3GPP TSG-RAN WG1 #87 **R1-1612117**
Reno, USA, Nov 14 – 18, 2016

Source: **Facebook**
Title: **On user and traffic distribution for extreme rural scenario**
Agenda Item: **7.1.6.3**
Document for: **Discussion and Decision**

1. Introduction

A fully connected society is expected in the near future resulting in a tremendous growth in connectivity, traffic volume and diversification of usage scenarios [1]. This generates the need to improve system efficiencies in terms of spectrum, energy, operation, etc. To better serve the needs of such a fully connected and networked society in the near future, in both developing and developed countries, TR 38.913 states typical usage scenarios and attributes associated with them [2]. The family of use cases include eMBB (enhanced Mobile Broadband), mMTC (massive Machine Type Communications) and URLLC (Ultra-Reliable and Low Latency Communications).

The main focus of this contribution is on scenario 6.1.6 titled "Extreme long distance coverage in low density areas". This scenario is characterized by long range (~100 km) macro cells which target low user density regions. Next generation technologies associated with this usage scenario is likely to result in a growth in the number of connected devices and consequently data traffic. To fully characterize this scenario, models related to user and traffic distribution are extremely important.

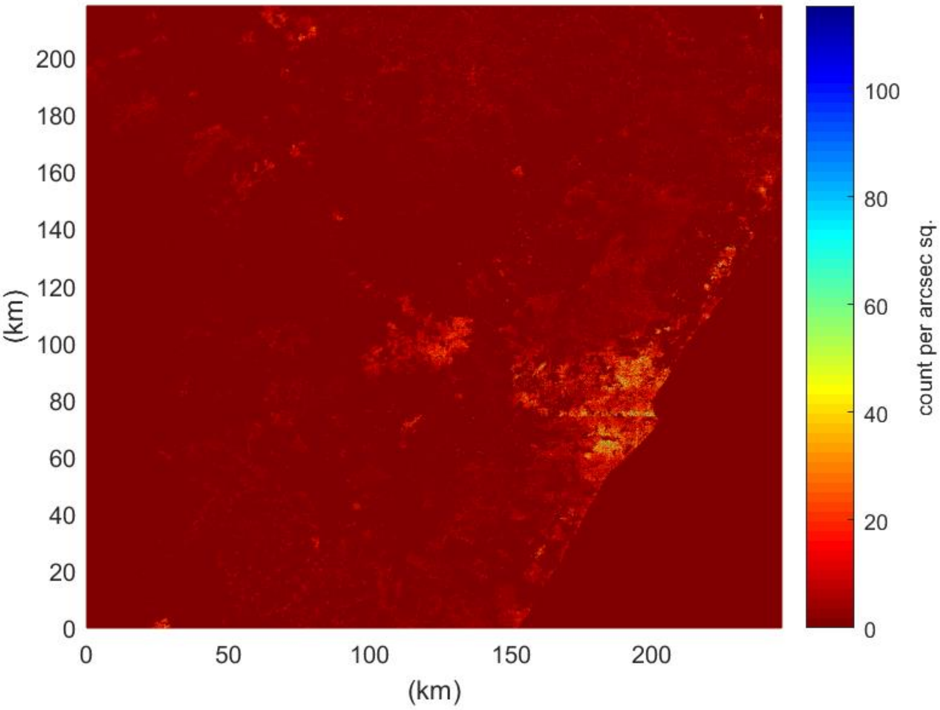
To this end, experimental studies and user distribution models have been proposed recently. For instance, [3] presents a statistical analysis of measurement data from a live 3G network and [4] proposes an empirical user distribution model based on these measurements. Specifically, it proposes that a uniform distribution assumption be replaced with a log normal distance distribution. While the model is important for service providers to gain an intuition into the traffic characteristics of the underlying network, it should be noted that this model is made based on the activity of users *already connected* to an existing network. The extrapolation of usage data of currently connected people to the spatial distribution of both connected and unconnected populations might not be completely accurate in practical scenarios. This highlights the need to evaluate and further tune these models.

In this contribution, we leverage new research results and data [5] – *population density data* (hereafter, *popdens*) – that provide a high-resolution view into global population density distributions. (These results and data were obtained through the application of state-of-the-art computer vision and machine learning technologies on high resolution satellite imagery combined with existing census data. This massive data set covers over 20 countries in the world).

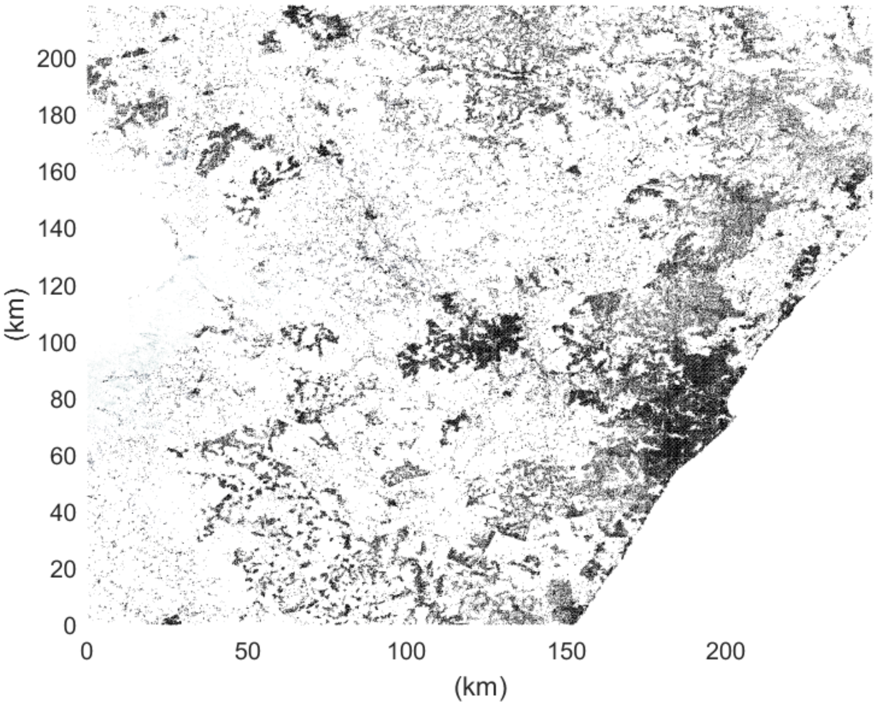
We conduct a rigorous user distribution model case study that focusses on a specific region. We use a subset of *popdens* to evaluate the accuracy of previously proposed models and specifications. We use demographic data that provides population distribution over a couple of regions whose size is roughly greater than the size that will be covered by future macro cells associated with scenario 6.1.6. We analyse this data to understand the traffic distribution associated with this region. Further, we compare the results of our analysis with those obtained from the distribution model proposed in [4].

2. Analysis of demographic and topographic data

With the objective of obtaining high-fidelity user and traffic distributions for the extreme long distance scenario, we conduct a network deployment case study for two regions in South Africa: one around Bloemfontein and the other around Durban. The regions are divided into grid cells of dimension 30 m x 30 m. Demographic data for each of these grid cells is obtained and is shown for Durban and Bloemfontein in fig. 1(a) and fig. 2(a) respectively. As a visual aid, a binary image of the demographic data is shown in fig. 1(b) and fig. 2(b) for these regions. The binary image is obtained by replacing the non-zero population in all the grids with a value of 1. This helps in visualizing the geolocation aspect associated with the demography data. Topography data for these regions is also obtained [6] and shown in fig. 3.

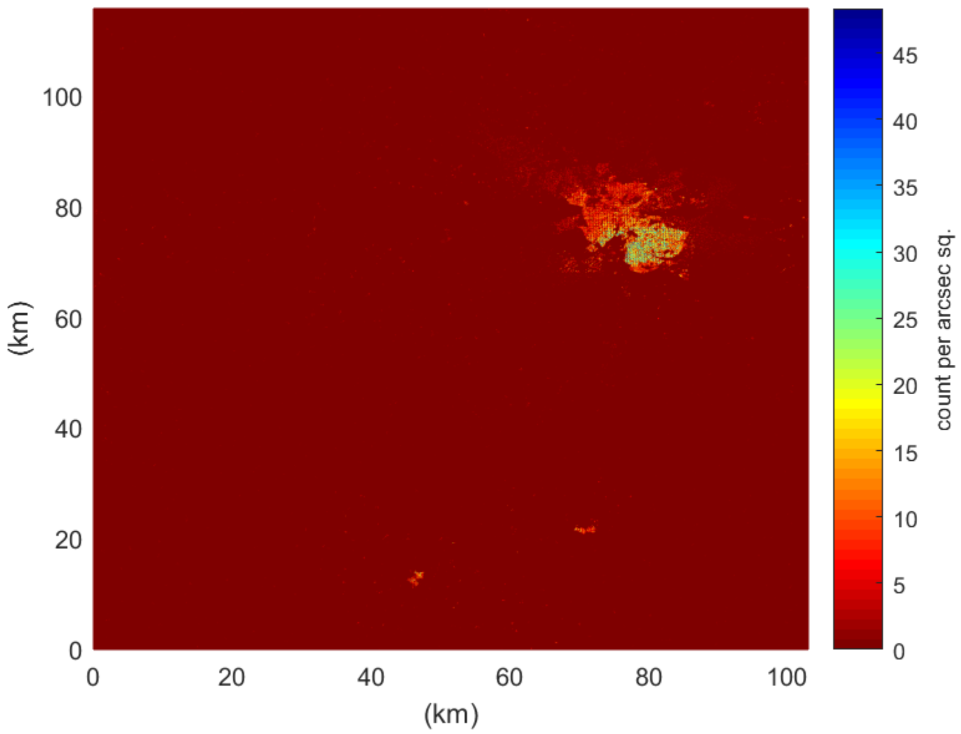


(a) Population density data for a region around Durban, South Africa

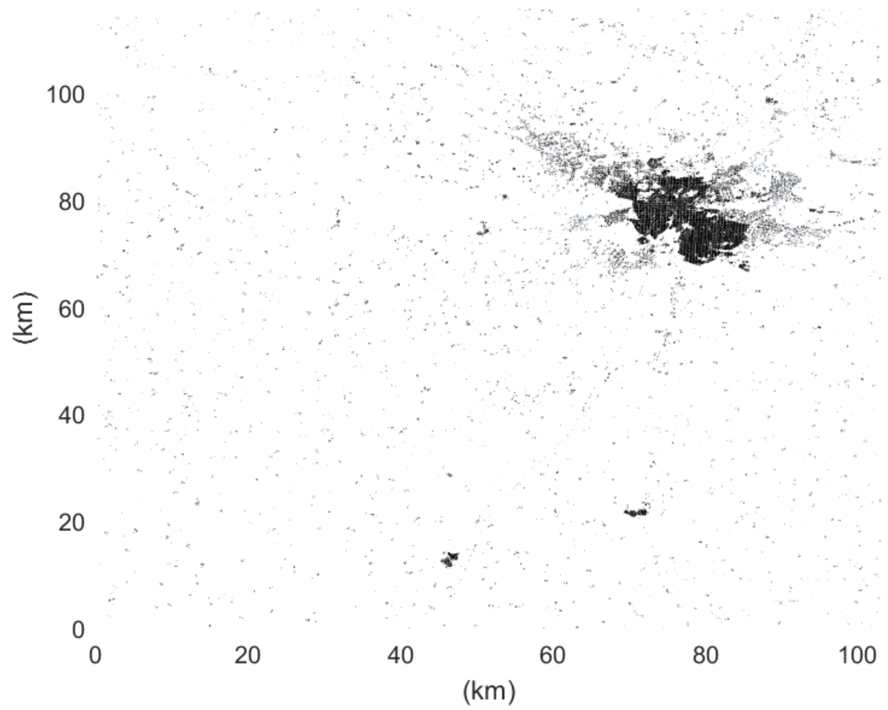


(b) Binary image of the population density for a region around Durban, South Africa

Figure 1 Demography data of a region around Durban, South Africa

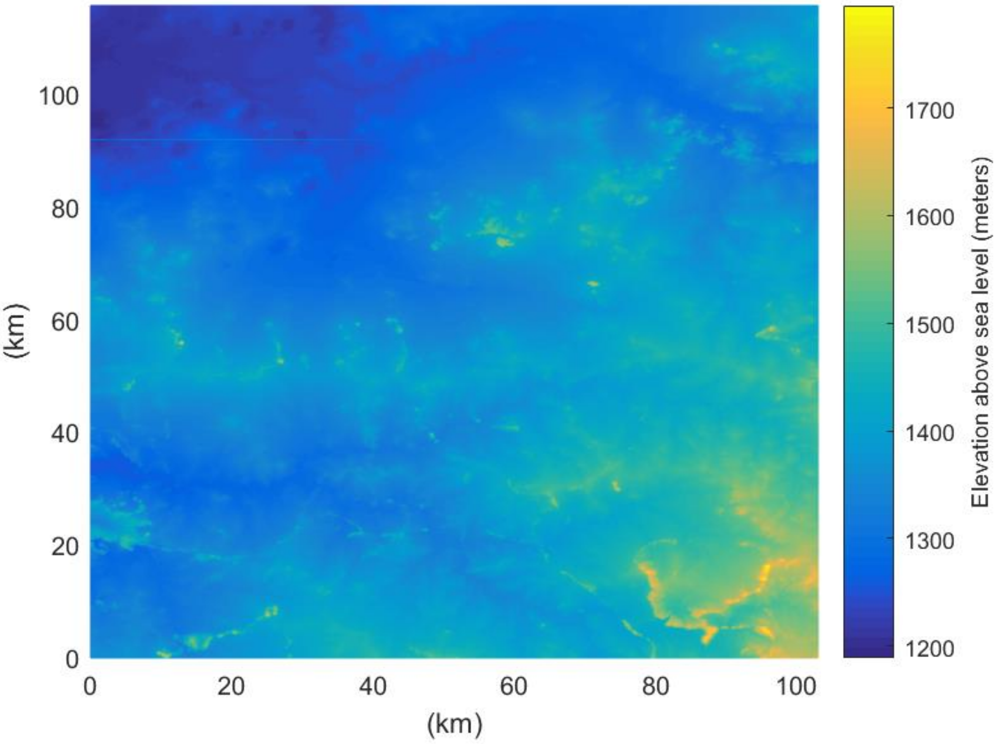


(a) Population density data for a region around Bloemfontein, South Africa

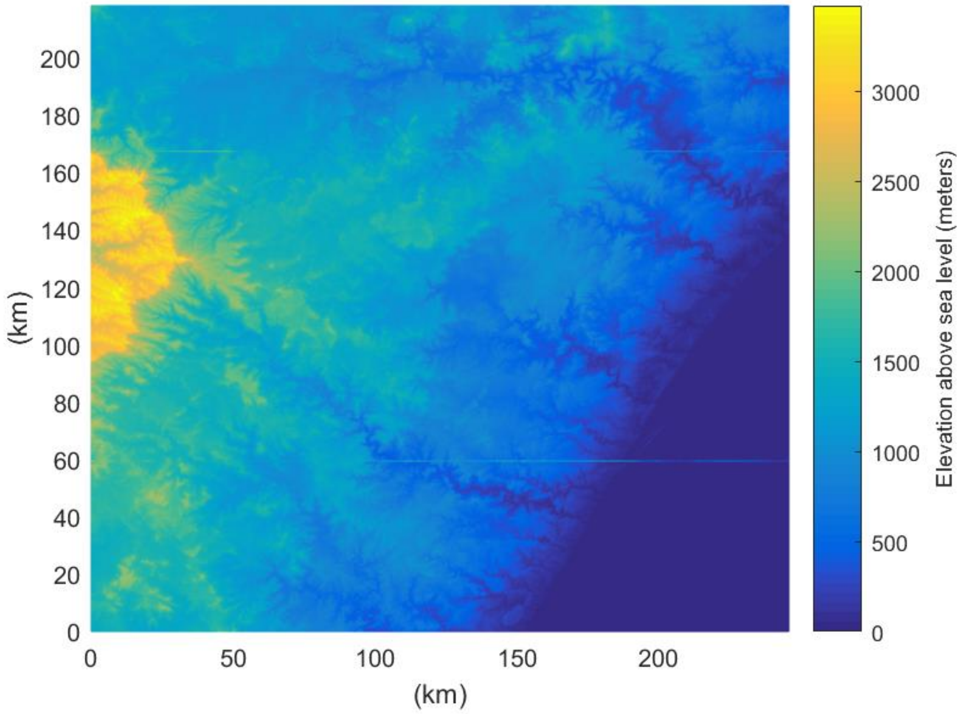


(b) Binary image of the population density for a region around Bloemfontein, South Africa

Figure 2 Demography data for a region around Bloemfontein, South Africa



(a) Topography data of the region around Bloemfontein, South Africa [5]



(b) Topography data of the region around Durban, South Africa [5]

Figure 3 Topography data of the regions under consideration in South Africa [5]

Using the demography data, we measure how population varies as a function of the distance from the nearest city centroid. In figure 4, the results from our empirical analysis are compared with those predicted by the user distribution model proposed in [4]. Figure 4 demonstrates that predictions from the model proposed by [4] deviates from the real distribution of users obtained from the analysis of the population density data by a great extent. Consequently, the model in [4] may not completely represent a typical user distribution for the extreme long distance scenario. This deviation indicates that further analysis is required to model user and traffic distributions accurately. On the other hand, the red curve in fig. 4 further validates the necessity for considering extreme long distance scenario in NR since there is a large percentage of population living far away from the city.

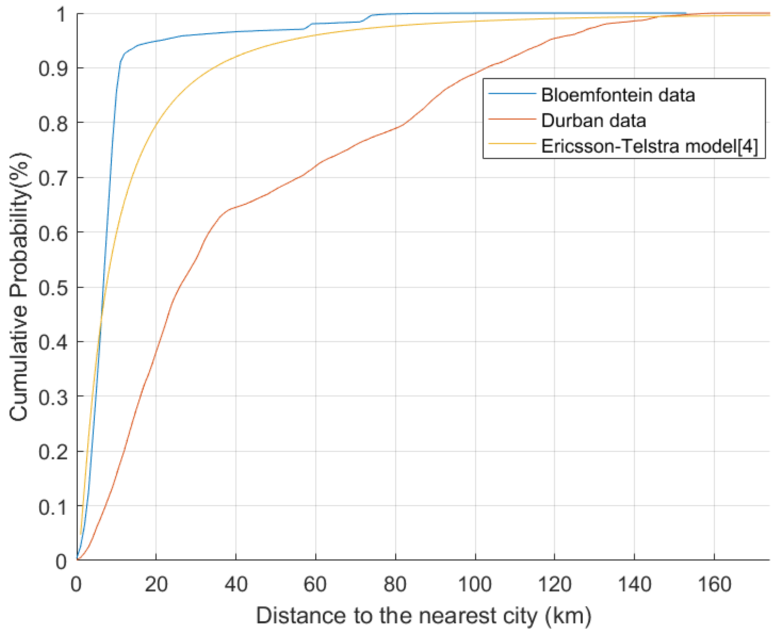


Figure 4 Analysis of population density and comparison with the model proposed in [4]

3. Proposals

Based on our analysis we make the following two proposals:

1. **For extreme rural scenario, high-resolution demographic data (i.e. *popdens*) may be a better tool to accurately model user and traffic distribution. Further work is needed to fully understand a typical user and traffic distribution for this scenario to provide a complete and accurate model for system level simulations.**

2. **User and traffic distribution models may be divided into classes based on the characteristics of the region under consideration.**

4. Conclusion

Topographic and demographic data are important to understand network deployments for the extremely long distance coverage scenarios. Two large regions in South Africa are analysed to serve as a case study for extremely long distance network deployments. The results are compared with those obtained from the user distribution model proposed in [4]. The study finds that the predictions of the model in [4] greatly varies from the user distribution in these two cases considered. Analysis based on high-resolution demographic data may be required to accurately model the user and traffic distribution for the extreme long distance scenario. Further, regions should be divided into multiple classes based on their demographic features and different models for each of these classes should be used for simulation of extreme long distance scenario.

5. References

[1] Recommendation ITU-R M.2083: IMT Vision – “Framework and overall objectives of the future development of IMT for 2020 and beyond” (September 2015).

[2] 3GPP, “Study on Scenarios and Requirements for Next Generation Access Technologies (Release 14),” TR 38.913

[3] R1-166559, “Measurements of extreme rural scenarios”, Telstra, Ericsson, Göteborg, Sweden, 22-26 Aug 2016.

[4] R1-167453, “User and traffic distribution for extreme rural scenario”, Ericsson, Telstra, Göteborg, Sweden, 22-26 Aug 2016.
[5] Andreas Gros, Tobias Tiecke, “Connecting the world with better maps,” Facebook, February 21st, 2016
[6] USGS (2004), Shuttle Radar Topography Mission, 1 Arc Second SRTM v003 scenes [S30E025, S30E026, S29E025, S29E026, S31E029, S31E030, S30E029, S30E030, S30E031, S29E029, S29E030 and S29E031], Unfilled Unfinished 2.0, Global Land Cover Facility, University of Maryland, College Park, Maryland, February 2000.