RECEIVED-FPSC

98 SEP -2 PM 4:35

RECORDS AND
REPORTING

Legal Department

J. PHILLIP CARVER
General Attorney

BellSouth Telecommunications, Inc.
150 South Monroe Street
Room 400
Tallahassee, Florida 32301
(404) 335-0710

September 2, 1998

Mrs. Blanca S. Bayó
Director, Division of Records and Reporting
Florida Public Service Commission
2540 Shumard Oak Boulevard
Tallahassee, FL 32399-0850

Re: Docket No. 980696-TP

Dear Ms. Bayó:

Enclosed is an original and fifteen copies of BellSouth Telecommunications, Inc.'s Rebuttal Testimony of Dr. Randall S. Billingsley, Dr. Robert M. Bowman, D. Daonne Caldwell, G. David Cunningham, Dr. Kevin Duffy-Deno, Georgetown Consulting Group, Peter F. Martin and Dr. William E. Taylor, which we ask that you file in the captioned matter.

A copy of this letter is enclosed. Please mark it to indicate that the original was filed and return the copy to me. Copies have been served to the parties shown on the attached Certificate of Service.

Sincerely,

J. Phillip Carver (RA)

J. Phillip Carver

RECEIVED & FILED

FPSC-BUREAU OF RECORDS

Enclosures

cc: All parties of record
A. M. Lombardo
R. G. Beatty
William J. Ellenberg II (w/o enclosures)

Billingsley
DOCUMENT NUMBER-DATE
09611 SEP-2 98
FPSC-RECORDS/REPORTING

Bowman
DOCUMENT NUMBER-DATE
09612 SEP-2 98
FPSC-RECORDS/REPORTING

Caldwell
DOCUMENT NUMBER-DATE
09613 SEP-2 98
FPSC-RECORDS/REPORTING

Cunningham
DOCUMENT NUMBER-DATE
09614 SEP-2 98
FPSC-RECORDS/REPORTING

Duffy-Deno
DOCUMENT NUMBER-DATE
09615 SEP-2 98
FPSC-RECORDS/REPORTIN

Georgetown
DOCUMENT NUMBER-DATE
09616 SEP-2 98
FPSC-RECORDS/REPORTING

Martin
DOCUMENT NUMBER-DATE
09617 SEP-2 98
FPSC-RECORDS/REPORTING

Taylor
DOCUMENT NUMBER-DATE
09618 SEP-2 98
FPSC-RECORDS/REPORTING

CERTIFICATE OF SERVICE
DOCKET NO. 980696-TP (HB4785)

I HEREBY CERTIFY that a true and correct copy of the foregoing was served via Federal Express this 2nd day of September, 1998 to the following:

Jack Shreve, Esquire
Charles Beck, Esquire
Office of Public Counsel
c/o The Florida Legislature
111 W. Madison Street, Rm. 812
Tallahassee, Florida 32399-1400
Tel. No. (850) 488-9330
Fax. No. (850) 488-4491

Michael Gross, Esquire (+)
Assistant Attorney General
Office of the Attorney General
PL-0 1 The Capitol
Tallahassee, Florida 32399-1050
Tel. No. (850) 414-3300
Fax. No. (850) 488-6589
Hand Deliveries:
The Collins Building
107 West Gaines Street
Tallahassee, FL 32301

Tracy Hatch, Esquire (+)
AT&T
101 N. Monroe Street, Suite 700
Tallahassee, Florida 32301
Tel. No. (850) 425-6364
Fax. No. (850) 425-6361

Richard D. Melson, Esquire
Hopping, Green, Sams & Smith, P.A.
123 South Calhoun Street
Tallahassee, Florida 32314
Tel. No. (850) 425-2313
Fax. No. (850) 224-8551
Atty. for MCI

Thomas K. Bond
MCI Metro Access Transmission
Services, Inc.
780 Johnson Ferry Road
Suite 700
Atlanta, GA 30342
Tel. No. (404) 267-6315
Fax. No. (404) 267-5992

Robert M. Post, Jr.
ITS
16001 S.W. Market Street
Indiantown, FL 34956
Tel. No. (561) 597-3113
Fax. No. (561) 597-2115

Charles Rehwinkel
Sprint-Florida, Inc.
1313 Blair Stone Road,
MC FLTHOO 107
Tallahassee, Florida 32301
Tel. No. (850) 847-0244
Fax. No. (850) 878-0777

Carolyn Marek
VP-Regulatory Affairs
S.E. Region
Time Warner Comm.
2828 Old Hickory Boulevard
Apt. 713
Nashville, TN 37221
Tel. No. (615) 673-1191
Fax. No. (615) 673-1192

Norman H. Horton, Jr., Esquire (+)
Messer, Caparello & Self P. A.
215 South Monroe Street
Suite 701
Tallahassee, Florida 32301
Tel. No. (850) 222-0720
Fax. No. (850) 224-4359
Represents e.spire™

David B. Erwin, Esquire
Attorney-at-Law
127 Riversink Road
Crawfordville, Florida 32327
Tel. No. (850) 926-9331
Fax. No. (850) 926-8448
Represents GTC, Frontier,
ITS and TDS

Floyd R. Self, Esquire
Messer, Caparello & Self, P.A.
215 South Monroe Street
Suite 701
Tallahassee, FL 32301
Tel. No. (850) 222-0720
Fax. No. (850) 224-4359
Represents WorldCom

Patrick Wiggins, Esquire
Donna L. Canzano, Esquire (+)
Wiggins & Villacorta
2145 Delta Blvd.
Suite 200
Tallahassee, Florida 32302
Tel. No. (850) 385-6007
Fax. No. (850) 385-6008

Kimberly Caswell, Esquire
GTE Florida Incorporated
201 North Franklin Street
16th Floor
Tampa, Florida 33602
Tel. No. (813) 483-2617
Fax. No. (813) 204-8870

Jeffry J. Wahlen, Esquire
Ausley & McMullen
227 South Calhoun Street
Tallahassee, Florida 32301
Tel. No. (850) 425-5471 or 5487
Fax. No. (850) 222-7560
Represents ALLTEL, NEFTC,
and Vista-United

Tom McCabe
TDS Telecom
107 West Franklin Street
Quincy, FL 32351
Tel. No. (850) 875-5207
Fax. No. (850) 875-5225

Peter M. Dunbar, Esquire
Barbara D. Auger, Esquire
Pennington, Moore, Wilkinson,
& Dunbar, P. A.
215 South Monroe Street
2nd Floor
Tallahassee, Florida 32301
Tel. No. (850) 222-3533
Fax. No. (850) 222-2126

Brian Sulmonetti
WorldCom, Inc.
1515 South Federal Highway
Suite 400
Boca Raton, FL 33432
Tel. No. (561) 750-2940
Fax. No. (561) 750-2629

Kelly Goodnight
Frontier Communications
180 South Clinton Avenue
Rochester, New York 14646
Tel. No. (716) 777-7793
Fax. No. (716) 325-1355

Laura Gallagher (+)
VP-Regulatory Affairs
Florida Cable Telecommunications
Association, Inc.
310 N. Monroe Street
Tallahassee, Florida 32301
Tel. No. (850) 681-1990
Fax. No. (850) 681-9676

Mark Ellmer
GTC Inc.
502 Fifth Street
Port St. Joe, Florida 32456
Tel. No. (850) 229-7235
Fax. No. (850) 229-8689

Steven Brown
Intermedia Communications, Inc.
3625 Queen Palm Drive
Tampa, Florida 33619-1309
Tel. No. (813) 829-0011
Fax. No. (813) 829-4923

Harriet Eudy
ALLTEL Florida, Inc.
206 White Avenue
Live Oak, Florida 32060
Tel. No. (904) 364-2517
Fax. No. (904) 364-2474

Lynne G. Brewer
Northeast Florida Telephone Co.
130 North 4th Street
Macclenny, Florida 32063
Tel. No. (904) 259-0639
Fax. No. (904) 259-7722

James C. Falvey, Esquire
e.spire™ Comm. Inc.
133 National Business Pkwy.
Suite 200
Annapolis Junction, MD 20701
Tel. No. (301) 361-4298
Fax. No. (301) 361-4277

Lynn B. Hall
Vista-United Telecomm.
3100 Bonnet Creek Road
Lake Buena Vista, FL 32830
Tel. No. (407) 827-2210
Fax. No. (407) 827-2424

William Cox
Staff Counsel
Florida Public Svc. Comm.
2540 Shumard Oak Blvd.
Tallahassee, FL 32399-0850
Tel. No. (850) 413-6204
Fax. No. (850) 413-6250

Suzanne F. Summerlin, Esq.
1311-B Paul Russell Road
Suite 201
Tallahassee, FL 32301
Tel. No. (850) 656-2288
Fax. No. (850) 656-5589

Kenneth A. Hoffman, Esq. (+)
John R. Ellis, Esq.
Rutledge, Ecenia, Underwood,
Purnell & Hoffman, P.A.
215 South Monroe Street
Suite 420
Tallahassee, FL 32301-1841
Tel. No. (850) 681-6788
Fax. No. (850) 681-6515

Paul Kouroupas
Michael McRae, Esq.
Teleport Comm. Group, Inc.
2 Lafayette Centre
1133 Twenty-First Street, N.W.
Suite 400
Washington, D.C. 20036
Tel. No. (202) 739-0032
Fax. No. (202) 739-0044

Joseph A. McGlothlin
Vicki Gordon Kaufman
McWhirter, Reeves, McGlothlin,
Davidson, Rief & Bakas, P.A.
117 South Gadsden Street
Tallahassee, FL 32301
Tel. No. (850) 222-2525

J. Phillip Carver (bd)

J. Phillip Carver

(+) Protective Agreements

ORIGINAL

BEFORE THE FLORIDA PUBLIC SERVICE COMMISSION

DOCKET NO. 980696-TP

REBUTTAL TESTIMONY
OF DR. KEVIN T. DUFFY-DENO
ON BEHALF OF BELLSOUTH TELECOMMUNICATIONS, INC.

SEPTEMBER 2, 1998

DOCUMENT NUMBER-DATE
09615 SEP-2 98
FPSC-RECORDS/REPORTING

Table of Contents

I. INTRODUCTION 2
II. CUSTOMER LOCATION 6
  A. HAI 5.0a Customer Location Methodology 6
  B. BCPM 3.1 Customer Location Methodology 17
III. CUSTOMER AGGREGATION 20
  A. HAI 5.0a Customer Aggregation Methodology 21
  B. BCPM 3.1 Customer Aggregation Methodology 22
V. DISTRIBUTION PLANT ESTIMATION 24
  A. HAI 5.0a Distribution Distance Estimation 24
  B. BCPM Distribution Distance Estimation 36
VIII. SUMMARY 41
EXHIBITS 45

REBUTTAL TESTIMONY
OF DR. KEVIN T. DUFFY-DENO
ON BEHALF OF BELLSOUTH TELECOMMUNICATIONS, INC.
BEFORE THE FLORIDA PUBLIC SERVICE COMMISSION
DOCKET NO. 980696-TP
SEPTEMBER 2, 1998

I. INTRODUCTION

Q. PLEASE STATE YOUR NAME AND BUSINESS AFFILIATION.

A. My name is Kevin T. Duffy-Deno. I am the Managing Director-Market Research at *INDETEC* International, a telecommunications consulting firm.

Q. ARE YOU THE SAME KEVIN T. DUFFY-DENO WHO FILED DIRECT TESTIMONY IN THESE PROCEEDINGS?

A. Yes.

Q. WHAT IS THE PURPOSE OF YOUR TESTIMONY?

A. The primary purpose of my testimony is to respond to Mr. Wood's assertion in his testimony of August 3, 1998 on page 20 that:

"By developing costs based on the actual locations of most customers, this release of the HAI Model provides a degree of precision in its results that simply cannot be duplicated by a model such as the BCPM which uses a more simplistic approach of arbitrarily distributing end users along roadways or within an artificial grid structure."

My testimony provides theoretical and empirical evidence that refutes Mr. Wood's assertion. This evidence consists of a relative evaluation of three key features of the HAI Model Release 5.0a (HAI 5.0a) and the Benchmark Cost Proxy Model Release 3.1 (BCPM 3.1): (1) the customer location methodology; (2) the customer aggregation methodology; and (3) a comparison of the minimum distance, as the crow flies, required to connect customers and the distribution plant provisioned in HAI 5.0a.

Q. PLEASE SUMMARIZE YOUR PRIMARY FINDINGS AND CONCLUSIONS.

A. The following summarizes key evidence that counters Mr. Wood's assertion that HAI 5.0a is more "precise" than BCPM 3.1.

- The rate of successful geocoding is extremely low in the rural, low-density areas of Florida. Consequently, the HAI Model customer location methodology is reduced to estimating the lion's share of customer locations in these areas. HAI simply places such customers on the perimeter of relatively large Census Blocks, ignoring the importance of placing customers along interior roads.
- The HAI's sponsors claim that the model accurately locates customers remains unsubstantiated because AT&T has refused to allow anyone access to the underlying geocoded and surrogate data to BellSouth for Florida.
- The rectangular HAI clusters to which the HAI model engineers plant, do not fully encompass the underlying geocoded and surrogate locations upon which these HAI clusters are based. The geocoded and surrogate locations themselves are not

used in the HAI model.

- An analysis of the Yankeetown wire center in Levy County indicates that BCPM's customer location methodology effectively identifies the actual distribution of customers within this wire center.

- An analysis of whether HAI 5.0a estimates the minimum distance needed to connect all of the customers in their main cluster locations identified by the model indicates that HAI 5.0a substantially underestimates this distance by 1,866 miles for BellSouth's Florida territory. In the lowest density zone, the model's estimated distribution distance (including drop and connecting cable) is less than this minimum connecting distance in 87% of its main clusters. Hence, HAI 5.0a's distribution plant substantially underestimates the requisite plant by a substantial margin to provide basic service, particularly in rural areas.

- In contrast to the pronounced internal inconsistency in HAI 5.0a determination of requisite distribution plant, a comparable analysis of BCPM 3.1 reveals that BCPM's modeling of distribution plant is internally consistent with BCPM's modeling intent. The minimum connecting distance analysis of BCPM 3.1 indicates that BCPM is only 465 miles short in the lowest density zone and short in only 32% of its ultimate grids.

Q. HOW IS YOUR REBUTTAL TESTIMONY ORGANIZED?

A. Section II provides an overview of HAI 5.0a's and BCPM 3.1's customer location methodology and an evaluation of the two methodologies. Section III provides similar information for the model's customer aggregation methodologies. The models' provision of distribution plant is addressed in Section IV. A summary of key points is provided in Section V.

Q. ARE THERE EXHIBITS TO YOUR TESTIMONY?

A. Yes. The following is a list of the Exhibits that accompany my testimony:

KDD-1	The Road Network in Dixie County, FL
KDD-2	Geocoded Locations in Dixie County, FL
KDD-3	Geocoded Locations in Levy County, FL
KDD-4	Geocoded Locations in Washington County, FL
KDD-5	Satellite Observations in the Yankeetown Wire Center, FL
KDD-6	Effect of Surrogate Point Placement On Minimum Spanning Tree Length
KDD-7	March 2, 1998 AT&T *ex parte* to the FCC
KDD-8	Concentric Ring Analysis of the Yankeetown Wire Center, FL
KDD-9	Figure 1. Yankeetown Wire Center: Distribution of Actual and BCPM predicted Counts.
KDD-10	BCPM Ultimate Grids in the Yankeetown Wire Center, FL
KDD-11	HAI Distribution Cable Requirements
KDD-12	HAI 5.0a Clusters in the Yankeetown Wire Center, FL
KDD-13	Figure 2. Stylized PNR Polygon Cluster and the HAI Equivalent-area rectangle (Access Database); Figure 3. Formation of the HAI 5.0a Rectangular Clusters
KDD-14	Using Minimum Spanning Trees to Estimate Subscriber Dispersion and Minimum Network Length
KDD-15	The "Shorter-Than-Minimum-Spanning-Tree" Fallacy

II. CUSTOMER LOCATION

A. HAI 5.0a Customer Location Methodology

Q. HOW DOES HAI 5.0a LOCATE CUSTOMERS?

A. As explained in the HAI Model Documentation, "address geocoding" is used to spatially locate customers. First, an address database is acquired from a source such as Metromail, which supplies addresses to the mass-mail marketing industry. These addresses are then input to geocoding software, which then determines the latitude and longitude of the address on a map of the road-network.

When customers cannot be accurately address-geocoded, their locations are placed uniformly on the perimeter of the Census Block in which they are located. These estimated customer locations are called "surrogate" locations.

Q. OF THE COMPLETE ADDRESSES METROMAIL PROVIDES, CAN THE LOCATIONS OF ALL CUSTOMERS BE ADDRESS-GEOCODED?

A. No. P.O. Box and Rural Route addresses cannot be accurately geocoded. Since P.O. Boxes and Rural Route addresses occur much more frequently in rural areas, this affects the ability to geocode in rural areas substantially more than it affects geocoding in the urban areas.

Failure to address-geocode may also result from incomplete information in the road network database. For example, consider a fictional Mrs. Emma Jones who lives at 120 Town Road. To accurately geocode Mrs. Jones' location, one needs three pieces of information in the road network database. First, the physical road segment Town Road, the portion of road between two intersections, needs to be in the database. Second, the physical road segment must be identified with the name "Town Road." Finally, the address range associated with "Town Road" must include "120."

The leading reason why customer locations in rural areas cannot be accurately address-geocoded is this road network information requirement. As an example, Exhibit KDD-1 shows the road network in Dixie County, Florida. Physical road segments are shown in black, named road segments are shown in blue, and named road segments with address ranges are shown in red. Customer locations can only be accurately geocoded to the red road segments. The portion of total road segments that are named and numbered is quite low. Less than 1% of the physical roads in Dixie County are named and have address ranges.

Q. WHAT SHARE OF CUSTOMER LOCATIONS COULD BE ADDRESS-GEOCODED IN FLORIDA?

A. The sponsors of HAI 5.0a filed with the FCC an *ex parte* on February 3, 1998 which presents the geocode rates obtained by the HAI Model developers, by density zone, for the 50 states. For the < 5 line per square mile density zone, the HAI Model developers could accurately address-geocode the locations of only 34% of customers in Florida. The national average was reported as being 15% for this density zone. Table 2 below shows all of the geocode rates for Florida.

Table 2. HAI 5.0a Address-Geocode Rates for Florida:

CBG Density Zone

Density Zone	MCI Reported Successful Geocode Rate
0 - 5	34%
5 - 100	62%
100 - 200	80%
200 - 650	85%
650 - 850	84%
850 - 2,550	78%
2,550 – 5,000	64%
5,000 – 10,000	46%
10,000 +	50%

Q. IS THERE ANOTHER WAY TO EXAMINE THE GEOCODE RATE IN FLORIDA OTHER THAN THAT PRESENTED IN TABLE 2?

A. Yes. Another set of geocode success rates has been provided by AT&T to the Fcc to support HAI 5.0a. These data are success rates by Florida wire center. These data, shown in Table 3, reveal that no residential customer locations could be successfully address-geocoded in 25 wire centers in Florida, or 5.3% of the total wire centers in Florida.

Table 3. Distribution of HAI Address-Geocode Success Rates for Florida Wire Centers.

Geocode Rate	WC Count	WC Share
0%	25	5.33%
0 - 10%	65	13.86%
10 - 20%	25	5.33%
20 - 30%	19	4.05%

30 - 40%	20	4.26%
40 - 50%	25	5.33%
50 - 60%	20	4.26%
60 - 70%	43	9.17%
70 - 80%	78	16.63%
80 - 90%	105	22.39%
90 - 100%	43	9.17%
100%	1	.21%
Total	469	100.00%

Another way to examine these wire center level data is to categorize wire centers into density zones using wire center level densities (density in Table 2 refers to Census Block Group density, the measure of density used by HAI 5.0a). This approach suggests that the address-geocode rate in the lowest density wire centers is lower than the 34% reported in Table 2. In fact, on average, the success rate in the less than 5 line per square mile density zone is 22%. These data for all HAI wire centers in Florida are shown in Table 4. Wire center area is taken from BCPM 3.1 as the HAI Access database does not provide these data.

Table 4. HAI 5.0a Address-Geocode Rates for Florida: Wire Center Density Zone

DZ	WC Count	Average Geocode Rate
< 5	19	22.43%
5 - 20	71	23.30%
20 - 100	91	46.83%
100 - 200	52	68.17%

200 - 650	79	72.78%
650 - 850	20	79.84%
850 - 2,550	62	70.16%
2,550 - 5,000	55	60.17%
5,000 - 10,000	18	40.87%
> 10,000	2	21.19%
Total	469	54.74%

Q. HAVE YOU EXAMINED THE ADDRESS-GEOCODE RATE FOR RURAL FLORIDA?

A. Yes, I have. Table 5 shows the 1995 Census housing unit count for three randomly selected rural Florida counties. Dixie and Levy Counties are located on the western coast of northern Florida while Washington County is located just east of Eglin Air Force Base. All three counties are characterized by low housing unit densities (i.e., less than 15 housing units per square mile). These counties were selected using a MapBasic random selection program from a list of the state's counties with densities less than 25 housing units per square mile and known to contain a BellSouth owned wire center. Wire centers containing Native American reservations, major state parks, or predominantly water were rejected if they were selected.

Also shown in Table 5, for each county is the number of Metromail complete addresses provided to INDETEC on July 11, 1998, the number of these addresses that can be geocoded, and hence, the share of 1995 Census housing units that can be geocoded.

Table 5. Address-Geocoding in Low-Density Counties of Florida

	1995 Census Housing Units	Metromail Complete Addresses	Geocodable Addresses	Census Count Geocodable
Dixie	7,361	216	0	0%
Levy	14,011	7,074	3,748	27%
Washington	8,461	3,794	2,253	27%

Table 5 clearly shows that the share of total customer locations (Census housing units) that can be geocoded varies across counties and can be extremely low, zero in fact, consistent with the HAI Model sponsor findings.

Q. YOU MENTIONED THAT THE ADDRESS-GEOCODE RATE DIFFERS BETWEEN RURAL AND URBAN AREAS. CAN YOU PROVIDE EVIDENCE OF THIS IN THESE RURAL FLORIDA COUNTIES?

A. Yes. The geocode rates shown in Tables 2 - 5 do not show the fact that customer locations in towns are much more likely to be geocoded than those out of town. As evidence of this, consider the three maps of wire centers in these counties provided as Exhibits KDD- 2, 3, and 4. These maps show, by red diamonds, the geocoded locations in these wire centers. No customer locations could be geocoded in Dixie County (KDD-2). Usually one sees that in rural counties, geocoded locations tend to occur in clusters, centered on towns. This is the case in both Levy (KDD-3) and Washington (KDD-4) Counties. In Levy County, the geocoded locations are clustered around the towns of Inglis, Williston, Bronson, and Chiefland. In Washington County, the geocoded locations are clustered around Chipley, at the intersection of Interstate 10 and route 77.

In fact, the 34% geocode rate for the lowest density zone in Florida reported by the sponsors of HAI 5.0a likely overstates the geocode rate in the truly rural areas for this reason. The density zones used to report these geocode rates likely contain both towns and out-of-town areas. Hence, an aggregate geocode rate is typically higher than what is true for the out-of-town areas.

Q. IS IT LIKELY THAT ADDRESS-GEOCODED LOCATIONS ACCURATELY REPRESENT THE TRUE DISTRIBUTION OF CUSTOMER LOCATIONS IN THESE WIRE CENTERS?

A. No. By examining actual locations relative to geocoded locations, one can see that indeed, geocoded locations tend to be only in and around towns, despite there being housing units scattered throughout the wire center.

Q. DID YOU EXAMINE A WIRE CENTER IN RURAL FLORIDA FOR THIS PHENOMENON?

A. Yes. Address-geocoded locations were obtained for the Yankeetown wire center in Levy County. In addition, actual customer locations were obtained through the analysis of a satellite image for this wire center.

Q. WHAT KIND OF SATELLITE IMAGE WAS USED FOR THE FLORIDA ANALYSIS?

A. The satellite image used is referred to as a "10-meter product". That is, one pixel equals 10 meters on a side. The image was taken on December 4, 1995 from an altitude of 520 miles. It was purchased from SPOT Image Corporation and analyzed by ERIM (Environmental Research Institute of Michigan).

Q. HOW WAS THE SATELLITE IMAGE ANALYZED BY ERIM?

A. Since the image is digitized, it can be loaded into a personal computer and enlarged on the computer monitor. ERIM's experienced imagery analysts then visually identified houses on a Census Block by Census Block basis.

Q. WHAT DID YOUR ANALYSIS REVEAL?

A. A map of the Yankeetown wire center Exhibit KDD-5 shows the locations of the houses that could be identified from the satellite image locations. Six hundred and thirty-three of the 2,119 housing units in this wire center could be geocoded to the HAI Model standards. It is clear that geocoding does not capture a significant portion of the customer locations in Florida low-density areas. Moreover, Exhibit KDD-5 shows that actual customers are dispersed throughout the wire center.

Q. CUSTOMERS WHOSE LOCATIONS CANNOT BE ADDRESS-GEOCODED ARE PLACED ON THE PERIMETER OF CENSUS BLOCKS. IS THERE EVIDENCE THAT CUSTOMERS ARE ACTUALLY LOCATED OTHER THAN ON THE PERIMETER OF CENSUS BLOCKS?

A. Yes there is. It is true that people tend to live along roads. It is also true that roads are not limited to the perimeter of Census Blocks. For example, in Florida, 44% of the populated roads in the low-density Census Blocks (densities greater than 0 but less than equal to 20 housing units per square mile) are "interior roads."

The share of populated road mileage that is interior to Census Blocks for the four lowest density zones in Florida is shown in Table 6.

Table 6. Florida Interior Roads

Density (HU / SQMI)	% of Populated Roads that are Interior to Census Block
< 5	48.2
5 - 20	39.5
20 - 100	38.3
100 - 200	32.7

In addition, when *INDETEC* geocoded customer locations in the counties of Levy and Washington we found that 32% and 27%, respectively, are located on interior roads. These findings are inconsistent with the placement of all non-geocodable customers on the perimeter of Census Blocks. Thus, HAI inappropriately disregards the fact that customers in rural areas live along both interior and perimeter roads.

Q. 'S THE PLACEMENT OF SURROGATE LOCATIONS ON THE PERIMETER OF CENSUS BLOCKS A "CONSERVATIVE" ASSUMPTION AS THE HAI PROPONENTS CONTEND?

A. No. By "conservative" I assume the reference is with respect to the *dispersion* of customer locations. Exhibit KDD-6 provides an example of where uniform placement of customer locations along roads both exterior and interior to a Census Block yields a *greater* dispersion (as measured by the Minimum Spanning Tree distance) than uniform placement along the Census Block boundary.

In addition, uniform placement along Census Block boundaries is not conservative if artificial clusters are formed along contiguous Census Block boundaries.

Q. HAVE THE DEVELOPERS OF HAI 5.0a PRESENTED AN ALTERNATIVE METHODOLOGY TO THE SURROGATE PLACEMENT YOU DISCUSSED ABOVE?

A. Yes. On March 2, 1998, AT&T filed with the FCC an *ex parte* that presents an "alternative methodology for determining the location of customers who were not geocoded to their precise street address location by the HAI Model, v5.0a." This *ex parte* is attached to my rebuttal testimony as Exhibit KDD-7.

Q. WHAT IS THIS ALTERNATIVE METHODOLOGY THAT HAI PRESENTED TO THE FCC?

A. The methodology discussed in this *ex parte* locates customers whose addresses cannot be accurately geocoded within a Census Block on the basis of both interior and boundary roads. This methodology uses the internal Census Block road network much in the same way that BCPM has used all along. The *ex parte* states, "We are currently using the same roads that are claimed to be used in BCPM3." (Emphasis added).

Q. IS IT TRUE THAT A MODEL WHICH ADDRESS-GEOCODES SOME CUSTOMER LOCATIONS IS NECESSARILY BETTER THAN ONE THAT DOES NOT USE ADDRESS GEOCODING?

A. No. First, the mere use of address-geocoding does not necessarily make a model's customer location methodology better than one which uses some other technique to locate customers. This argument is especially suspect in the low-density areas where the address-geocode rate is extremely low. Consequently, the assertion of accuracy of HAI's placement of customers in rural areas depends critically upon the erroneous assumption that customers live on only perimeter roads.

Second, the degree to which a model uses address-geocoding needs to be determined. For example, as discussed later, the address-geocoded and surrogate locations are used only to define the perimeter of the PNR polygon clusters in the HAI preprocessing stage. Once HAI transforms the PNR clusters, generating new HAI clusters that encompass a different geographic area than the PNR clusters, the customer latitude and longitude information is discarded. This information in no way enters the Access database used by HAI 5.0a.

Q. WHAT IS YOUR OVERALL ASSESSMENT OF THE HAI CUSTOMER LOCATION METHODOLOGY?

A. First, the HAI customer location methodology *is* severely limited in its ability to use geocoded data, especially in rural areas. Since the rate of successful address-geocoding is low in rural low density areas, this methodology relies heavily on an inadequate estimate of customer locations. This estimation places customers on the perimeter of Census Blocks, disregarding the fact that customers live along interior roads as well.

Secondly, despite claims by the HAI proponents that the HAI customer location methodology more accurately locates customers than BCPM, particularly in the low-density areas, this conclusion is counterintuitive given the limitations just described. Furthermore, AT&T has not provided any quantitative evidence to substantiate this claim, nor has it provided the underlying data for the geocoded and surrogate locations as requested by BellSouth in discovery, to permit such an analysis.

B. BCPM 3.1 Customer Location Methodology

Q. WOULD YOU PLEASE BRIEFLY REVIEW BCPM'S CUSTOMER LOCATION METHODOLOGY?

A. BCPM 3.1 assumes that customers are located on or near roads and uses detailed road-mileage information to allocate U.S. Census housing units counts within Census Blocks. Specifically, a "fishnet" of microgrids, each roughly 1,500' by 1,700', is placed over a wire center. Census Block housing unit counts are then allocated to each microgrid based on each microgrid's share of total Census Block road mileage. The end result is a statistical distribution of customer locations across the microgrids of a wire center. That is, the process yields the *likely* (estimated) location of customers within a wire center.

Q. HOW ARE HOUSING UNITS DISPERSED WITHIN A MICROGRID?

A. The customer location methodology results in a housing unit count for each microgrid. However, BCPM effectively assumes, for purposes of estimating distribution cable distances, that housing units are evenly distributed along the roads within a microgrid.

Q. DID YOU COMPARE BCPM's CUSTOMER LOCATION PREDICTIONS WITH ACTUAL CUSTOMER LOCATIONS?

A. Yes. A key test of any customer location methodology is whether the model's estimated customer locations are consistent with actual customer locations. This is of paramount importance in the rural, low-density area since Census Blocks are quite large in these areas.

The first step was to choose a BellSouth - Florida wire center in a low-density area. As described earlier, this selection was made randomly and resulted in the Yankeetown wire center in Levy County. ERIM then analyzed two satellite photographs that covered this wire center and identified house locations. These locations (latitudes and longitudes) were then digitized with the result being the map presented as Exhibit KDD-5. As Exhibit KDD-5 shows, house locations are scattered through out the wire center.

The next step is to overlay this map with concentric circles each with a radius 1-mile greater than the previous circle's. This yields "rings" around the central office "bull's eye" with a width of 1 mile. The idea is to count the number of actual houses that fall within each "ring." These counts are summed and then plotted against the ring's outer-edge distance from the central office. The result is the distribution of actual houses as measured against distance from the central office.

The map shown in Exhibit KDD-8 (with the concentric rings) is next overlaid with BCPM's microgrids. As noted earlier, housing units are allocated to the microgrids in the wire center based on each one's share of livable road mileage. Using the centroid of the microgrid, each microgrid is assigned to an appropriate ring and the number of BCPM predicted housing units is summed for each ring. This step yields the distribution of BCPM predicted housing units as measured against the distance from the central office.

The actual house and BCPM housing unit distributions for Yankeetown are shown graphically in KDD-9, Figure 1. As one would expect, the majority of houses (62%) is actually located within 3 miles of the central office with the distribution having a "long tail." Figure 1 also shows that the actual and BCPM distributions are a very close match. Since the "actuals" are single, detached-houses and the "predicted" are all housing units, there cannot be an exact one-to-one match. What we are looking for is the tendency of actual locations to lie where BCPM predicts them to be.

For example, 62% of actual locations are within 3 miles of the central office. The comparable figure for BCPM's predicted housing unit locations is 66%. At 10 miles, the percentages are 86 and 88. Moreover, the simple correlation between the actual house counts and BCPM's predicted housing unit counts across the rings is 0.99. Hence, BCPM's customer location methodology, using this benchmark, accurately identifies the actual distribution of customers within this wire center.

Q. DID YOU PERFORM A SIMILAR EVALUATION OF THE HAI CUSTOMER LOCATION METHODOLOGY?

A. No. BellSouth requested in discovery that AT&T provide the customer location data necessary to perform this analysis. AT&T claimed that the information is proprietary and refused to produce it. Thus, AT&T has refused to provide the data needed to conduct a comparable test of the Hatfield model.

Q. WHAT IS YOUR OVERALL ASSESSMENT OF THE BCPM CUSTOMER LOCATION METHODOLOGY?

A. Since the rate of address-geocoding is extremely low in the areas of primary interest for universal service, most, if not all, customer locations must be estimated in the low-density areas. Using road information is a logical approach for estimating customer locations. Not only is the relationship between Census Block road mileage and housing unit counts empirically verifiable but the methodology is based on a comprehensive database. That is, road data are reasonably complete for every Census Block in the country. Address databases are not.

Moreover, the soundness of BCPM's approach has been validated by comparing the customer locations predicted by the BCPM model with real-world customer locations. As presented above, such a test of BCPM's road-based methodology indicates that it effectively predicts the actual distribution of houses, as a related to distance from the central office, in the Yankeetown wire center.

III. CUSTOMER AGGREGATION

Q. HOW DO THE COST PROXY MODELS USE THE CUSTOMER LOCATION INFORMATION?

A. The next step in the modeling process is to aggregate customers into telephone serving areas. These serving areas are the fundamental units that are served by the wire-based network. A brief presentation of the models' aggregation process is necessary as it bridges my discussion of the customer location and distribution plant methodologies.

A. HAI 5.0a Customer Aggregation Methodology

Q. HOW DOES HAI 5.0a FORM ITS TELEPHONE SERVING AREAS?

A. Once the address-geocoded and surrogate customer locations are determined, a process developed by PNR and Associates (PNR) determines clusters of customers. This process is described in the HAI Model Documentation in section 5.5. The documentation indicates that there are several criteria used to determine the ultimate size of a cluster. These stated criteria are: (1) no point in a cluster may be more than 18,000 feet distant (based on right angle routing) from the cluster's centroid; (2) no cluster may exceed 1,800 lines in size; and, (3) no point in a cluster may be farther than two miles from it's nearest neighbor. The end result of this process is a set of irregularly shaped polygon clusters.

Q. WHAT ARE OUTLIER CLUSTERS?

A. The process described above applies to the "main" clusters, which consist of 5 or more locations. PNR also identifies very small clusters, called outlier clusters, which consist of 4 or less locations. These outlier clusters are "homed" on a parent main cluster and are strung together in HAI 5.0a by T1 road cable. In BellSouths's Florida service territory, there are 5,948 main clusters and 210 outlier clusters. The main clusters account for 99.99% of the locations and 99.99% of the lines identified by HAI 5.0a.

In the discussion that follows, "serving areas" in HAI 5.0a are synonymous with "main clusters."

Q. VISUALLY, WHAT DO THE PNR POLYGON CLUSTERS LOOK LIKE?

A. Given that AT&T refused to provide BellSouth the necessary data when it was requested through the discovery process, it is not possible to graphically depict the actual PNR polygon clusters for a wire center in Florida.

B. BCPM 3.1 Customer Aggregation Methodology

Q. PLEASE BRIEFLY REVIEW BCPM'S CUSTOMER AGGREGATION METHODOLOGY?

A. Once housing units and business lines are allocated among the microgrids in a wire center, microgrids (along with the estimated locations within each microgrid) are aggregated into telephone Carrier Service Areas (CSAs), referred to as "ultimate grids." Ultimate grids range in size from a single microgrid (in the high-density areas) to approximately 12,000 feet by 14,000 feet, roughly 6 square miles, in the low-density areas.

In rural, low-density areas, a BCPM ultimate grid situated away from the edge of the wire center is typically a rectangle that is 8 contiguous microgrids wide by 8 contiguous microgrids tall.

Q. VISUALLY, WHAT DOES THE BCPM 3.1 ULTIMATE GRID NETWORK LOOK LIKE?

A. Exhibit KDD-10 shows the Yankeetown wire center with actual locations, overlaid with the BCPM ultimate grids. Also shown is the number of housing units predicted to reside in each ultimate grid. There are 51 ultimate grids in this wire center. The maximum sized grid is 8.3 square miles. BCPM 3.1 places 2,392 housing units (1,865 households) in this wire center and 350 business locations.

Q. ONCE "ULTIMATE GRIDS" ARE FORMED, HOW ARE CUSTOMER LOCATIONS TREATED WITHIN THE ULTIMATE GRID?

A. Customers are still located within the ultimate grid in the micrograds to which they were originally assigned.

Q. HOW DOES THE BCPM CUSTOMER AGGREGATION METHODOLOGY DIFFER FROM THAT USED BY HAI 5.0a?

A. The PNR methodology is a "nearest neighbor" methodology whereby a cluster is formed from the "bottom up." Distance to the nearest neighbor is a primary guide in this process. The BCPM methodology starts with a macrogrid, a 1/25th of a degree latitude and longitude grid consisting of, at the most, 64 microgrids, and seeks to determine if this area can be broken into smaller serving areas. Hence, the BCPM methodology is a "top down" approach. Density, or concentrations of lines, is the primary guide in the BCPM process. Both methodologies yield serving areas of varying sizes, with larger areas serving the lower-density zones.

V. DISTRIBUTION PLANT ESTIMATION

Q. WHAT IS THE NEXT STEP IN THE MODELING PROCESS ONCE CUSTOMERS ARE AGGREGATED INTO SERVING AREAS?

A. The next step is to design a distribution network to serve these areas from the current location of the central office. My focus in this section is on whether the models estimate enough "distribution" plant to serve customers in the locations assumed by the models.

A. HAI 5.0a Distribution Distance Estimation

Q. HOW DOES HAI 5.0a ESTIMATE THE AMOUNT OF DISTRIBUTION CABLE DISTANCE NEEDED TO SERVE CUSTOMERS IN THE LOCATIONS WITHIN THE PNR POLYGON CLUSTERS?

A. This is a multiple step process. The first step is a transformation of the irregularly shaped PNR polygon clusters into rectangles. The second step is placement of customers within these rectangles. The last step is the design of a branch and backbone network to serve these customers.

Q. HOW DOES HAI 5.0a TRANSFORM THE PNR CLUSTERS?

A. HAI 5.0a converts PNR's irregular polygons into the model's rectangular serving areas in two steps. First, for each of PNR's polygon clusters, HAI 5.0a forms a "minimum bounding rectangle," a rectangle that exactly bounds the cluster's "convex hull," by enclosing the polygon's four most northerly, southerly, easterly and westerly coordinates. (See Exhibit KDD-11 for an illustration.) This minimum bounding rectangle has a North-South, East-West orientation.

Next, HAI 5.0a converts each minimum bounding rectangle into an "equivalent-area" rectangle. The model performs this second step by forming a rectangle with the same area as the underlying PNR polygon cluster but with the "aspect ratio" of the minimum bounding rectangle. An aspect ratio is the ratio of a rectangle's height to its width. HAI 5.0a uses the resulting equivalent-area rectangles as the telephone serving areas *internal to HAI 5.0a*. That is, these are the areas to which the HAI model "builds plant."

Q. WHAT DO THE MAIN, "EQUIVALENT-AREA" RECTANGULAR CLUSTERS LOOK LIKE IN FLORIDA?

A. Exhibit KDD-12 shows the Yankeetown wire center and the rectangular clusters as derived from the cluster Access database accompanying HAI 5.0a. In this wire center, HAI 5.0a assumes there are 15 main clusters and 3 outlier clusters. Ninety-nine point eight percent of the locations assumed to exist in this wire center are placed into the main clusters. The largest main cluster is 13.8 square miles. In the State as a whole, the largest HAI 5.0a cluster is 20.2 square miles in size.

Q. ONCE THE RECTANGULAR MAIN CLUSTERS ARE FORMED, FOR MODELING PURPOSES, HOW ARE CUSTOMERS LOCATED WITHIN EACH RECTANGULAR CLUSTER?

A. HAI 5.0a assumes that customer lots are, essentially, evenly distributed within each cluster.

Q. HOW DOES HAI 5.0a DESIGN THE DISTRIBUTION NETWORK WITHIN THE MAIN, RECTANGULAR CLUSTERS?

A. Distribution plant is modeled in a simple branch and backbone configuration. HAI 5.0a assumes customer lots are essentially evenly distributed within each main cluster. Each lot is assumed to be twice as tall as it is wide. The size of each lot is simply the area of the polygon cluster divided by the number of locations. If the model determines that more than one DLC is needed, then connecting cable is also placed to connect the centroid of the main cluster (where the subfeeder terminates) with the DLCs.

Q. DO THE EQUIVALENT-AREA, RECTANGULAR MAIN CLUSTERS CONTAIN ANY INFORMATION ON THE LOCATION OF THE ADDRESS-GEOCODED AND SURROGATE LOCATIONS USED TO DEFINE THE PNR POLYGON CLUSTERS?

A. No. The equivalent-area rectangles are a modeling tool used by HAI 5.0a to estimate the amount of distribution cable needed to serve customers in the locations within the associated PNR polygon clusters. The address-geocoded and surrogate locations are used only in the determination of the PNR polygon clusters. Once the shape and area of the PNR polygon clusters are determined, the information on the geocoded and surrogate locations is no longer used by HAI 5.0a.

A visual representation may help. KDD-13, Figure 2 shows a stylized PNR

Q. CAN YOU PROVIDE A VISUAL DEMONSTRATION OF THIS ISSUE?

A. Certainly. KDD-13, Figure 3 shows a cluster of customer locations, some geocoded, some surrogate. This polygon cluster is transformed by HAI 5.0a into a rectangle that is used in the estimation of distribution plant. Although HAI 5.0a constrains the area of the rectangular cluster to the area of the PNR polygon cluster, the resulting rectangular cluster may bear little resemblance to the shape of the underlying PNR polygon cluster of customer locations. The original customer locations as well as the original distance between these locations are not preserved in the transformation process.

Q. DO YOU HAVE A CONCERN WITH THE HAI 5.0a DISTRIBUTION NETWORK DESIGN WITHIN THE MAIN RECTANGULAR CLUSTERS?

A. Yes. There is an assumption that reinforces the effect on the estimated distribution distance caused by the compression of customer dispersion discussed above. This assumption concerns the placement of the branch and backbone cable within the main rectangular clusters.

After producing the customer lots, HAI 5.0a places backbone distribution cable vertically and branch cable horizontally. Because branch and backbone cable extends to within one lot width (depth) from each rectangle's boundary, low-density rectangles are characterized by locations (i.e., structures) that must be compressed around the interior lots in order to be reached. Now this is not a problem in clusters that are densely populated. However, in sparsely populated clusters, the assumed lots are very large and the compression around the interior lots is much greater. The total effect of the transformation process coupled with this assumption concerning branch and backbone length is a tendancy to underestimate the distribution distance. Again, Exhibit KDD-11 illustrates how this underestimation can occur.

Q. WHAT MEASURE CAN BE USED TO QUANTIFY THE EXTENT TO WHICH THE HAI 5.0a UNDERSTATES DISTRIBUTION DISTANCE?

A. The Minimum Spanning Tree ("MST") can be used to provide an appropriate lower bound for quantifying customer dispersion. The MST is the most conservative measure of the minimum distance required to connect all customer locations. As such, it provides a measure of customer dispersion.

Simply, the MST of a set of points is that set of connecting line segments whose total length is the *shortest possible* for this set of points. The attached paper, "Using Minimum Spanning Trees to Estimate Subscriber Dispersion and Minimum Network Length" (Exhibit KDD-14) provides further rationale for the usefulness of the MST. The attached paper also provides a step-by-step example of how a MST is calculated.

Q. IN REALITY, ARE NETWORK DISTRIBUTION DISTANCES LIKELY TO EXCEED THE MST DISTANCE?

A. Yes, for the simple reason that actual distribution distances likely exceed the MST distance. For example, actual distribution paths must adhere to rights of way (e.g., streets). The MST ignores any such constraints and simply measures the shortest way to connect houses with a straight line. As such, a MST segment will traverse straight across a lake rather than follow a road around the lake to reach the other side.

Q. CAN YOU PROVIDE AN ANALOGY TO HELP EXPLAIN THE MST CONCEPT?

A. Yes. Suppose that an interstate highway is to be constructed directly between Gainesville and Jacksonville. We know that as the crow flies, the aerial distance between these two cities is approximately 65 miles. Clearly, the constructed interstate that connects these two cities cannot be shorter than 65 miles. If it were then cars would have to "fly" over the gaps in the highway. Realistically, the amount of interstate highway distance constructed would be greater than the "crow" distance as natural barriers, rights-of-way, and other obstacles would have to be factored into the routing of the highway.

Hence, the MST distance should be considered as a "reality check," not as the amount of distribution distance that a model should estimate. A model should estimate a distribution distance that *exceeds* the MST distance.

Q. SHOULD THE MINIMUM SPANNING TREE DISTANCE BE CONSIDERED A 'LOWER BOUND' FOR A REQUIRED AMOUNT OF DISTRIBUTION DISTANCE?

A. The MST should not be considered as a "lower bound" for a required amount of distribution distance. Such a lower bound likely *exceeds* the MST for the reason given above. Our analysis is based on the premise that if a model's calculated distribution distance is less than the MST distance, then it is less than the minimum distance required for a functional distribution network.

Q. IS IT TRUE THAT THE MST DISTANCE MAY NOT BE THE SHORTEST DISTANCE CONNECTING A SET OF POINTS?

A. Theoretically speaking, yes. By adding points (nodes) one *may* be able to reduce, under certain conditions, the distance needed to connect the original set of points. However, in most cases of interest, i.e., ***greater than 5 locations***, it is very difficult to find a connecting distance that is less than the MST distance. Exhibit KDD-15 discusses this in more detail.

Q. DOES THE MST TEST THAT YOU ARE PROPOSING CONSIDER ACTUAL, I.E., "REAL-WORLD," CUSTOMER LOCATIONS?

A. No. It is important to realize that the test I am proposing is one for examining whether HAI 5.0a estimates enough distribution cable distance to connect the customers in the locations ***assumed by HAI 5.0a, i.e., in the PNR clusters***, not in their "real-world" locations. A ***comprehensive*** database on the real-world locations of ***all*** customers is ***not*** available. Hence, this is a test of a model's "internal consistency."

Q. DID YOU USE THE MST TO DETERMINE IF HAI 5.0a UNDERESTIMATES DISTRIBUTION DISTANCE FOR BELLSOUTH'S FLORIDA SERVICE TERRITORY?

A. Yes. We first calculated the MST distance for each PNR irregular polygon falling within BellSouth's wire centers in Florida. The MST distance represents the minimum distance required to connect the geocoded and surrogate coordinates encompassed by each polygon. For each *corresponding* equivalent-area, rectangular main cluster formed by HAI 5.0a, we then compared the MST distance with the distribution route distance calculated by HAI 5.0a. In making this comparison, we added drop lengths and connecting cable lengths to the distribution route distance calculated by HAI 5.0a.

Q. DID YOU ACQUIRE THE COORDINATES FOR THE GEOCODED AND SURROGATE LOCATIONS FROM THE ACCESS DATABASE THAT ACCOMPANIES HAI5.0a?

A. No. As discussed earlier, the Access database that accompanies the HAI model does not contain any information on the original locations in the PNR polygon clusters. A data request was made of AT&T to obtain the MST distance, based on a program supplied to AT&T by StopWatch Maps. We received for each HAI 5.0a cluster the MST distance, but was not provided any geocoded or surrogate locations.

Q. HOW ARE YOU DEFINING "UNDERSTATEMENT OF DISTRIBUTION DISTANCE"?

A. An understatement or "shortage" occurs if the MST distance is greater than the distribution route distance calculated by HAI 5.0a. Again, this does not imply that the MST is a lower bound for a required amount of distribution distance. It simply means the model is not providing for enough distribution distance to connect all the customer locations identified by PNR in the underlying polygon cluster ***using the shortest distance configuration that is theoretically possible***.

Q. WHAT DID YOUR CALCULATIONS OF THE PERTINENT MINIMUM SPANNING TREES REVEAL?

A. Using the HAI 5.0a default drop lengths, we calculated the difference between the MST distance and the distribution route distance calculated by HAI 5.0a for each main cluster. Table 9 presents a summary of our findings, again by density zone. Table 9 shows the cumulative amount by which the HAI 5.0a calculated distribution route distance falls short of the MST distance ("shortage"), the cumulative MST for the clusters that are short, the average shortage, the number of main clusters that are short, the number of main clusters in each density zone, and the percentage of main clusters that are short.

HAI 5.0a does not use the 5 - 20 and 20 - 100 density zones but considers only the aggregate 5 - 100 density zone. To provide greater detail for low-density areas, we provide data for these two subcategories.

Table 9. HAI 5.0a Distribution Route Distance Understatement: Default Drop Lengths, BellSouth Florida

Data for Only Main Clusters That Are Short

DZ	HAI MC Dist Route Feet Shortage	MST for Short MC	% Short	Number of MC Short	Number of MC in DZ	Number of MC Short in DZ (%)
< 5	2,784,677	6,569,067	42.39%	136	157	86.62%
5 - 20	4,491,981	15,795,651	28.44%	265	396	66.92%
20 - 100	1,793,590	7,124,473	25.18%	142	415	34.22%
100 - 200	300,093	1,384,879	21.67%	31	227	13.66%

200 - 650	192,303	687,053	27.99%	32	604	5.30%
650 - 850	10,600	46,356	22.87%	5	216	2.31%
850 - 2,550	163,312	1,099,637	14.85%	43	1,491	2.88%
2,550 - 5,000	64,046	624,684	10.25%	31	1,376	2.25%
5,000 - 10,000	35,165	291,621	12.06%	24	832	2.88%
> 10,000	18,648	130,309	14.31%	15	234	6.41%
	9,854,415	33,753,930	29.19%	724	5,948	12.17%

As Table 9 indicates, HAI 5.0a significantly underestimates the required distance to simply connect the customers, as the crow flies, to the network. The understatement by HAI 5.0a of distribution distance is greatest in the lower density areas, specifically, zones with fewer than 20 lines per square mile. Generally, the understatement declines as density rises. Estimated distribution distances that are short of the MST distance characterize 87% of the main clusters in the lowest density zone. This shortage in the lowest density zone is, on average, 42%. For BellSouth's entire Florida service territory, HAI 5.0a understates distribution distance by at least 9.9 million feet (1,866 miles) using the HAI 5.0a default drop lengths.

Q. IS IT LIKELY THAT THE PLACEMENT OF SURROGATE LOCATIONS ON THE PERIMETERS OF CENSUS BLOCKS LEADS TO AN OVERSTATEMENT OF THE MST DISTANCES FOR THE PNR POLYGON CLUSTERS?

A. No. Exhibit KDD-6 shows that a placement of locations on interior and boundary roads can lead to greater dispersion than placement just on the Census Block perimeter. Hence, this counters the argument that the MST distances calculated for the PNR clusters are "too long," and the shortage in distribution distance is overstated, because of the location of the surrogate points along the perimeter of the Census Block boundaries.

Q. IS IT MORE APPROPRIATE TO FOCUS ON THE GROSS SHORTAGE OR NET SHORTAGE IN DISTRIBUTION DISTANCE?

A. It is more appropriate to focus on the gross shortage in distribution distance. First, a definition of terms is in order. A gross shortage is the total shortage that occurs across main clusters when only the distribution distance shortages are added together. A net shortage is the total shortage that occurs when both shortages and "surpluses" are added together across main clusters.

Now, the shortage in one cluster (for which the MST distance exceeds the distribution distance calculated by HAI 5.0a) cannot be offset by another cluster for which the opposite is true. There are two reasons. First, the MST is not a "lower bound" distribution distance for a functional network. Second, and more fundamentally, distribution cable is not fungible across distribution areas. Because a *physical* network is being modeled, 100 feet of distribution distance beyond the MST amount in cluster X cannot be used to offset a 100 feet deficiency in distribution distance in cluster Y. Each and every cluster should have an appropriate amount of distribution distance so that *everyone* on the modeled network can "talk," not just the "average" customer.

Q. BUT IF THE OBJECTIVE IS A COST ESTIMATE, THEN WHY DOES IT MATTER THAT THE MODEL IS SHORT IN SOME CASES IF THERE ARE

CABLE DISTANCE NEEDED TO SERVE CUSTOMERS IN THEIR MICROGRID LOCATIONS WITHIN THE BCPM SERVING AREAS?

A. BCPM employs two modeling tools in this estimation. First, each ultimate grid is divided into 4 potential "distribution quadrants," with the "cross hairs" being at the road-centroid of the ultimate grid. Subfeeder then extends into each ultimate grid to the road-centroid of the ultimate grid. In low-density areas, this is where the DLC is located. Horizontal and vertical connecting cable extend from the DLC to each ***populated*** distribution quadrant of the ultimate grid. The connecting cable terminates at the road-centroid of each populated distribution quadrant.

Q. HOW IS THE AMOUNT OF BRANCH AND BACKBONE CABLE DISTANCE NEEDED TO SERVE THE CUSTOMERS IN EACH POPULATED DISTRIBUTION QUADRANT DETERMINED?

A. This is determined with the aid of another modeling tool. An area equal in size to 1,000' times the amount of road mileage within a populated distribution quadrant is conceptualized. This area is assumed to be a square consisting of equal sized customer lots. Branch and backbone cable is then "laid" to serve each lot.

Q. HAVE YOU APPLIED THE MST REALITY TEST TO BCPM IN FLORIDA?

A. Yes, I have. I performed a test on BCPM 3.1 for BellSouth's service territory in Florida. The relevant unit of analysis in BCPM 3.1 is the Carrier Serving Area or "ultimate grid." The MST is computed for each ultimate grid based on the assumption that customer locations are evenly distributed along roads.

Q. HOW SHOULD THE TERM "DISTRIBUTION" BE USED TO ANALYZE

BCPM'S DISTRIBUTION NETWORK USING THE MST TEST?

A. The issue is whether BCPM is estimating enough cable distance to connect customers to each other *and* to the network. Hence, "distribution" cable should include all cable on the customer's side of the subfeeder termination point in the serving area, i.e., ultimate grid. This distance includes branch, backbone, drop, and connecting cable distance. For the purpose of the MST test, connecting cable is always defined as "distribution" cable regardless of the location of the FDI.

Q. WHAT ARE YOUR FINDINGS FOR BCPM?

A. The findings are presented in Table 10.

Table 10. BCPM 3.1 Distribution Route Distance Understatement: Default Drop Lengths BellSouth Florida

Data for Only Grids That Are Short

DZ	BCPM Dist Route Feet Shortage	MST for Short Grids	% Short	Number of Grids Short	Number of Grids in DZ	Number of Grids Short in DZ (%)
< 5	1,136,087	5,387,477	21.09%	256	806	31.76%
5 - 20	621,726	3,991,302	15.58%	106	703	15.08%
20 - 100	349,609	770,058	45.40%	22	751	2.93%
100 - 200	82,343	205,984	39.98%	8	536	1.49%
200 - 650	86,867	177,997	48.80%	12	1,931	0.62%
650 - 850	18,399	19,563	94.05%	4	836	0.48%
850 - 2,550	109,886	224,708	48.90%	16	4,975	0.32%
2,550 - 5,000	9,634	35,370	27.24%	4	1,223	0.33%
5,000 - 10,000	26,607	26,607	100.00%	1	40	2.50%
> 10,000	12,958	12,958	100.00%	1	5	20.00%

2,454,016	10,851,924	22.61%	[illegible]	11,808	3.84%

In Table 10, the data are for the ultimate grids for which the MST distance exceeds the amount of distribution cable estimated by the model (i.e., "short" grids). In addition, BCPM 3.1 does not use the 5 - 20 and 20 - 100 density zones but considers only the aggregate 5 - 100 density zone. To provide greater detail for low-density areas, we provide data for these two subcategories.

Q. WHAT DOES TABLE 10 SHOW?

A. In the areas of interest for universal service, i.e., the two lowest density zones, the data in Table 10 show that BCPM 3.1 does not estimate enough distribution distance to connect customers in their estimated locations in 24% of its ultimate grids. Considering the entire BellSouth Florida service territory, BCPM's estimated distribution distance falls short of the MST distance in 4% of the ultimate grids. The total "shortage" is at least 2.5 million feet or 465 miles of distribution distance.

Q. WHAT IS YOUR OVERALL ASSESSMENT OF BCPM'S DISTRIBUTION DISTANCE ESTIMATION PROCESS?

A. The results indicate that BCPM is much more internally consistent than HAI 5.0a. That is, BCPM more effectively estimates a minimum required distribution distance (i.e., the MST distance) to connect customers in the locations estimated by the model.

Q. CAN ONE COMPARE THE BCPM MST RESULTS WITH THOSE OF THE

HAI MODEL MST TEST?

A. Yes, but it is important that one keep in mind what the MST test represents. The test is a test of a model's internal consistency, in other words, whether the respective model does what it purports to do, assuming that one accepts its particular modeling assumptions.

With respect to the HAI model, the test addresses whether the HAI model estimates the minimum amount of cable distance, via the rectangular ***main clusters***, to connect customers in the locations identified by the model, i.e., in the corresponding PNR ***main clusters***.

With respect to BCPM, the test addresses whether BCPM estimates the minimum amount of cable distance, via the road-reduced areas and connecting cable configuration, to connect customers in the locations identified by the model, i.e., in the microgrids that comprise an ultimate grid.

Hence, the conclusion one can make is that BCPM is more internally consistent than HAI 5.0a. That is, BCPM is much more likely to estimate the minimum amount of distribution distance needed to connect customers in *its* serving areas, i.e., ultimate grids, than is HAI 5.0a to connect customers in *its* serving areas i.e., main PNR polygon clusters.

Q. DO THE RELATIVE RESULTS OF THE TWO MODELS' MST TESTS CHANGE IF THE DEFINITION OF A "SERVING AREA" IN THE HAI MODEL IS EXPANDED TO INCLUDE THE ASSOCIATED OUTLIER

CLUSTERS?

A. Not substantially. Table 11 presents the results of the HAI MST test, in the same format as Tables 9 and 10, for HAI serving areas defined in this manner. As Table 11 indicates, the addition of the outlier clusters reduces by 0.89 million feet (169 miles or 9%) the total shortage for BellSouth's Florida territory. In the lowest density zone, < 5 lines per square mile, the share of "servings areas" that are short declines from 87% to 76%. The comparable figure for BCPM 3.1 (from Table 10) is 32%. Including outliers improves the HAI model's showing in this test because the T1 road cable distance between the outliers is estimated assuming rectangular routing while the MST is the straight-line distance.

Table 11. HAI 5.0a Distribution Route Distance Understatement: Default Drop Lengths, Expanded Serving Area Definition, BellSouth Florida

Data for Only Serving Areas That Are Short

DZ	HAI SA Dist Route Feet Shortage	MST for Short SA	% Short	Number of SA Short	Number of SA in DZ	Number of SA Short in DZ (%)
< 5	2,314,677	6,789,656	34.09%	120	157	76.43%
5 - 20	4,016,334	15,756,075	25.49%	256	396	64.65%
20 - 100	1,697,831	6,980,286	24.32%	138	415	33.25%
100 - 200	295,974	1,360,514	21.75%	30	227	13.22%
200 - 650	187,645	740,964	25.32%	32	604	5.30%
650 - 850	19,973	137,864	14.49%	6	216	2.78%
850 - 2,550	250,752	1,380,601	18.16%	48	1,491	3.22%
2,550 - 5,000	80,714	661,603	12.20%	31	1,376	2.25%
5,000 - 10,000	35,165	291,621	12.06%	24	832	2.88%
> 10,000	64,757	176,762	36.64%	16	234	6.84%
	8,963,823	34,275,946	26.15%	701	5,948	11.79%

VIII. SUMMARY

Q. PLEASE SUMMARIZE THE MAIN POINTS OF YOUR REBUTTAL TESTIMONY.

A. There are three points I wish to emphasize that pertain respectively to the Hatfield models' customer location, customer aggregation, and provision of distribution plant.

First, the rate of successful address-geocoding in the rural areas of Florida is very low. In fact, not a single location could be geocoded in 25 wire centers in Florida. HAI 5.0a relies on an estimation process for those locations that cannot be address-geocoded. Due to the limited ability to address-geocode customers in rural areas, HAI 5.0a's customer location methodology is reduced essentially to placing customers along the perimeter of Census Blocks.

The proponents of the HAI model have not provided any quantitative analysis of the predictive accuracy of the geocode-surrogate methodology relative to actual, real-world customer locations. In comparison, it has been demonstrated in this testimony that BCPM yields a reasonably accurate depiction of the distribution of customers across the randomly chosen Yankeetown wire center.

Second, the degree to which a model uses address-geocoding needs to be determined. For example the address-geocoded and surrogate locations are used only to define the perimeter of the PNR polygon clusters in the HAI preprocessing stage. Once these clusters are formed, the customer latitude and longitude information is discarded. This information never enters the Access database used by HAI 5.0a.

Third, a key validation test is whether the models estimate enough distribution cable distance to at least connect customers as the crow flies, in the locations identified by the models.

Once customers have been located and aggregated into serving areas, HAI 5.0a and BCPM use different modeling tools in the estimation of the distribution distance needed to connect customers to each other and to the network. The focus should not be on the assumptions behind these tools but on the estimated distances that result from the application of these tools. Specifically, the focus should be on whether the models estimate enough distribution cable distance to connect customers in the locations identified by the models. In the case of HAI 5.0a, these are the geocoded and surrogate locations within the PNR polygon clusters. In the case of BCPM 3.1, these are the microgrids within the ultimate grids.

The minimum spanning tree (MST) test, offered in my testimony, is a test of a model's internal consistency in this regard, i.e., whether it does what its purports to do based upon its own modeling assumptions. When applied to HAI 5.0a and BCPM 3.1, the test indicates that the HAI 5.0 contains a substantial shortfall. In the lowest density zone, the model's estimated distribution distance (including drop and connecting cable) is less than its MST distance in 87% of its main clusters. For the same density zone, BCPM 3.1's estimated distribution distance (including drop and connecting cable) is less than its MST distance in substantially fewer ultimate grids. Overall, the HAI 5.0a shortfall totals at least

1,866 miles while that of BCPM totals at least 465 miles.

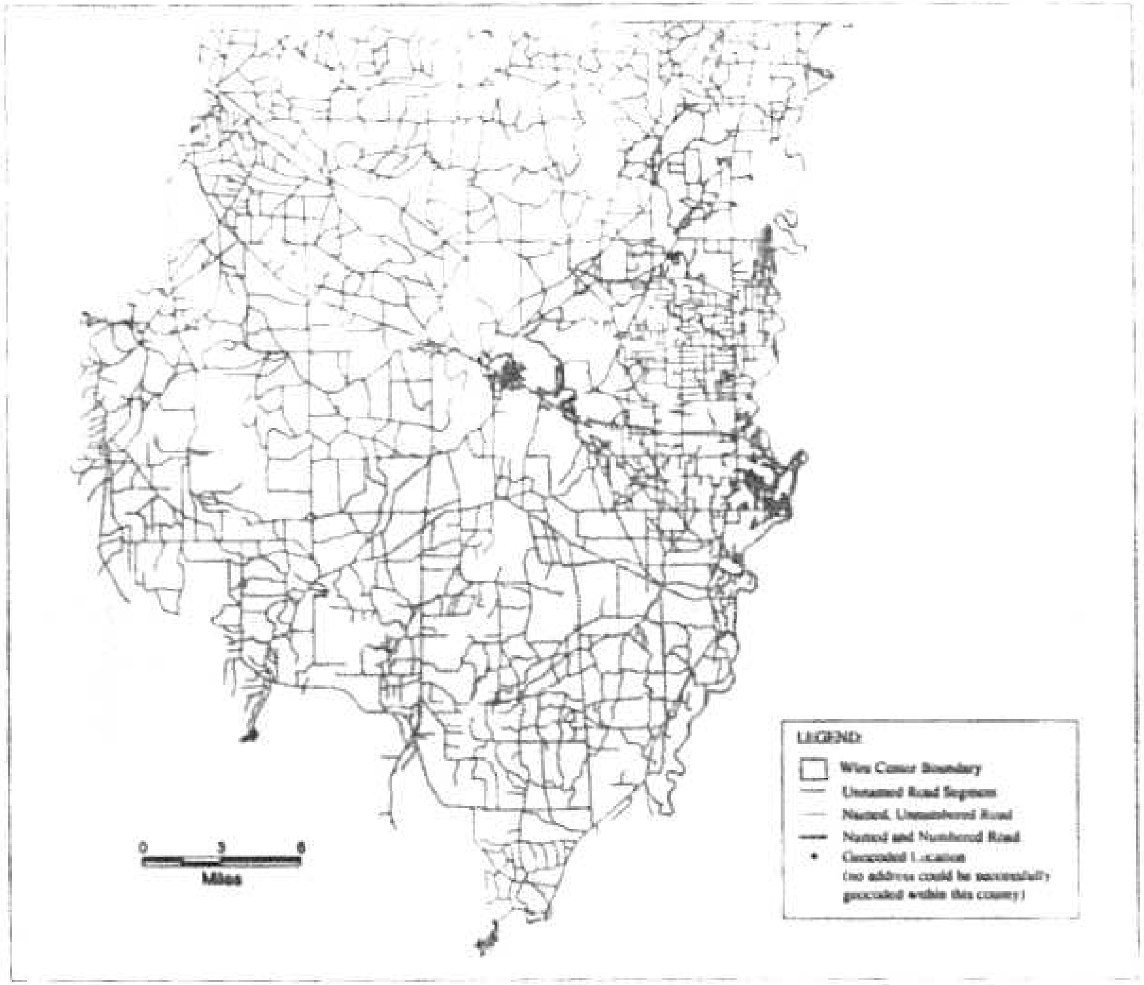
Q. DOES THIS CONCLUDE YOUR TESTIMONY?

A. Yes.

1

EXHIBITS

Dixie County, Florida
Road Network

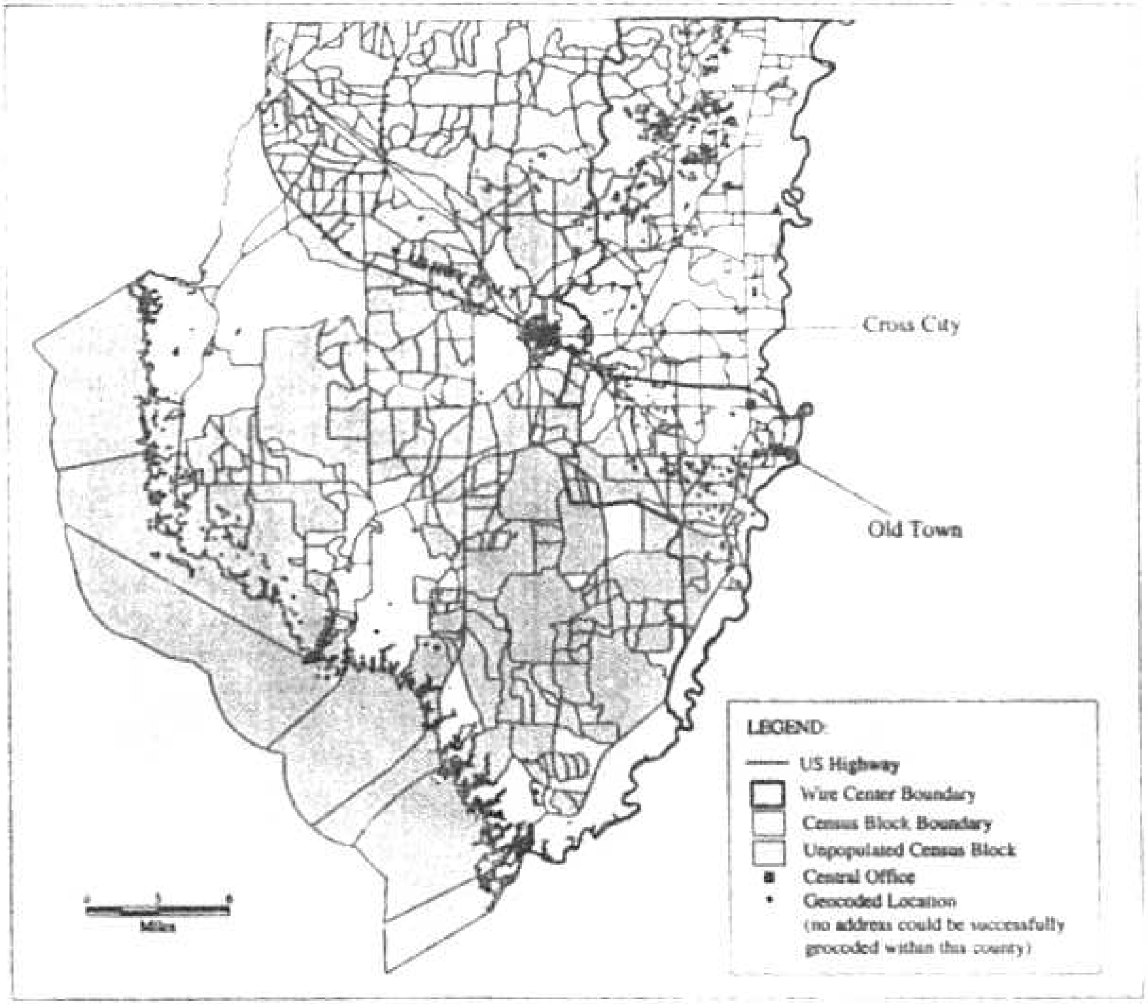


Less than 1% of the Dixie County roads (shown in red) are named and numbered and are therefore, geocodable.

Exhibit KDD-1

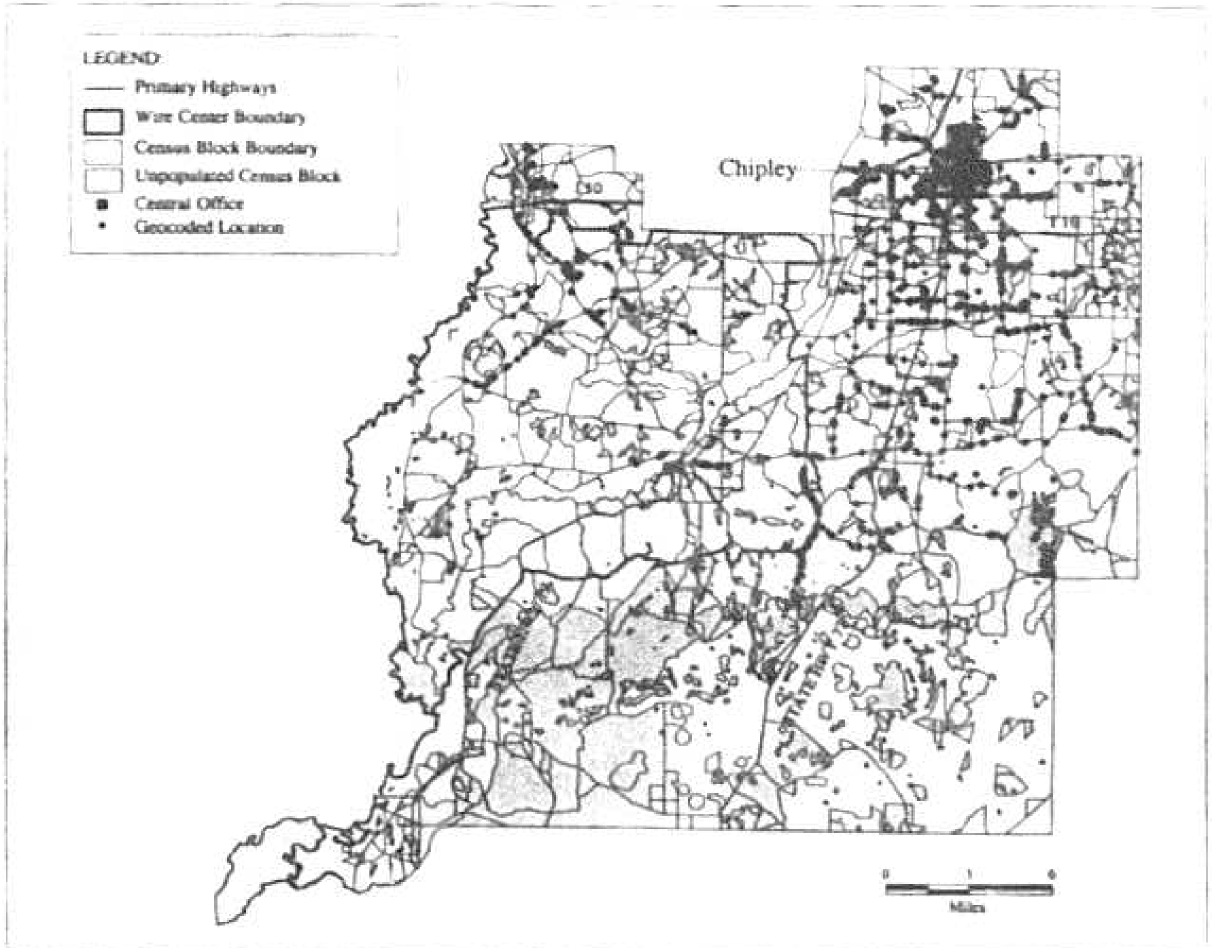
Dixie County, FL

Geocoded Customer Locations



Wire Center FL 07914 01411
CLLI OLTWFLLN
Exhibit KDD-2

Washington County, FL
Geocoded Customer Locations

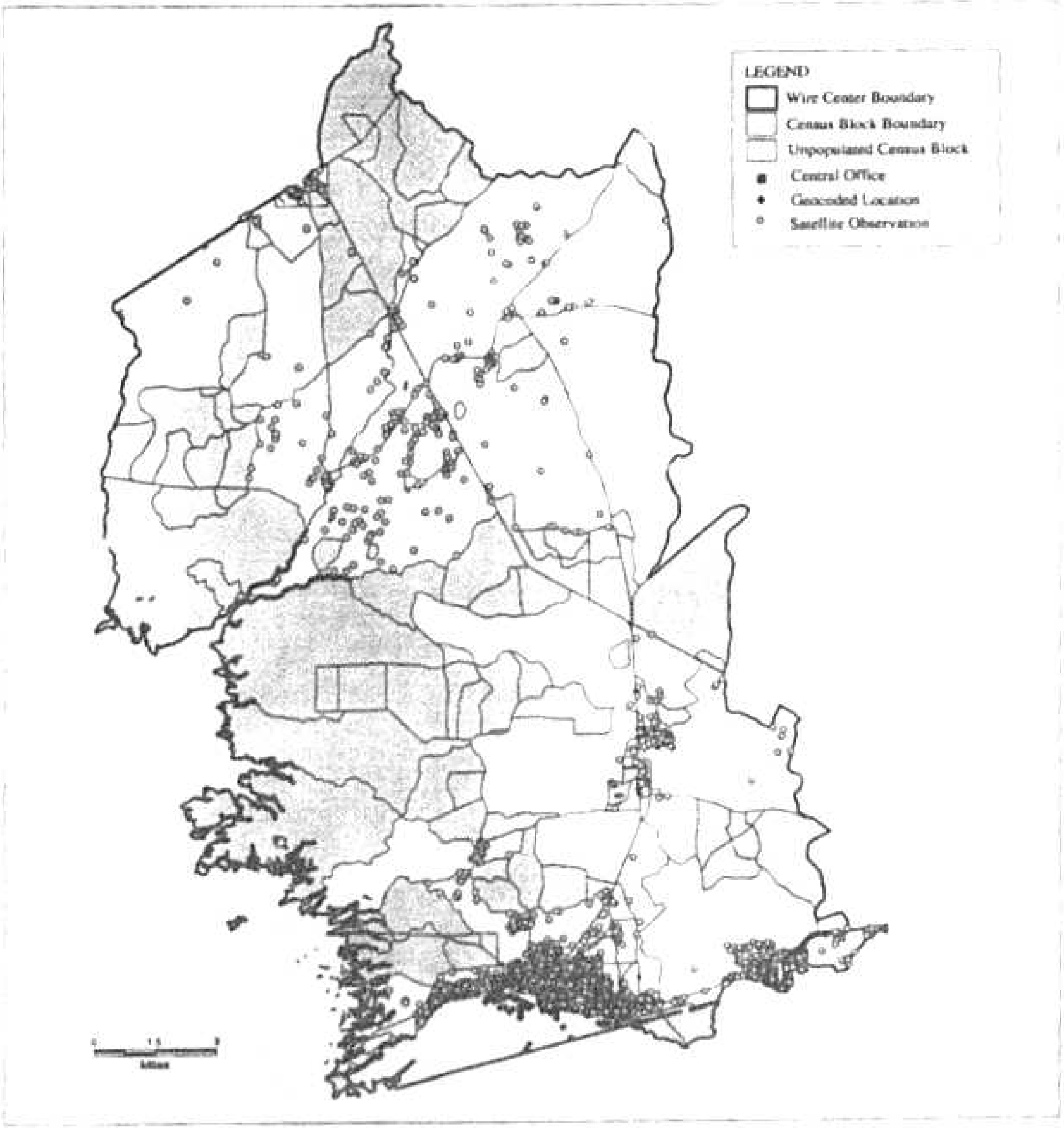


Wire Center FL 07974 01972
CLLI VERNFLMA
Exhibit KDD-4

Yankeetown Wire Center

Levy County, FL

Satellite Observations and Geocoded Customer Locations



Wire Center FL 07991 01303
CLLI YNTWFLMA
Exhibit KDD-5

Rebuttal Testimony of
Kevin Duffy-Deno
Exhibit KDD-6

Effect of Surrogate Point Placement On Minimum Spanning Tree Length

By Phil Bolian, Stopwatch Maps
For *INDETEC* International

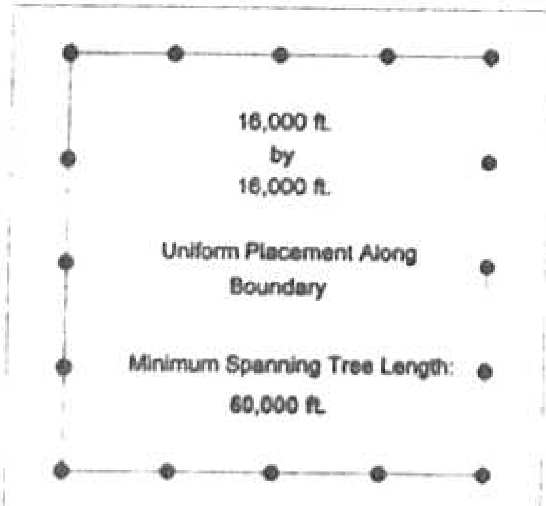
The documentation of the HAI Model Version 5.0a claims that the placement of surrogate points uniformly about the periphery of a Census Block causes those points to be "maximally separated from one another" [Section 5.4.4, first paragraph]. The documentation claims that this placement is highly conservative ... that is, that it causes the *greatest dispersion* of points possible.

In fact, it does *not* cause the greatest dispersion of points. This paper will illustrate this by placing the same number of surrogate points in two *other* configurations:

- Uniformly *within* a Census Block
- Uniformly along *interior roads* as well as the periphery

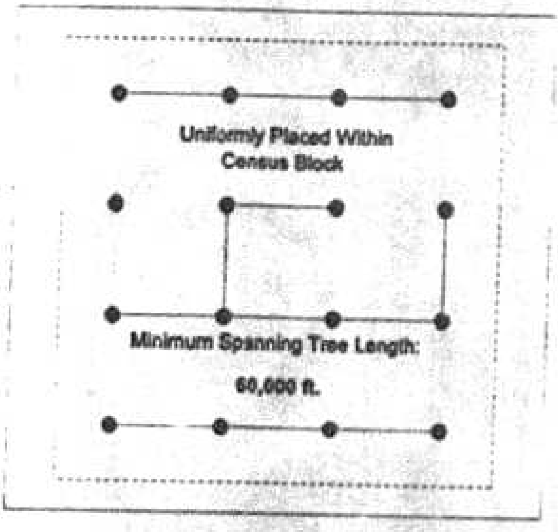
We will then determine the dispersion (as measured by a Minimum Spanning Tree) of each of the newly placed sets of points, then compare each to the Minimum Spanning Tree for points placed about the periphery of the Census Block. We will find that the surrogate points in these new placements are either *just as dispersed as* or *more dispersed than* in the original placement about the periphery.

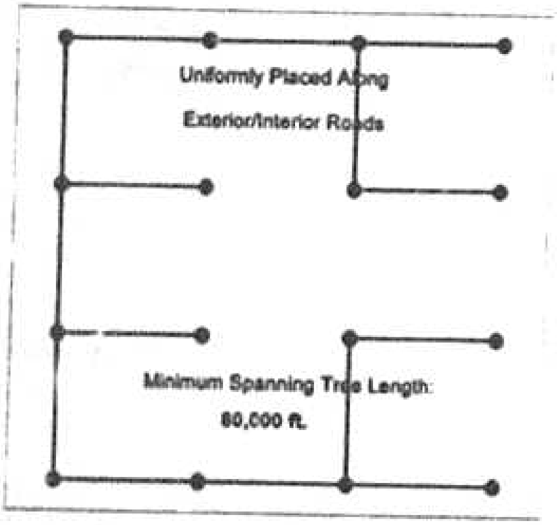
For every case, let us construct a square Census Block, conveniently (for calculation) exactly 16,000 ft. by 16,000 ft. Let us place 16 subscriber locations as surrogate points in that Census Block.



In the first case, we place these points uniformly along the periphery of the Census Block, exactly as is done for the current HAI Model. When we calculate the Minimum Spanning Tree of this set of points, we find it to be 60,000 ft., the length of the full perimeter *minus* the distance between two adjacent points.

Suppose, instead, we were to place our points uniformly distributed *within* our square Census Block. One might think that this would make them less dispersed. But then there is a set of "inner" connections to make. On the next page, as the first figure, we see one of the possible configurations of Minimum Spanning Tree for this uniform placement *within* the square (but, of course, *every* configuration of Minimum Spanning Tree for that placement of points has exactly the same length). Surprisingly or not, it is *again* 60,000 ft.





Then what of a placement along interior as well as exterior roads? In the figure at the right, above, imagine that there are two east-west and two north-south interior roads, and that the bounds of the Census Block are also roads. Then, if we place these points uniformly along *all* roads, we find that the dispersion of the points has *grown*, not diminished. The Minimum Spanning Tree of this configuration is 80,000 ft.

In other words, the placement of surrogate points uniformly on the periphery of a Census Block is *not* a *more* dispersed configuration of points than the other two placements we have investigated here. In fact, it is *less* dispersed than the second alternative. Said yet another way, neither of the two alternative placements presented here would reduce the Minimum Spanning Tree of these points … One would even extend it.

We have examined the dispersion of uniformly placed surrogate points *in a single* Census Block, and found that the placement for surrogates used by HAI 5.0a is *not* the most conservative placement available. We do not even address the fact that if two *adjacent* Census Blocks have surrogate points placed along their peripheries, the points along a common boundary will be far closer together than if they had been spread throughout the areas of each Census Block.

Exhibit KDD-7

AT&T

Michael Lieberman

Room [illegible]
295 North Maple Avenue
Basking Ridge, New Jersey 07920
(908) 221-5467

March 2, 1998

EX PARTE OR LATE FILED

RECEIVED
MAR - 2 1998
FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

Ms. Magalie Roman Salas
Secretary
Federal Communications Commission
1919 M. St., NW, Room 222
Washington, D.C. 20554

RE: Ex Parte Presentation -- Proxy Cost Models
CC Docket No. 96-45

Dear Ms. Salas:

Attached to this submission are two items. The first is a brief description of an alternative methodology to determine the location of customers who were not geocoded to their precise street address location by the HAI Model, v5.0a. The second is a diskette indicating by wire center, the success rate of the HAI Model at geocoding residential addresses to their precise street location.

Two copies of this Notice are being submitted to the Secretary of the FCC in accordance with Section 1.1206(a)(1) of the Commission's rules. A copy of the diskette is being provided to ITS.

Sincerely,

Michael Lieberman /ha

Michael Lieberman

Attachments

cc: Bob Loube
Brad Wimmer
Chuck Keller (w/o diskette)
Natalie Wales (w/o diskette)
Sheryl Todd (w/o diskette)

Recycled Paper

Exhibit KDD-7

Description of Alternative "Road" Surrogating Methodology

The current PNR procedure followed to determine customer locations is to:

A. Geocode each customer's location to its precise street address; and, if this is impossible

B. Assign that customer a "surrogate" geocode by assuming that unlocated customers actually are located uniformly along the perimeter of their Census Block.

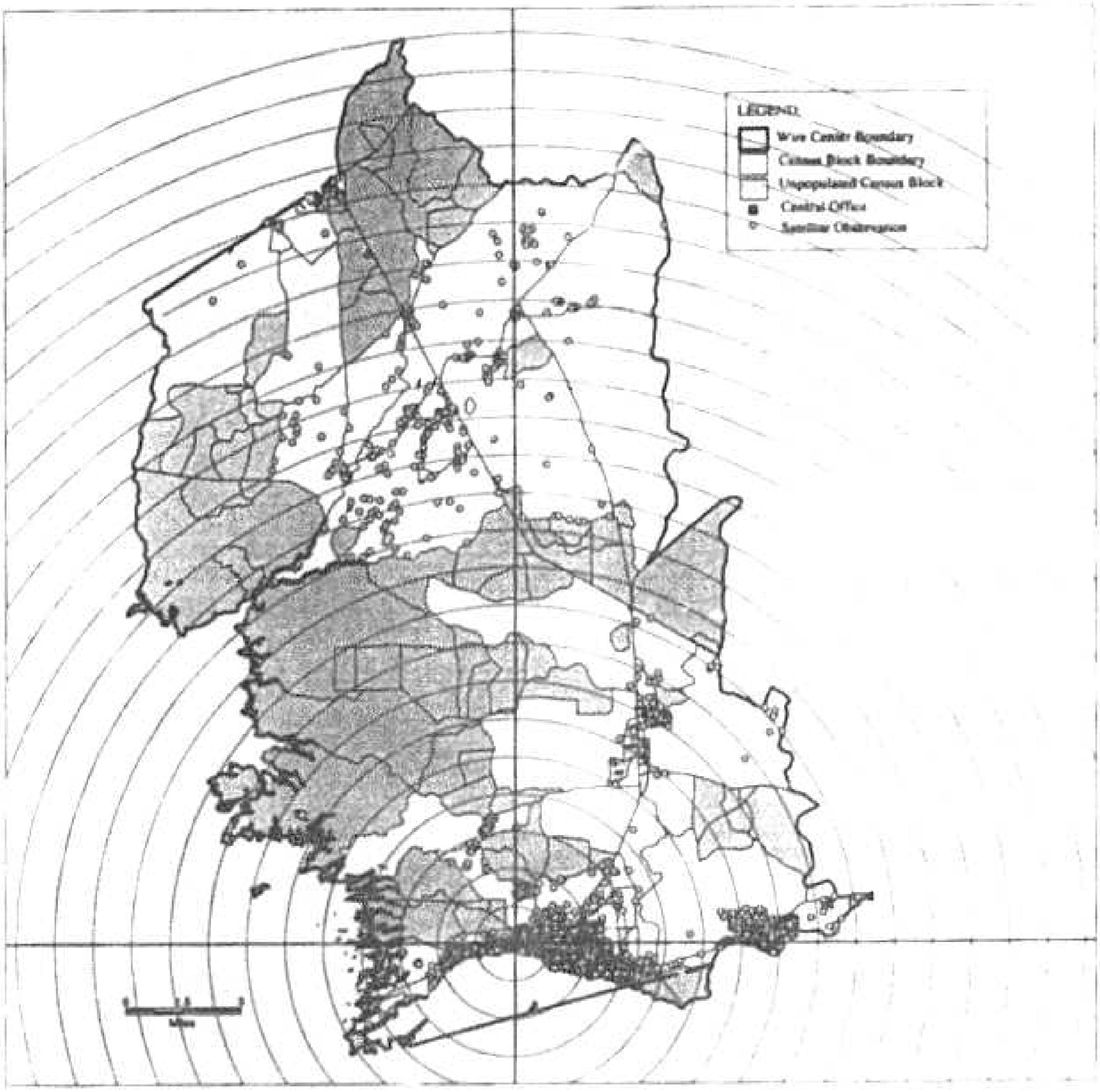
In general, over 70 percent of all customer locations are determined by stage A, and the remaining 30 percent or less are determined by the "surrogate" location process.

The following routine could be employed as an alternative to stage B of the current surrogating methodology. Rather than assuming the geocodes of unlocated customers to be uniformly along the perimeter of their Census Block, it assumes that these unlocated customers are located in specified density or proportions (not necessarily uniform) along a specified selection of road types inside of and bordering their Census Block.

1) Find all the roads of the type that houses could be located on. We are currently using the same roads that are claimed to be used in BCPM3: A21, A24, A25, A28, A31, A34, A35, A38, A41, A44, A45, A48, A51, A52, A64
2) Sort by segment centroid and then remove duplicate segments. Duplicates occur because the same segment can occur twice if the road has 2 different names.
3) String together all of the segments with common TLIDs (Tiger Line Identifications).
4) Build an ordered list of all segments. This is done by ordering all of the polylines created in step 3.
5) Using boundary roads, starting in the south west, continuing along connected lines, using the remaining far south west segment when breaks occur.
6) Using interior roads, starting in the south west, continuing along connected lines, using the remaining far south west segment when breaks occur.
7) Compute the total equivalent distance of the segments.
    - Double the distance of interior roads because houses can be located on either side of these roads – as opposed to only on one side of exterior roads.
    - Inflate or deflate the assumed distance of all segments of a particular road type based on the assumed relative density of customer locations along this type of road versus the average type of road.
8) Based on the number of surrogate points to be inserted, compute the distance between points to be inserted as

   DBP = Number of Surrogates / Total Equivalent Road Distance.
9) Starting at a distance of DBP/2 into segment 1, and for every DBP of road thereafter, save the latitude and longitude as a surrogate customer location point.

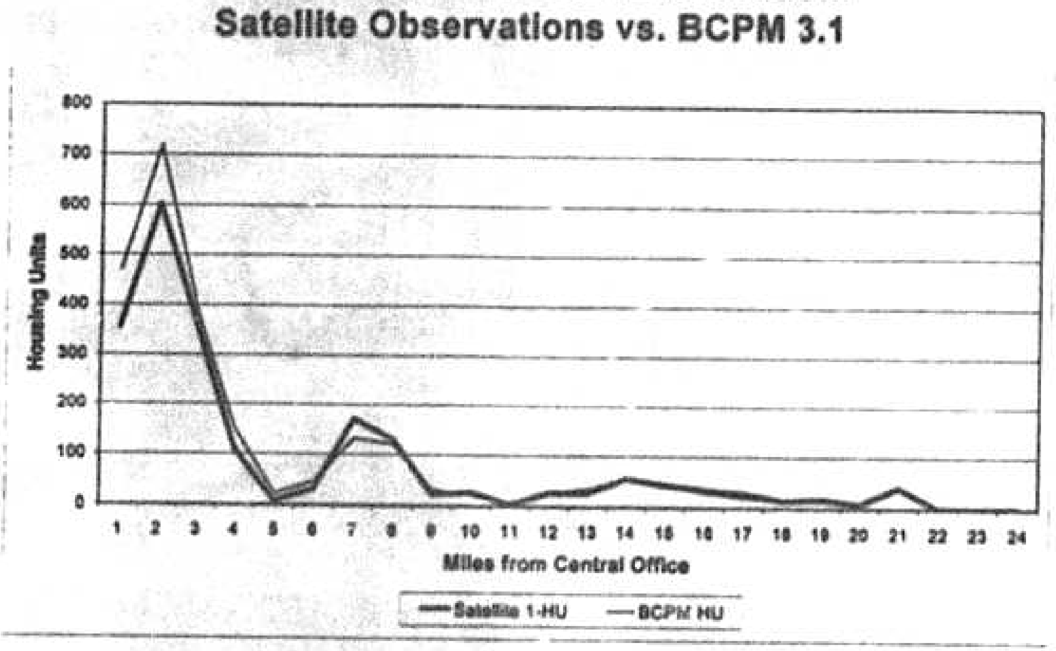
Yankeetown Wire Center
Levy County, FL
Concentric Ring Analysis



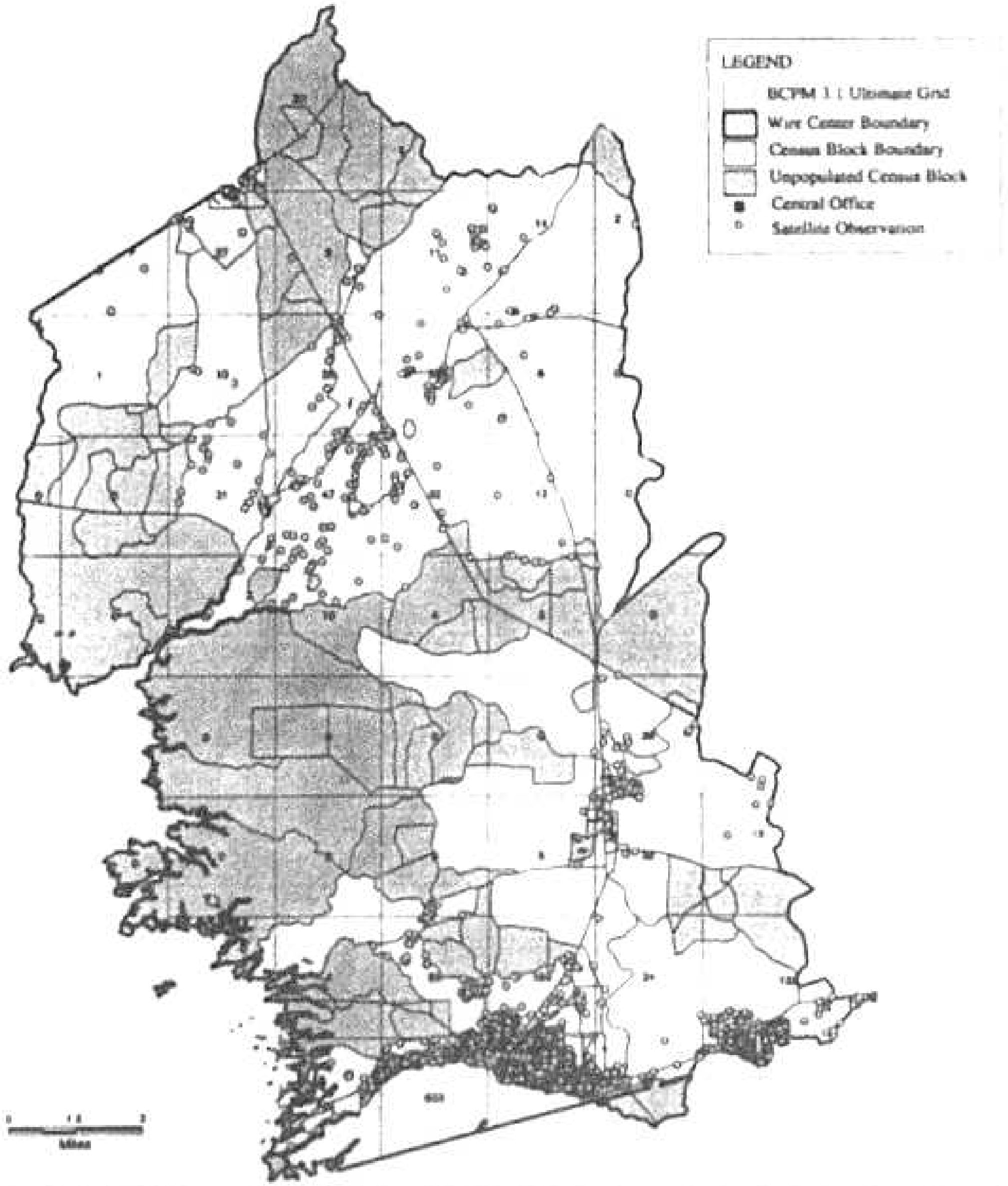
Wire Center FL 07991 01303
CLLI YNTWFLMA
Exhibit KDD-8

Rebuttal Testimony of
Kevin Duffy-Deno
Exhibit KDD-9

Figure 1. Yankeetown Wire Center: Distribution of Actual and BCPM Predicted Counts



Yankeetown Wire Center
Levy County, FL
BCPM 3.1 Ultimate Grids Labeled with Housing Units and Satellite Observations



Wire Center FL 07991 01303
CLLI YNTWFLMA
Exhibit KDD-10

Rebuttal Testimony of
Kevin T. Duffy-Deno
Exhibit KDD-11

HAI Distribution Cable Requirements

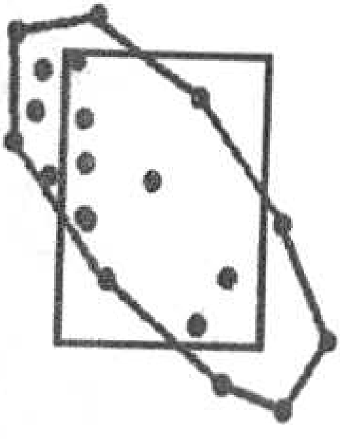
Issue: Whether the distribution plant modeled by HAI 5.0a is adequate to serve customers in their "actual" locations as identified by PNR and Associates (PNR).

Finding: The distribution route miles modeled by HAI 5.0a are too few to serve the customers in the *convex hull* clusters of geocoded and surrogate locations that underlay the rectangular clusters. The rectangular clusters are used in HAI 5.0a in the design of the network.

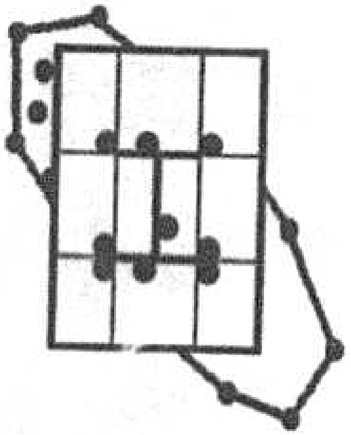
Hence, HAI 5.0a's estimate of the required investment in rural, low-density areas is too low.

Discussion: The customer locations assumed by HAI 5.0a for the purpose of "building" plant are inconsistent with the "actual" locations in the underlying polygon (convex hull) clusters.

The figure below shows a hypothetical convex hull cluster of geocoded and surrogate locations. The rectangle shown is derived from the North-South, East-West aspect ratio and area of the convex hull cluster. Specifically, the rectangle has the aspect ratio of the rectangle that just covers the convex hull cluster (a *minimum bounding rectangle*) and the area of the convex hull cluster itself. The rectangle cluster is what is directly used by HAI 5.0a in its design of the network.



HAI 5.0a assumes that customer locations (i.e., lots) are evenly distributed within the rectangular cluster. For simplicity, assume there are 9 locations. This yields the following figure.



HAI 5.0a subtracts off two lot depths from the cluster North-South length to determine the length of the backbone cable. It also subtracts off two lot widths from the East-West cluster length to determine the length of the branch cable. In the figure shown above, there are two branch cables. Backbone and branch cable is laid in only the middle lot. A drop serves the house in each lot.

Since the default drop length in the lowest density area is 150 feet, the house in each lot must be 150 feet from a branch cable. That is, the houses are concentrated toward the center of the rectangular cluster as indicated in the figure.[1]

This has an important implication for whether the model is providing for a realistic amount of cable. Assume that the area of the convex hull is 15 square miles. Hence, the area of the rectangle is the same and the area of each lot is roughly 1.67 square miles. Lots are assumed to be twice as deep as they are wide. Each lot is 1.83 miles deep (9,640') and 0.91 miles wide (4,820'). Thus, the total distance of cable, including the 150' drops, in this cluster = 9,640' + 2*4,820' + 9*150' = 20,630' or 3.91 miles.

Examining the underlying convex hull cluster of geocoded and surrogate locations strongly suggests that this amount of cable is much too little to serve customers in their "actual" locations. That is, the placement of customers for determining cable lengths within the rectangular clusters is inconsistent with where PNR locates customers in the underlying polygon

---

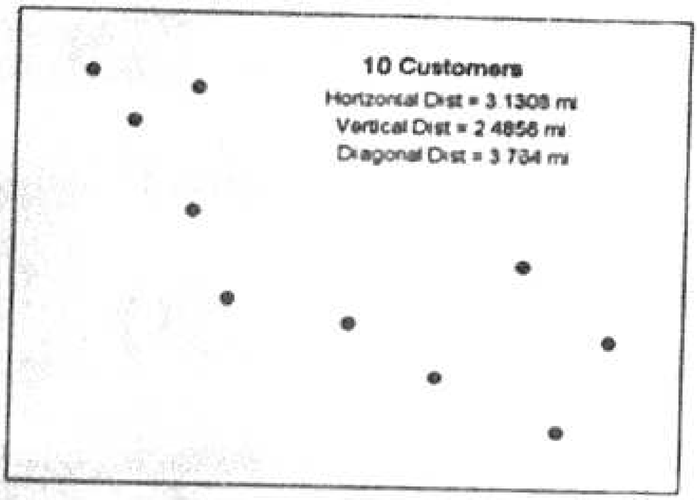
[1] As modeled by HAI 5.0a, it is only the distance from the cluster center to the edge of the middle lot (in this example) that matters for determining whether multiple DLCs are needed.

clusters. In reality, customers are more widely dispersed. Not only will more cable be required but also the 18-kft copper criterion will likely be violated more often, thus requiring additional electronics.

Analysis: A determination of whether HAI 5.0a is not modeling enough distribution plant in its rectangular clusters can be made in the following manner. First, the distribution plant route miles modeled by HAI 5.0a for a specific rectangular cluster is found. Then, the "minimum spanning tree" distance in the underlying polygon cluster is calculated.[2] If the amount of distribution plant route miles modeled by HAI 5.0a is less than the minimum spanning tree amount, then we conclude that HAI 5.0a is not building enough plant to reach customers in the "actual" locations identified in the polygon clusters.[3]

Theoretical Examples: **Example #1**

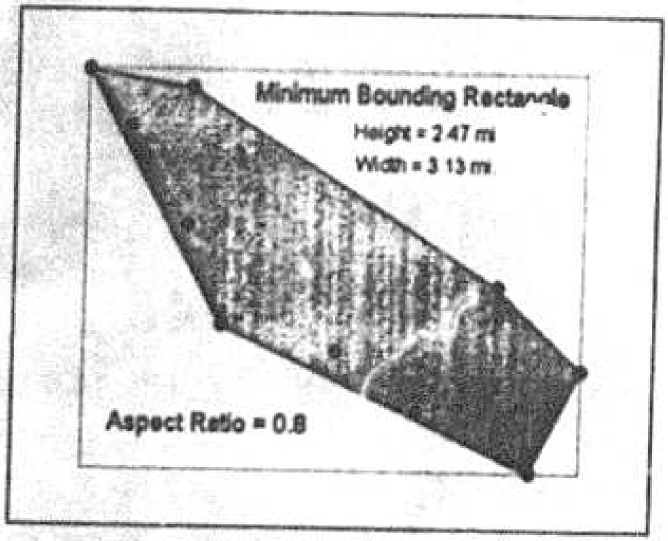
HAI 5.0a groups a set of "actual" customer points into a *cluster*, according to a set of aggregation rules. The two key aggregation criteria are that no customer in the cluster be more than 2 miles from its nearest neighbor and that no customer is more than 18-kft from the centroid of the cluster, measured rectilinearly. Below is shown a hypothetical cluster that meets these criteria.



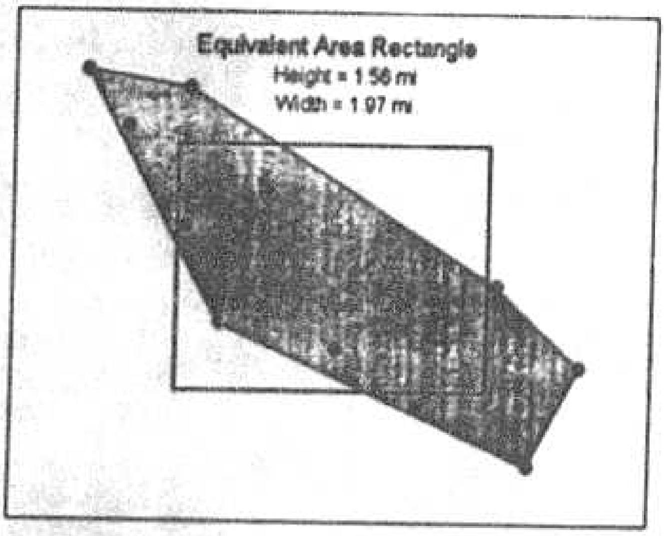
---

[2] A minimum spanning tree distance is the mathematically determined shortest distance that connects all of the customers within a given area.

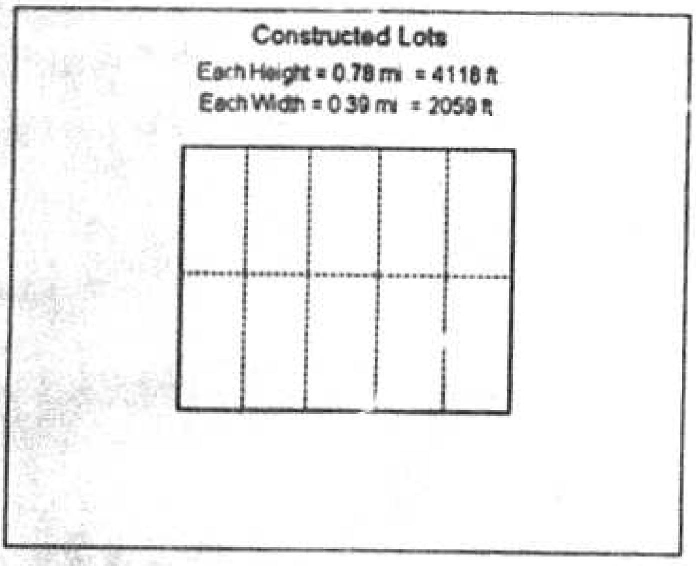
[3] Actual is in quotes to indicate that this refers to PNR's location of customers using geocoding or its surrogate methodology. The surrogate locations likely are not customers' true spatial location.



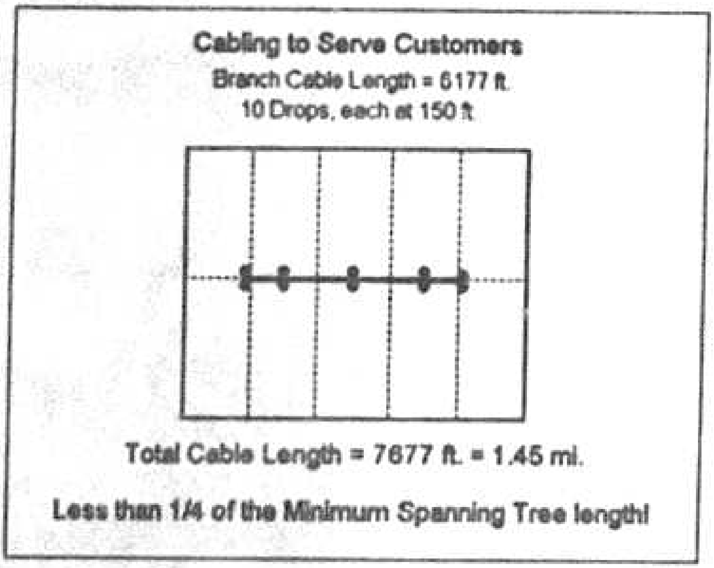
HAI then constructs a *rectangle* with the above aspect ratio; the *size* of that rectangle is determined by its *area* ... and that area is set to be the *area of the convex hull* ... in this case, 3.07 square miles.



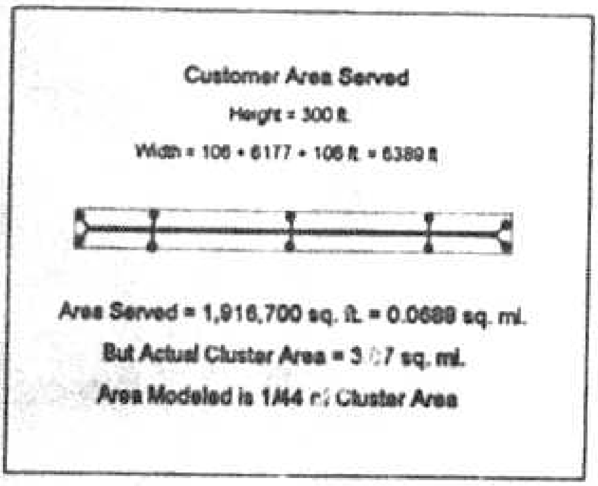
HAI then constructs *lots* within this constructed rectangle. Each lot is twice as high as it is wide.



In this example, there is no backbone cable, only a branch cable. The DLC site is at the centroid of the rectangular cluster. 150-ft. drops connect to the customers.



But note how closely the customers are squeezed toward the branch cable. The arrangement is unrealistic, both from the standpoint of cable length *and* from the standpoint of area served.



Customer Area Served

Height = 300 ft.

Width = 106 + 6177 + 106 ft. = 6389 ft

Area Served = 1,916,700 sq. ft. = 0.0688 sq. mi.

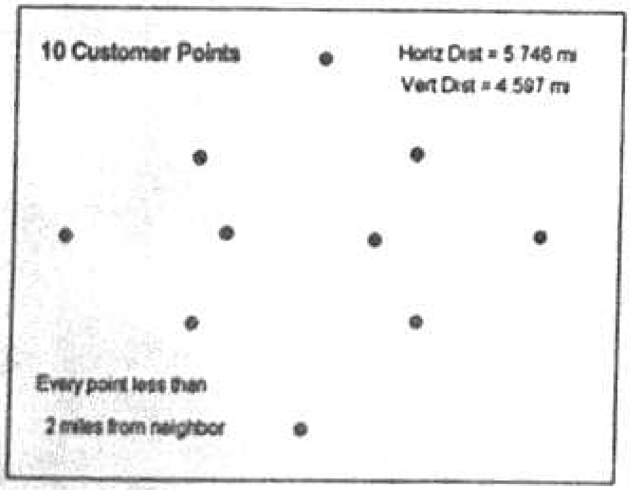
But Actual Cluster Area = 3.07 sq. mi.

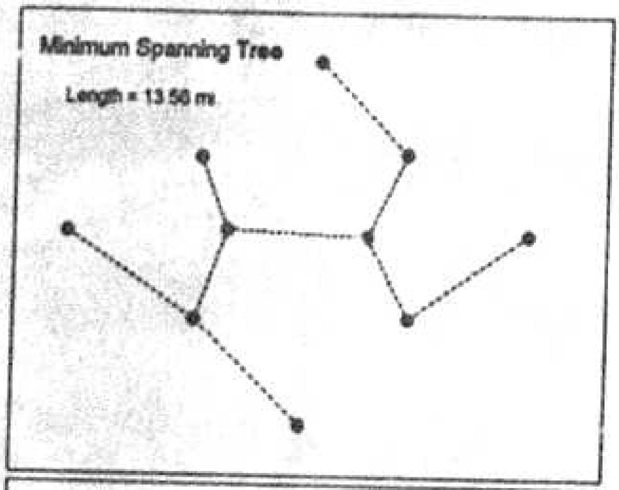
Area Modeled is 1/44 of Cluster Area

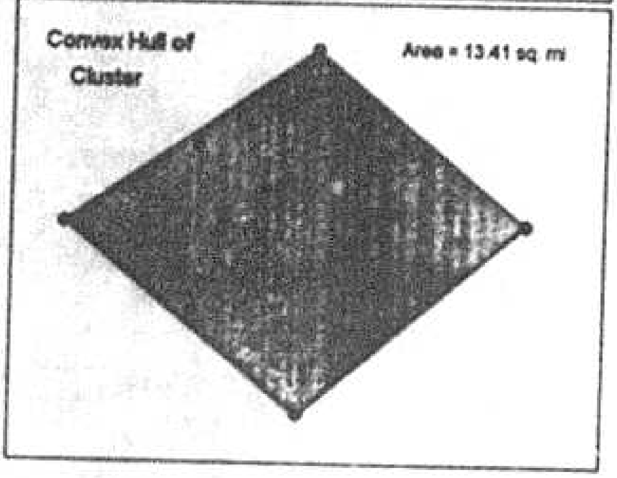
Hence, for this example, the distribution plant route miles modeled by HAI 5.0a are only 25 % of the minimum amount required to connect the 9 customers in their "actual" locations. Moreover, the area modeled as containing distribution plant is only 2 % of the area of the polygon (convex hull) cluster.

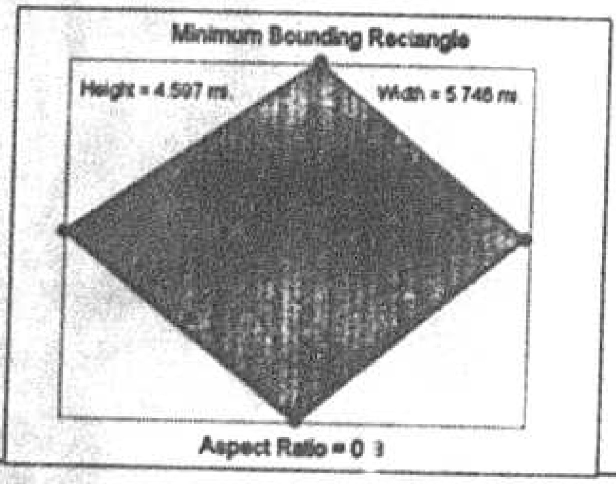
Example #2

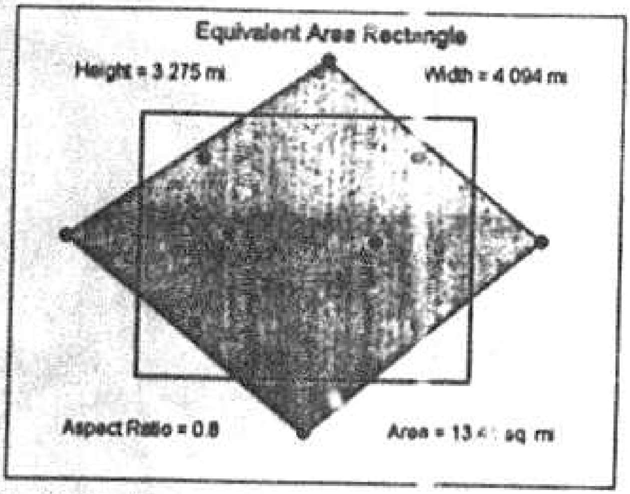
The next example considers a much larger cluster, similar in size and density to which HAI 5.0a models in low-density areas.

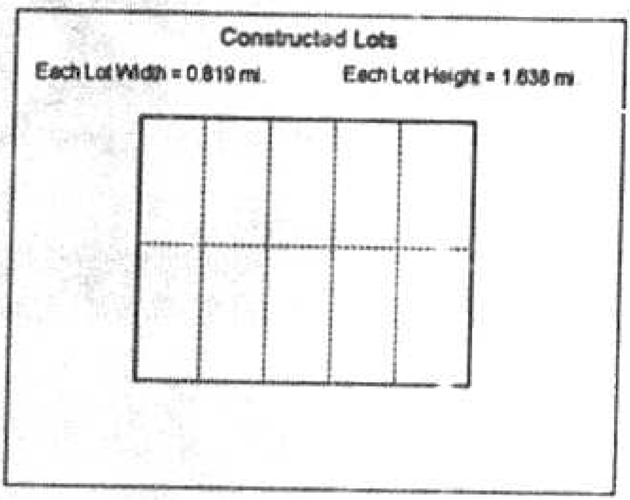



Minimum Spanning Tree
Length = 13.50 mi


Convex Hull of Cluster
Area = 13.41 sq. mi

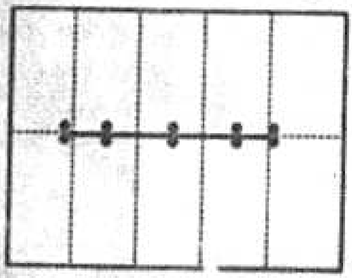

Minimum Bounding Rectangle
Height = 4.597 mi.
Width = 5.746 mi
Aspect Ratio = 0.8


Equivalent Area Rectangle
Height = 3.275 mi
Width = 4.094 mi
Aspect Ratio = 0.8
Area = 13.41 sq mi


Constructed Lots
Each Lot Width = 0.819 mi.
Each Lot Height = 1.638 mi

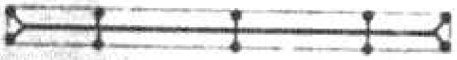
Constructed Customer Locations and Cabling

Each customer at 150 ft. drop from branch cable



Total cable length = (3 x 0.819) + (10 x 150/5280) = 2.74 mi.
(Minimum Spanning Tree is 13.56 mi.)

Total Customer Area Served



300 ft. x (106 ft. + (2.457 mi. x 5280) + 106 ft.) = 3,955,488 sq. ft.
= 0.142 sq. mi.

Area of the Cluster is 13.41 sq. mi.

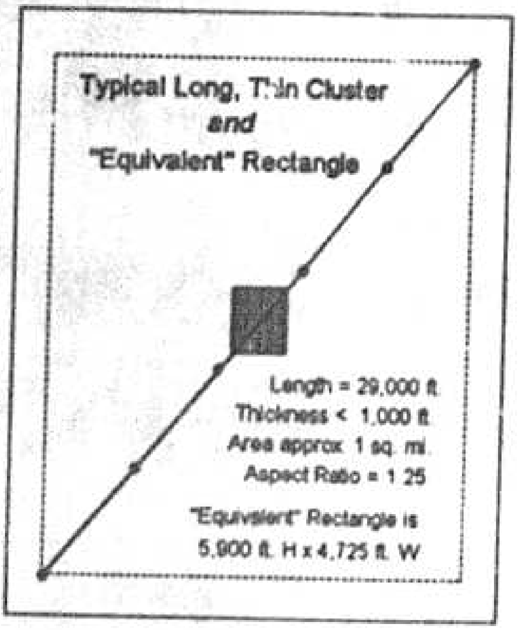
Area Served is 1/94 of the Cluster Area

Hence, in this example, the distribution plant modeled by HAI 5.0a is only 20 % of the minimum amount necessary to serve these 9 customers in their "actual" locations. Moreover, the area that contains distribution plant represents only 1 % of the total area of the polygon cluster of "actual" locations.

Example #3

An extreme case occurs when the convex hull cluster is long and thin. This commonly occurs in rural areas where Census Blocks tend to be large and the roads tend to be long. Thus, the distance constraints employed by the HAI clustering algorithm tend to group together *strings* of subscribers along a several mile segment of road. Sometimes the road is straight, sometimes it is curved, and sometimes it bends. But very typically, the convex hull of the resulting cluster is long and skinny.

The figure below shows a long and thin convex hull cluster that can occur in rural areas. The cluster consists of 6 locations strung out along a relatively straight line (road). The length of this string is 29,000′ with a width of less than 1,000′. The minimum bounding rectangle for this cluster is also shown and is assumed to have an aspect ratio of 1.25. In this example, the equivalent area rectangle has an area of approximately 1 square mile.



Assuming 6 locations in this cluster yields 6 plots, each 0.17 square miles in size. The HAI distribution module algorithm then assumes each lot is twice as deep as it is wide. This yields lots that are 3,048′ deep and 1,524′ wide.[4]

HAI 5.0a conceptually models this cluster as consisting of 2 rows of lots (East-West). Since twice the lot depth exceeds the North-South dimension

---

[4] Note that the HAI algorithm is not consistent with respect to the aspect ratio of lots versus the aspect ratio of the equivalent area rectangular cluster. The aspect ratio of a lot is independent of the aspect ratio of the rectangular cluster and is always 2. Thus, in this example, the sum of the lot depths (3,048′ x 2 = 6,096′) exceeds the "depth" of the rectangular cluster (5,900′).

of the cluster, HAI 5.0a defaults to no backbone cable with to two East-West branch cables emanating from the DLC. The cable extends for only 1,524', the width of one lot. Assuming 150 ' drops yields a total route distance of 2,424'.[5]

In other words, HAI 5.0a assumes that only 2,424' of cable is required to serve 6 customers who are actually identified by HAI as being strung out along a road 29,000' in length. Since the 6 customers are assumed to be essentially in a straight line, 29,000' *is* the minimum spanning tree distance. Hence, HAI 5.0a places only 8.4 % of the cable necessary to serve these customers in their locations within the convex hull.

Summary: Our analysis indicates that there are two effects that work together to lower the amount of distribution plant calculated by HAI 5.0a in rural, low-density areas.

The first effect results from the distortion of the original polygon cluster of "actual" customer locations caused by the formation of the rectangular clusters. The distortion results from the rectangular clusters having the aspect ratio of the minimum bounding rectangle of the polygon cluster and the area of the polygon cluster.

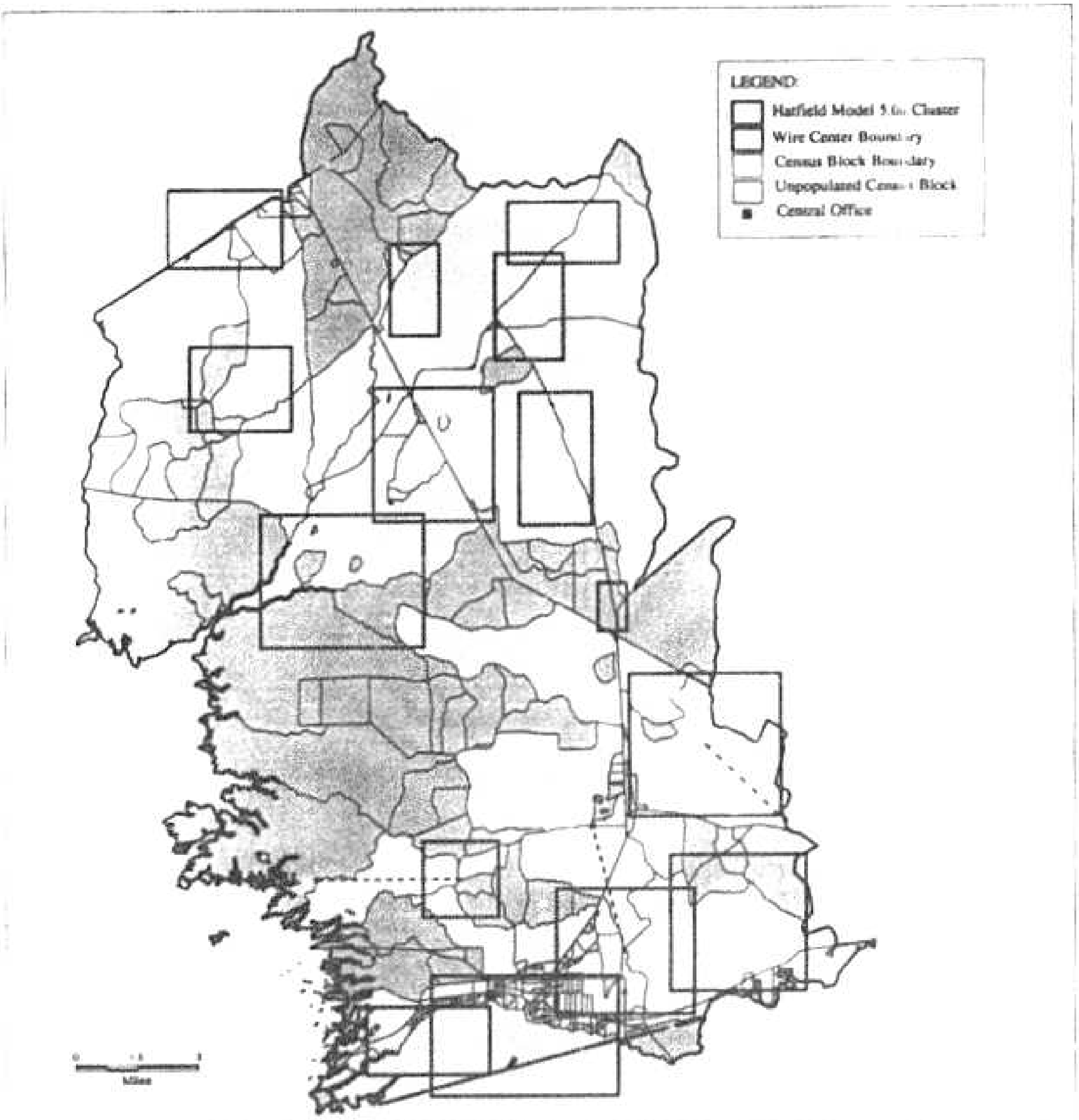
The second effect results from the branch and backbone cable length algorithm that essentially forces customer premises to be concentrated around the center lot(s) of the cluster. This results from the requirement that the backbone and branch cables extend no further than one lot depth (width) from the rectangle cluster's boundary. This constraint has the greatest effect on distribution route distance in large, low-density clusters where the individual lots are very large.

The bottom line conclusion is that HAI 5.0a is not placing enough distribution cable to serve customers in their "actual" locations, as identified by PNR's polygon clusters. This underplacement appears to be the most severe in the low-density clusters.

---

[5] HAI 5.0a actually models 1,674' of branch cable for this cluster. In calculating the branch cable length, HAI 5.0a refers to the aspect ratio for the rectangular cluster despite its inconsistency with the lot aspect ratio of 2 (see Distribution Module.xls, Calculations Sheet, column W).

Yankeetown Wire Center
Levy County, FL
Hatfield Model 5.0a Clusters



Wire Center FL 07991 01303
CLLI YNTWFLMA
Exhibit KDD-12

Rebuttal Testimony of
Kevin Duffy-Deno
Exhibit KDD-13

Figure 2. **Stylized PNR Polygon Cluster and the HAI 5.0a Equivalent Area Rectangle (Access Database)**

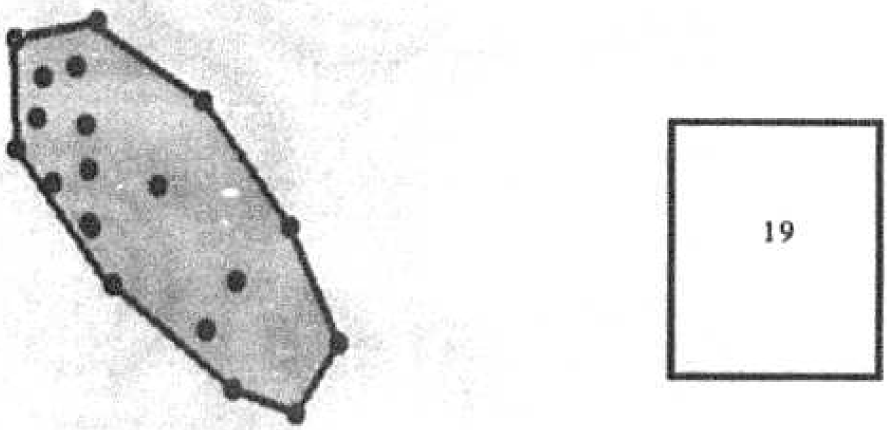
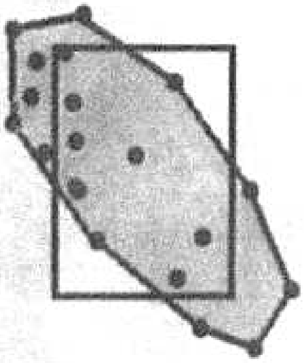


Figure 3. **Formation of the HAI 5.0a Rectangular Clusters**



Rebuttal Testimony of
Kevin T. Duffy-Deno
Exhibit KDD-14

Using Minimum Spanning Trees to Estimate Subscriber Dispersion and Minimum Network Length

Phil Bolian, Stopwatch Maps
For *INDETEC* International

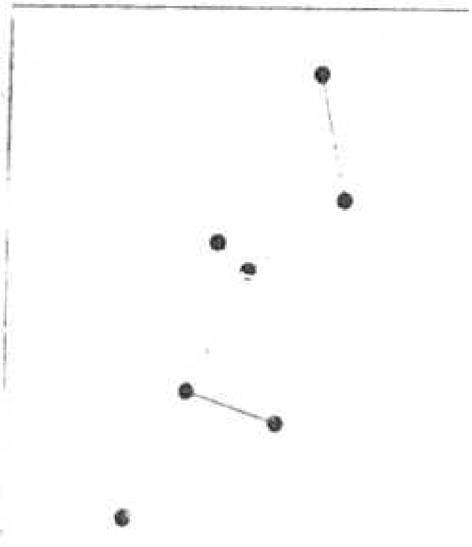
I. Background

A *Minimum Spanning Tree* is a construct from graph theory. It is commonly used in network design as a measure of the *dispersion* of the points to be served by a network, and as a benchmark for the *shortest possible* length of a network to serve those points.

For a set of *points* (we would say "subscriber locations"), a *Spanning Tree* is a set of straight line segments that connect *every* point (subscriber), simply drawing a line from one point to another, using no excess lines. If there are *N* points, there will necessarily be *N - 1* of these line segments.

The *Minimum* Spanning Tree of a set of points is that set of connecting line segments whose total length is the *shortest possible* for this set of points.

If you know the distance from every point to every other point in a set, it is not difficult to construct, and to determine the length of, the Minimum Spanning Tree of those points. The famous algorithm for calculating it, published in 1957 by R.C. Prim of Bell Labs[1], uses this simple logic:



A Minimum Spanning Tree

- First, find the two points that are closest to each other and connect them
- Then repetitively, until all points have been connected, find the shortest distance between any already-connected point and any not-yet-connected point, and connect those points

As Prim pointed out in his paper, there is one and only one *shortest total length.*

---

[1] R. C. Prim, "Shortest Connection Matrix Network and Some Generalizations," *Bell System Technical Journal*: 36, 1389-1401, November 1957

While the Minimum Spanning Tree seems a very satisfying measure of the degree of *dispersion* of a set of points, there are two objections we would make to its use in estimating a minimum possible *telephone network*:

- First, telephone networks are not constructed by chaining together one subscriber to another. Rather, a set of cables is run along as optimal a path as possible, and short drops from *terminals* connect those cables to subscribers. (Those terminals represent additional *points* in the network, introduced at will by the designer.) Perhaps one could construct a *shorter* network than a Minimum Spanning Tree when using this method.
- On the other hand, the line segments of a Minimum Spanning Tree run directly from one point to another. If those points represent real subscribers, these lines could possibly run across back lots and cow pastures, and through lakes, mountains, and tall buildings. Surely the Minimum Spanning Tree is a significant *understatement* of the realistic routing of network cable.

Both points have merit. Let's take them in order.

The Minimum Spanning Tree construct does not allow the introduction of additional points. That's what keeps the construct simple, and easy to calculate. The construct that attempts minimum total length by adding additional points as necessary is known as a *Steiner Minimum Tree*, named for the mathematician Jakob Steiner who posed this construction problem in designing road networks two centuries ago.

There are not many configurations of original points for which adding additional points (forming a Steiner Minimum Tree) will connect with less total length than a Minimum Spanning Tree, but there are *some*. Even in those special cases, however, there is an absolute limit to the improvement. In a paper published in 1990, D. Z. Du and Frank Hwang (Hwang is of Bell Labs) proved that adding extra interconnection points cannot reduce the total length of the tree by more than about 13 percent[2].

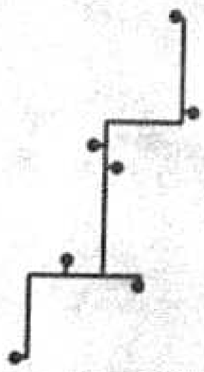
The calculation of a Steiner Minimum Tree for a large number of points is known to be a monstrous effort, taking immense amounts of computer time. Because it seldom improves on a Minimum Spanning Tree's length, and even then only slightly, the simple Minimum Spanning Tree calculation is regularly used as a benchmark for *shortest theoretical length*.

The second objection has greater significance, and illustrates why the Minimum Spanning Tree is simply a *benchmark* for, and not a realistic measure of, the shortest possible network. Because a

---

[2] D. Z. Du & Frank Hwang, "A Proof of Gilbert-Pollak's Conjecture on the Steiner Ratio", Publication 90-72 of the *Center for Discrete Mathematics & Theoretical Computer Science of Rutgers University*, 1990

Minimum Spanning Tree has no respect for rights-of-way, and a telephone network *must* respect them, the Minimum Spanning Tree regularly understates the minimum practical network length.



A More Realistic Tree

In this figure, we have constructed a more nearly realistic part of a network, running along what would be streets or roads. Even having tailored this sub-network to this exact set of points, we find the length of the tree in this figure to be 18% greater than the length of the Minimum Spanning Tree for those same points. To account for future growth, real telephone networks can not be tailored so tightly to a static set of customers, and are therefore even *less* efficient of length than in the illustration at the right.

We know that a common rule-of-thumb factor used by telephone engineers to convert arbitrary straight line distances (such as are used in a Minimum Spanning Tree) to realistic cable runs is the square root of 2, or 1.414. It would be no great leap to consider that a reasonable minimum network would be something like 1.414 times the length of the Minimum Spanning Tree of the points served.

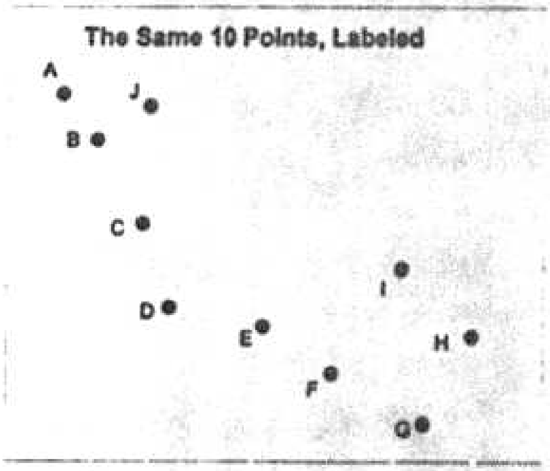
II. How a Minimum Spanning Tree Is Formed

The principal reason that a Minimum Spanning Tree is so much used as a measure of dispersion of a set of points is that it is a relatively easy metric to calculate.

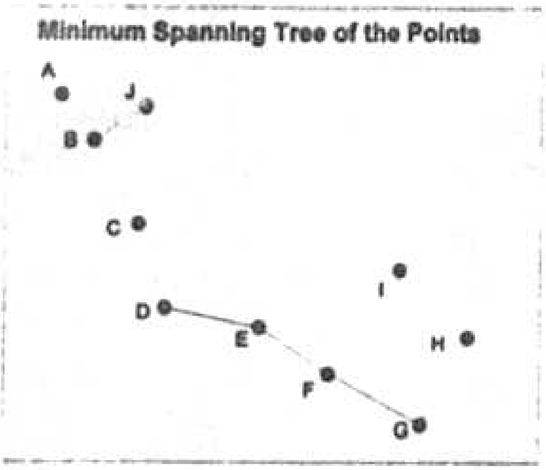
This section illustrates the calculation of a Minimum Spanning Tree for the ten points shown at the right, step by step.

So that we will be able to identify those points in this discussion, let's label each with a letter, as we show directly below.

Ten Customer Points



The Same 10 Points, Labeled



Minimum Spanning Tree of the Points

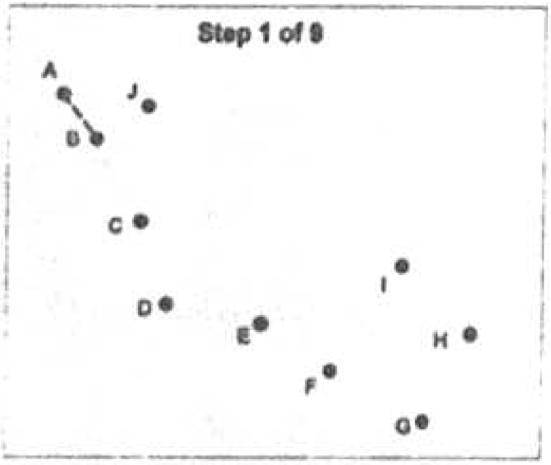
We've also shown, to the right above, the *resulting* Minimum Spanning Tree that we have calculated for these points. Even before we show the steps that get us to this tree, let's remember what a Minimum Spanning Tree *is* … it is the *shortest* set of line segments that can connect *all* the points of a group, using only those points themselves (not introducing any additional points).

The procedure for determining that shortest set of line segments is really very simple:

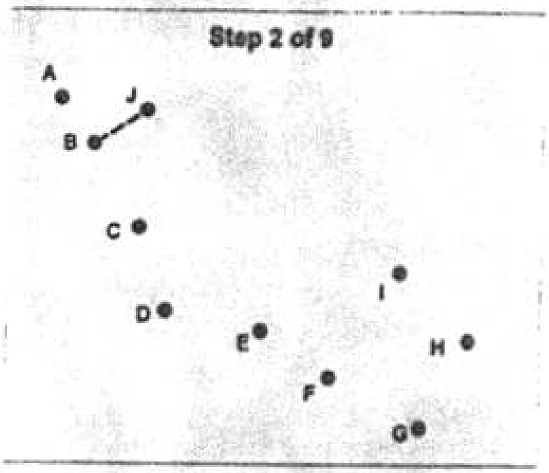
- First, find the *shortest of all distances* between any two points, and connect those two points
- Then, until all points have been connected, repeat the following: Determine the shortest remaining distance any *connected* point and any *not-yet-connected* point, and connect those two points

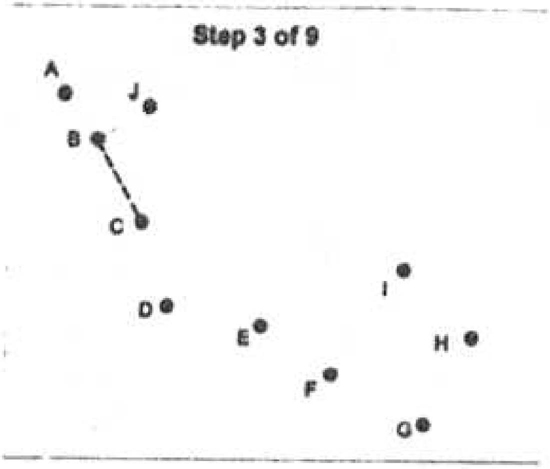
We haven't shown the actual distance numbers here, but the shortest distance between any two of these points is between A and B. So we'll begin by connecting those two.

The next step, the one we repeat over and over, requires us to determine the shortest distance between any already connected point and any not-yet-connected point. A and B are the already-connected points. The shortest distance from either of them to any other point is from B to J. So we'll follow the procedure and connect them, as we see directly below.



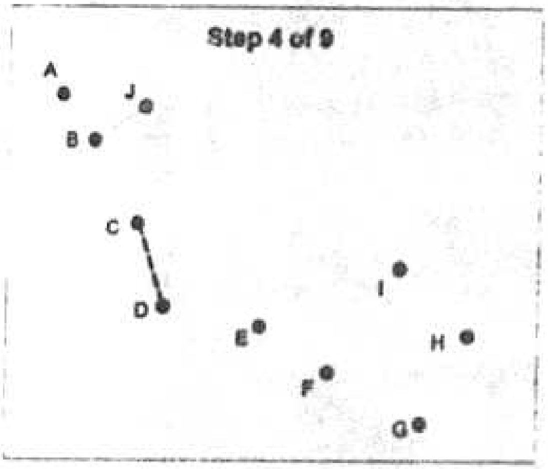
Step 1 of 9

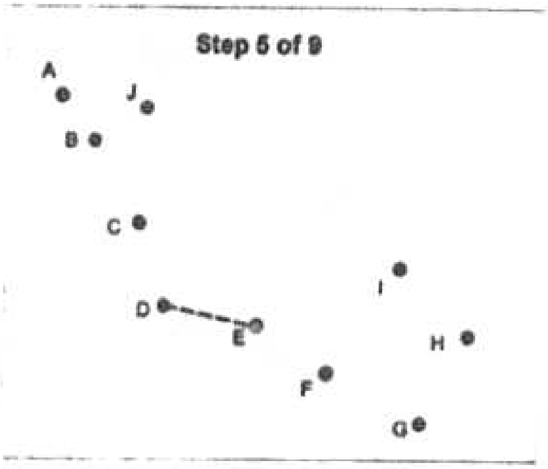

Step 2 of 9


Step 3 of 9

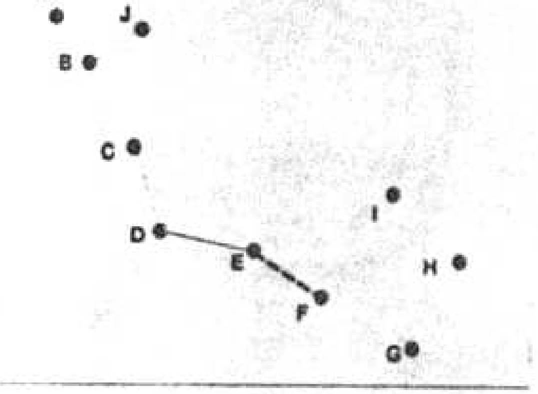
From **A, B,** and **J,** the shortest connection to any other point is from **B** to **C.** So we'll connect them, as seen on the right, above.

The process continues following the same rules until all points have been connected. We show the complete sequence below.

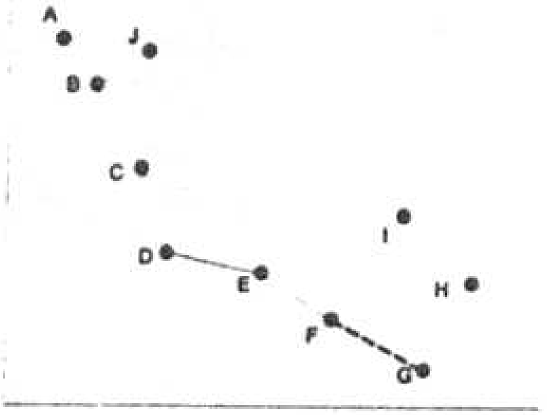

Step 4 of 9


Step 5 of 9

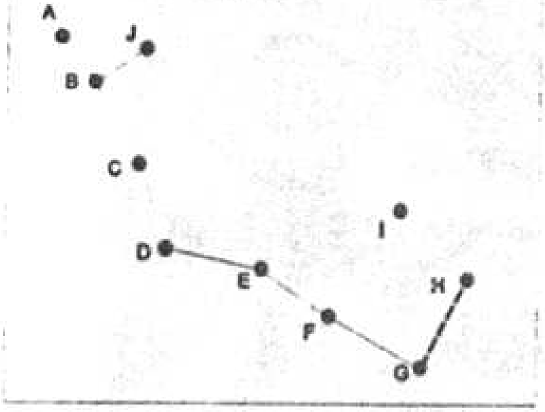
Step 6 of 9



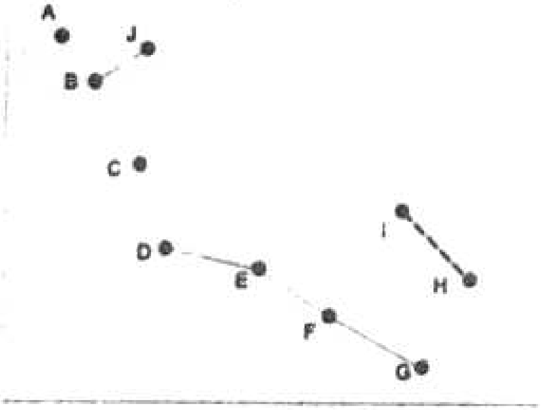
Step 7 of 9



Step 8 0f 9

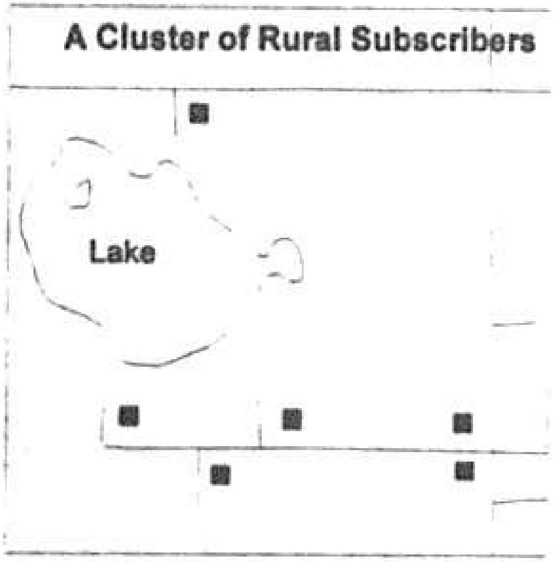


Step 9 of 9

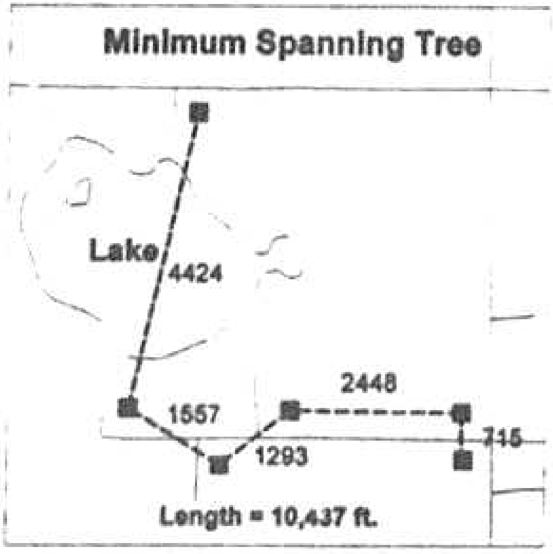


III. Minimum Spanning Tree vs. Actual Cable Route

Here is an example of the relation of Minimum Spanning Tree and a possible cable route to serve a cluster of subscribers in a rural area.



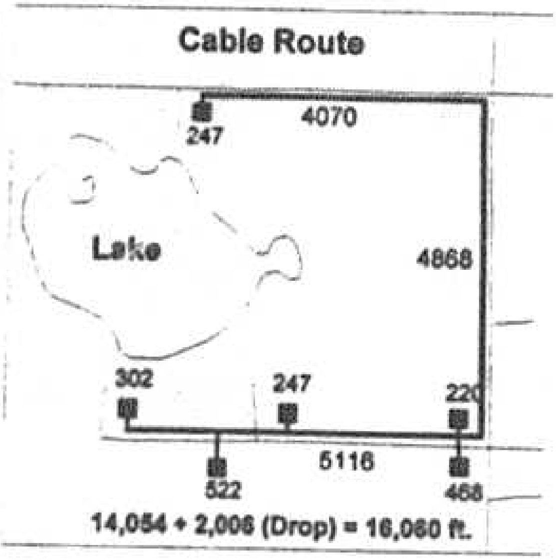
We must remember that Minimum Spanning Tree is an arbitrary, mathematical measure that has no respect for natural obstacles nor humanly restricted of rights-of-way. It simply measures the straight-line distance from one subscriber point to another, using the shortest set of straight lines possible. If that should lead through a cow pasture, a body of water, or a high mountain, the calculation does not care. And it certainly does not consider that cables basically run along roads … the calculation makes use of nothing other than the location of each of the points, and the distance of each point from each other.



So the Minimum Spanning Tree that would be produced for this configuration of subscribers is as shown at the right. The line segments connect the points from one to another, always with a straight line, and always using the shortest set of line segments possible. The fact that several of these line segments run obliquely across a road is natural … the calculation is not even aware of roads. And the fact that one of the segments runs across a lake is, once again, a natural result of a mathematical procedure that always seeks the shortest straight-line distances and knows nothing of obstacles.

Here we have shown the length, in feet, of each of the line segments of the Minimum Spanning Tree. The total length is 10,437 feet. We will be hard pressed to devise a realistic cabling route that can match that length, because cable routes – unlike abstract mathematical procedures – are compelled to honor natural and man-made restrictions.

The cable route is compelled to follow roads. In this case, we have run the cable along the side of the road that favors the largest number of subscriber points. We show here the length of each length of distribution cable, and the length of each drop. We find that to correspond to the connections of the Minimum Spanning Tree, we must use 14,054 feet of distribution cable and 2,006 feet of drops, a total of 16,060 feet.



Clearly this length is greater than that of the Minimum Spanning Tree for this set of points, just as we would expect it to be. In this case, the 16,060 feet is 1.54 times the Minimum Spanning Tree length of 10,437 feet, a significant multiplier.

The multiplier will vary with different configurations of subscribers in different natural and man-made settings. But it should be clear that except in the most trivial of circumstances the route distance is certain to be more than 1.0 times the Minimum Spanning Tree length.

Rebuttal Testimony of
Kevin Duffy-Deno
Exhibit KDD-15

The "Shorter-Than-Minimum-Spanning-Tree" Fallacy

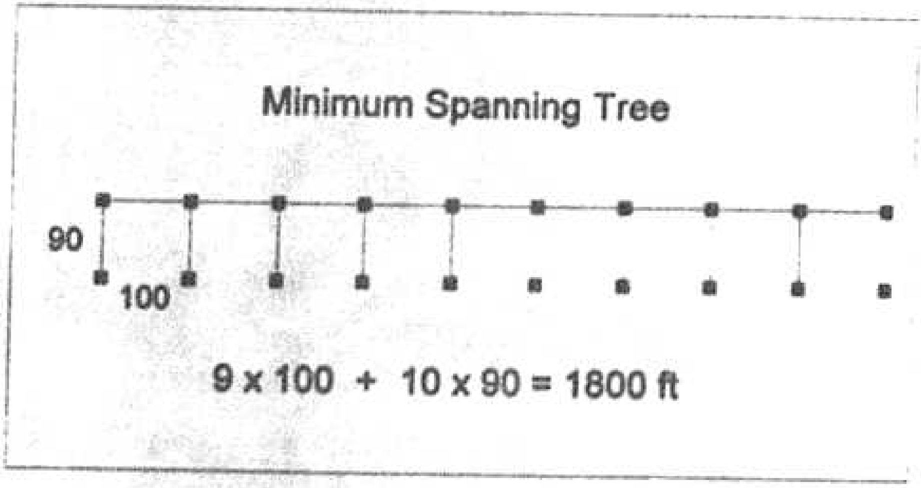
By Phil Bolian, Stopwatch Maps
For *INDETEC* International

It is certainly true that the classic Minimum Spanning Tree construct allows branches only at the existing nodes of a graph. It is also true that – in a few very special cases – the deliberate insertion of additional nodes might produce a slightly shorter tree than the Minimum Spanning Tree. In a telephone network, additional nodes may be introduced at will. Thus one might argue that it is at least *conceivable* that some cabling in a telephone network could be slightly shorter than the measure of a Minimum Spanning Tree. That argument would certainly require an example to illustrate the case. However, such examples are difficult to develop.

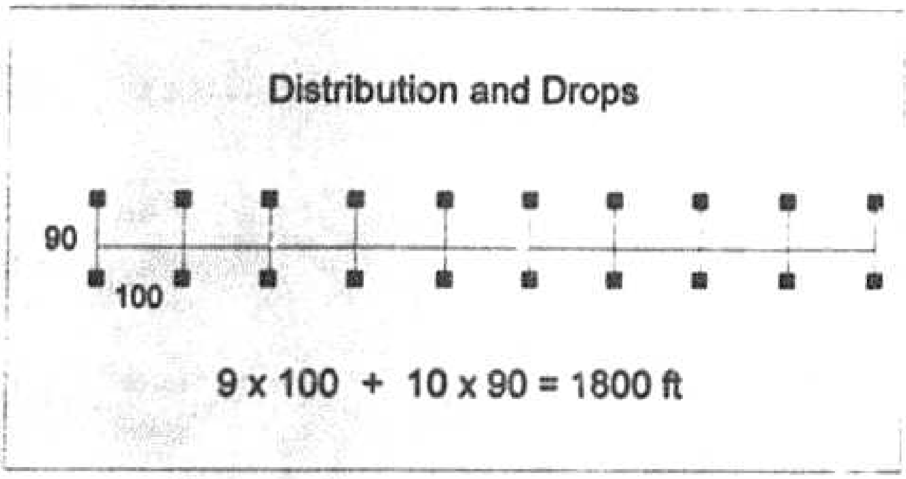
In a June 10, 1998 ex parte to the FCC, AT&T and MCI present an example purportedly illustrating part of a telephone network that uses less cable footage than the measure of the Minimum Spanning Tree for the subscribers to be served. The example is based on the premise that on a typical suburban street, running cable down one side (or the middle) of the street, and extend drops to each house, will yield less DRD [Distribution Route Distance] than the Minimum Spanning Tree distance.

Unfortunately for AT&T and MCI, the example they cite does *not* prove their point. In fact, it proves them wrong. Let's examine the circumstances AT&T and MCI cite.

Imagine a suburban block, with ten houses on either side of the street. Imagine them evenly spaced. In this first example, let the lot sizes be 100 feet, and let the distance from the front of one house to its cross-street neighbor be 90 feet (in a later example we'll reverse those distances). The Minimum Spanning Tree length for these original locations is 1,800 feet.

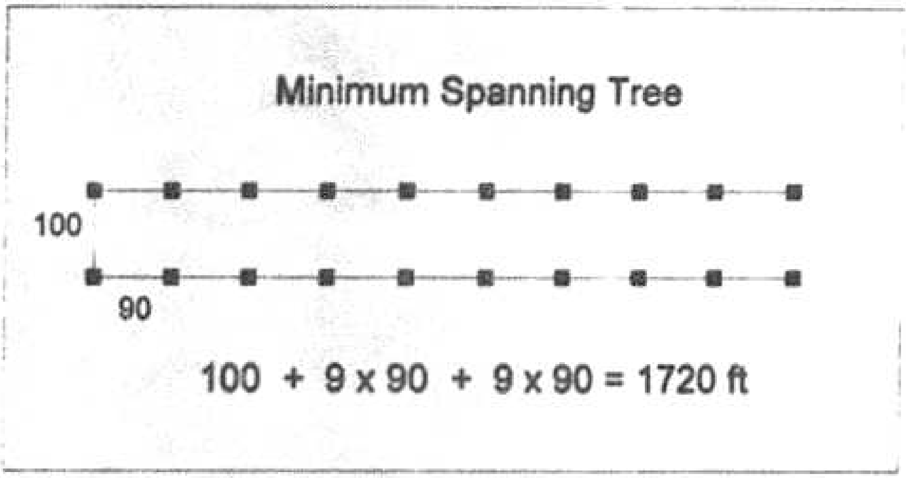


Now, if a single cable is run down one side (or the middle) of the street, and drops are extended to each house, the following configuration results. In this case, the DRD is *identical* to that for the Minimum Spanning Tree.


Distribution and Drops

9 x 100 + 10 x 90 = 1800 ft

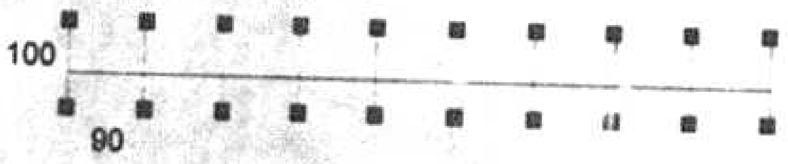
Now, let's reverse the numbers, such that the lot size is 90 feet and the distance to a cross-street neighbor is 100 feet.


Minimum Spanning Tree

100 + 9 x 90 + 9 x 90 = 1720 ft

The Minimum Spanning Tree by necessity runs the full block length through the houses on both sides of the street. In this case, when we construct the distribution and drop configuration we find that it is *longer*, not shorter, than the Minimum Spanning Tree. The Minimum Spanning Tree is, to be exact, 5% shorter than the configuration AT&T and MCI cite.

Distribution and Drop



9 x 90 + 10 x 100 = 1810 ft

Hence, it is quite difficult to improve upon the Minimum Spanning Tree distance.