Detailed Analysis of Selected Wire Centers

Q. Can you please explain some of the other steps you took in order to evaluate the most appropriate geographic inputs to use with the FCC model for KUSF purposes?

A. Yes. As a result of our work in other jurisdictions, our firm has long known that customer locations, rights of way, and other geographic factors strongly influence the amounts of cable deployed, and costs generated, by forward looking cost models. In dealing with these issues we have found that it is helpful to prepare a very detailed analysis of a few wire centers, rather than relying exclusively upon less detailed analyses of all wire centers. By gathering detailed information about the actual rights of way and other geographic characteristics of selected Kansas wire centers, one can learn a great deal about the results being produced by a forward looking model. Judgments could then be made about how well the model is matching actual Kansas conditions. This knowledge can then be used to refine the inputs into the model, to ensure that it develops costs that more closely reflect actual conditions in the serving area in question.

Q. Could you discuss the types of information you employed in this detailed examination of selected wire centers?

A. Geographic information was obtained from a variety of different sources. For this docket, in addition to the embedded sheath data described above, we looked at geocoded white page listings from Select Phone, analogous geocoded E911 data obtained from SWBT and Sprint, Census TIGER road segment files, population data, census block boundaries, and other data obtained from the Census Bureau, exchange boundaries obtained from the Commission, and wire center boundaries obtained from

GDT–the same data relied upon by PNR in developing its customer location data sets, as well as the road surrogate data provided by PNR. In addition, we obtained detailed maps of the actual customer locations and embedded network configuration provided by Sprint for its Green, Kansas wire center. We analyzed these various data sets in an effort to determine how well the FCC model performs in modeling conditions in Kansas, and to determine whether any of the geographic inputs to the model should be modeled in order to better reflect cost conditions in the state.

Q. Could you further explain the Select Phone data you mentioned earlier?

A. Select Phone provides electronic copies of all white page listings in the country. Their data base contains not only names and phone numbers but also addresses and in many cases the longitude and latitude of the customers' locations. The listed numbers are also distinguishable by residence or business telephone numbers. This data is contemporaneous with, or slightly more recent than that produced by PNR. While the Select Phone geocoded data is not identical to that used by PNR, it offers the important advantage of being readily available for review and analysis. The FCC rejected use of PNR's geocoded data in part because it wasn't available for review and analysis. In sharp contrast, Select Phone provides us with access to all the underlying data, including the longitude and latitude of individual customer locations throughout the state.

Q. What E911 data sets have you reviewed?

A. E911 data was provided by Sprint for all of its wire centers, and by SWBT for Valley Center and Colwich. These data sets are similar to Select Phone in that they accurately pinpoint actual customer locations. However, the E911 data is presumably more comprehensive since it includes unlisted numbers.

Q. What Census data have you used in your analysis?

A. Census TIGER road segment files are continually being updated and made available to the public. This data is used in developing all of the geocoded data sets just mentioned. Customer locations are geocoded by comparing customer addresses with corresponding address ranges contained in the TIGER files. As they are updated, the TIGER files include more and more roads with detailed address ranges which can be used for address matching, and thus the various geocoded data bases tend to improve over time. In 1992 there were 168,955 miles of road segments in the Kansas TIGER files, of which just 12,071 miles or 7.1% had corresponding address ranges which can be used for geocoding. The situation has improved since then, and the 1998 TIGER files contain 173,658 miles of road segments of which 31,996 miles have address ranges, or 18.4%. Thus, over this six year span the percentage of roads with address ranges has nearly tripled, but it remains fairly low. In both years most of the road segments with address ranges are located within towns and urban areas, and the vast majority of the road segments in rural Kansas do not have address ranges assigned in the TIGER files.

Q. Can you please elaborate on the obstacles to using actual geocoded customer locations in rural Kansas?

A. Geocoding is typically accomplished by having a computer look at a customer's street address and determining the corresponding mapping coordinates, latitude and longitude, for the specified location. A typical address consists of several items that can be used for identifying the location or area including street number and street name, city, state and ZIP code or a ZIP+4 code. The address information is matched to existing digital mapping information (derived from the TIGER files) to determine the mapping coordinates. An address can be matched to map information using address number and street names, ZIP+4 codes or ZIP codes. The street number and name approach will produce the most precise mapping coordinates. The precision decreases with the use of ZIP+4 and ZIP codes. In general one can expect to achieve between 60% and 70% success rate when geocoding to street addresses.

Several conditions are necessary for precise geocoding to work. First, a legitimate address which includes a street number and street name must be used. Post office box and rural route delivery addresses can not be used to determine customer locations. (The geocoding software will locate the post office itself, but that isn't especially helpful.) Second, mapping information must be available to match with the address. The mapping information generally consists of street maps with the street name and address ranges, ZIP+4 centroid points, or 5 number ZIP code centroid points. As mentioned a moment ago, the primary source of street name and address range data is the U.S. Census bureau TIGER files which are used to map census data. TIGER files include street name and address range information in those areas where the Postal Service uses street name and numbers for addressing. In those areas without Postal street addresses street name and address numbers are not included in the data.

Maps 11, 12 and 13 visually demonstrate the problem. Map 11 shows the road system in Rawlins County, Kansas. It includes two wire centers, McDonald and Atwood. The roads without address range information are shown in green while the roads with address ranges are shown in blue. The only area with address information usable for geocoding are in the town of Atwood, as indicated by the small concentration of blue road segments.

Maps 12 and 13 depict the analogous information for the state as a whole. The larger number of blue roads in Map 13 relative to Map 12 reflects with improvement which occurred from 1992 to 1998. Given the lack of address ranges in the current TIGER files, it is not currently possible to develop accurate geocoded locations in rural areas like Rawlins County using existing data bases. The only way geocoded customer locations can be developed is to physically go to Rawlins county and gather the needed data in the field (using a GPS data collection unit). While this could be done almost immediately, it would be a time consuming and costly process to gather this data. The only other alternative is to wait until customer lists, TIGER files and other data bases have been updated to include meaningful street addresses that can be matched to address ranges.

Q. Would you please explain what information you received from Sprint for its Green wire center?

A. Certainly. Sprint provided us with uniquely detailed information concerning its Green wire center. Sprint personnel analyzed their network blueprints and manually located every customer and every cable route on its actual network, then placed this information into mapping software that can be analyzed in a computer and compared with the results of the FCC modeling process. This offers a unique opportunity to compare the FCC's model results with an actual network, including the specific routes used in connecting cable from actual customer locations back to the wire center.

Q. Would you please elaborate upon some of the other analytic techniques you relied upon in analyzing the appropriate geographic inputs to use with the FCC model for KUSF purposes?

A. As alluded to earlier, we analyzed the distances cable would need to follow along actual rights of way, in order to connect customers to the SAI and to connect the SAIs to the wire center. For this purpose, we used the SAI locations developed by the FCC model, which were "snapped", i.e., moved to, the nearest road intersection within a specified distance, preferring four way intersections. Customers were connected to the SAI in order, based upon their distance from the nearest SAI, beginning with customers closest to each SAI. As additional customers were mapped to the SAI, the computer attempted to minimize the cost of the distribution system by looking for the least costly path, considering the total distance covered by alternative routes, giving preference to those rights of way that were already being used to serve other customers.

In this way we simulated in the computer the type of least-cost routing decisions that are made by network engineers when designing a distribution system. The least cost path is the one that covers the shortest distance, except where a longer routing may be less costly because it allows sharing of cable and structure costs with other customers. The service distribution layout grows in a branching pattern as customers are added to the system further and further from the SAI. Customers that are far away from the SAI may or may not be connected to the nearest SAI.

The feeder system from the SAIs to the wire center was developed along rights of way using a similar procedure. The closest SAI was connected to the wire center, and then other SAIs were connected to the wire center, in the order of their distance from the wire center. Each SAI was connected using the shortest potential route along the available rights of way (roads), except where cost savings might be achieved by slightly more circuitous routing along feeder routes that have already been established to serve another SAI.

Q. This modeling of a forward looking network along actual rights of way sounds very data intensive. How does it compare to the more simplified approach used by the FCC model?

A. Both approaches start with the same SAI locations (except that the right of way analysis "snaps" the SAIs to a nearby road intersection). The FCC model generates the SAI locations by clustering customer locations, so the SAIs tend to be located in the midst of natural clusters of customers, to the extent such clusters are reflected in the customer location data used in running the FCC model. The FCC model then takes a rectilinear distance approach to connecting customers to SAIs. A rectilinear approach sends feeder cable north-south and east-west. At least in Kansas, this works reasonably well, since most roads are laid out on right angles following the compass orientation. Potential north-south and east-west routes are identified throughout the wire center and then customers are connected to all potential routes in their area. The system is then pruned so that only the best choices remain, and each customer is ultimately connected to only one cable leading to one SAI.

The feeder system from the SAIs to the wire center is also developed in the FCC model using a rectilinear distance approach. Two primary trunks extend from the wire center, north-south and east-west, and the SAIs are connected to these trunks by branches that extend from junction nodes. The branches are optimized to allow multiple SAIs to share a branch in many cases. When this happens, the branches are not entirely rectilinear, so that the resulting feeder network has less total distance than a strict rectilinear approach would dictate. The resulting SAI locations, distribution line layouts, and feeder system layouts reflect simplified geometric assumptions and they do not reflect geographic features including road layout and barriers like rivers, lakes, mountains, and military reserves within each particular wire center.

In contrast, our more detailed modeling of cable routes along rights of way does take these types of unique features into account. Our detailed modeling of the selected wire centers is useful for several purposes, including an examination of how well the FCC model comports with the unique geographic conditions within individual wire centers. The major limitation is that it is time consuming and data intensive; hence, we only applied this technique to a small group of wire centers. Due to the time consuming nature of this work, we were not able to update our analysis to reflect the baseline version of the model, which was not finalized until a few days before the deadline for filing this testimony. The results I will be discussing are based upon the PNR road surrogate data and the FCC default inputs prior to the most recent revisions.

By comparing the detailed right of way results with the FCC default results, a much clearer picture can be obtained concerning how much cable feeder would actually be required to connect the SAIs in the FCC model to the wire center, and this can be useful in reaching a conclusion concerning the appropriate value for the feeder routing input. The same type of comparison can be made for distribution cable. One can see how much distribution cable would actually be needed to connect a given set of customer locations (e.g. those shown in the PNR road surrogate data set) to the SAIs, and compare that with the amount of distribution cable generated by the FCC model.

Q. Would you please describe your right of way analysis?

A. Yes. We analyzed 14 Kansas wire centers; 10 from SWBT's service area and 4 from Sprint's service area. The 10 SWBT wire centers are Atwood, De Soto, Ellsworth, Harper, McDonald, Oakley, South Topeka Zone, Lecompton Zone, Colwich, and Valley Center. The 4 Sprint wire centers are Downs, Centropolis, Green and Onaga. These wire centers were selected to encompass a variety of different wire center characteristics. We selected wire centers of widely differing size and geographic location within the state. Two of the selected wire centers are in the Topeka area (Lecompton and South Topeka) and two are in the Wichita area (Colwich and Valley Center), but we did not analyze any truly urban wire centers. We focused exclusively on high cost, rural wire centers, since these are the types of wire centers which have the greatest potential impact on the KUSF.

Schedule 11 compares cable quantities from three separate sources: embedded cable quantities provided by SWBT and Sprint; cable quantities generated by the FCC model and cable quantities using actual rights of way. The latter two data sets are directly comparable; they rely upon the same SAI locations (latitude and longitude) generated by the FCC model and the same road surrogate customer locations. Page 1 of this schedule compares the embedded company sheath feet to the FCC default route feet. Page 2 compares the FCC default route feet to the analogous route feet following actual the rights of way (roads).

As I explained earlier in my testimony, when the FCC model is run using road surrogate data and the default routing inputs, it tends to produce more route feet than the actual sheath feet–despite the fact that the latter data reflects multiple sheaths along individual routes and it may also include some interoffice cable. This general pattern can be seen by comparing the column labeled "Total Embedded Sheath Feet" to the column labeled "Total FCC Route Feet". In every wire center except Downs the FCC model deploys significantly more route feet than actual sheath feet.

A similar, though lesser, difference is revealed when comparing the FCC default model route feet to the analogous route feet following actual rights of way. This can be seen by comparing the columns labeled "Total FCC Route Feet" to the columns labeled

"Total ROW Route Feet." In total, the FCC model generates more route feet than required by the actual rights of way in 12 of the 14 wire centers.

Q. Would you please elaborate on the right of way results for feeder?

A. Yes. The feeder quantities generated by the FCC model are generally quite similar to those we developed following actual rights of way. In 8 wire centers the right of way results were somewhat higher, and in 6 wire centers they were somewhat lower, suggesting no overall bias to the FCC default results. Furthermore, in all but 2 of the wire centers the two feeder quantities were within 15% of each other.

These results are not particularly surprising. The FCC model's "right angle" feeder routing algorithms would be expected to work particularly well in a state like Kansas, where the road system tends to follow a rectilinear grid with north-south-east-west orientation. The default feeder inputs may not be as successful in other geographic areas, where roads meander in a more curvilinear fashion, and where mountains, lakes and other obstacles require more circuitous routing from SAIs to the wire center. I would not be surprised to find the FCC model consistently under-deploying feeder in some states, and thus a feeder routing variable greater than 1 might be more appropriate in those states.

Q. Would you please elaborate on the right of way results for distribution?

A. Yes. Overall, the FCC model generated 14,248,489 distribution route feet, whereas the analogous right of way calculations indicated just 11,911,335 feet was actually needed. Stated differently, the road surrogate customer location points can be networked together with 16% less distribution cable than the FCC model estimates.

Although the specific percentage differences varied, this general pattern held true for 13 of the 14 wire centers we studied in detail.

It is important to keep in mind that this comparison isolates a problem with the FCC distribution cable quantities which is separate from the problems previously discussed relating to the road surrogate data set. The observed differences between the amount of cable deployed by the FCC model and the amount needed using actual rights of way are entirely attributable to flaws in the FCC model's default routing input when applied to the actual geographic conditions in these wire centers. Since both sides of this particular comparison -- the FCC default results and the right of way results -- are generated using the same customer locations from the PNR road surrogate data set, these calculations don't deal with the amount of excess cable resulting from use of the road surrogate points.

Q. Have you prepared any additional analyses which are helpful in evaluating the appropriate distribution routing input for Kansas?

A. Yes. We performed some additional analyses for the Green, Valley Center, Atwood and Oakley wire centers.

As I mentioned earlier, Sprint provided us with the latitude and longitude of each customer in the Green wire center. We calculated the amount of distribution cable needed to connect these locations to the SAIs generated by the FCC model along actual rights of way, and compared these results to the analogous quantities using the road surrogate locations. The actual customer locations are shown as blue dots on Map 12, along with the cable routes used to connect these locations to the SAIs. The analogous PNR road surrogate locations are shown as green dots on Map 13, along with the corresponding cable routes. A comparison of these two maps shows that there are small, but significant differences in the dispersion of the blue and green dots. The green dots (generated by the PNR road surrogate algorithm) are more widely distributed than the blue dots (the actual customer locations). For instance, PNR placed 12.7% of the customers within 500 meters of the wire center boundaries; in contrast, just 5.1% of the actual Sprint customer locations are this close to the wire center boundaries.

A visual comparison of the two maps confirms that the problems with the road surrogate technique are not extreme in this wire center. Nevertheless, these subtle differences translate into a significant difference in the amount of cable required to connect customers. More specifically, connecting the actual Sprint customer locations to the SAIs requires 718,389 route feet of distribution cable. In comparison, 809,610 route feet of distribution cable are needed to connect the road surrogate locations. The difference between these two figures--nearly 100,000 feet–results from the excessive dispersion of the PNR data, as discussed earlier.

However, as discussed earlier, the road surrogate data is not the only reason the FCC model deploys too much cable. The actual locations can be connected with far less cable than is generated by FCC model using the default routing inputs. Our analysis suggests that the distribution routing variable should be reduced from the default value of 1 to a value more like .8–at least in Green.

We performed a similar analysis of the Valley Center wire center using the actual customer locations provided by Select Phone, which were derived from white page address listings. In Valley Center when we connected the road surrogate locations to the SAIs along rights of way it took 588,648 route feet of distribution cable. In comparison, using the Select Phone customer locations just 405,938 route feet of distribution cable was required. At the time we performed this analysis, the corresponding quantity estimated by the default version of the FCC model was 1,013,383 feet. These comparisons confirm that regardless of whether road surrogate or actual location data is used, it would be appropriate to reduce the distribution routing variable well below 1. Since the Select Phone locations can be connected with less than 70% of the distribution route feet generated by the FCC default input, this analysis suggests that, at least in Valley Center, the distribution routing variable could appropriately be reduced to a value more like .7.

Finally, we also looked at an alternative to the road surrogate approach which placed the customers in each census block at a road adjacent to the centroid of the census block. This provides an alternative view of the census data, in which customers are assumed to be clustered to the maximum degree possible–the logical opposite of the approach used by PNR, which disperses customers as widely as possible. When we applied this approach in Green, the result was 598,776 distribution route feet. This compares to 809,610 distribution route using PNR's road surrogate data. As expected, these two results bracket the requirements using actual customer locations -- 718,389 route feet.

When we applied this same centroid-based approach in Atwood, the result was 1,691,286 distribution route feet. This is substantially less than the 2,162,457 feet developed using road surrogate location points. Since we don't have the latitude and longitude of the actual customer locations in Atwood, we can't know precisely how much cable would be required to connect these locations. However, it is reasonable to conclude that the quantity using actual customer locations would fall somewhere between the centroid approach and the road surrogate approach. If the cable requirement using actual locations were midway between these two alternatives this would equate to 1,926,872 distribution route feet, which is approximately 15% less than the 2,254,725 feet generated by the FCC model using the default input of 1. Accordingly, this analysis would suggest that, at least in Atwood, the distribution routing input could appropriately be reduced to a value more like .85.

Finally, we applied the centroid-based approach in Oakley, the wire center in which the FCC model default results were very close to the right of way results using the road surrogate data. The centroid based result was 1,010,731 distribution route feet, which is about 10% less than the 1,126,634 feet developed using road surrogate location points. While we don't know exactly where the customers are located in the Oakley wire center, it is reasonable to conclude that if we did have actual customer locations, the corresponding route feet would fall somewhere between the quantities generated using the centroid locations and the road surrogate locations. If the cable requirement using actual locations were exactly midway between these two results, it would equate to 1,068,682 distribution route feet, which is approximately 4% less than the 1,115,506 feet which was generated by the FCC model using the default input of 1. Accordingly, this Oakley analysis still suggests that the distribution routing input should be reduced below the default level, but in this particular case the data would support a somewhat higher value, more like .96.

Q. Why are the Oakely results using the centroid approach so similar to the results using the road surrogate approach?

A. In essence, Oakley is a good example of how the spread of the estimated customer locations has more influence than anything else. Density per se has relatively little impact, because it takes the same amount of trenching to link 10 customers along a particular road as it takes to link the one customer at the distant end of the road. In measuring the required amount of trenching, the customers between the most distant customer on a particular route and the SAI have little or no impact. The total quantity of cable needed to network customers together is driven primarily by customer dispersion, not density. It happens that in Oakely the amount of cable needed to reach all the census block centroids is nearly as great as the amount needed to reach the road surrogate locations. Thus, the additional dispersion introduced by the road surrogate algorithms is minimal in this case.