southern regions for which analysis could be performed with the greatest degree of accuracy. Each symbol represents at least one housing unit. However, adjacent houses may occasionally appear as only one symbol if the proximity of the units is so close as to cause the symbols to stack upon one another.

Wirecenter TX 08326 04567
CLLI VERNTXLI

Albany, Texas Wirecenter
Customer Location
Observed and Geocoded Housing Units



*Please note: Observed housing unit locations are the result of satellite imagery and residential customer locations provided by Southwestern Bell. The analysis of satellite images allows for truly accurate locations to be assigned. Each symbol represents at least one customer. However, adjacent houses may occasionally appear as only one symbol if the proximity of the units is so close as to cause the symbols to overlap.

Wirecenter TX 08614 4459
CLLI ALBYTXPC

Cardinal®

Benchmark Cost Proxy Model
Release 3.0

Model
Methodology

DECEMBER 11, 1997 EDITION

Developed by
BellSouth, *INDETEC* International,
Sprint and U S WEST

CONTENTS

	PAGE
1.0 Background and Highlights of BCPM 3.0	9
2.0 BCPM 3.0 Attains the FCC's 10 Criteria	9
3.0 The Network	15
3.1 Description of the Local Exchange Network	
3.2 Technical Capabilities of the BCPM 3.0 Network	
4.0 Overview of the BCPM 3.0 Model	19
4.1 Model Structure	
4.2 Model Inputs	
4.3 Model Flexibility	
5.0 Customer Location	21
5.1 Introduction	
5.2 BCPM 3.0 Enhancements	
5.3 Methodology	
6.0 Outside Plant Methodology	33
6.1 Overview	
6.2 Engineering Standards	
6.3 Feeder Design	
6.4 Subfeeder Design	
6.5 Feeder Equipment	
6.6 Feeder Cable Requirements	
6.7 Distribution Plant Design	
6.8 Distribution Equipment	
6.9 Distribution Cable Requirements	
6.10 Loop Length Calculation and Special Considerations	
6.11 Terrain	
6.12 Additional Features in the Model	
6.13 Data Input File	
7.0 Switching	49
7.1 Introduction	
7.2 BCPM 3.0 Enhancements	
7.3 Switching Overview	
7.4 Switch Model Methodology	

8.0 Transport 67
- 8.1 Introduction
- 8.2 BCPM 3.0 Enhancements
- 8.3 SONET Overview
- 8.4 Transport Model Methodology

9.0 Signaling 76
- 9.1 Introduction
- 9.2 BCPM 3.0 Enhancements

10.0 Support Plant 78
- 10.1 Introduction
- 10.2 Support Investment Methodology

11.0 Capital Costs 80
- 11.1 Introduction
- 11.2 Annual Cost Factors
- 11.3 Applying Cost Factors to Investment Accounts
- 11.4 User Adjustable Inputs

12.0 Operating Expenses 82
- 12.1 Introduction
- 12.2 Operating Expense Methodology

13.0 Report Module 84
- 13.1 Introduction
- 13.2 Report Example

APPENDICES

APPENDIX A	Exhibit 1- Satellite Map Clustering for Rural Areas	87
	Exhibit 2- Comparison of BCPM 1.1, Hatfield 4.0, and BCPM 3.0	88
	Exhibit 3- Comparison of BCPM 1.1, Hatfield 4.0, and BCPM 3.0	89
	Exhibit 4- Improved Wire Center Boundaries	90
	Exhibit 5- Census Block Mapping for Waukon, Iowa	91
	Exhibit 6- Variable Size Grids for Waukon, Iowa	92
APPENDIX B	BCPM 3.0 Data Specifications	93
	BCPM 3.0 Processing Steps	100
	Grid Aggregation--General Rules	109
APPENDIX C	Summary of Major Changes fromBCPM 1.1 to BCPM 3.0	111

SECTION 1.0

BACKGROUND AND HIGHLIGHTS OF BCPM 3.0

The FCC and State Commissions are at a critical juncture in deciding the appropriate cost proxy model to use for determining Universal Service Funding. The FCC's May 8, 1997 order regarding Universal Service requires states that elect to conduct their own forward-looking cost study as the basis for calculating federal universal service support in their states, to file the cost study with the FCC by February 6, 1998.[1] In that order the FCC adopted criteria appropriate for determining federal universal service support "to guide the states as they conduct those studies."[2] The FCC explicitly indicated in their order that cost studies submitted by the states will be approved only if they meet the FCC criteria. Section 2.0 outlines the FCC criteria and describes how the enhanced Benchmark Cost Proxy Model (BCPM), Release 3.0, attains each of the 10 criterion.

In addition, the FCC concluded in the Order that they anticipate choosing a specific model to use as the platform for developing a forward-looking cost methodology for non-rural carriers by December 31, 1997.[3] The FCC intends to select a complete mechanism, including inputs, by August 1998 with an implementation date of January 1, 1999.

In order for the Federal and State Universal Service programs to achieve their objective of ensuring virtually ubiquitous access to basic telecommunications service, an objective reiterated in the 1996 Telecommunications Act, it is imperative that a cost proxy model locates customers effectively and constructs adequate facilities to provide basic service to high cost customers. BCPM 3.0's customer location algorithm appropriately locates customers in rural areas. Furthermore, BCPM 3.0's engineering of

---

[1] FCC Report and Order, "In the Matter of Federal-State Joint Board on Universal Service," CC Docket no. 96-45, released May 8, 1997, paragraph 248.

[2] Ibid.

[3] Ibid., paragraph 245.

outside plant estimates a network and costs that network based on an efficient, forward-looking design.

The BCPM team has incorporated enhancements to BCPM 1.1 in two stages. Using BCPM 1.1 as the base, substantial changes to the customer location and outside plant design modules were first implemented in BCPM 2.0. The current model, BCPM 3.0, includes the customer location and outside plant changes incorporated in BCPM 2.0 and supplements these modules with new switching, transport, signaling, capital cost, and expense modules, and a new user interface.

Previously BCPM 1.1, based customer location on Census data at the Census Block Group (CBG) level. BCPM 3.0's customer location algorithm uses housing and business line data at the Census Block (CB) level to more precisely locate customers. On average, there are 30 CBs within a CBG. By overlaying microgrids upon CBs, BCPM 3.0 takes into account the actual road network to more accurately reflect the location of customers within a CB. This enhances accuracy because customers and rights of way for provisioning telecom cables are most frequently found along roadways. Utilizing all of this data, BCPM 3.0 models clusters of customers where they are indeed clustered, and models sparsely populated areas where customers are, in fact, dispersed. This is all done while still retaining the shape and relative cable design of the wire center territory.

BCPM 1.1 assigned CBGs to wire centers based on the centroid, i.e. geographic center, of the CBG. This resulted in a significant number of misassignments of customers to wire centers, as well as misassignments of customers to their respective local exchange carrier. BCPM 3.0's assignment of customers to the appropriate wire center and local exchange carrier is quite accurate. It achieves this accuracy by utilizing wire center boundaries specified by Business Location Research (BLR), and determining the CBs located within that wire center boundary.

BCPM 3.0 integrates more precise information regarding customer location with a customer location algorithm that establishes an optimal grid size based on an efficient network design. Thus, the optimal grid size is determined by adhering to sound engineering practices that reflect forward looking, least cost technology for basic service.