Exhibit No. CMD–1
Docket No. UT–023003

BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

In the Matter of the Review of: Unbundled Loop and Switching Rates; the Deaveraged Zone Rate Structure; and Unbundled Network Elements, Transport, And Termination	Docket No. UT-023003

REPLY DECLARATION OF CHRISTIAN M. DIPPON FILED ON BEHALF OF VERIZON NORTHWEST INC.

APRIL 27, 2004

TABLE OF CONTENTS

I. INTRODUCTION

A. Qualifications

Q. PLEASE STATE YOUR NAME, BUSINESS ADDRESS, AND TITLE.

A. My name is Christian M. Dippon. My business address is One Front Street, Suite 2600, San Francisco, CA 94111. I am an economist and Senior Consultant of National Economic Research Associates, Inc. ("NERA"), an international economic consulting firm.

Q. ON WHOSE BEHALF ARE YOU PRESENTING TESTIMONY IN THIS PROCEEDING?

A. I am presenting testimony on behalf of Verizon Northwest Inc. ("Verizon NW").

Q. PLEASE DISCUSS YOUR EDUCATION AND EXPERIENCE.

A. I received a B.S. in Business Administration from the California State University in 1993, and an M.A. in Economics with a concentration in microeconomics and econometrics from the doctoral program of the University of California in 1995. Before joining NERA in 1996, I was an analyst at BMW in Bangkok, Thailand.

I have provided economic consulting services and written testimony in numerous state regulatory proceedings, and have submitted affidavits and declarations to the Federal Communications Commission ("FCC"). I have also testified before state regulators on Unbundled Network Element ("UNE") and Universal Service Fund ("USF") issues.

Over the past few years, I have been retained numerous times to analyze and address the economics of UNE pricing with a particular focus on determining and locating demand for local telephone service. I have analyzed numerous versions of the HAI Model (formerly the Hatfield Model), and have submitted written and oral testimony with respect to this cost model on at least 20 occasions. Recently, I analyzed the HAI Model, Release 5.3 ("HM 5.3" or "Model") and presented my findings to the Public Utilities Commission of the State of California ("California PUC") on behalf of SBC California Bell Telephone Company ("SBC CA").[1]

My publications include articles on estimating costs for local telephone service, regulation, and price setting in the United States and other countries. A copy of my curriculum vitae is attached as Exhibit CMD-2.

B. Purpose of the Reply Testimony

Q. WHAT IS THE PURPOSE OF YOUR REPLY TESTIMONY?

A. The purpose of my Reply Testimony is to describe and present my analyses of the processes used by AT&T Communications of the Pacific Northwest, Inc. ("AT&T") and WorldCom, Inc. ("MCI") (collectively "AT&T/MCI") and their subcontractor, TNS Telecoms ("TNS"), to yield HM 5.3's customer location database.[2] I will explain why this database is conceptually, technically, and

[1] *See* Before the California Public Utilities Commission, Application Nos. 01-02-024, et al., *Reply Declaration of Christian M. Dippon on behalf of SBC California Bell Telephone Company,* (Feb. 7, 2003) ("Dippon SBC Declaration").

[2] TNS Telecoms is part of Taylor Nelson Sofres, a company that provides market information services.

factually flawed, and why HM 5.3 cannot properly be used by the Washington Utilities and Transportation Commission ("Commission") to estimate Verizon NW's forward-looking costs of UNEs. I will demonstrate how HM 5.3 derives UNE cost estimates based on a fantasy network design that enjoys unrealistic economies of scale, violates the most basic engineering constraints, and defies common sense. In addition, I will address the claims of AT&T/MCI's witnesses Dr. Robert A. Mercer and Mr. John Donovan regarding the development and accuracy of HM 5.3's customer location database specifically, and HM 5.3 in general.[3] I will prove theoretically, visually, and statistically that Dr. Mercer's much-touted "highly sophisticated costing tool" is nothing more than an artifice to obtain Verizon NW's UNEs at a fraction of their forward-looking cost.[4] Finally, I will provide clear evidence that Verizon NW's proposed cost model, VzCost, employs a far superior modeling approach to outside plant.[5]

C. Summary of Findings

Q. BRIEFLY SUMMARIZE YOUR REPLY TESTIMONY.

A. HM 5.3 and its cluster input database are flawed in at least three respects: conceptually, technically, and factually. Conceptually, the Model develops cost

[3] *See* Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Supplemental Direct Testimony of Dr. Robert A. Mercer on behalf of AT&T Communications of the Pacific Northwest, Inc.*, (Jan. 23, 2004) ("Mercer Supplemental Direct Testimony"); Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Direct Testimony of John C. Donovan on behalf of AT&T Communications of the Pacific Northwest, Inc. WorldCom, Inc., and XO Washington, Inc.,* (June 26, 2003) ("Donovan Direct Testimony").

[4] Mercer Supplemental Direct Testimony at p. 5.

[5] The modeled network in VzCost is created by VzLoop, VzCost's loop investment calculator.

estimates for a utopian network that even AT&T/MCI admit cannot realistically be built. HM 5.3 assumes that Verizon NW's customers are uniformly spread in rectangular-shaped distribution areas—an assumption that is entirely divorced from reality. Each of these rectangular-shaped distribution areas is assumed to contain lots of equal-size and shape that are uniformly dispersed within the distribution area. Further, the Model also assumes that each of these lots has the same line demand and an identical dispersion of equal-sized distribution terminals. HM 5.3 ignores the numerous cable types and sizes deployed in real-world networks, employing generally only two types of cables and cable sizes to serve the lots in its distribution areas. The Model does not take into account rights-of-way, and disregards entirely physical obstacles and manmade obstructions (such as rivers, highways, freeways, and mountains) when it places outside plant. These overly simplistic and arcane modeling techniques ignore crucial cost drivers and yield unrealistic economies of scale—the result being insufficient investment and artificially low UNE cost estimates.[6]

HM 5.3 does not fair any better technically than it does conceptually. HM 5.3's cost estimates are almost entirely insensitive to important cost drivers. For example, one would expect the number of modeled distribution areas to have a significant effect on costs, since it has a direct impact on feeder and serving area interface ("SAI") investment. However, HM 5.3's cost estimates change only minimally when this number is increased three-fold. Even more significant

[6] A "cost driver" is a variable that has a significant affect on total costs. That is, a change in a cost driver should change the total cost of a related cost object.

is the fact that HM 5.3's cost estimates often *decrease* as the number of clusters increases. The same counter-intuitive result occur when the clustering algorithm is replaced with the flawed clustering rule used in Version 2.2.2 of the HAI Model ("HM 2.2.2"). That is, replacing TNS's much-touted and closely-guarded clustering algorithm with a clustering rule that is known to be inaccurate yields only slightly different cost results. HM 5.3's minimal sensitivity to this and other cost drivers does not mean that these variables are not important; rather, it illustrates how HM 5.3's cost estimates are predominately driven by the overly simplistic and arcane modeling assumptions embedded in the modules that determine HM 5.3's outside plant network (i.e., preprocessing, feeder, and distribution modules). In addition to these counter-intuitive results, the Model's cluster input database suffers from sloppy data preparation. For instance, customers are assumed to live along freeways, including the off- and on-ramps, and along private driveways and service roads.[7]

Finally, the Model's factual failings are best illustrated with maps of the outside plant network modeled by HM 5.3. These maps expose the network upon which AT&T/MCI base their UNE cost estimates, and demonstrate conclusively that this network is entirely quixotic, and produces cost estimates that are entirely unrealistic. Verizon NW's witnesses Dr. Timothy Tardiff, Mr. Francis

[7] Obviously, there are no customers that live along any freeways or their on- and off-ramps. Similarly, customers are rarely, if ever, located along service roads, which are typically not zoned for residential or commercial use. By definition, private driveways are exclusive to one owner and will generally not have multiple houses or businesses along them.

Murphy, and Mr. Willett Richter identify scores of additional reasons why HM 5.3 should be rejected.[8]

Contrasting Verizon NW's proposed cost model VzCost's method of modeling outside plant (contained in VzLoop) to HM 5.3's method of cost modeling leaves no doubt that VzCost is superior to HM 5.3. Most notably, maps of VzLoop's modeled outside plant demonstrate how VzCost, unlike HM 5.3, follows feasible network routes by generally avoiding physical obstacles and boundaries, accounting for rights-of-way, and thereby producing representative investment estimates of a forward-looking network in the State of Washington. I therefore recommend that this Commission adopt Verizon NW's cost model instead of HM 5.3.

D. Outline of the Reply Testimony

Q. HOW IS YOUR REPLY TESTIMONY STRUCTURED?

A. My Reply Testimony is structured as follows. Section II describes the objective of and the general processes that yield HM 5.3's cluster input database. Section III explains why the cluster input database is one of the most important aspects of HM 5.3's modeling of outside plant and, as such, has a significant influence on HM 5.3's estimated loop cost. Section IV addresses the errors and

---

[8] *See* Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Reply Testimony of Timothy J. Tardiff on behalf of Verizon Northwest Inc.* (April 26, 2004) ("Tardiff Reply Testimony"); Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Reply Testimony of Francis J. Murphy on behalf of Verizon Northwest Inc.* (April 26, 2004) ("Murphy Reply Testimony"); Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Reply Testimony of Willett G. Richter on behalf of Verizon Northwest Inc.* (April 20, 2004) ("Richter Reply Testimony").

fundamental modeling flaws embedded in the cluster input database. This section also discusses the cost impact of these errors, and illustrates why these errors constitute much more than simple modeling inaccuracies. Section V contrasts HM 5.3's outside plant modeling efforts to those used in VzCost, and demonstrates that VzCost's modeling of outside plant is superior to HM 5.3. Section VI presents my conclusions.

Q. PLEASE LIST THE EXHIBITS TO YOUR TESTIMONY.

A. There are a number of exhibits to this testimony. These include:

- Exhibit CMD-2: Curriculum Vitae of Christian Dippon
- Exhibit CMD-3: Letter from Christopher Huther, Esq., counsel for Verizon NW to Gregory Kopta, Esq., counsel for AT&T
- Exhibit CMD-4: Email correspondence between Verizon NW and AT&T regarding the authenticity of the TNS preprocessing data
- Exhibit CMD-5: HM 5.3 Preprocessing Sensitivity Analyses
- Exhibit CMD-6: CD-ROM Containing Maps of HM 5.3's and VzCost's Modeled Outside Plant Network
- Exhibit CMD-7: Christian M. Dippon and Kenneth E. Train, "The Cost of the Local Telecommunications Network, A Comparison of Minimum Spanning Trees and the HAI Model," *Telecommunications Policy,* Vol. 24, No. 3 (April 2000)

II. HM 5.3's Cluster Input Database

Q. WHAT IS YOUR UNDERSTANDING OF THE PURPOSE OF THE CLUSTER INPUT DATABASE?

A. The cluster input database serves as the starting point for the so-called modeling done by HM 5.3. It contains much, if not all, of the information HM 5.3 uses to model a network. If cost is the product of quantity and price, the cluster input database contains most of the critical data that HM 5.3 uses to determine quantities. This information is hard-coded in the cluster input database, and is the result of an enormous amount of unverifiable, largely undocumented, and convoluted preprocessing steps that are done outside the Model by TNS and AT&T/MCI. From this perspective, the modeling done by HM 5.3 is merely the final stage of an obscure process that essentially starts with the modeled network plant already in place.

Dr. Mercer touts HM 5.3's cluster input database as one of the most profound improvements over previous versions of the Model. Specifically, Dr. Mercer claims:

> The processes for locating and clustering customers to form such serving areas for HM 5.3 represent state-of-the-art modeling technology developments that have profoundly improved the accuracy of HM 5.3.[9]

As I will show, this statement is misleading at best. HM 5.3 does not contain any actual customer locations, does not model any plant to a single Verizon NW customer, and hardly uses state-of-the-art modeling techniques.

Q. HAVE YOU ANALYZED ALL ASPECTS OF HM 5.3'S CLUSTER INPUT DATABASE?

A. No. Pursuant to the ALJ's and Commission's orders granting Verizon NW's Motion to Compel, AT&T/MCI were directed to produce all of the processes yielding the HM 5.3 cluster input database, including the source code to the clustering algorithm, when they filed the revised version of HM 5.3 on January

[9] Mercer Supplemental Direct Testimony at p. 21.

26, 2004.[10] However, the January 26th cost model submission failed to include the preprocessing data the Commission had ordered and AT&T/MCI had agreed to produce. It was not until Verizon NW's counsel reminded AT&T/MCI of their obligation to produce these data that AT&T agreed to do so.[11] On March 4, 2004—almost six weeks after the information was to have been provided—Verizon NW received a DVD from AT&T/MCI containing 675 files that were allegedly responsive to Verizon NW's unanswered requests. After analyzing some of these files, I discovered that these files *did not* yield the cluster input database used in the version of HM 5.3 filed on January 26th, thus precluding me from undertaking my intended analyses of HM 5.3's preprocessing. When notified, AT&T and TNS repeatedly insisted that the data on the DVD were the files that yielded the January 26th HM 5.3 cluster input database.[12] Eventually, AT&T and TNS conceded that the data were flawed, and provided Verizon NW with what purported to be a corrected DVD on April 8, 2004. This DVD was not in a readable format. It was not until April 12, 2004, that Verizon NW obtained a corrected and working version of the DVD—11 weeks after AT&T/MCI were ordered to produce the files. Administrative Law Judge Theodora Mace briefly extended the filing date because of this delay, but I still have not had adequate time to review the data in detail. Moreover, the

---

[10] *See* Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Fourteenth Supplemental Order: Denying Petition for Review of Interlocutory Order, Granting Motions to Compel* (Oct. 14, 2003) at p. 8.

[11] *See* Letter from Christopher Huther, Esq., counsel for Verizon NW to Gregory Kopta, Esq., counsel for AT&T (Feb. 13, 2004), attached hereto as Exhibit CMD-3.

[12] *See* Email correspondence between Verizon and AT&T, attached hereto as Exhibit CMD-4.

corrected version of the DVD contained files that were minimally documented, thereby making it extremely difficult, even for an experienced user, to decipher the sequence of the processes let alone understand what each of the steps entailed. Thus, Dr. Mercer's claims that HM 5.3 is state-of-the-art, that its inputs are extensively documented, and that the Model is straightforward to use are misleading to say the least.[13]

Q. DID THE CORRECTED DVD CONTAIN ALL THE FILES NECESSARY TO REVIEW HM 5.3'S PREPROCESSING?

A. No, not even the corrected DVD obtained on April 12, 2004, contained all the necessary files to review the Model's preprocessing. Most notably, AT&T/MCI refused to provide access to the clustering algorithm's source code, claiming that it was not in their possession, custody, or control as it is the intellectual property of TNS and commercially available to Verizon.[14] Without the source code for the clustering algorithm, however, it is impossible for Verizon NW, the Commission, or any other party to this proceeding to review HM 5.3's clustering algorithm, perform sensitivity tests that require algorithm modifications, or validate whether HM 5.3 performs as AT&T/MCI describe. Access to the clustering algorithm's source code is particularly important as the Model's limited documentation is clearly at odds with how the clustering algorithm

---

[13] *See* Mercer Supplemental Direct Testimony at p. 32.

[14] *See* Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Joint Responses of AT&T & MCI to Verizon's First Set of Data Requests* (July 24, 2003) at Response Nos. 1-10, 1-11, 1-12, 1-13, 1-20 ("Joint Responses to Verizon's First Set of Data Requests"). Although AT&T/MCI say that this product is commercially available, I do not know of any way to purchase it or of anyone else ever being able to purchase it.

seems to function. In short, critical aspects of HM 5.3's preprocessing cannot be verified and, in this respect, HM 5.3 remains a "black box."

Q. PLEASE DESCRIBE YOUR UNDERSTANDING OF HM 5.3'S PREPROCESSING.

A. My understanding of HM 5.3's preprocessing is limited to what has been made available to me. As stated, I have been denied access to the clustering source code, and in general have had to rely on very limited information as to the functioning and objectives of many other files. Nevertheless, it is clear that the cluster input database is one of the most important cost drivers in HM 5.3, as it lays the foundation for the network being modeled. If that foundation is inaccurate, the cost estimates produced and the corresponding UNE prices will be inaccurate as well. Accordingly, reviewing and validating the cluster input database is a necessary prerequisite to any analysis of HM 5.3.

Based on a review of Dr. Mercer's declaration and the data provided by AT&T/MCI in discovery, my understanding of HM 5.3's preprocessing is as follows:

1. AT&T/MCI and their subcontractor TNS start with three Verizon NW customer service address databases.
2. TNS then removes test circuits, duplicate addresses, and records associated with wire centers outside Verizon NW's territory. The resulting database contains 1.166 million records.
3. Verizon NW offers approximately 150 different service types. TNS groups these service types into 11 broad categories.

Type	Description
A	ADSL Shared Lines
B	Business Lines
C	ADSL Dedicated Lines
E	Individual NS Lines
I	ISDN Lines
P	Public Lines
R	Residential Lines
T	NS DS-1 Service
U	SW DS-1 Service
V	HC Optical Services
W	DS-3 Optical Services

4. Next, TNS attempts to geocode (i.e., assign a longitude and latitude to) the 1.166 million circuits using Centrus Desktop Version 4.01 and a version of the GDT street reference database.

5. TNS then builds wire center boundary files for each wire center.

6. Using various programs, TNS determines which customer records were successfully geocoded and which ones need to be located using a surrogate method (i.e., spread along roads within a wire center or Census Block ("CB")).

7. The customer records in the above step that cannot be geocoded are placed using a surrogation algorithm. It appears that TNS uses a version of the TIGER street reference database (as opposed to the GDT street reference database used for geocoding) to surrogate these locations.[15]

8. The 1.166 million geocoded and surrogated customer records are then collapsed by longitude and latitude and by service type. That is, all circuits with the same longitude and latitude and that belong to the same service type are treated as one record, with demand (number of lines) equal to the sum of the records collapsed. The collapsed residential records are referred

[15] TIGER, or Topologically Integrated Geographic Encoding and Referencing, is the name for the system and digital database developed by the U.S. Census Bureau.

to as “households” and the collapsed business records are referred to as “businesses.” Thus, in this step, TNS seems to determine the number of households and businesses in Verizon NW’s territory. This procedure reduces the 1.166 million circuits to 656,119 locations.

9. The 656,119 locations are then clustered using TNS’ clustering algorithm. It appears that TNS uses Version 5.0 of its clustering program. The clustering program seems to do more than just clustering. In addition to assigning customers to distribution areas (or clusters), it also appears to: (1) calculate the strand distance (or rectilinear minimum spanning tree), (2) determine the locations of the SAIs, and (3) through a number of steps, make each distribution area rectangular.[16] Locations that exceed 536 lines (a hard-coded value in the undisclosed source code of the clustering algorithm) are assumed to form a “high-rise” cluster.

10. Using three FoxPro programs, TNS then converts the resulting files into a format that can be read by the next step, which is the PointCode process.[17]

11. In what appears to be the final step, the output from the FoxPro programs is run through a collection of seven Microsoft Access databases, each containing a number of queries. Importantly, in the first stage of this process, TNS removes all locations that are not residential or business locations. That is, all “non-R “and “non-B” service types listed in step 3 above are removed from the household and business counts for the cluster, although the demand remains. This reduces the number of locations from 656,119 to 579,375 households and businesses. Specifically, TNS estimates that there are 132,535 businesses and 446,840 residential locations in Verizon NW’s serving territory.

[16] AT&T/MCI and TNS use the terms “clusters” and “distribution areas” interchangeably.

[17] Microsoft’s Visual FoxPro is an application development tool for building database applications.

12. The output from the PointCode program serves as the input into HM 5.3. As recently revealed, AT&T further manipulates this file before it serves as the cluster input database for HM 5.3. This final manipulation by AT&T ostensibly serves as a "line true up" and is primarily manual, thus it cannot be easily replicated.[18]

In HM 5.3's distribution module, the 579,375 locations (from step 11) are further reduced by assigning them to lots. AT&T/MCI estimate the number of housing types (i.e., single-detached houses, duplexes, multidwelling units, etc.), each of which contains numerous households. Then, the Model somehow determines that households in some housing types occupy one lot, while others occupy half a lot, and still others are on a quarter of a lot. These lot allocations are entirely unsupported; and, as a result of this conversion, the 579,375 locations are ultimately assigned to 437,027 lots. HM 5.3 models outside plant to these 437,027 lots, all of which are entirely divorced from the actual locations of Verizon NW customers.[19]

In short, the preprocessing module and HM 5.3 reduce Verizon NW's 1.166 million customer accounts to 437,027 lots, which are uniformly distributed within rectangular-shaped distribution areas. As I will explain further below, this

---

[18] *See* Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Joint Responses of AT&T and MCI to Verizon's Ninth Set of Data Requests* (Mar. 26, 2004) at Response No. 9-62.

[19] In particular and as described later, HM 5.3 uses information on the number of lots to determine the initial lengths and numbers of branch and backbone cables, which are subsequently overridden by then constraining total lengths within a cluster to equal the strand-distance measure included in the TNS database.

method of modeling outside plant is absurd and leads to a significant understatement of Verizon NW's forward-looking costs.

III. THE CLUSTER INPUT DATABASE IS ONE OF THE MOST IMPORTANT ASPECTS OF HM 5.3'S OUTSIDE PLANT MODELING

A. Most Aspects of Outside Plant Are Determined in the Preprocessing

Q. WHY IS THE CLUSTER INPUT DATABASE CRITICAL TO HM 5.3'S COST ESTIMATES?

A. The cluster input database developed by TNS is unarguably one of the most important aspects of HM 5.3's modeling and costing of outside plant. It serves as the foundation of the Model from which most, if not all, costs components of outside plant are derived. Dr. Mercer has characterized the importance of the cluster input database as follows:

> HM 5.3 dramatically improves the way in which it determines customer locations and configures outside plant to serve those customers, using data that have been provided by Verizon itself that contain customer locations for all types of loops. Using that data allows HM 5.3 to model not only a network used to provide POTS and other narrowband services, but also additional loop types, such as DS-1, DS-3, and ADSL loops … As a result, HM 5.3 models the network more accurately and more completely, and more accurately assigns joint costs of outside plant structure, network operations, and general support to this wider range of loop elements than could prior versions of the Model.[20]

Similarly, MCI economist, Dr. Michael Pelcovits, views the cluster input database as the "key initial driver:"

[20] Mercer Supplemental Direct Testimony at p. 7.

> The HAI model constructs a "bottom up" estimate of the UNE costs based upon detailed data describing demand quantities, network component prices, operational costs, network operations costs, and other factors affecting the costs of providing local service. The model's demand data, particularly data describing customer locations, line demand, and traffic volumes, serve as key initial drivers.[21]

There are few, if any, values in the modeling of outside plant that are not either directly determined by the preprocessing, or at least significantly impacted by it. Among the cost drivers directly determined by the preprocessing module are:

- Number of distribution areas
- Density zone designation, which in turn determines the structure mix and structure sharing percentages
- Size of distribution areas
- Size of modeled network
- Number of business and residential locations
- Number of drop cables
- Number and location of SAIs
- Number of indoor SAIs
- Number of high-rise buildings in Verizon NW's territory
- Demand by distribution area
- Demand distribution (i.e., uniform distribution of demand within clusters)
- Number of households by distribution area
- Number of firms by distribution area

[21] Before the Federal Communications Commission, WC Docket No. 03-173, *Declaration of Michael D. Pelcovits on Behalf of MCI* (Dec. 16, 2003) at p. 40 ("Pelcovits Decl.").

These cost drivers, and potentially more, are all hard-coded in the Model, and predetermine much of HM 5.3's UNE cost estimates for Verizon NW. While Dr. Mercer might (incorrectly) claim that the cluster input database captures Verizon NW's forward-looking network more completely and accurately, all this is done outside the Model, and outside of the review of other parties and the Commission. Moreover, AT&T/MCI have repeatedly failed to provide any support (let alone proof) for their assertion that the Model, in fact, produces more accurate and complete results. As I will demonstrate, the cluster input database, while certainly a major cost driver, is severely flawed, and thus renders HM 5.3's cost estimates useless.

Q. IF MOST DECISIONS ABOUT OUTSIDE PLANT MODELING ARE CONTAINED IN THE CLUSTER INPUT DATABASE, WHAT IS THE PURPOSE OF HM 5.3?

A. HM 5.3 merely fills-in-the-blanks; that is, the Model simply produces cost estimates for distribution areas that have been predetermined in the cluster input database. In terms of outside plant, HM 5.3 starts out with a detailed, albeit incorrect and unrealistic, blueprint of the modeled network derived from the TNS preprocessing, and then attempts to determine what it costs to build that network. AT&T/MCI give the impression that the outside plant modeled by HM 5.3 can be changed with user-adjustable inputs. This is wrong. The preprocessing module largely determines the layout (and hence the costs) of the modeled network, and there is not a single user-adjustable input contained in HM 5.3 that is capable of curing the network design produced by the Model's extensive preprocessing.

Q. WHAT USE ARE THE OVER 2,100 USER-ADJUSTABLE INPUTS IN HM 5.3?

A. Dr. Mercer touts the flexibility of HM 5.3 by stating that there are over 2,100 user-adjustable inputs.[22] At least in terms of the outside plant configuration, the user-adjustable inputs are of little to no use. None of the major components of the modeled network (e.g., size of distribution areas, number of indoor SAIs, and density designations) are user-adjustable.[23] As I will discuss below, what Dr. Mercer calls the Model's "groupings of customer locations that have a realistic correlation to efficient distribution areas" is also not user-adjustable.[24] Thus, even with 2,100 user-adjustable inputs, HM 5.3 is inflexible in terms of the fundamental layout of the modeled outside plant.

B. HM 5.3's Modeled Network Does Not Model to a Single Verizon NW Customer Location

Q. DOES HM 5.3 MODEL OUTSIDE PLANT TO ACTUAL VERIZON NW CUSTOMER LOCATIONS?

A. Absolutely not. The HM 5.3 Model Description claims:

> Customer locations are determined through geocoding, augmented as necessary by a surrogate location process for those customers whose geocoded location are not known. A clustering algorithm is used to develop groupings of customer

[22] *See* Mercer Supplemental Direct Testimony at p. 29.

[23] Similarly, a new variable to HM 5.3 that supposedly limits the size of the SAI and splits clusters if necessarily does not address any of the fundamental design parameters hard-coded in the cluster input database.

[24] Mercer Supplemental Direct Testimony at Exhibit RAM-4 (HAI Model Release 5.3 Model Description ("Model Description"), p. 3.

> locations that have a realistic correlation to efficient distribution areas…. Using these data, the Model calculates required network investment by detailed plant category.[25]

Similarly, Dr. Mercer claims: "Based on the customer location data, and detailed and granular information as to the existing demand for services, the Model then constructs a network to serve the identified demand."[26] These statements leave the false impression that HM 5.3 constructs plant to actual Verizon NW customer locations. This is not true. While HM 5.3's preprocessing starts out with Verizon NW's customer locations, after clustering these locations all the information about actual customer locations is discarded. Only marginal information (i.e., the area and aspect ratio of the convex hull surrounding a cluster and the location of the modeled SAI) is *initially* retained. Much of this information, however, is also discarded when the Model develops its erroneous distribution route distances by overwriting the results that were originally based on these data with the strand-distance multiplier. Thus, very little information gained from the elaborate geocoding and surrogating exercise, if any at all, is used to derive the final UNE cost estimates. The clustering process, the altering of clusters into rectangular shapes, the assumption that demand is uniformly distributed within a cluster, the reliance on only dominant Census Block Group ("CBG") information for density information, and the use of the strand-distance true-up, all cause any information about the original location of

[25] Mercer Supplemental Direct Testimony at RAM-4 (Model Description) at p. 3.
[26] Mercer Supplemental Direct Testimony at p. 12.

customers to be lost. Clearly, HM 5.3 does not rely on Verizon NW's customer locations.

Q. DO YOU AGREE WITH DR. MERCER'S DESCRIPTION OF HOW HM 5.3'S DISTRIBUTION MODULE WORKS?

A. No. Dr. Mercer states that:

> The Distribution Module lays distribution plant directly over the rectangular areas where customer clusters are located. This plant extends from the SAI location (or locations) to the customer premises in the cluster.[27]

There are at least two inaccuracies in this statement. First, as illustrated in the maps of HM 5.3's modeled network, attached hereto as Exhibit CMD-6, the rectangular distribution areas do not reflect any actual customer locations. Rather, the Model assumes that customer locations are situated in so-called lots, which in turn are distributed uniformly in a rectangular cluster. Further, instead of placing the SAI at the line-weighted centroid of the convex hull of actual customer groupings, AT&T/MCI place the SAI at half the distance between the two farthest customer locations and assume that this location is the "centroid" of the rectangular cluster. While I will elaborate on this below, this incorrect determination of the centroid's location further distorts the distribution of demand because it then moves customer locations around a false center point. While some of the geocoded and surrogated customer locations might fall into this rectangular area, many others do not. In fact, the rectangular clusters often fall outside wire center boundaries and/or overlap with other

[27] Mercer Supplemental Direct Testimony at RAM-4 (Model Description) at p. 34.

clusters. Further, even if the rectangular area contained customer locations, these customer locations are often assigned to entirely different clusters.

Second, Dr. Mercer claims that the outside plant extends from the SAI to the customer premises. This is incorrect, because the outside plant modeled by HM 5.3 is not going to any Verizon NW customer locations. It is going to lots that are entirely unrelated to the actual locations of Verizon NW's customers. In fact, HM 5.3 models hypothetical loops that serve no actual customers. The cost model inputs are based on a series of assumptions that bear absolutely no resemblance to actual network topography.

Q. DOES HM 5.3 INCORPORATE STATE-OF-THE-ART MODELING TECHNIQUES?

A. No, it does not. Contrary to Dr. Mercer's claim that age is a major advantage of HM 5.3, it is likely the major *disadvantage*.[28] The HAI Model was developed over ten years ago when many of the current, more sophisticated modeling techniques were not available. Rather than keep up with the development in modeling techniques, the HAI Model developers built on this outdated foundation. The result is a model that builds a rectangular representation of what are purported to be distribution areas, and then compounds this problem by spreading customer locations and demand uniformly within these rectangular

[28] A technical workshop in the Verizon Unbundled Network Element phase of Rulemaking 93-04-003 was held in San Francisco, California, on January 13–15, 2004, where the two parties (Verizon and AT&T/MCI) submitting cost models were allowed to give an overview of their cost studies and the modeling process. Hereinafter the discussion at this workshop will be referred to as the "Verizon CA Workshop Transcript." The accompanying slides will be referred to as "Verizon CA Workshop Slides." *See* Verizon CA Workshop Slides at slide 59.

distribution areas rather than reflecting actual customer dispersion. Although the long, elaborate process of geocoding and surrogating customer locations when forming clusters is relatively new, the distribution area modeling process employed by HM 5.3 is almost the same as it was eight years ago. In HM 2.2.2, released in 1996, customers were uniformly spread throughout a square representation of a distribution serving area. In HM 5.3, however, customers are uniformly distributed throughout rectangular distribution serving areas. All that has changed in these eight years is that the Model's distribution areas are now rectangular instead of square. What has not changed is the Model's failure to account for actual customer locations. As shown below, the geocoding and clustering process employed by HM 5.3 produces clusters that are not much different than they were years ago.

Q. WHAT IS AT&T/MCI'S EXPLANATION OF WHY ACTUAL CUSTOMER LOCATIONS, AS WELL AS PHYSICAL AND MAN-MADE OBSTACLES ARE IGNORED?

A. Dr. Mercer blames technology for the fact that HM 5.3 still ignores real customer locations, as well as physical and man-made boundaries (e.g., mountains and actual roads). In a recent cost model workshop, Dr. Mercer claimed:

> [N]either Verizon nor HM 5.3 has advanced technology enough to say what would really be cool to know is it the red road or is it a bunch of blue roads like this, and actually reroute, do what the telephone company engineer does, and actually lay this out. Neither model does that and no model yet has gotten good

> enough to be able to do that. That is just beyond the technology. So far, we've gone a long way, we're not that far.[29]

Dr. Mercer was responding to my question as to why HM 5.3 ignores customer locations and instead builds to uniformly distributed lots. While I do not claim that a model can (or needs to) produce an exact blueprint of an actual telephone network in order to produce valid cost estimates, I disagree with Dr. Mercer's statement about what is possible today. With today's technology, far more is possible than HM 5.3's overly simplistic rectangular distribution areas with backbone cable running along one dimension with a few uniformly sized branch cables perpendicular to it. In fact, it has been possible for years to model an outside plant network along the potential routes a telephone network can take (e.g., along roads) using actual customer locations and data about the services purchased by customers at each location. TELRIC models can and have modeled networks within those constraints that meet forward-looking engineering criteria and mirror reasonable economies of scale and scope. For instance, the BellSouth TELRIC loop model, BSTLM, models a network that serves existing customer locations, taking into account available rights-of-way and the location of roads.

I also disagree with Dr. Mercer's characterization of Verizon NW's cost model. VzCost is undoubtedly superior in network design when compared to HM 5.3. VzCost's loop investment calculator, VzLoop, models network plant along

[29] Verizon CA Workshop Transcript at pp. 3534-35.

existing network nodes, which usually fall along roads.[30] The basic idea is the same as in BellSouth's BSTLM loop cost model—VzLoop models a network along feasible network routes, properly reflecting rights-of-way and generally avoiding physical boundaries, such as freeways, highways, and rivers. Verizon NW's cost model does not rely on unsupported guesses as to what is the right amount of plant needed to serve an area. Instead, VzLoop properly begins with verifiable starting and ending points in a network; that is, switch locations and network nodes, including customer distribution terminals. Per MCI's own expert, routing along potential network routes is possible, and:

> [A]llows the design of a network that serves the existing customers, taking account of the constraints regarding available rights of way for placing loop plant, using computational methods that are available and can be tested and confirmed by all interested parties.[31]

HM 5.3 does none of that, thereby contradicting Dr. Mercer's explanation as to why HM 5.3 is allegedly more sophisticated.

Q. HOW CAN THIS COMMISSION COMPARE THE REASONABLENESS OF HM 5.3'S AND VZCOST'S MODELED OUTSIDE PLANT?

A. I have prepared maps of the outside plant for all wire centers as modeled by HM 5.3 and VzLoop. Comparing these maps and contrasting feeder and distribution routes against roads and physical boundaries, such as freeways, highways, and bodies of water, provides an excellent tool to evaluate the

[30] A network node can be a distribution terminal, a cross-box, a technical point of interface, a feeder terminal, or a DLC.

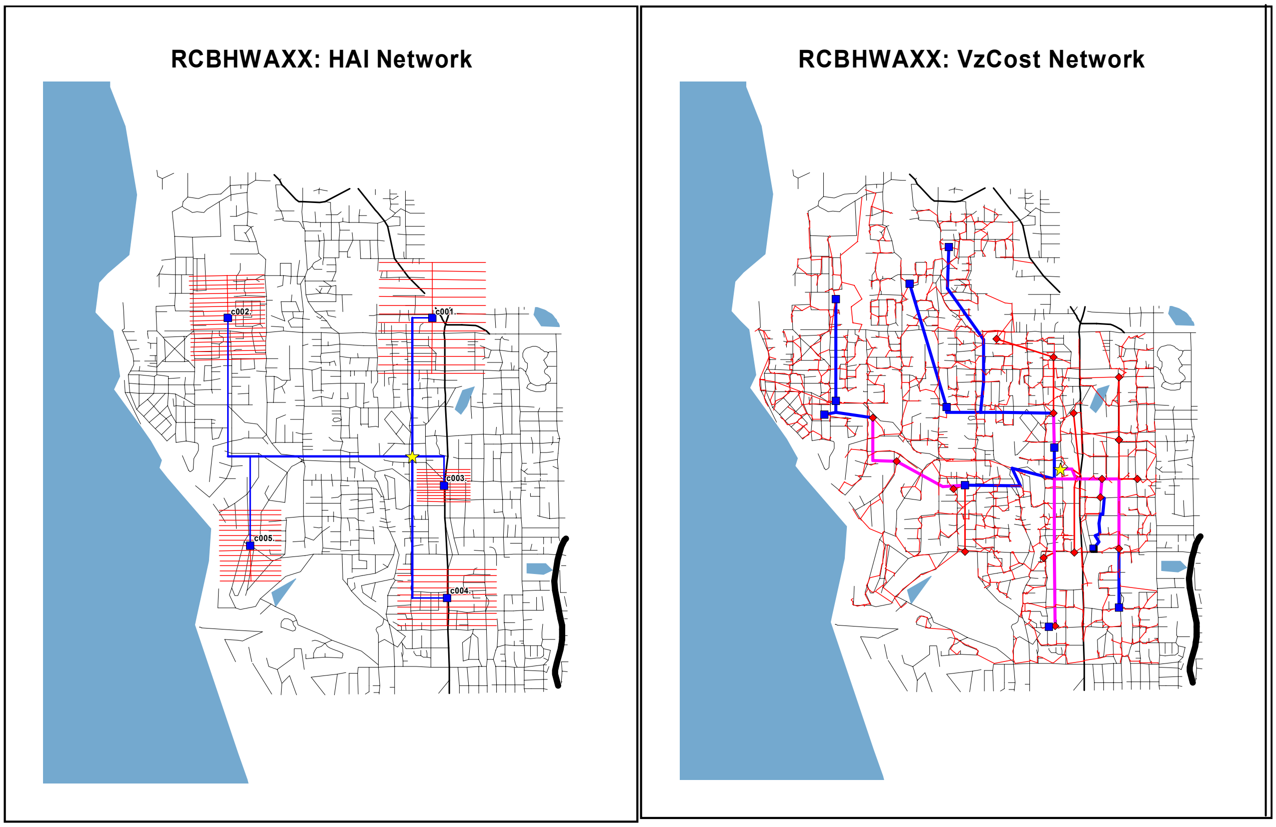
[31] Pelcovits Decl. at p. 37.

reasonableness of each model's outside plant modeling assumptions. Image files of these maps are included herewith as Exhibit CMD-6. While I will discuss these maps in detail later, they clearly illustrate the superior modeling techniques employed by VzLoop. HM 5.3 simplistically models plant throughout Washington without taking into account any physical boundaries or existing roads. HM 5.3's distribution and feeder plant intersect freeways, are routed through lakes, and are placed in the ocean. Furthermore, its distribution routes are entirely divorced from street layouts and topography, and overlap, in some instances, almost perfectly with other distribution plant. More important, HM 5.3 does not route its plant to any existing customer locations. Instead, the distribution cable is routed deliberately and uniformly in a configuration that resembles a television antenna centered on the cluster centroid. The customer lots are then "hung" uniformly along the branches of the antenna. However, no matter how precisely or imprecisely this is done, no distribution cable is routed to actual customer locations.

VzCost, on the other hand, contains a much more realistic representation of the routes along which a forward-looking outside plant network would be built. Its loop investment calculator, VzLoop, routes outside plant mostly along roads, and for the most part does not cross physical boundaries, avoiding lakes and other obstacles. Furthermore, rather than placing all demand uniformly across the wire centers' serving areas, VzLoop routes its plant to actual distribution terminal locations and sizes its cable according to demand at a particular network node. One does not need to be a network engineer to quickly realize that VzLoop's modeling techniques are superior to HM 5.3's.

To illustrate some of the differences, Map 1 below compares how HM 5.3 and VzLoop model the outside plant for the Richmond Beach wire center. The red lines indicate the modeled distribution plant, the blue lines indicate the feeder network, and the black lines indicate the road network in the wire center's serving area. As can be clearly seen, VzLoop closely approximates the road network in Richmond Beach. HM 5.3, on the other hand, sporadically models rectangular distribution areas that bear no resemblance to the road network, and fail to serve the majority of the customers in this wire center. The maps for all of Verizon NW's wire centers are included in Exhibit CMD-6.

Map 1 Comparison of HM 5.3 and VzCost Outside Plant Network



Q. ARE PHYSICAL OBSTACLES AT LEAST ACCOUNTED FOR IMPLICITLY BY HM 5.3?

A. No. Although AT&T/MCI claim otherwise, HM 5.3 does not account for physical obstacles, implicitly or explicitly. In response to a Verizon NW data request asking AT&T/MCI to explain: "How do HM 5.3's rectangularized clusters account for physical obstacles such as freeways, highways, rivers, waters, etc.," AT&T/MCI responded:

> The customer location clusters used by the HAI Model do not explicitly account for physical obstacles such as those described in the request. However, because clusters are formed based on actual geocoded customer locations and on customer locations

> located on the road network, accounting for such physical obstacles is implicit in the customer location data. Customers typically are not located in rivers, lakes, freeways, *etc.*[32]

This answer is interesting in several respects. First, it clearly documents the fact that HM 5.3's clusters ignore physical boundaries. Consequently, the Model ignores important cost components, which is confirmed by the testimony of MCI's own economist, Dr. Pelcovits.[33] Second, the explanations offered as to how physical boundaries are in some manner accounted for make no sense. Clustering based on actual customer locations does not guarantee that physical boundaries are not crossed. If clustering does not account for physical boundaries, it makes no difference whether actual, surrogate, or random customer locations are being clustered.

Q. WHAT ARE THE CONSEQUENCES OF IGNORING PHYSICAL BOUNDARIES?

A. Since working around physical boundaries and rights-of-way is costly for real-world local exchange carriers, ignoring them causes the Model to produce significantly understated cost estimates.

Q. DOES RECTILINEAR ROUTING TAKE PHYSICAL BOUNDARIES INTO ACCOUNT?

A. No, it does not. It is interesting how AT&T/MCI search for possible answers as to how the Model somehow accounts for physical boundaries. In the answer

[32] Joint Responses to Verizon's First Set of Data Requests at Response No. 1-35 (emphasis in original).

[33] "[E]xisting rights of way and actual customer locations are ignored, even though these factors would obviously have a significant effect on loop lengths." Pelcovits Decl. at p. 23.

above, AT&T/MCI recited one list of excuses; however, in the recent SBC CA UNE proceeding, Dr. Mercer tried to explain away the Model's deficiency by relying on the Model's strand-distance multiplier. In response to a question by SBC CA regarding whether HM 5.3 recognizes highways and rail tracks in routing the modeled outside plant, Dr. Mercer explained: "It does not do that explicitly. It does that by providing route miles to take care of obstacles like that."[34] Dr. Mercer was referring to the strand-distance multiplier gross-up that increases or decreases the backbone and branch cables in a cluster to supposedly match the length of the rectilinear minimum spanning tree, or what AT&T/MCI refer to as the strand distance. As discussed below, the strand-distance multiplier is nothing but a Band-Aid. It does not address the fact that HM 5.3 ignores physical boundaries, and brings with it an entirely new set of problems. Cost is much more than a function of length. It is also a function of cable material, feeder and distribution mix, cable size, and distribution structure mix. None of these cost components are addressed by HM 5.3's strand-distance multiplier.

To illustrate the problems associated with the Model's strand-distance multiplier, consider cluster c008 in the Chelan (CHLNWAXX) wire center. As depicted in the maps contained in Exhibit CMD-6, in HM 5.3 this cluster spans Lake Chelan. The SAI (and thus the ending point for the feeder cable) is located in the middle of the lake. By grossing up the distribution route distance,

[34] The California PUC held a meeting with SBC and AT&T/MCI regarding cost models for UNEs on June 25, 2003, hereinafter the "SBC Model Workshop Transcript," p. 709.

HM 5.3 simply increases the length of the backbone and branch cables; however, they still span the lake. Dr. Mercer claims that this enlargement of distribution route distance somehow accounts for the fact that people do not live in Lake Chelan. However, the Model ignores entirely the real problem with the modeling of this particular situation. A more realistic network likely would have routed feeder cable along Highway 97, which runs along the lake for some distance, with distribution cable branching off to serve customers and likely sharing some of the trenches with feeder cable. Thus, the problem caused by ignoring the lake is not that the distribution route distance is too short, but that the feeder cable is not long enough.[35] The gross-up allows HM 5.3 to increase the distribution route distance, which likely was already too long prior to the gross-up; however, the feeder cable is left unchanged, and falls far short of the required distance.[36] In short, using rectilinear distances to somehow account for physical obstacles does not address the problem, and in many situations makes it worse.

Q. DOES VZCOST ACCOUNT FOR PHYSICAL OBSTACLES AND RIGHTS-OF-WAY?

A. Yes, it does. HM 5.3 and VzCost both begin with the same premise—they both attempt to geocode the location of customer demand. HM 5.3 geocodes and surrogates customer addresses, while VzCost geocodes and surrogates

[35] As Mr. Murphy explains, HM 5.3 categorically understates feeder investment. *See* Murply Reply Testimony at pp. 56-79.

[36] As I will elaborate further below, this at least partially explains HM 5.3's overestimation of distribution cable and underestimation of feeder cable. In fact, HM 5.3 models 41 percent more distribution route distance and 33 percent less feeder route distance than VzCost.

customer distribution terminals and other network nodes. The main difference between the two models is that HM 5.3 later ignores the geocoded results and reverts to almost the same method employed eight years ago in HM 2.2.2; it models to rectangular-shaped distribution areas (as opposed to square-shaped areas in the older version) and assumes that demand is uniformly distributed within these areas. VzLoop, on the other hand, models plant to the actual location of distribution terminals and network nodes, which are located along or near roads. Consequently, and as depicted in the maps in Exhibit CMD-6, VzCost's modeled network routes, both for feeder and distribution cable, generally follow roads, avoid physical obstacles, and reflect rights-of-way.

Q. YOU DESCRIBE HM 5.3'S MODELING AS ARCANE, BUT DIDN'T THE FCC INCORPORATE PART OF THE HAI MODEL INTO UNIVERSAL SERVICE MODEL?

A. First, the FCC never accepted the HAI Model's approach to network modeling. In fact, the FCC concluded that:

> [T]he customer location and outside plant platform of the federal mechanism should consist of a synthesis of the best ideas presented by the model proponents, including HAI's use of geocoded customer location data, BCPM's use of the road network to estimate the locations of customers for whom no geocode data are available, HCPM's approach to identifying customer serving areas based on natural clusters of customers, and HCPM's ability to design plant to the precise customers locations within each serving area.[37]

[37] Fifth Report and Order, *In re Federal-State Joint Board on Universal Service, In re Forward-Looking Cost Mechanism for High Cost Support for Non-Rural LECs*, 13 FCC Rcd 21323 (rel. Oct. 28, 1998) at ¶ 26.

Second, the FCC's Synthesis Model was developed in 1997, when much of today's modeling technology was not available. Thus, AT&T/MCI's claim that even the FCC adopted part of their model is at best outdated.

C. The Clustering Algorithm in the Cluster Input Database Presents Significant Problems

Q. WHAT IS THE PURPOSE OF THE CLUSTERING ALGORITHM?

A. The clustering algorithm is an important aspect of the Model in that it determines the number and characteristics of Verizon NW's distribution areas. Specifically, it determines how many distribution areas will be modeled, their size, aspect ratio, density zone, the SAI location, and the rectilinear minimum spanning tree to connect the customer locations in a cluster. The cluster assigned density, although done incorrectly, determines important cost drivers such as structure type (aerial, buried, or underground) and sharing percentages. Thus, the clustering algorithm does much more than merely cluster customer locations; it determines many other aspects of the HM 5.3 modeled distribution and feeder network.

Dr. Mercer hails the clustering algorithm for developing "groupings of customer locations that have a realistic correlation to efficient distribution areas."[38] The HM 5.3 Model Description claims that "clusters developed pursuant to this process are likely to be the most closely representative of actual telephone distribution areas as determined by outside plant engineers."[39] Neither of these

[38] Mercer Supplemental Direct Testimony at RAM-4 (Model Description), p. 3.
[39] *Id.* at p. 21.

claims has any merit. Mr. Murphy explains the importance and complexity of properly designing distribution areas.[40] Thus, undoubtedly the clustering algorithm is a key component of HM 5.3 that needs to be thoroughly reviewed and validated.

Q. HAVE AT&T/MCI PROVIDED ANY EVIDENCE AS TO THE ACCURACY OF THE CLUSTERING ALGORITHM?

A. No, AT&T/MCI claim that the clustering algorithm has a high degree of accuracy, but have offered no proof. In fact, when asked to state all the facts upon which Dr. Mercer relied for his claim that clusters "are likely to be most closely representative of actual telephone distribution areas," and to produce all studies and documents that would support that statement, AT&T/MCI answered:

> The referenced statement is intended to mean that serving areas (clusters) based strictly upon engineering criteria will more closely resemble actual serving areas than earlier modeling techniques, which were based on arbitrary geographical areas, such as census block groups or arbitrary grids.[41]

To validate AT&T/MCI's claim the new clustering technique should improve the Model's accuracy when compared to previous versions of the HAI Model that relied on CBGs. I have replaced the clustering algorithm with a rule that forms a cluster for each CBG. As I will discuss in detail below, this replacement, however, had only a minor impact on the Model's cost estimates. Thus,

---

[40] *See e.g.*, Murphy Reply Testimony at pp. 38-40, 60 -62. *See also* Richter Reply Testimony at pp 12-15.

[41] Joint Responses to Verizon's First Set of Data Requests at Response No. 1-31.

AT&T/MCI's claim regarding HM 5.3's improved accuracy is entirely unsupported.

As this analysis demonstrates, Dr. Mercer and AT&T/MCI have no support for the statement that TNS's clustering algorithm produces realistic distribution areas, as they apparently have never taken the time to validate the results of this process against specific distribution areas contained in Verizon NW's network. I have attempted to validate HM 5.3's clusters and found that the clusters produced by this process utterly fail all validation tests.

Q. HAVE YOU BEEN ABLE TO REVIEW THE PRECISE FUNCTIONING OF THE CLUSTERING ALGORITHM?

A. No, I have not. In order to (1) determine how the clustering in HM 5.3 functions and whether it functions correctly, (2) perform sensitivity analyses on the Model's hard-coded values, and (3) modify some of these hard-coded values, one needs access to the source code that runs the clustering software. However, despite repeated requests and consecutive orders by the ALJ and the Commission, AT&T/MCI refuse to produce the source code for the clustering algorithm. AT&T/MCI acknowledged in discovery that they never checked the accuracy of the source code.[42] Moreover, TNS and AT&T/MCI refuse to allow third-parties to review the source code, which means that the foundation of HM 5.3's modeling process is unknown and untested. Therefore, at least in this

[42] *See e.g.*, Joint Responses to Verizon's First Set of Data Requests at Response No. 1-7 and 1-8.

important respect, HM 5.3 remains a “black box,” as neither the Model sponsors nor third-parties can truly attest to its accuracy.

In the recent SBC CA UNE proceeding, Dr Mercer claimed that my ability to rerun the clustering module, produce maps of HM 5.3’s modeled network, and examine the preprocessing “with a ‘microscope’ …is hardly something one would expect an independent party to be able to do with a ‘black box’ missing vital information.” Dr. Mercer was implying that I was not hindered in my analysis by AT&T/MCI’s steadfast refusal to produce the clustering source code.[43] However, Dr. Mercer misses an important point of my review of HM 5.3. Being able to reproduce something by rerunning a piece of software or following some predetermined steps is no substitute for being provided complete and comprehensive access (with documentation) to the fundamental underpinnings of the clustering process. I can drive a car, but that does not mean that I understand what makes it runs. Being able to rerun the clustering algorithm software does not enable one to understand the logic behind it. Dr. Mercer’s comment is nothing more than an attempt to side-step what is unarguably true: an important part of HM 5.3 is and remains a black box.

Q. PLEASE DESCRIBE YOUR REVIEW OF THE CLUSTER OUTPUT.

A. Since I was unable to review what the clustering source code actually does, I had to limit my analysis to the output of the clustering algorithm, that is, the

[43] *See* Before the California Public Utilities Commission, Application Nos. 01-02-024 et al., *Declaration of Robert A. Mercer in Support of Joint Applicants’ Rebuttal Comments* (Mar. 12, 2003) at p. 6 (“Mercer SBC Rebuttal Decl.”).

resulting main and outlier clusters. This review revealed three types of errors. First, apparently due to a design error, the clustering algorithm does not employ a nearest-neighbor procedure or any other clustering procedure from graph theory.[44] Instead, it uses a nearest-neighbor technique in a first pass, but then "chops up" all the resulting clusters that exceed a predetermined line limit until they meet this limit. Second, a visual inspection of the clustering results reveals a number of anomalies and many instances where cluster assignments of customer locations defy common sense and engineering practices. Finally, the clustering algorithm incorrectly incorporates the 17,000-foot copper-length threshold, yielding clusters that violate AT&T/MCI's own copper-length limit. I will discuss each of these errors further below.

Q. PLEASE DESCRIBE THE VARIABLES THAT ARE HARD-CODED IN THE CLUSTERING ALGORITHM AND COULD NOT BE REVIEWED.

A. The clustering source code contains a number of hard-coded values that have a significant impact on the modeled network. Without access to the clustering source code, none of these variables can be modified or even analyzed. First, and foremost, AT&T/MCI seem to treat all locations that have more than 536 lines as high-rise clusters. It is assumed that these clusters contain indoor SAIs and form their own cluster. This threshold number is hard-coded and cannot be

[44] The clustering of points is not a new topic and many acceptable clustering techniques have been developed. Among them are the nearest-neighbor technique, which AT&T/MCI falsely claim to have used, an agglomerative method, and a divisive method. *See generally* Brian S. Everitt, *Cluster Analysis* (Arnold: London, 3rd ed. 1993). The FCC's Synthesis Model offers all three types of clustering algorithms.

changed without proper access to the clustering source code, and, as Mr. Murphy explains, yields a significant understatement of indoor SAIs.[45]

Second, although TNS states that the clustering process starts "within a 150 foot radius of the center of the initial cell," the HM 5.3 Model Documentation is silent as to where that initial cell is located.[46] Since the location of the initial cluster has a significant impact on the clustering output, at a minimum, it must be identified and made user-adjustable.

Third, the second phase of the process splits the minimum-bounding rectangle of an oversized cluster into equally sized geographic halves. As noted above, this procedure seems to be entirely arbitrary and defeats sound engineering and modeling techniques. Users must have sufficient access to the Model to review and modify this process.

IV. HM 5.3's Preprocessing Is Conceptually Flawed

A. Problems with Data Limitation, Modeling Assumptions, and Forecasting Power

Q. MUST A MODEL BE FLAWLESS IN ORDER TO BE ACCURATE?

A. No, absolutely not. Dr. Mercer has not attempted to portray the many errors in HM 5.3 as mere approximations "instead of replicating each nook and cranny of the 'actual'…network," and has claimed that I "would apparently not be satisfied with anything short of modeling perfection."[47] Nothing could be further from the

[45] *See* Murphy Reply Testimony at pp. 27-30, 67-76.
[46] Mercer Supplemental Direct Testimony at RAM-4 (Model Description), p. 22.
[47] Mercer SBC Rebuttal Decl. at p. 4.

truth. Any attempt to model a real-world process, be it a cost model or any other model for that matter, must make some simplifying assumptions. However, there are important differences between simplifying assumptions, omitting relevant aspects of a process, and out-and-out modeling errors. Therefore, the criteria for assessing whether a model is accurate, and thus valid, cannot be that it makes no assumptions. Rather, the relevant question is whether those simplifying assumptions appropriately capture the relevant aspects of real-world network design processes (e.g., modeling to actual customer locations, following feasible network routes, bypassing physical obstacles, and accounting for rights-of-way). HM 5.3 clearly fails on this front as its problems have nothing to do with "replicating each nook and cranny of the 'actual' network," as Dr. Mercer contends, but rather have everything to do with a fundamental misrepresentation of the very foundation upon which real-world networks are built. HM 5.3 does not just approximate; it completely ignores important cost drivers and critical network design assumptions.

Q. DO YOU CONTEND THAT A COST MODEL'S DOCUMENTATION MUST DESCRIBE EACH FUNCTION IN DETAIL IN ORDER FOR THE MODEL NOT TO BE LABELED A BLACK BOX?

A. No, I do not expect a model description to explain every function in absolute detail. However, the major cost drivers (e.g., the cluster input database) must be completely explained. This is not the case for HM 5.3. HM 5.3's preprocessing is hardly documented, and many files have no explanation whatsoever as to their purpose or how the columns and variables are defined.

Specifically, the preprocessing, although consisting of 675 files, comes with only nine pages of vague and confusing descriptions. Second, only five of the some 300 pages filed by Dr. Mercer in this proceeding address preprocessing, one of the most important aspects of HM 5.3.

B. The Most Significant Error in HM 5.3 Is the Overly Simplistic and Improper Modeling of Outside Plant.

Q. WHAT IS HM 5.3'S MOST SIGNIFICANT FLAW?

A. By far, HM 5.3's most significant flaw is the Model's inability to model and cost a network that is built to actual customer locations. Devoid of any information regarding actual geocoded customer locations (because the Model developers mysteriously have decided to discard the results of this exercise), the Model is left with simplistic rectangles and uniformly distributed demand on customer lots rather than actual geocoded customer locations. Specifically, it is AT&T/MCI's assumption that Verizon NW's customers are uniformly distributed on rectangular lots in rectangular distribution areas. One does not need to be a telephone engineer or modeling expert to know that Verizon NW's customers: (1) are not uniformly spread over a distribution area; (2) do not live on adjacent lots that are twice as wide as they are deep; (3) do not share the same lot size in each distribution area; and (4) do not live in rectangular distribution areas. Common sense also dictates that it is far more complex for Verizon NW to build cable to a real-world distribution area than simply placing a thicker cable in the middle of a rectangular-shaped cluster with a few thinner cables perpendicular to it. Finally, logic tells us that Verizon NW cannot place cables across highways or through impenetrable natural or manmade structures, and it cannot ignore rights-of-way. Nevertheless, this is what HM 5.3 assumes, thereby rendering the Model itself and the cost estimates it produces useless.

Q. AREN'T THESE OMISSIONS MERELY THE RESULT OF MODELING LIMITATIONS?

A. No, they are not. AT&T/MCI, and particularly Dr. Mercer, would like this Commission to believe that HM 5.3 is state-of-the-art and that these omissions are simply the result of limitations in the art of modeling. This is not true, as more advanced and sophisticated modeling techniques are available today than were available when the Model was originally developed. The HAI Model simply has not incorporated new modeling techniques to correct old problems. By way of contrast, VzLoop is a good example of what modeling techniques are available today. VzLoop, unlike HM 5.3, is able to model plant to customers using appropriately sized distribution and feeder cables, and deploy the plant along feasible network routes, such as roads. The ability to model along tangible network routes may not have been readily available ten years ago when the Hatfield Model (now the HAI Model) was introduced; however, it is available now. Nevertheless, AT&T/MCI have not updated HM 5.3's most fundamental modeling assumptions. The basic premise of uniform customer distribution has remained unchanged for at least eight years; and, although AT&T/MCI now claim that HM 5.3 relies on accurately pinpointing demand and customer clustering, the majority of these results are not being used. In this regard, HM 5.3 still models distribution areas very much the same as HM 2.2.2 did many years ago.

Q. WHY DO YOU BELIEVE THAT HM 5.3 DOES NOT USE THE GEOCODING AND CLUSTERING RESULTS?

A. As explained above, HM 5.3 does not build plant to a single actual customer location. It does not even attempt to approximate the dispersion of customer demand. A simple review of HM 5.3 reveals that the geocoding and surrogating results are ignored entirely.

Determining the impact (or lack thereof) of TNS's faulty clustering algorithm, however, is not as straightforward because the clustering source code has never been made available for review and analysis. In order to assess the impact of the clustering algorithm, I replaced it with a trivial clustering rule—the same clustering rule AT&T/MCI used in HM 2.2.2, which is that each CBG forms its own cluster.[48] Everything else was left unchanged. That is, in my modified HM 5.3, the Model still rectangularizes the clusters, still builds to uniformly distributed lots that are twice as wide as deep, and still builds distribution plant in the grill fashion described above. By AT&T/MCI's own account, the distribution areas derived for HM 5.3 should "more closely resemble actual serving areas than earlier modeling techniques, which were based on arbitrary geographical areas, such as *census block groups* or arbitrary

[48] *See* Hatfield Associates, Inc., "Model Description, Hatfield Model Version 2.2, Release 2" (Sept. 4, 1996) at p. 14.

grids."[49] Thus, it is reasonable to expect that replacing the current and allegedly more accurate and sophisticated clustering algorithm with an algorithm based on CBGs would yield significantly different cost estimates. However, this is not the case. HM 5.3 produces nearly identical results whether the supposedly highly accurate current clustering algorithm is used or whether distribution areas are formed using arbitrary geographic areas, such as CBGs.

Specifically, when I replaced the clustering algorithm with a rule that assumes each CBG forms its own cluster, the number of distribution areas (clusters) in HM 5.3 increases from 1,019 to 2,517, an increase of 147 percent. However, at the same time, HM 5.3's estimated monthly per-line loop cost merely increases 10 percent, from $7.87 to $8.66.[50] Even more telling, if the TNS clustering algorithm is run at a maximum-line threshold of 900, the program generates 2,570 clusters (a number similar to the number of CBGs) and a monthly loop cost of $8.41. Thus, not only does clustering have only a marginal impact on cost, it also does not seem to matter much what clustering procedure is used—either way, the Model generates almost identical results.

This lack of sensitivity illustrates a number of important issues. First, it shows that the current clustering algorithm is not more accurate than an arbitrary rule, which assumes that each CBG forms one cluster. Replacing TNS's

[49] Joint Responses to Verizon's First Set of Data Requests at Response No. 1-31 (emphasis added).

[50] Note that the benchmark for this comparison is $7.87 and not $7.64 as filed by AT&T/MCI on April 9, 2004. This is because AT&T/MCI made manual adjustments to the TNS preprocessing result, which cannot be replicated. These adjustments reduce monthly loop costs from $7.87 to $7.64 and reduce the number of clusters from 1,019 to 1,018. This comment applies throughout this testimony.

much-guarded and alleged commercially-valuable clustering algorithm with an admittedly imprecise clustering rule from an old version of the HAI Model produces cost estimates that differ only slightly. Second, it illustrates an important and fatal error in HM 5.3—its insensitivity to the number of distribution areas. I will discuss this error with other analyses later in this testimony. This example clearly demonstrates that the total monthly loop cost remains nearly unaffected, regardless of how many distribution areas HM 5.3 models. This is because with each increase in the number of clusters, the Model assumes an approximately equal-sized decrease in the investment per cluster. For instance, if the number of clusters is doubled, the investment per cluster is decreased by approximately 50 percent, resulting in only a marginal change in total investment. For example, with 1,019 distribution areas, HM 5.3 produces an investment of approximately $856,000 to serve each distribution area. Yet, with 2,517 distribution areas, it estimates only $382,000 to serve each distribution area. That is, more than doubling the number of clusters (actually increasing the number by 2.5) decreases the per-cluster investment by almost 55 percent, effectively offsetting any cost impact. These results are clearly counter-intuitive. One would expect the per-cluster investment to drop as the number of clusters increases. It makes no sense for an increase in the number of clusters to be offset by a corresponding decrease in investment. According to this logic, one large cluster requires the same investment as two clusters each of which is half the size of the large cluster. Such an outcome is highly unlikely, defeats common sense, and as Mr. Murphy explains, is contrary to sound engineering principles and expectations.[51]

Q. DOES HM 5.3'S INSENSITIVITY TO THE NUMBER OF CLUSTERS MEAN THAT THE CLUSTERING ALGORITHM IS NOT IMPORTANT AND SHOULD NOT BE USED?

A. No, that would be the wrong conclusion to draw from this analysis. The determination of distribution areas is one of the most crucial components in modeling outside plant. The fact that it does not affect HM 5.3's costs in a logical fashion does not mean that it is not important, but that other modeling flaws (i.e., uniform distribution of demand on equal-sized lots that reside in rectangular-shaped distribution areas) override the results of the clustering algorithm. That is, HM 5.3 is insensitive to the clustering algorithm because AT&T/MCI model outside plant incorrectly, not because the clustering algorithm is unimportant. Even AT&T/MCI seem to agree with the importance of the clustering algorithm. AT&T/MCI believe that the clustering algorithm is state-of-the-art, and Dr. Mercer claims that it has "profoundly improved the accuracy of . . . HM 5.3."[52] If the clustering algorithm is state-of-the-art and if it has improved the accuracy of the Model so profoundly, why then when it is replaced with a simplistic rule does the Model produce almost identical cost estimates? The answer seems simple; the results of the clustering algorithm are lost when transforming the clustering results to rectangular clusters and by

---

[51] *See e.g.*, Murphy Reply Testimony at pp. 56-58.
[52] Mercer Supplemental Direct Testimony at p. 21.

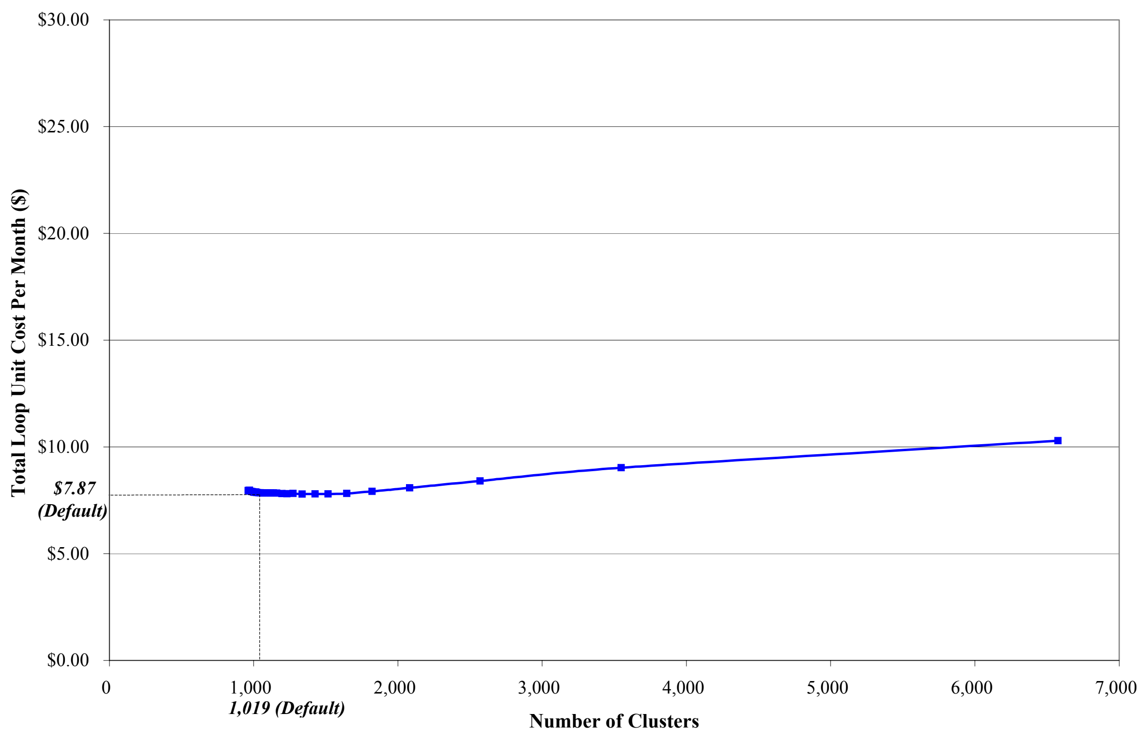
the assumption that demand is uniformly distributed. Not surprisingly, this odd method of modeling outside plant yields equally odd sensitivity analyses, such as the one illustrated above where, regardless of how many clusters serve an area, the Model produces approximately the same amount of required investment.

Q. IS THE MODEL AT LEAST SENSITIVE TO CHANGES IN CLUSTERING ASSUMPTIONS?

A. No. Changes in the number of clusters have little impact on the cost estimates produced by HM 5.3. In order to test the sensitivity of HM 5.3 to the number of clusters, I modified the maximum line count per cluster.[53] I started with a maximum line size limit of 300, and raised this limit by 300 until I reached 9,000 lines. I then ran the clustering algorithm 30 times for all wire centers in Verizon NW's serving territory. Subsequently, I processed the 30 clustering run results first through the FoxPro conversion programs, and then through PointCode to produce 30 cluster input databases for HM 5.3. I then ran these 30 cluster input databases through HM 5.3. The results of these runs are attached in Exhibit CMD-5a and 5b and summarized in Chart 1 below.

[53] Per AT&T/MCI, this variable is defined as follows: "The maximum lines in a cluster is the greatest number of lines that may be contained in a cluster." *See* Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Supplemental Joint Responses of AT&T & MCI to Verizon's First and Fifth Sets of Data Requests* (Nov. 5, 2003) at Supplemental Response No. 1-11.

Chart 1: Monthly Loop Cost



The horizontal axis shows the number of clusters as determined by each rerun of the preprocessing module. For instance, the default maximum-line threshold for a cluster is 6,451 lines. This setting creates 1,019 clusters. The vertical axis shows the monthly loop cost estimates per line.

Q. WHAT CONCLUSIONS DO YOU DRAW FROM THIS CHART?

A. This chart demonstrates HM 5.3's insensitivity to the number of distribution areas. That is, regardless of how many distribution areas are generated to serve Verizon NW's service territory, HM 5.3's investment and monthly loop cost remain practically unchanged. Regardless of the number of clusters (be it 962 or 3,548), the Model's cost estimates vary by less than 15 percent. For instance, at a line-size limit of 600 lines, the Model increases the number of clusters approximately 250 percent, yet monthly loop costs increase by less than 15 percent. Only when the maximum line count is set at 300 does the Model produce a change (31 percent) in monthly loop cost.

The chart also illustrates that HM 5.3 is not capable of modeling a network that incorporates critical engineering decisions. Per HM 5.3, a distribution area with 600 lines or 9,000 lines produces little difference in loop costs. As Mr. Murphy explains in his reply testimony, the number of living units in a distribution area usually is in the range of 200–600. The upper limit should be found in dense distribution serving areas to improve feeder efficiency and to economically minimize the number of interfaces. In relatively sparse distribution serving areas, one should find the lower limit to avoid wasting money building excessive lengths of distribution cables. The Model's task should be to balance distribution cable costs and feeder interface efficiency to form optimally sized distribution serving areas.[54]

Thus, rather than applying HM 5.3's overly-simplistic "one-size-fits-all" approach in which all distribution areas have the same line limit, a proper model designs distribution areas based on the number of living units *in a given geographic area*. Mr. Murphy refers to AT&T's Practice Standard that specifies that the number of living units in a distribution serving area generally ranges between

[54] *See* Murphy Reply Testimony, pp. 40-41.

200 and 600.[55] Assuming a 20 percent penetration of second lines produces a range of 240 to 720 working lines per distribution serving area, as opposed to HM 5.3's assumption of 6,451-targeted lines. Therefore, HM 5.3 does not balance distribution cable costs and feeder interface efficiency to form optimally sized distribution areas.

These sensitivity runs, once more, confirm that HM 5.3 does not model a realistic network, as it is not sensitive to the number of clusters or the clustering procedure except when clusters are sized in accordance with AT&T's own distribution serving area sizing guidelines. HM 5.3 produces practically identical cost results for virtually all cluster sizes beyond those associated with the guidelines, regardless of the maximum number of lines in a cluster or the clustering algorithm used. These runs also confirm that each increase in the number of clusters is offset by a same-size decrease in the investment per line. This problem is illustrated in Exhibit CMD-5b.

Q. PLEASE EXPLAIN EXHIBIT CMD-5B WITH AN EXAMPLE.

A. Consider a sensitivity run where the maximum number of lines in a cluster was set at 1,200 lines. Using the 1,200-line threshold, the Model generates 2,082 clusters with an average investment per cluster of $440,897. Using the default 6,451-line threshold, the Model generates 1,019 clusters with a per cluster investment of $855,754. The monthly loop cost estimates for these two runs

[55] *See* Exhibit No. FJM-1T, p. 40.

are $8.09 and $7.87, respectively. Thus, although the number of clusters more than doubled, monthly loop costs increased less than 3 percent.

Column E in Exhibit CMD-5b titled “cluster ratio” shows the ratio by which the amount of clusters has increased (2,082/1,019 = 2.04). Column F titled “investment ratio” shows the ratio by which the total investment per cluster has increased (855,754/440,897 = 1.94). If cost is the product of quantity and price, then multiplying quantity (or the number of clusters in this example) by 2.04 and then dividing price (or the investment per cluster in this example) by 1.94 yields a net effect of merely 1.05 (2.04/1.94). This number is reported in Column G, and indicates that, although the number of clusters has more than doubled, total cost only increased by 5 percent. As can be seen in Exhibit CMD-5b, with two exceptions, the net effect of the 30 sensitivity runs is always below 10 percent.

Q. WHAT DOES THIS MEAN IN PRACTICAL TERMS?

A. This means, in effect, that according to HM 5.3 there are very limited economies of scale associated with local exchange networks. In other words, it does not matter if an area is served by one or four distribution areas; there is hardly any difference in cost. This is clearly wrong.

Q. IS THIS THE FIRST TIME THAT YOU DISCOVERED THIS ERROR?

A. No, I discovered this error in SBC CA’s UNE proceeding. In that proceeding, for the first time I was able to rerun the preprocessing. Similar to the analyses that I performed here, I chose to decrease the maximum cluster line size, which more than doubled the number of clusters with only a marginal change in loop cost estimates.[56] AT&T/MCI attempted to rebut my analysis by claiming that "the model works as it should."[57] They claimed that, for the particular sensitivity run that I had done, an increase in feeder and concentrator investment resulted in an equally sized decrease in distribution investment, thus eliminating any cost impact.[58]

While Dr. Mercer's explanation as to why HM 5.3 is not sensitive to the number of clusters was questionable in the SBC CA proceeding, it clearly does not explain why HM 5.3 still produces almost the same result when the clustering algorithm is replaced with the simple rule used in HM 2.2.2. Further, it also cannot serve as an explanation of why 30 different runs performed on a different HM 5.3 version for a different state and with different inputs still do not produce significantly different cost estimates. These sensitivity runs simply confirm what I found in the SBC CA proceeding and other previous analyses—HM 5.3 ignores customer locations and clustering results.

The Model does not build to customers or clusters, but assumes perfectly rectangular-shaped distribution areas where demand is uniformly distributed. These arcane modeling assumptions abandon the modularity of a real-world network, and instead create an almost linear relationship where, for each

[56] Specifically, I decreased the maximum line size from 6,451 lines to 1,800.
[57] Mercer SBC Rebuttal Decl. at p. 25.
[58] *Id.*

increase in the number of clusters, there is an offsetting decrease in the estimated investment per cluster.

Q. HOW DIFFICULT IS IT TO RERUN HM 5.3'S PREPROCESSING MODULE?

A. Rerunning the preprocessing module for HM 5.3 is extremely difficult. First, the process is not well-documented. Likely recognizing this failing, TNS recently produced a document that describes the general flow of the preprocessing. Although falling far short of what I would consider a complete description, it enables the user to at least understand the general flow of databases, processes, algorithms, and models used to develop the cluster input database, and to blindly rerun the module. Fully understanding the process, however, remains very difficult as the source code has not been explained (let alone produced), and many of the variables are not described. Even when a description is provided, it does not allow a user to understand why a decision was made and why a process or step functions the way it does.

Q. HOW LONG DOES IT TAKE TO RERUN HM 5.3'S PREPROCESSING MODULE?

A. Rerunning the many programs and models that make up the preprocessing module is very labor intensive and not at all user-friendly. To rerun the clustering algorithm alone (which is only one program in the preprocessing module) takes approximately 21 hours in computer run time; and that does not include the manual intervention that is required or the time required to run other programs.[59] It takes about two to three days to perform a simple sensitivity test.[60]

Q. HOW DO YOU KNOW THAT THE RECTANGULAR CLUSTERS AND UNIFORM DISTRIBUTION OF DEMAND MAKE THE MODEL COST INSENSITIVE?

A. As I have shown, HM 5.3 does not rely either on customer locations or on the results of its clustering process. Instead, it clumps demand points into lots and assumes everything in a cluster is uniform. Specifically, HM 5.3 assumes that:

1. Cluster lots are uniform in size
2. Demand on these lots is uniform
3. The distance between the lots is uniform
4. The location and size of the distribution terminal are uniform
5. The drop investment in a cluster is uniform

In addition, HM 5.3 assumes that all these perfectly spread locations are served with an SAI that can be placed in the geometric center of this cluster. When conducting a sensitivity run, these fundamental modeling assumptions cannot be changed. While each sensitivity run will form new clusters, the customers are again moved and spread uniformly within the new rectangular clusters. Thus, no matter how the clusters are modified, HM 5.3 spreads the customers in the exact same fashion as before.

[59] This is the total computer running time that AT&T/MCI needed to run the clustering algorithm for the model database included in the November 3, 2003 filing.

[60] At a maximum line count per cluster of 9,000, the clustering algorithm alone takes about 32 hours to run.

Essentially, HM 5.3's preprocessing moves the customer locations each time the clustering algorithm is rerun in order to reestablish perfectly flat demand conditions, thereby holding costs steady by losing the discreteness of the customer locations before the network is modeled. Consequently, sensitivity runs do not have an impact, as these unrealistic demand conditions are automatically recreated.

Q. WHAT IS NEEDED TO CORRECT HM 5.3'S INSENSITIVITY TO THE NUMBER OF DISTRIBUTION AREAS?

A. The aforementioned modeling assumptions are deeply engrained in the Model, and cannot be changed without redesigning the Model entirely. The Model developers would have to start with the geocoded customer locations obtained from TNS, and model the network by following realistic network routes, such as roads, and by maintaining the geocoded customer locations throughout. Essentially, HM 5.3 would have to follow the example of cost models such as VzCost, which follow feasible network routes.

Q. COULD HM 5.3'S MODELING ASSUMPTIONS STILL YIELD ACCURATE RESULTS?

A. Not unless it was by pure coincidence. As previously mentioned, to illustrate the results of the arcane modeling approach employed by HM 5.3, I have mapped its outside plant network. For each wire center, I mapped the locations of the SAIs and the routing of each cluster's branch, backbone, and feeder cable. This mapping exercise is straightforward as it simply takes the data supplied by TNS and turns it into a picture, strictly following the model description provided by Dr. Mercer. It is on this network that HM 5.3 bases its cost estimates. The maps are contained in the attached CD-ROM labeled Exhibit CMD-6. Even a cursory look at these maps reveals that HM 5.3's modeled network is nothing but an array of cables that are intermingled with each other and routed irrespective of feasible network routes, physical boundaries, and rights-of-way. Given the network it models, it is highly unlikely and perhaps impossible for HM 5.3 to produce a "reliable and accurate estimation of Verizon's economic costs for substantially the entire Verizon local exchange service network," as claimed by Dr. Mercer.[61]

On the odd chance that the Model can somehow yield accurate results, I asked AT&T/MCI in the SBC CA UNE proceeding to provide an electronic copy of all documents concerning, referring, or relating to any external validation tests or studies that have been performed on HM 5.3. In response to this data request, AT&T/MCI admitted that other than some route distance comparisons between an old version of the HAI Model, the BellSouth model, and the Benchmark Cost Proxy Model, no such validations were conducted.[62] Similarly, in the current proceeding when responding to a virtually identical data request, AT&T/MCI claimed to have undertaken extensive efforts to validate the clustering

[61] Mercer Supplemental Direct Testimony at p. 5.

[62] *See* Before the California Public Utilities Commission, Application Nos. 01-02-024 et al., *Responses of AT&T Communications of California Inc. and WorldCom Inc. to SBC Pacific Bell Telephone Company's Fourteenth Set of Data Requests* (Nov. 11, 2002) at Response No. 14-74.

process.[63] Yet, when pressed to produce their results, AT&T/MCI admitted that, other than the pre-processing data (which contained no such validation efforts), they "found no other responsive documents."[64] AT&T/MCI's response makes one thing clear: they have not validated the results of the Model and have no idea whether the modeled network resembles an actual outside plant network or produces reasonable estimates of Verizon NW's forward-looking UNE costs. Thus, not only is the Model outdated in its modeling approach, it defies common sense when inspected visually. Moreover, AT&T/MCI cannot provide a single instance in the Verizon NW serving area where they compared their modeling results to actual outside plant network designs.

Dr. Mercer seems to be confused as to what such a validation effort would entail, as he stated in the recent SBC CA UNE proceeding that "Mr. Dippon calls for external validation based on some unspecified factual, verifiable source that does not exist."[65] Dr. Mercer is wrong; such data do exist. Further, Dr. Mercer states: "What Mr. Dippon does not address is that one would need to have a reasonably accurate TELRIC model to test the validity of the outputs of a TELRIC model."[66] Again, Dr. Mercer is wrong. In order to verify HM 5.3, or any part of HM 5.3, one does not need to compare it against another model but against observable real-world benchmarks. Dr. Tardiff has offered a number of

---

[63] *See* Joint Responses to Verizon' First Set of Data Requests at Response No. 1-10.

[64] *See* Before the Washington Utilities and Transportation Commission, Docket No. UT-023003, *Additional Supplemental Joint Responses of AT&T & MCI to Verizon's Third and Fifth Sets of Data Requests* (Dec. 18, 2003) at Supplemental Response and Additional Supplemental Response No. 3-14.

[65] Mercer SBC Rebuttal Decl. at p. 29.

[66] *Id.*

such validation tests. For example, Dr. Tardiff compared HM 5.3's estimated investment and expenses to Verizon NW's ARMIS results.[67] While Dr. Tardiff does not claim that there needs to be a dollar-for-dollar match, the analysis clearly shows that HM 5.3 produces only a small fraction of what Verizon NW has on its books. Thus, combining HM 5.3's out-of-the-ordinary outside plant routing, the results of Dr. Tardiff's external validation tests, and Dr. Mercer's failure to offer any support to the contrary, there is powerful evidence that HM 5.3's cost estimates are inaccurate.

Q. HAVE THE MODEL SPONSORS ATTEMPTED TO VALIDATE ANY OF THE OUTPUTS GENERATED BY PREVIOUS MODEL VERSIONS IN OTHER JURISDICTIONS?

A. The only validation the HAI Model developers attempted was at the direction of the Georgia Public Service Commission, which ultimately rejected Version 4.0 of the HAI Model. That effort entailed a comparison of the amount of distribution feet estimated by an old version of the HAI Model to the amount estimated in a detailed forward-looking design by the HAI Model engineers of new facilities in ten specific CBGs selected by the Georgia Commission Staff. AT&T/ MCI, who sponsored that version of the HAI Model, concluded that the Model estimated only 92.5 percent of the distribution plant required for the ten CBGs in a forward-looking environment. However, as it turned out this figure was incorrect. In presenting their results to the Georgia Commission, AT&T/MCI miscalculated the distances of the Model's connecting cable routes.

[67] *See* Tardiff Reply Testimony at pp. 39-40.

For CBGs where the Model provided fiber connecting cable, which is therefore considered to be part of the feeder, AT&T/MCI incorrectly included the routes as part of the distribution plant produced by the Model. For CBGs where the connecting cables were copper, and therefore properly included as distribution plant, AT&T/MCI "double counted" the associated routes.

When these errors were corrected, that version of the HAI Model estimated only 70 percent of the distribution plant the HAI Model engineers believed would be required under idealized conditions. In other words, the HAI Model, Version 4.0, in addition to understating cable amounts as a result of improper utilization assumptions, underestimated necessary route feet by 30 percent in the Georgia Commission's validation effort.

Q. PLEASE DESCRIBE WHAT THE MAPS IN EXHIBIT CMD-6 SHOW AND HOW YOU MAPPED HM 5.3'S OUTPUT.

A. The CD-ROM (included as Exhibit CMD-6) contains a map of HM 5.3's modeled outside plant network for each wire center. Each map shows the main feeder, subfeeder, SAI, and distribution cable for each modeled distribution area. Main feeders are shown with thick blue lines, subfeeders are thinner blue lines, copper feeders connecting outliers to main clusters are shown with thick red lines, and distribution cables (consisting of backbone and branch cables) are shown with thin red lines.

I mapped the distribution areas for these wire centers using the data provided to Verizon NW by AT&T/MCI and TNS through discovery and certain intermediate output obtained from HM 5.3. I used MapInfo 7.5 (a type of mapping software) to create the maps. Since I am not sure that all parties have a license for this mapping software, the attached CD-ROM contains images (.pdf files) of these maps.

Q. WHAT CONCLUSIONS DO YOU DRAW FROM YOUR ANALYSIS OF THE MAPS?

A. One can only speculate why AT&T/MCI have refused for years to provide access to the model's extensive preprocessing. However, after conducting the analyses discussed herein and illustrated by the maps contained in Exhibit CMD-6, I can finally understand why. When HM 5.3's network is illustrated in map form, it becomes clear that the Model produces entirely unrealistic distribution areas, resulting in absurd outside plant investment estimates.

Q. COULD IT BE THAT HM 5.3 ATTEMPTS TO MODEL AN ABSTRACT NETWORK RATHER THAN A REAL, FUNCTIONING NETWORK?

A. No. Dr. Mercer has argued that HM 5.3 attempts to model a real, functioning network. For example, Dr. Mercer claims that the Model:

1. "determines customer locations and configures outside plant to serve those customers;"[68]
2. "estimates in a consistent fashion the forward-looking economic costs that Verizon would incur to *build* a complete forward-looking network, including a defined set of UNEs;"[69] and
3. "[b]ased on the customer location data, and detailed and granular information as to the existing demand for services,

---

[68] Mercer Supplemental Direct Testimony at p. 7.
[69] Mercer Supplemental Direct Testimony at p. 9 (emphasis added).

> … then *constructs a network* to serve the identified demand."[70]

After being shown the first maps of HM 5.3's modeled outside plant network, however, AT&T/MCI changed their approach and now claim: "HM 5.3 is not a model that builds a network. It's a costing model, and it produces costs."[71] Further, they claim that they are "getting the amount of cable right."[72] This retreat to the world of the abstract is merely an attempt to obfuscate the many engineering and economic criticisms that can be levied against the Model. However, it raises an important point. During the ten years of the HAI Model's existence, AT&T/MCI have never provided any solid proof that the random array of distribution and feeder cable modeled by HM 5.3 produces the correct amount of cable and the correct TELRIC results. As shown throughout this testimony and the testimonies of Dr. Tardiff and Mr. Murphy, HM 5.3 clearly does not produce accurate cost estimates.

Q. HAVE AT&T/MCI EVER PRODUCED THEIR OWN MAPS OF THE MODELED DISTRIBUTION AND FEEDER PLANT?

A. For reasons that are obvious after reviewing the maps in Exhibit CMD-6, AT&T/MCI have never explicitly mapped the distribution and feeder plant of the network modeled by HM 5.3. However, a few years ago, AT&T/MCI included in their model description a map of a cluster (or what AT&T/MCI referred to as

[70] Mercer Supplemental Direct Testimony at p. 12 (emphasis added).
[71] Verizon CA Workshop Transcript at p. 3623.
[72] Verizon CA Workshop Transcript at p. 3624.

"distribution architecture employed by the Hatfield Model").[73] The maps of HM 5.3, attached in Exhibit CMD-6, are fundamentally no different than AT&T/MCI's map.

Q. HAVE YOU DONE THE SAME ANALYSIS FOR VZCOST?

A. Yes, I have. Realizing that this Commission is faced with a choice of two models, I performed the same exercise for VzCost. Specifically, taking the output from its ARC table, I mapped VzCost's modeled network using the same mapping software that I used to map HM 5.3's modeled network.[74] The attached CD-ROM (Exhibit CMD-6) contains a second file labeled "VzCost Maps." This file has the same type of maps for VzCost as were produced for HM 5.3, including the same color legend. Unlike HM 5.3, the maps for VzCost can be created directly from the model's output and do not require additional data from Verizon NW.

Q. HOW DO THE MAPS FOR VZCOST COMPARE TO THOSE FOR HM 5.3?

A. Although I do not claim to be an expert on the entire VzCost model, I believe VzLoop's ability to model along feasible network routes is one of the single most powerful arguments for adopting VzCost over HM 5.3. HM 5.3 is simply incapable of modeling outside plant in a reasonable, realistic fashion. On this basis alone, HM 5.3 should be rejected. These maps should provide this

[73] Hatfield Model, Release 3.1, Model Description, Hatfield Associates Inc. (Feb. 28, 1997) at p. 29. While I understand that HM 3.1 modeled outside plant differently than HM 5.3, both models rely on the same fundamentally flawed principles – that is, customers were and still are assumed to be distributed within equal-sized lots in a rectangular distribution area.

[74] The ARC table contains the architecture of the network modeled by VzCost.

Commission with enough evidence to support rejecting HM 5.3 in favor of VzCost.

C. The Strand-Distance Multiplier Is No Miracle Fix

Q. DOES THE STRAND-DISTANCE MULTIPLIER ENSURE THAT HM 5.3 CALCULATES THE CORRECT OUTSIDE PLANT INVESTMENT?

A. Absolutely not. In 1998 when the minimum spanning tree analysis was first introduced to this type of proceeding by Sprint and GTE, AT&T/MCI dismissed it, claiming that it had nothing to do with how telephone plant was modeled.[75] Sprint and GTE's motivation for introducing this type of analysis was simple—it showed that the HAI Model produced less cable than was needed to connect customer locations using the shortest distance possible.[76] The minimum spanning tree analysis was simply a tool used to illustrate that the HAI Model's modeling assumptions resulted in a network configuration that was not possible in the real world. In that sense, the minimum spanning tree served as a type of external validation test—it was the *floor* for the amount of distribution cable required in a cluster. Any amount of cable less than the amount produced using a minimum spanning tree was simply implausible.

[75] *See* Letter from Chris Frentrup, MCI, to Magalie Roman Salas, FCC (April 23, 1998); Letters from Richard Clarke, AT&T, to Magalie Roman Salas, FCC (May 5, 1998 and June 8, 1998); HAI June 22 *ex parte* cited in an FCC Public Notice, Common Carrier Bureau Seeks Comment on Model Platform Development, DA-98-157 (rel. Aug. 7, 1998).

[76] *See* Christian M. Dippon and Kenneth E. Train, "The Cost of the Local Telecommunications Network, A Comparison of Minimum Spanning Trees and the HAI Model," *Telecommunications Policy,* Vol. 24, No. 3 (April 2000), attached hereto as Exhibit CMD-7.

AT&T/MCI have since changed their minds, and now include rectilinear minimum spanning trees as distribution-route-distance benchmarks to which the Model's estimates of distribution route distances can be normalized.[77] Thus, the rectilinear spanning tree now serves, inappropriately, as a *ceiling*. While AT&T/MCI may claim that this ensures that sufficient cable is contained in the modeled distribution plant, the strand-distance normalization does not fix any of HM 5.3's problems. First, by normalizing HM 5.3's route distances to the rectilinear minimum spanning trees, AT&T/MCI miss the basic point that the minimum spanning tree distance should be the *minimum* against which the modeled route distance is judged. Second, and perhaps more fundamentally, rather than remedy any of HM 5.3's modeling errors, AT&T/MCI use the strand-distance multiplier to simply overwrite its results. However, overwriting fundamentally flawed results does not make the process any more accurate. In fact, as can be seen by reviewing the route distance results prior to the strand-distance gross-up, HM 5.3 still produces cable lengths that are shorter than a minimum spanning tree requires.

Second, outside plant investment is a function of more than just distribution route length. Preprocessing determines numerous, additional aspects of the outside plant network that impact the cost estimates produced by the Model. Even though the minimum spanning tree analyses demonstrate that HM 5.3's modeling of outside plant is incorrect, AT&T/MCI continue to have the Model

[77] *See* Mercer Supplemental Direct Testimony at RAM-4 (Model Description) at p. 37.

use the same inaccurate processes to produce the SAI investment, cable size, cable type, structure mix, and structure sharing. Moreover, *the strand-distance multiplier does not correct any of these errors.* Instead, it grosses-up what we already know is incorrect. Thus, the strand-distance multiplier is far from being a solution to HM 5.3's many problems.

Q. DOES HM 5.3 PRODUCE MORE CABLE THAN VZLOOP?

A. Yes, it does. HM 5.3 models a total loop route distance of 95,642,749 feet, or 18,114 miles. VzLoop, on the other hand, models a total loop route distance of 79,486,330 feet, or 15,054 miles—17 percent less than HM 5.3. Similarly, HM 5.3 models 80,659,622 feet of distribution cable, while VzLoop models 57,086,648 feet.[78] In the recent workshop before the California PUC, Dr. Mercer argued that looking at HM 5.3's modeled route distance somehow proves that the rectangular distribution areas and the uniformly distributed demand in HM 5.3 produce sufficient outside plant.[79] Looking at route distance, or any distance measure for that matter, does not validate the use of these obscure modeling techniques and assumptions. First, loop route distance does not indicate what portion is feeder and what portion is distribution. This is important as a route foot of feeder is more expensive than a route foot of distribution cable.[80] Moreover, HM 5.3 uses different sharing and support

[78] Note that some of this difference is explained by the two models' different demand assumptions. VzCost models 891,788 business and residential lines, while HM 5.3 models 1,000,929 lines.

[79] *See* Verizon CA Workshop Transcript, p. 3536-37.

[80] *See* Murphy Reply Testimony, p. 60-62.

structure assumptions for feeder cable. This means that depending on what fraction of the loop route distance is feeder, cost results will change dramatically. Notably, HM 5.3 models 33 percent less feeder than VzLoop (14,983,127 feet and 22,399,682 feet, respectively). Second, total route distance also does not indicate what type of cable the Model installs. That is, it provides no information on cable size, type (copper or fiber), and whether the supporting structure is aerial, buried, or underground. Thus, the fact that HM 5.3 uses more cable than VzLoop does not mean that it places sufficient outside plant.

Q. WHAT CAN BE LEARNED FROM COMPARING ROUTE DISTANCE STATISTICS BETWEEN THE TWO MODELS?

A. There is only one useful fact that emerges from comparing route distance statistics between the two proposed models. That is, HM 5.3 places more distribution and less feeder cable than VzLoop. As I have found in previous proceedings, HM 5.3 severely underestimates the number of distribution areas. Based on data from its existing network, Verizon NW models over 3,300 distribution areas. HM 5.3, on the other hand, bases its cost estimates on merely 1,018 distribution areas, 70 percent less than Verizon NW.[81] Since the wire center serving areas are fixed, this means that HM 5.3 produces fewer and larger distribution areas than VzCost. This, in turn, means that HM 5.3 underestimates the number of SAIs, DLCs, and most importantly feeder cable

[81] This conservatively assumes that even outlier clusters in HM 5.3 clusters form independent distribution areas.

route distance. In essence, large distribution areas produce more of the less expensive distribution cable relative to feeder cable, thereby understating the cost estimates produced.[82]

V. HM 5.3's Preprocessing Is Riddled with Technical Errors

Q. CONCEPTUAL ERRORS ASIDE, HAVE YOU IDENTIFIED ANY OTHER SPECIFIC ERRORS?

A. Yes. There are many technical errors and flaws that I discovered during my review of the cluster input database. For instance, TNS placed Verizon NW's customer along freeways and their on- and off-ramps. These errors are somewhat different from the ones discussed above, as they seem to be the result of flawed coding rather than poor modeling. A number of these errors are repeats from previous proceedings. Although AT&T/MCI have made numerous changes to the Model and its preprocessing, and claim that the Model has been subjected to the "refiner's fire," they apparently decided it was not necessary to correct these many, blatant errors.[83]

It should be emphasized that, even if these technical errors were corrected, this, by no means, would correct the numerous conceptual errors (such as the assumption that customers are uniformly distributed among equal-sized lots in rectangular-shaped distribution areas and that there are no physical obstacles or rights-of-way) discussed above. These technical errors simply add yet

[82] Mr. Murphy discusses a variety of errors that cause HM 5.3 to understate feeder route distance, including oversized clusters, misclassification of feeder plant as distribution plant, and failure to recognize most indoor SAIs. *See* Murphy Reply Testimony at pp. 57, 59-62, 70-76.

[83] Mercer Supplemental Direct Testimony at p. 32.

another wrinkle to the ever-growing critique of AT&T/MCI's cost model. To the extent possible, I will discuss the potential impact of the errors on the Model's cost estimates.

Q. DOES IT MATTER WHETHER AN ERROR OVERESTIMATES OR UNDERESTIMATES TELRIC COSTS?

A. Not if the goal is accuracy. AT&T/MCI often seek to explain away an error by claiming that it has either minimal impact or overestimates cost.[84] This response is incorrect for several reasons. First, as I discuss above, major cost drivers, like the number of distribution areas, have virtually no impact on the cost estimates produced by HM 5.3. This, however, does not mean that clustering is not important and that it can be done incorrectly since it has "virtually no effect on the results."[85] The reason that many of the corrections have no effect on the cost results is because the Model is conceptually flawed. Second, AT&T/MCI are quick in identifying errors that increase costs, and stating that Verizon NW should not be concerned because the error, in effect, works in Verizon NW's favor. What AT&T/MCI neglect to mention is that many of these errors affect other components of the Model, and thus cannot be viewed in isolation. Rather, to properly determine the manner in which an error affects costs, the error must be corrected and the Model must be rerun. Third, many errors would require changing fundamental modeling assumptions, and thus a complete rewrite of the Model. Since this would require substantial

[84] *See* Mercer SBC Rebuttal Decl. at p. 5
[85] Mercer SBC Rebuttal Decl. at p. 5.

resources and still would not address the more severe conceptual flaws discussed earlier, this exercise would be highly unproductive. Even if all the technical errors were corrected, HM 5.3 would still not produce accurate cost estimates. Finally, since AT&T/MCI are proposing HM 5.3 and touting it as a state-of-the-art model that yields highly accurate costs, they should be responsible for correcting the Model's errors and conceptual flaws, rather than attempting to explain them away as supposed non-issues.

A. HM 5.3's Clustering Algorithm Creates Illogical Clusters

Q. DOES THE CLUSTERING ALGORITHM PERFORM AS DESCRIBED IN THE MODEL'S DOCUMENTATION?

A. No, it does not. The HM 5.3 Model Documentation describes the clustering as a nearest-neighbor procedure that clusters subject to three constraints: (1) no point can be more than 17,000 feet from the cluster centroid; (2) no point can be more than two miles from its nearest neighbor; and (3) no cluster can exceed the 6,451-line size.[86] Initially only two of these constraints are followed. It is not until a later stage of the clustering process that the third constraint is incorporated. Thus, instead of clustering subject to three constraints, the clustering algorithm clusters subject to two constraints and then "chops up" the results until it meets the third constraint.

More specifically, the clustering algorithm initially ignores the third constraint (cluster line size limit) and produces clusters that exceed the maximum cluster

[86] *See* Mercer Supplemental Direct Testimony at RAM-4 (Model Description), p. 21-22.

line size. It then divides (into two equal-sized parts based on the area) the oversized clusters along the shortest axis of the minimum-bounding rectangle of the cluster in question until the maximum line count restriction is met. This method of clustering is contrary to what the HM 5.3 Model Documentation describes, and does not make any modeling or engineering sense. No real-world outside plant engineer would design distribution areas in such a manner.

Further, given the vast amount of literature on clustering, it is surprising that TNS uses such a simplistic method to group customer locations. What seems even more surprising is that Dr. Mercer claims that the resulting clusters have "a realistic correlation to efficient distribution areas."[87]

Q. HOW DOES THIS ERROR IMPACT THE CLUSTERING RESULT?

A. The following figures illustrate HM 5.3's clustering process and how it impacts the cluster results. Figure 1 is an illustrative distribution of customers in a square wire center serving area. Assume the wire center is in the lower right hand corner of the serving area. The diagonal in the serving area is 17,000 feet. There are 63 customer locations, each with one line. These customers are spread somewhat uniformly within the serving area, and no customer lives more than two miles from its nearest neighbor. There are also 16 locations, located in the upper left corner (the diamond symbols) with approximately 403 lines apiece and 6,451 in total. Thus, the entire wire center serving area has

[87] Mercer Supplemental Direct Testimony at RAM-4 (Model Description) p. 3

6,514 lines. Since these 16 locations do not exceed the 536-line threshold, they are not considered as high-rises by HM 5.3.

Figure 1: Wire Center Serving Area

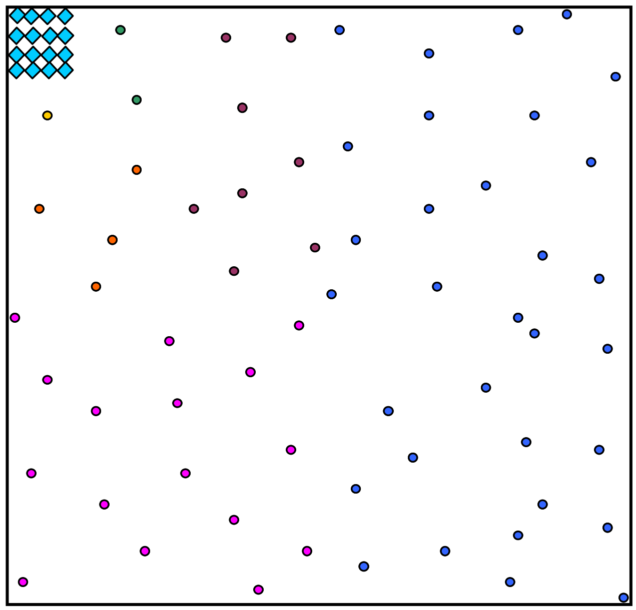
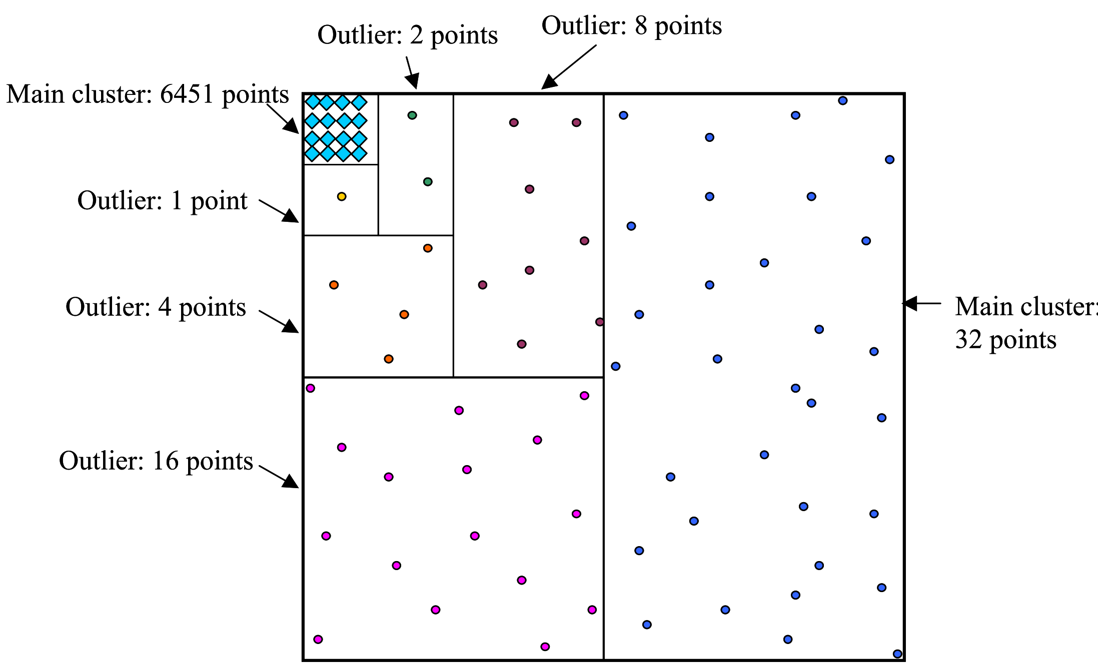


Figure 2 illustrates how HM 5.3's preprocessing would cluster such a situation. In the first pass of the clustering procedure, the Model would group all customers into one large cluster, as there are no customers that live more than 17,000 feet from the wire center and no customer lives more than two miles from its nearest neighbor. In the second pass, the Model takes into account that the cluster has exceeded the 6,451-line constraint (since there are 6,514 lines) and splits the cluster along its shortest axis. This split creates two clusters, one with 32 lines and another with 6,482 lines. Since the second cluster still exceeds the 6,451-line constraint, the Model then splits the second cluster along its shortest axis. As a result, there are now three clusters: one with 32 lines (unchanged from the first split), one with 16 lines, and one with 6,466 lines. The splitting process continues until the Model has created seven clusters, two main clusters and four outlier clusters. These clusters and their associated lines are illustrated in Figure 2 below.

Figure 2: Distribution Areas within the Wire Center Serving Area



Had the clustering algorithm functioned as described and clustered subject to all three constraints simultaneously, the clustering results most likely would not have produced two main clusters and four outliers. Rather, two main clusters (one with 6,451 lines and another with 63 lines) or three main clusters (one with 6,451 lines, another with 32 lines, and a third with 31 lines) would have been produced.

Although hypothetical, this example is merely intended to illustrate that had TNS clustered using a more reasonable approach, or at least one consistent with its own documentation, the resulting clusters would have been significantly different. Because of this error, the Model will produce outliers when it should have included the lines in the main clusters, and sometimes it will produce outliers when it should have produced more main clusters.[88]

B. HM 5.3's Clusters Violate AT&T/MCI's Own 17,000-Foot Copper Length Threshold

Q. HOW DOES HM 5.3 VIOLATE THE 17,000-FOOT COPPER LOOP LENGTH THRESHOLD?

A. Dr. Mercer explains: "No point in a cluster may be more than 17,000 feet distant (based on right angle routing) from the cluster's centroid."[89] To achieve this, the clustering software checks whether each individual customer location is more than 17,000 feet from the cluster's centroid. However, this initial constraint is later overridden by HM 5.3, causing the Model to produce copper loops that clearly exceed the 17,000-foot (or even the Revised Resistance Design Standard's 18,000-foot)[90] threshold. The cause of this violation of AT&T/MCI's own threshold is the strand-distance multiplier. As discussed, this multiplier expands or contracts HM 5.3's distribution route distance by forcing it to match the length of the rectilinear minimum spanning tree. AT&T/MCI, however, fail to

[88] If the Model is not sensitive to a small increase in cost, then it should not be sensitive to a small decrease.

[89] Mercer Supplemental Direct Testimony at RAM-4 (Model Description), p. 21-22.

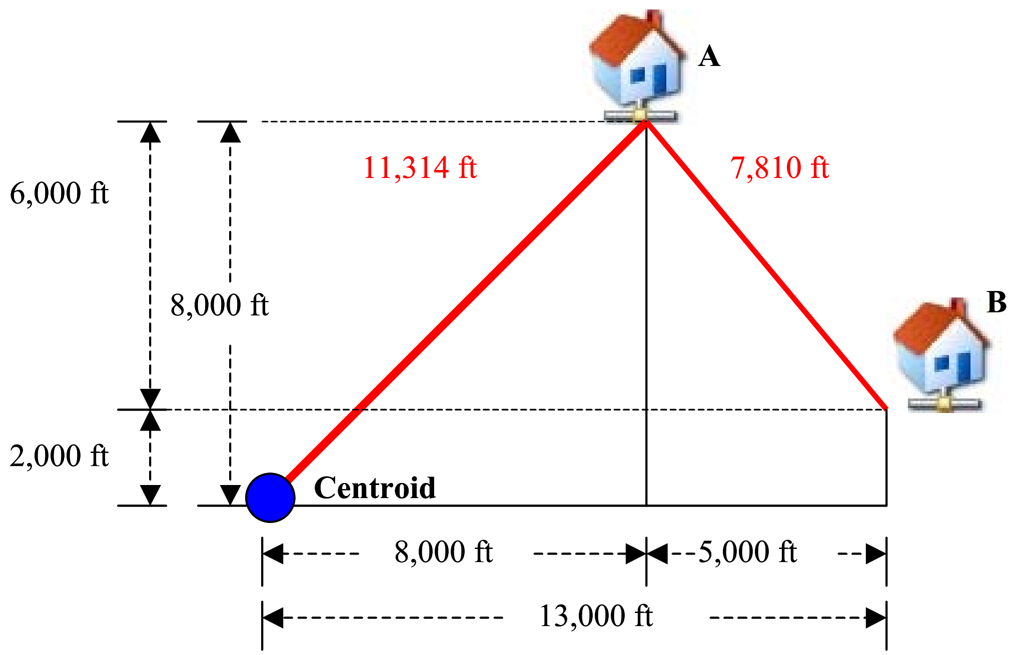
[90] Mr. Murphy discusses this standard on pages 45-46 of his Reply Testimony.

constrain the Model's distribution route distance *after* the strand distance gross-up to 17,000 feet. This omission leads to numerous occasions where copper loops exceed 17,000 feet. In fact, as Mr. Murphy points out, some clusters in HM 5.3 have copper loop lengths as long as 38,000 feet.[91]

To illustrate this point, consider the following example. There are two neighboring customer locations, A and B, which the clustering algorithm checks to see if they can be included in a cluster. Location A is 16,000 feet and location B is 15,000 feet from the centroid, based on right-angle routing. The clustering algorithm would include these two locations in the same cluster. However, when the preprocessing module calculates the strand distance (to which the distribution route distance is grossed-up), it determines the rectilinear minimum spanning tree, which in this case extends from the cluster centroid to location A and then to location B. This results in a strand distance of 27,000 feet. If for instance, HM 5.3 computes a total distribution route distance of 19,000 feet for such cluster, with the farthest customer at 14,000 feet, then HM 5.3 would gross up all distances by a factor of 1.42 (27,000/19,000). This gross-up would increase the length to the farthest customer from 14,000 feet to approximately 20,000 feet and thus exceed the 17,000 (and 18,000 foot) threshold. This example is illustrated in Figure 3 below.

[91] *See* Murphy Reply Testimony at pp. 45-46.

Figure 3: The 17,000-Foot Threshold Is Exceeded



Dr. Tardiff further addresses this issue and provides additional evidence as to HM 5.3's violation of its own copper threshold.[92]

Q. WHAT IS THE IMPACT OF THIS ERROR?

A. By exceeding the 17,000-foot threshold, it is clear that there are too few distribution areas modeled in HM 5.3, which results in an underestimation of SAIs, DLCs, and feeder cable. It also means that the strand-distance multiplier is incorrect, producing too much distribution cable. This is exactly what I had found previously by comparing HM 5.3 feeder and distribution route distances

[92] *See* Tardiff Reply Testimony at p. 77-78.

to VzLoop. Without access to the source code of the clustering algorithm, this error cannot be corrected easily.

C. Converting Households and Firms to Lots Is Nonsensical

Q. IS IT REASONABLE TO CONVERT HOUSEHOLDS AND BUSINESSES INTO LOTS?

A. No, it is not. In fact, this is one of HM 5.3's principal failings. It does not model outside plant to actual customer locations but to hypothetical lots. As discussed above, as part of the preprocessing, TNS estimates the number of households and businesses by summing unique geocoding locations with the same line type. Through this process, TNS determined that there were 579,375 unique households and businesses in Verizon NW's territory. By unique, I mean that these locations have either a different longitude or latitude or both, and therefore are located at different street addresses. In HM 5.3, however, these 579,375 unique records are then further converted to lots using the assumption that households and businesses occupy either a whole, half, or quarter lot. Through this process, the Model reduces the 579,375 unique household and business locations to 437,027 lots. Then, HM 5.3 models its network to these 437,027 lots and determines UNE cost estimates. This reduction to customer lots makes no sense as TNS has already determined that there are 579,375 unique locations. This example illustrates, once more, how HM 5.3 overwrites previously determined information in favor of its quixotic network layout.

Q. WHAT ARE THE COST CONSEQUENCES OF THIS ERROR?

A. By converting households and businesses to lots, AT&T/MCI conveniently reduce the size of the modeled network, from a network that would serve 579,375 *locations* to one that serves 437,027 *lots*. Since a smaller network means less investment and because demand remains unchanged, overall per-loop costs necessarily will be understated.[93] In order to assess the exact cost differential, one would have to recode the Model and employ a modeling technique similar to the one used by VzLoop where outside plant is built to actual customer locations rather than lots that bear no relation to where customers are located.

D. The Modeled SAI Locations Are Incorrect

Q. PLEASE ELABORATE ON YOUR CONCERN REGARDING THE PLACEMENT OF THE SAIS.

A. The SAI locations modeled by HM 5.3 are wrong. As part of the clustering algorithm, the preprocessing module places the SAI halfway between the two farthest points in a cluster. While Mr. Donovan and Dr. Mercer might argue that this is merely a surrogate for the actual location of an SAI, it simply is not where an engineer would place an SAI. It is also not the centroid of the convex hull of a cluster, as AT&T/MCI continue to claim in their documentation.[94] There are three acceptable approaches for locating the centroid of the convex hull that ultimately is converted into a rectangular-shaped distribution area: geographic

[93] *See* Tardiff Reply Testimony, n. 115.
[94] *See* Mercer Supplemental Direct Testimony at RAM-4 (Model Description), p. 14.

area, point-location weighted, and line weighted. The geographic approach ignores customer locations (or lines) and simply calculates the center of the boundary formed by the convex hull. A location-weighted approach considers customer locations, but tends to locate the centroid in more populous areas. The line-weighted approach uses customer lines as its basis. A line-weighted centroid, similar to that used by the FCC's Synthesis Model, is the best method for cost estimation purposes as it actually determines the center of the mass.[95]

HM 5.3 ignores all these acceptable approaches and instead places the SAI at half the distance between the two furthest points in a cluster. As a result, SAIs are placed in areas far removed from customers, outside wire center boundaries, in rivers, on top of freeways, and in the middle of downtown areas—none of which are places where an engineer would place an SAI.

Q. WHAT ARE THE CONSEQUENCES OF MISPLACING THE SAIs?

A. Misplacing the SAIs has serious consequences. First, since the SAI serves as the center point of the rectangular cluster, misplacing the SAI shifts demand to the area surrounding the SAI, regardless of whether the SAI is even close to actual customer locations. Second, since the feeder-distribution demarcation point is at the SAI, misplacing the SAI will result in inaccurate estimates as to how much of a given loop is feeder and how much is distribution. Third, because the clustering algorithm calculates the 17,000-foot constraint from the

[95] *See* C. A. Bush, D. M. Kennet, J. Prisbrey and W. W. Sharkey of the Federal Communications Commission and Vaikunth Gupta of Panum Telecom, LLC, The Hybrid Cost Proxy Model, Customer Location and Loop Design Modules (Aug. 19, 1998), p. 1.

centroid, misplacing the SAI will result in incorrect cluster sizes (i.e., they would be measured from the wrong point).

Dr. Mercer responded to my criticism in SBC CA's UNE proceeding by claiming: "This calculation is intentional and appropriate because it is the point at which the maximum distance from the DLC, or SAI is minimized, thereby creating more efficient and better quality telephone service."[96] Dr. Mercer missed the point. I did not question the use of the SAI location as the center point; rather, I questioned how that center point was determined. Dr. Mercer never explained why his method for locating the center point produced accurate results. Moreover, he continues to refer to the point that is half the distance between the two furthest points in a cluster as the cluster centroid, although he has already admitted that this point is not a true centroid.[97]

To determine the cost consequences of placing an SAI at one-half the distance between the two farthest points, one would need to correct the code that determines this point to properly reflect a cluster centroid, using one of the three concepts described above. Because this would require access to the clustering source code, which has not been made available to Verizon NW, I am unable to perform this analysis. However, I do expect that, once corrected, remedying this error will result in a different set of clusters yielding different cost results.

[96] Mercer SBC Rebuttal Decl. at p. 22.
[97] *Id.*

E. HM 5.3's Cluster Input Database Contains Many Questionable Entries

Q. PLEASE ELABORATE ON YOUR CONCERNS REGARDING THE HOUSING UNIT ENTRIES IN THE CLUSTER INPUT DATABASE.

A. The unrealistic nature of the outside plant network modeled by HM 5.3 becomes particularly clear when reviewing the cluster input database. For instance, Dr. Mercer explained in the recent California UNE workshop:

> There's information in each cluster record that indicates what type of households there are, meaning are they single, detached homes, duplexes, five or less residential buildings et cetera. So you have a record of the kind of households which will matter when it comes to placing drops and how big the drops are.[98]

Similarly, the HM 5.3 Model Description states: "For residences, the census database supplied by TNS identifies the number of households located in various types of buildings."[99] However, when reviewing these entries in HM 5.3's cluster input database, it becomes evident that they must be wrong. For instance, consider the Newport wire center (NWPTWAXX), main cluster c006. According to TNS and AT&T/MCI, this cluster contains one household in a 50+ unit building. HM 5.3 then assumes that this household is on a lot one-fourth the size of a single-family detached residence lot. This is wrong on many levels. First, it is unknown how AT&T/MCI support this seemingly arbitrary assumption. It is similarly unclear why other multidwelling housing types should occupy land equal to half the size of a single-family detached

[98] Verizon CA Workshop Transcript at pp. 3458-59.
[99] Mercer Supplemental Direct Testimony at RAM-4 (Model Description), p. 35.

residence lot. Second, and more generally, it is unrealistic to derive the lot size of a business or multi-tenant building by assigning each household in these housing types to a fraction of a single-family detached residence lot.[100] Third, it is similarly unrealistic to assume that 50+ unit buildings exist with only one tenant (or household). In HM 5.3's world, however, Verizon NW's serving area in Washington has many multidwelling buildings that are occupied at a small fraction of their capacity. Finally, in the preprocessing module, TNS has already identified the number of unique residences and businesses. Thus, by definition, the locations in the cluster input database are *unique* and thus cannot fall on the same lot. Nevertheless, AT&T/MCI entirely ignore these results and cram residences and firms on lots that bear no resemblance to the demand distribution faced by Verizon NW.

Q. IS THIS THE ONLY EXAMPLE OF ILLOGICAL HOUSING UNIT ENTRY IN THE CLUSTER INPUT DATABASE?

A. No, it is not. I want to stress that the problems associated with the conversion to lots is not an isolated issue applying to select observations. Quite the opposite, this problem extends throughout the database and impacts the large majority of households and firms. In fact, the vast majority of household counts per housing type suffers from the same fundamental error as the example discussed above. What is most disturbing is the fact that TNS has the street

[100] For instance, per HM 5.3, if a single-family residence is determined to occupy an area of 10,000 square feet, then: (1) all other single-family detached residences in the same cluster have the same lot size; (2) firms occupy an area equal to 5,000 square feet times the number of firms in a building; and (3) large multi-tenant buildings occupy an area equal to 2,500 square feet times the number of households in that building. These assumptions make absolutely no sense.

addresses for all these households and firms, their longitude and latitude, and their estimated number of lines. Yet, HM 5.3 totally ignores this information and instead spreads households and firms over lots that bear absolutely no resemblance to the customer locations determined by TNS. This is yet another illustration of how information is ignored because of modeling techniques that are not only divorced from reality, but produce results that make absolutely no sense.

Q. WHAT ARE THE COST CONSEQUENCES OF THIS ERROR?

A. Short of rewriting the entire preprocessing and distribution modules, I cannot estimate the cost impact of this particular modeling decision. However, in general, smaller networks tend to have lower total costs and lower unit costs. In addition, the more customers are equally distributed, the more the modeled network deviates from the actual network. Again, there is no excuse for this arcane modeling technique, as the technology to model along feasible network routes is currently available, and has been for a number of years.

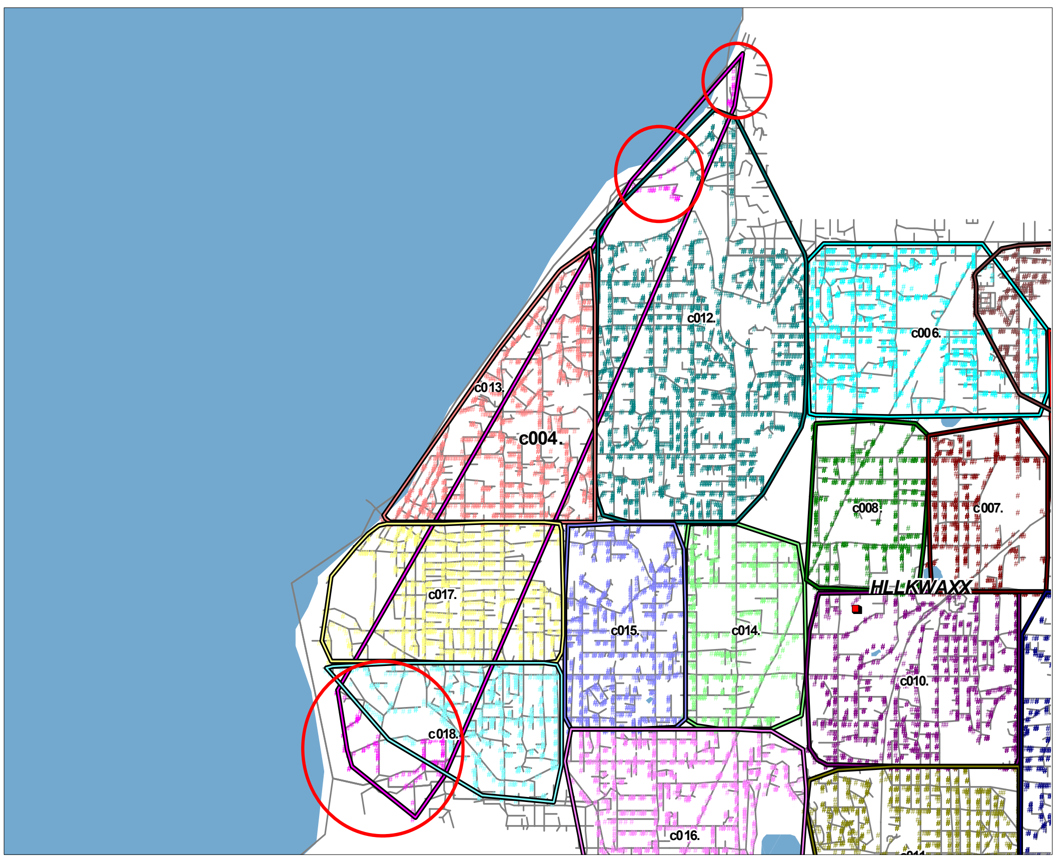
F. The Clustering Algorithm Produces Visual Anomalies

Q. PLEASE ELABORATE ON YOUR CONCERN REGARDING THE CLUSTERS PRODUCED BY THE TNS CLUSTERING ALGORITHM.

A. While I have already described the various conceptual and factual errors contained in the clustering algorithm, actually seeing the obviously incorrect clusters modeled by HM 5.3 provides even further evidence that the Model's clustering algorithm is fundamentally flawed. Included below is a small sampling of the numerous clustering anomalies I encountered during my analysis.

Anomaly 1: Map 2 below illustrates some of the anomalies identified in the Halls Lake (HLLKWAXX) wire center.

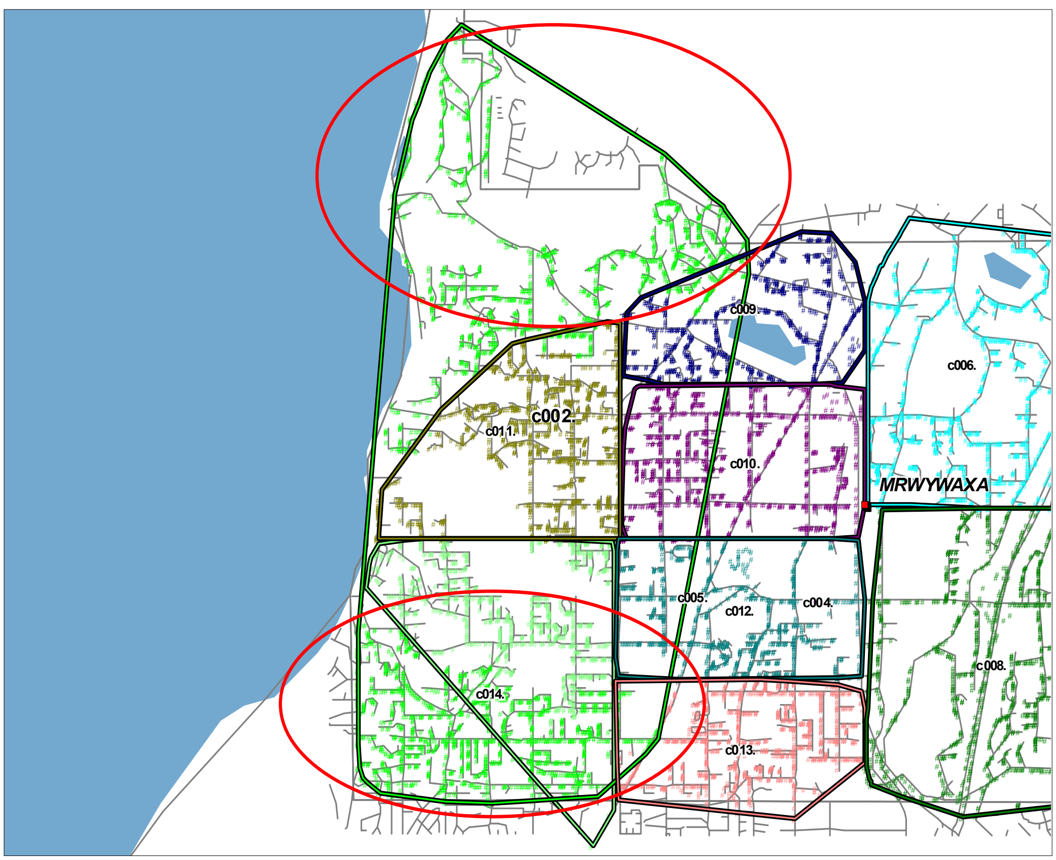
Map 2: Anomaly 1



As can readily be seen in this map, cluster c004 does not comport with realistic groupings of customers. Specifically, the pink customer points of this cluster circled in red in the southern extent of the map should not be in the same cluster as the same colored locations in the north. These locations should be clustered with their nearest locations. For instance, it would be more reasonable to group the circled customer locations in the southern end of the map with customer locations from cluster c018 since they are in closer proximity than the northern customer locations of cluster c004.

Anomaly 2: Map 3 below illustrates some of the anomalies identified in the Manor Way (MRWYWAXA) wire center.

Map 3: Anomaly 2



This map provides one (of many more) examples of apparent clustering anomalies. Again, it is unclear why cluster c002 is intersected by clusters c011 and c014. Rather, it appears to be more realistic to group the southern part of cluster c002 with cluster c014, and the northern part of cluster c002 with cluster c011.

Q. WHAT ARE THE COST CONSEQUENCES OF THESE ERRORS?

A. As I have demonstrated, there is no truth to Dr. Mercer's claim that the "input development process next identifies all customer locations within a wire center's boundaries that are close enough together to be efficiently engineered as a single telephone plant serving area."[101] The clustering algorithm is fundamentally flawed; it produces loop lengths that exceed the modelers' own threshold, and clusters that defy common sense and sound engineering principles. Absent the ability to review and edit the clustering algorithm, there is (unfortunately) no way to put a price tag on the significant errors I have identified.

G. The Dominant CBG Methodology Employed Over-Densifies

Q. PLEASE ELABORATE ON YOUR CONCERNS REGARDING THE USE OF DOMINANT CBGS.

A. Because many input values are determined by density zone, the density zone designation of a cluster is very important. HM 5.3's preprocessing module designates one CBG per cluster as the representative CBG for the cluster (i.e., the "dominant CBG"), and assumes that all other CBGs bear the same density, as well as geographic and demographic attributes. Supposedly, the dominant

[101] Mercer Supplemental Direct Testimony at RAM-4 (Model Description), p. 21.

CBG is the CBG in a cluster with the most lines. Thus, although a cluster might have most of its lines in one set of CBGs, if there is one particular CBG that has the most lines (perhaps due to a high-rise building in the cluster), then all other CBGs will bear the same characteristics as the dominant CBG. In doing so, AT&T/MCI again ignore actual, readily available information regarding line densities and geographic and demographic parameters. AT&T/MCI and TNS know in which CBG (and even which CB) a customer location falls, and thus can determine the exact density zone into which that CBG would fall. Nevertheless, AT&T/MCI have chosen to simplify the modeling process and again disregard actual, reliable data. Although I raised this issue before in the SBC CA UNE proceeding, apparently AT&T/MCI did not deem it necessary to adjust their Model.[102]

Q. WHAT ARE THE COST CONSEQUENCES OF THIS SIMPLIFICATION?

A. In response to my criticism regarding the use of dominant CBGs, Dr. Mercer stated that I did not "suggest exactly how this should be done," and "the alleged error…actually results in *lower* loop costs."[103] I disagree with Dr. Mercer's assertions and note that it is not Verizon NW's responsibility to correct HM 5.3. A possible way to capture more accurately the density of distribution areas is to use a line-weighted average of all CBGs or even CB densities in a cluster. Surely, either Dr. Mercer and/or Mr. Kevin Landis (of TNS) must know how to do this, or they would not have been able to perform a run and conclude that it

---

[102] *See* Dippon SBC Declaration at p. 19.
[103] Mercer SBC Rebuttal Decl. at p. 12 (emphasis in the original).

supposedly lowered costs. Moreover, Dr. Mercer's alleged lower loop costs are suspicious and might simply be another example of the illogical results and erroneous sensitivities produced by the Model. By using the dominant CBG to determine the density for a cluster, HM 5.3 ignores the fact that many clusters include lines that are in CBGs other than the dominant one. These non-dominant CBGs are generally less dense as is demonstrated in Table 1 below.

Table 1: Line Count by Density Zone

Density Zone Range			
Lines per Square Mile	**Actual**	**HM 5.3**	**Difference**
0–5	20,582	20,253	1.62%
6–100	140,468	103,940	35.14%
101–200	70,509	59,425	18.65%
201–650	117,632	127,464	-7.71%
651–850	42,550	32,694	30.15%
851–2,550	272,007	225,663	20.54%
2,551–5,000	242,451	268,329	-9.64%
5,001–10,000	84,314	141,425	-40.38%
10,000+	6,047	17,367	-65.18%
Total	**996,560**	**996,560**	

The column entitled "HM 5.3" in Table 1 shows the total lines per density zone as determined in the preprocessing module using dominant CBGs. The column entitled "Actual" contains the lines per density zone calculated using actual CBG density zones rather than dominant CBG representations. As can be seen, using dominant CBGs, instead of the (available) CBG data, tends to over-densify Verizon NW's serving area. For instance, using dominant CBG densities instead of actual CBG densities artificially places roughly three times the amount of lines in the highest density zone—leading to a general over-densification of lines.

As a result, inaccurate density zones are applied throughout the estimation process, resulting in imprecise outside plant configurations, erroneous sharing percentages, and so on. The sensitivity result that Dr. Mercer seems to be referring to is suspicious, since, as illustrated above, HM 5.3's UNE cost estimates barely change, even when radical changes in modeling assumptions are made.

H. HM 5.3 Suffers from a Low Geocoding Success Rate

Q. PLEASE ELABORATE ON THE GEOCODING SUCCESS RATE IN HM 5.3'S PREPROCESSING MODULE.

A. The geocoding success rate in HM 5.3's preprocessing module is surprisingly low. TNS was only able to successfully geocode 42 percent of the business locations and 85 percent of the residential locations. For other service types, such as SW DS-1 (switched) and NS DS-1 (nonswitched), TNS could only geocode 19 percent and 22 percent, respectively. Although HM 5.3 only relies marginally on the results of this placement exercise, the low geocoding success rates cast further doubt on the overall accuracy of the Model's customer placements. These extremely low geocoding success rates further distort the cost estimates produced by HM 5.3.

I. Customers Are Incorrectly Surrogated

Q. DOES HM 5.3 SURROGATE VERIZON NW CUSTOMERS ON APPROPRIATE ROADS?

A. No, it does not. In its surrogation process, TNS seems to have included roads where no customers are or would be located. For instance, it appears that TNS included roads such as limited-access highways (TIGER type A1x), the ramps to these highways (TIGER Type A63), as well as private driveways and service roads (TIGER Type A74). By including these types of roads, TNS has surrogated Verizon NW's customers erroneously along highways, on- and off-ramps, private driveways, and service roads. In the Bothell (BOTHWAXB) wire center, TNS placed surrogate points along Interstate 405 and State Highway 522, both limited-access highways where no customers are located. In the same wire center, TNS placed business and residential locations along private driveways and service roads.

J. Customer Locations Are Erroneously Dropped from the Cluster Input Database

Q. WERE ALL OF VERIZON NW'S SERVICE ADDRESSES CORRECTLY PREPROCESSED?

A. No, they were not. In my review of the TNS geocoding exercise, I noticed 2,533 customer locations with geocoding results of 0,0; that is, a longitude of zero and a latitude of zero.[104] These geocoding results are clearly incorrect, as they

[104] The Everett Casino (EVRTWAXC) wire center contains 1,429 of these incorrect geocoding results; 53 are in the EVRTWAXCCG1 wire center and 1,051 are customer locations for which wire center information is missing.

would place the location at the equator. It appears that due to the erroneous geocoding results, these locations are also dropped from the cost modeling exercise. While it is unclear whether the demand for these locations is also removed, the impact of this error is likely a further underestimation of the plant required to serve Verizon NW's customers in Washington.

Similarly, there are other inexplicable data discrepancies. For instance, according to the TNS's preclustering process, the Juanita (JUNTWAXA) wire center contains a location with 1,055 business lines. According to Dr. Mercer's definition of a high-rise cluster, this location should be a high-rise cluster. Yet, when reviewing the clusters designated as "high-rises" by HM 5.3, it is clear that there is no high-rise cluster in this wire center. In fact, this location appears to be dropped from the analysis, as it does not seem to be carried over from the preclustering to the clustering portion of the TNS preprocessing module.

K. Distribution Terminal Investment is Incorrectly Determined

Q. PLEASE EXPLAIN WHY THE TERMINAL INVESTMENT IN HM 5.3 IS INCORRECT.

A. Terminal investment in HM 5.3 is determined by dividing the number of lots by four; the assumption is that four lots share one distribution terminal. This is incorrect for several reasons. First, as discussed, HM 5.3's concept of modeling to lots instead of actual customer locations is one of its principal failings. Thus, any estimate from this flawed approach will yield incorrect numbers. Second, it is unrealistic to assume that a distribution terminal will always be shared by four lots. Per HM 5.3, regardless of how many distribution areas are formed, distribution terminal investment remains unchanged. This is because the Model assumes that there will always be four lots sharing a distribution terminal, regardless of how far apart these lots might be.

Q. WHAT IS THE IMPACT OF THIS ERROR?

A. This error leads to an underestimation of terminal costs for at least two reasons. First, by assuming distribution terminals serve lots instead of customer locations and then cramming customer locations onto lots, the absolute number of distribution terminals is understated. Second, by assuming that each of these terminals serves four lots, the terminal investment is further underestimated as in reality customers might live quite distant from each other and four customers may not always be able to share a distribution terminal.

L. The TNS Preprocessing Results Are Manually Manipulated by AT&T/MCI Prior to their Inclusion in HM 5.3.

Q. PLEASE EXPLAIN HOW AT&T/MCI MANUALLY MANIPULATE THE TNS PREPROCESSING RESULT.

A. Although TNS explains in its brief documentation that the end result of its preprocessing is used in HM 5.3, Verizon NW recently learned that this is not true. AT&T/MCI have now disclosed that Mr. Douglas Denney (of AT&T) has further manipulated the file received from TNS and performed a so-called true-up of lines. While the file for this true-up was provided, AT&T/MCI have not explained how and why the TNS end-result was manipulated. More important, the file provided by Mr. Denney can only be used for the default scenario and cannot be incorporated in a sensitivity analysis as it is unclear how and why items were adjusted. This issue further illustrates the difficulties faced when trying to understand and review HM 5.3's preprocessing.

VI. DECIDING BETWEEN TWO COST MODELS

Q. WHAT CRITERIA WOULD YOU ADVISE THIS COMMISSION TO USE IN DECIDING BETWEEN VZCOST AND HM 5.3?

A. In deciding between the two proposed cost models, I recommend that the Commission check for completeness and accuracy. Completeness guarantees that a cost model properly account for *all* the costs that a local exchange carrier incurs. Focusing strictly on outside plant, I find that VzCost is superior to HM 5.3, as it explicitly takes into account all the important cost drivers, such as the costs incurred by having to route around natural and manmade obstacles and accommodating rights-of-way. In terms of accuracy, I also find that VzCost is superior to HM 5.3. Unlike HM 5.3, VzCost does not discard existing customer locations, but models to actual distribution terminals that serve actual customer locations. Moreover, VzCost does not use a surrogation process that places customers along private driveways or on- and off-ramps. Instead, it relies on the average costs of the network modeled to connect those distribution terminals that it has located. VzCost uses feasible network routes, and accurately and properly reflects costs that a local exchange carrier can expect to incur. HM 5.3, on the other hand, constructs an impossibly ideal network, using arcane modeling techniques that no local exchange carrier would ever build. Further, HM 5.3 seems to be internally flawed, as many of its cost drivers do not function properly. With respect to completeness and accuracy, VzCost is much better than HM 5.3.

Q. WHAT ARE THE ECONOMIC CONSEQUENCES OF SELECTING AN INACCURATE MODEL?

A. One of the most important reasons for the implementation of price regulation based on the use of cost proxy models is related to the need to send the appropriate signals to the market, and foster efficient entry, adequate investments, and innovation. Relying on the cost estimates produced by HM 5.3 would distort important market signals. Specifically, it would distort a potential new entrant's perception of the marketplace and discourage efficient entry. It would force the incumbent carriers to cross-subsidize their competitors, thereby severely hampering (if not thwarting completely) facilities-based competition. In short, relying on HM 5.3 would defy the very objective of the Telecommunications Act of 1996.

VII. CONCLUSION

Q. WILL YOU BRIEFLY SUMMARIZE THE MOST SERIOUS FLAWS IN HM 5.3 THAT LEAD YOU TO BELIEVE THAT VZCOST IS A BETTER MODEL?

A. Yes. HM 5.3 produces an outside plant network that is entirely removed from reality. It does not make sense theoretically, because in the real world customers are not uniformly distributed on equally sized lots, distribution areas are rarely rectangular, and there are physical obstacles and rights-of-way that must be taken into consideration. The network modeled by HM 5.3 also makes no sense when illustrated with maps, since even a cursory look at the maps depicting HM 5.3's network shows that the array of distribution and feeder cable produced cannot and does not yield accurate cost estimates. Finally, the Model has obvious internal defects, as sensitivity analyses simply make no sense—costs decrease when they should increase or they do not change at all even when a fundamental modeling assumption is changed or an important cost driver modified. Add this to the countless errors in the database and other significant miscalculations discussed by Dr. Tardiff and Mr. Murphy and you have a clear picture: HM 5.3 must be rejected.

Q. DOES THIS CONCLUDE YOUR REPLY TESTIMONY?

A. Yes, it does.