3. The HAI Model

Q. Please turn to the third section of our testimony, concerning the cost estimates presented by AT&T in this docket. Are you familiar with the cost model used by AT&T for this purpose?

A. Yes. Over the past few years I have examined several versions of this model, formerly known as the Hatfield Model and now known as the HAI model. Members of my firm and I have recently examined version 5.0a in Idaho and Nevada and we are generally familiar with the process used by in developing the UNE cost estimates that were submitted by AT&T in this proceeding.

Q. Would you please summarize your understanding of the general properties of the HAI model and its approach to modeling UNE costs?

A. The HAI Model shares many general characteristics with other models used in estimating telecommunications costs. It provides hundreds of user-adjustable input choices, processes these inputs through numerous different algorithms, and generates outputs that can be summarized in several different ways. It relies primarily on publicly available data rather than proprietary data, places its core logic, or algorithms, in one or more Excel spreadsheets, and uses Visual Basic to automate various functions. It attempts to satisfy the FCC's 10 criteria for acceptable universal service cost models and relies upon the Bellcore's LERG data base in identifying the location of wire centers and it uses data provided by BLR to estimate the location of wire center boundaries.

Q. Is HAI "user friendly"?

A. Yes. Compared to SCIS and some of the other models that have historically been used by incumbent local exchange carriers, HAI is very user friendly. Although bugs continue to be found in it (as in all such models), HAI can be used effectively by a reasonably informed analyst. Inputs are readily accessible and fairly straightforward. It is possible to change assumptions and source data quickly and rerun the model, immediately seeing the effect on model outputs. This ability contrasts starkly with the situation with some ILEC models, where analysts working for regulatory commissions or other parties must ask ILEC personnel to input changes, rerun the model(s), and send the results to the analyst.

Q. Does HAI have some drawbacks?

A. Yes. HAI is primarily a national model, and any national modeling effort is fraught with difficulty. The unique geographic and demographic attributes of the various states make it difficult, if not impossible, to develop a single cost study to accurately model every part of every state. This is particularly true for a national proxy model attempting to maintain nationwide consistency of methodology and data sources and working within time and budget constraints. As one narrows the focus to individual wire centers within each state, it becomes apparent that the HAI "one size fits all" approach, uniformly applied across the entire country, involves limitations and compromises which reduce the usefulness and accuracy of the resulting cost estimates.

Q. But wasn't HAI customized for purposes of the NH TELRIC studies?

A. Yes, to a limited degree. Customization was of two kinds--data specific to individual BA-NH wire centers, and "scenario inputs" modifying HAI's default settings. Unlike the Company's cost studies, which focused on just 3 archetypal wire centers, the HAI model attempts to estimate costs for numerous individual wire centers within the state, using state-specific geographic and customer location data. model. The "scenario inputs" were confined to 35 inputs, listed in an attachment to the testimony of AT&T witness Stephen E. Siwek, relating to capital costs, depreciation lives, and salvage values adopted by the Commission in DE-96-252. As I discussed earlier, the Company used similar inputs in its cost studies, as a result of discussions with the Staff.

Q. You spoke of HAI using data relating to individual BA-NH wire centers. What is the function of this data?

A. As I just indicated, this data is primarily geographic in nature. It is used to enumerate and locate end-use customers, to specify the type of terrain, and to specify the size and configuration of the cable network that comprises each wire center. This data determines the number of working links, the sizing of switches and cabling, and the extent and kind of supporting structures within each wire center.

Customer locations and other geographic data have a significant impact on link costs, since these factors determine loop length, terrain-related costs (e.g. additional labor required to dig trenches or install poles in rocky soil) and the extent to which customer density allows economies of scale to be exploited. Loop length potentially helps determine the presence and nature of loop electronics as well as the total quantity of cable and supporting structures (e.g. poles).

In developing telecommunications cost estimates, the accuracy of the results depends heavily on the quality of the customer location and other geographic data used in the cost model, as well as the specific assumptions and process used in interpreting this data. In this regard, it is noteworthy that simplifying assumptions are necessarily used in all cost models, particularly with regard to distribution costs. However, the HAI model uses a less simplified, more accurate approach to the geographic aspects of cost modeling than the BA-NH model presented in this proceeding. This is particularly noticeable with regard to distribution costs, where the BA-NH approach is greatly over simplified, and results in a substantial underestimate of distribution cable investment.

Q. How does HAI estimate customer lines and locations?

A. As reported by Mr. Siwek, HAI uses various data sources to estimate the number of telephone lines and their location, including geocoding of customer addresses, company data on wire center, tandem, and STP locations, recent (1996) ARMIS data, and user-adjusted inputs [pp. 22-23.]

Q. Are address geocodings from public sources totally reliable for this purpose?

A. No. Discrepancies are bound to occur, because the phone books, mailing lists, and other data sources are not identical to the ILEC subscription data. For example, unpublished numbers are by definition excluded from the white pages and must be estimated. Mailing lists include people with unlisted phones, but they also include people who do not have a phone. Census data can be helpful in areas where address listings do not provide an adequate data source, but the 1990 census data are now nearly a decade old. The available updates are merely estimates--and not necessarily very accurate. Although white pages data exclude unpublished numbers, they are generally more accurate than census data in capturing the impact of differential growth rates and movement of customers within the state. Hence, precise address listings are preferable where they are available. However, in locations where public data sources for customer addresses fail to precisely locate a large fraction of the customers, estimating techniques, or alternative data sources (e.g. census data) must necessarily be relied upon.

The developers of the HAI model have gone to considerable effort and expense in an attempt to accurately deal with the geographic aspects of the cost modeling problem. And, it is fair to say that the latest version of the HAI model is improved in this regard, relative to earlier versions of the Hatfield model. Nevertheless, the HAI model still exhibits weaknesses in this area, which are probably a result of the compromises that are inherent cost model that is primarily designed for uniform nationwide application.

A comparison of estimated and actual line counts for individual exchanges can be an indicator of the accuracy of the geographic modeling process, if this comparison is accomplished prior to normalizing or conforming the model output to match the data set being compared. Where substantial discrepancies exist in a raw line- count comparison (e.g., a mismatch of more than 50% in the total number of lines within a specific wire center), this would indicate room for further improvement in the modeling of that wire center. For example, the HAI model estimates 97 lines in the Waterville Valley wire center, which is far less than the *****Begin proprietary*** ***End proprietary***** actual lines in this location. As shown on schedule 3 of my exhibit, the HAI model violates the 50% criterion in 23 of the 117 New Hampshire wire centers. In an additional 28 wire centers, the HAI line counts differ from the actual line data by more than 25%, but less than 50%.

This discrepancy in line count can be eliminated at the wire center level by "conforming" the data to match the actual line counts. However, when the wire center data is "conformed" or "reconciled" in this manner, it tends to mask or exacerbate the impact of errors within each wire center. For example, recall that HAI estimates 97 lines for the Waterville Valley wire center instead of the *****Begin proprietary***** *****End proprietary***** lines actually in service. Such a large error is quite likely the result of a complete failure to identify customers located within one or two specific geographic areas served by that wire center, not an across-the-board error affecting the entire wire center.

Thus, if the 97 line figure is "corrected" by simply scaling up the HAI results by *****Begin proprietary***** % *****End proprietary***** this procedure will achieve the overall wire center target, but the HAI cost estimates will be further distorted. An across-the-board reconciliation of this type will create nonexistent lines in neighborhoods where the HAI line count estimates are already valid, and it won't necessarily insert enough lines into the specific geographic areas where HAI has grossly underestimated the number of customers (increasing 0 lines by *****Begin proprietary***** % *****End proprietary***** still results in 0 lines in that particular area).

Q. You mentioned geocoded address data. Does HAI use precise geographic location data throughout the cost modeling process?

A. No. To the extent it is available, HAI 5.0a uses geocoded address data at the earliest stages of the cost modeling process, in assigning customers to "clusters." There are two types of clusters, "main" and "outlier." An outlier cluster has 1-5 lines; a main cluster has 6 lines or more. According to Mr. Siwek,

> In order to be considered members of a cluster, customer locations must meet the following criteria:
>
> - no point in a cluster may be more than 18,000 feet distant from the cluster's centroid (based on right angle routing);
> - no cluster may exceed 1,800 lines in size;
> - no point in the cluster may be farther than two miles from its nearest neighbor in

the cluster. [p. 30.]

While the gathering of customers into "clusters" is a key step in the HAI 5.0a cost modeling process, the HAI model does not use the shape of each cluster, nor does it consider the actual location of customers within each cluster. Instead, the HAI Distribution Module substitutes a simplified rectangular grid of equivalent size for each of the main clusters, and an algorithm assigns end-users to hypothetical rectangular lots within the overall rectangle. "Lots are arranged within clusters to reflect the [height-to-width] ratio determined for each cluster [by the data development process]." [Id., p. 23.]

Thus, even where the HAI model starts with precise information concerning customer locations within each wire center, it fails to fully utilize this information. It simplifies away or ignores important aspects of the geographic data that could potentially be used in the cost modeling process. For instance, the HAI model makes no use of road data, despite the fact that distribution cable generally is located along road rights of way, and it assumes customers are neatly arranged on rectangular lots, regardless of whether or not this is actually the case. The effect of these simplifications is to reduce the reliability of the HAI model in estimating distribution costs. In some cases the HAI model may overestimate and in some cases it may underestimate the amount of distribution cable required to reach customers.

Q. How do the HAI results compare with BA-NH's results?

A. In a side by side comparison, HAI's total statewide costs for individual UNEs are generally lower than the corresponding BA-NH costs as shown on the Company's revised SGAT pages. With regard to the cost of a basic 2 wire analog link, the HAI model estimates a statewide average cost of $16.71 per month. This compares to the BA-NH estimate of $17.99 per month. While these overall estimates are relatively similar, the components of these estimates are quite different. For instance, the BA-NH estimate includes $8.89 of feeder costs, including fiber electronics, while the analogous HAI estimate is $6.04. The BA-NH estimate includes $4.34 of customer termination and distribution costs, while the analogous HAI estimate is $10.19.