Digital Divide: A Perspective From Base Station Distributions

Chuanting Zhang

Contents

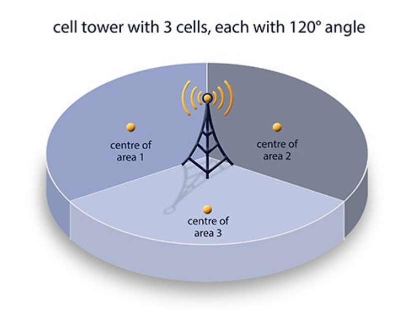
- Data explanation
  - BS data
  - Population data
- Digital divide
  - Country-wise
  - City-wise
- Explanations on imbalance index calculation
- Validation (comparison) on BS coverage of France

OpenCellId Dataset Explanation

	radio	mcc	net	area	cell	unit	lon	lat	range	samples	changeable	created	updated	averageSignal
0	UMTS	208	10	18013	8653445	0	2.272110	48.980484	1000	3	1	1290000321	1290000321	0
1	UMTS	208	10	16505	47046446	0	2.540156	50.485447	1000	151	1	1290033687	1406341488	0
2	GSM	208	20	80	17939	0	-1.229193	49.292793	8732	63	1	1290277096	1290465250	0
3	GSM	208	20	80	18060	0	-1.214537	49.290525	10858	77	1	1290277096	1290464523	0
4	UMTS	208	1	29191	26486747	0	7.172216	43.753281	1342	27	1	1290406958	1290406958	0

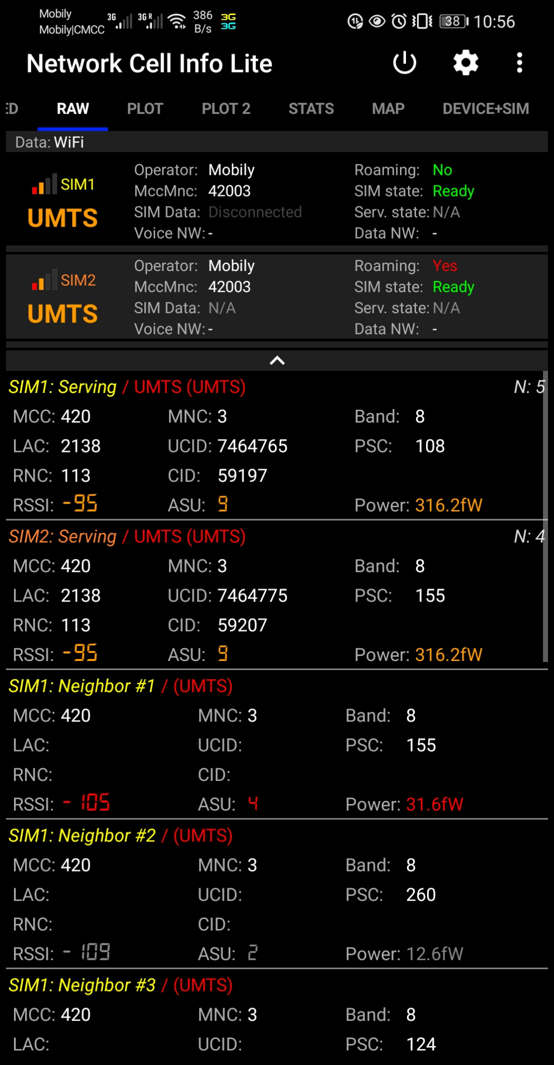
- ✓ **Radio:** radio access technologies: GSM, UMTS, CDMA, LTE, and NR
- ✓ **MCC:** mobile country code
- ✓ **Net:** network operators
- ✓ **Area:** network tracking area
- ✓ **Cell:** cell id
- ✓ **Location:** estimated cell location
- ✓ **Range:** how accurate of the loc
- ✓ **Samples:** # of samples

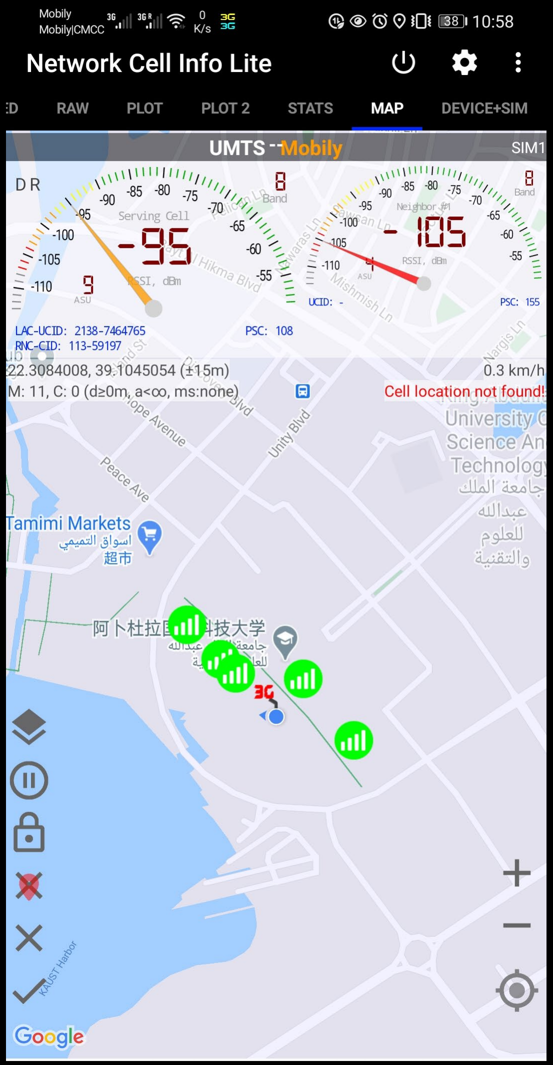
MCC	MNC	ISO	Country	Country Code	Network
289	88	ge	Abkhazia	7	A-Mobile
289	68	ge	Abkhazia	7	A-Mobile
289	67	ge	Abkhazia	7	Aquafon
412	88	af	Afghanistan	93	Afghan Telecom Corp. (AT)
412	80	af	Afghanistan	93	Afghan Telecom Corp. (AT)
412	01	af	Afghanistan	93	Afghan Wireless/AWCC
412	40	af	Afghanistan	93	Areeba/MTN
412	50	af	Afghanistan	93	Etisalat

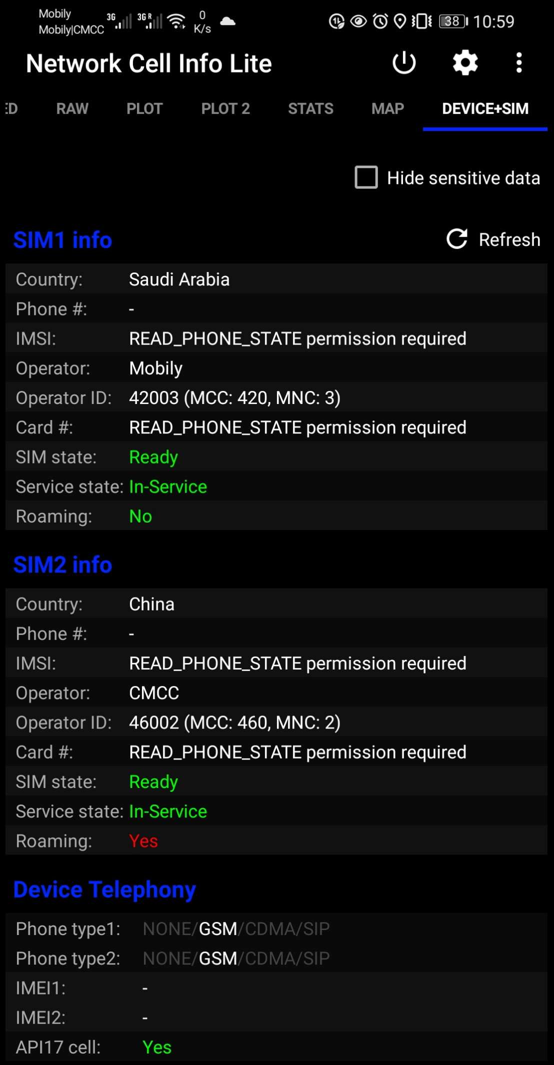


This is a unique number used to identify each Base transceiver station or sector of BTS

How They Get These Data: User Contribution







App: Network Cell Info Lite

http://wilysis.com/images/cellinfo/pdf/Network_Cell_Info_Manual_v2_170308.pdf

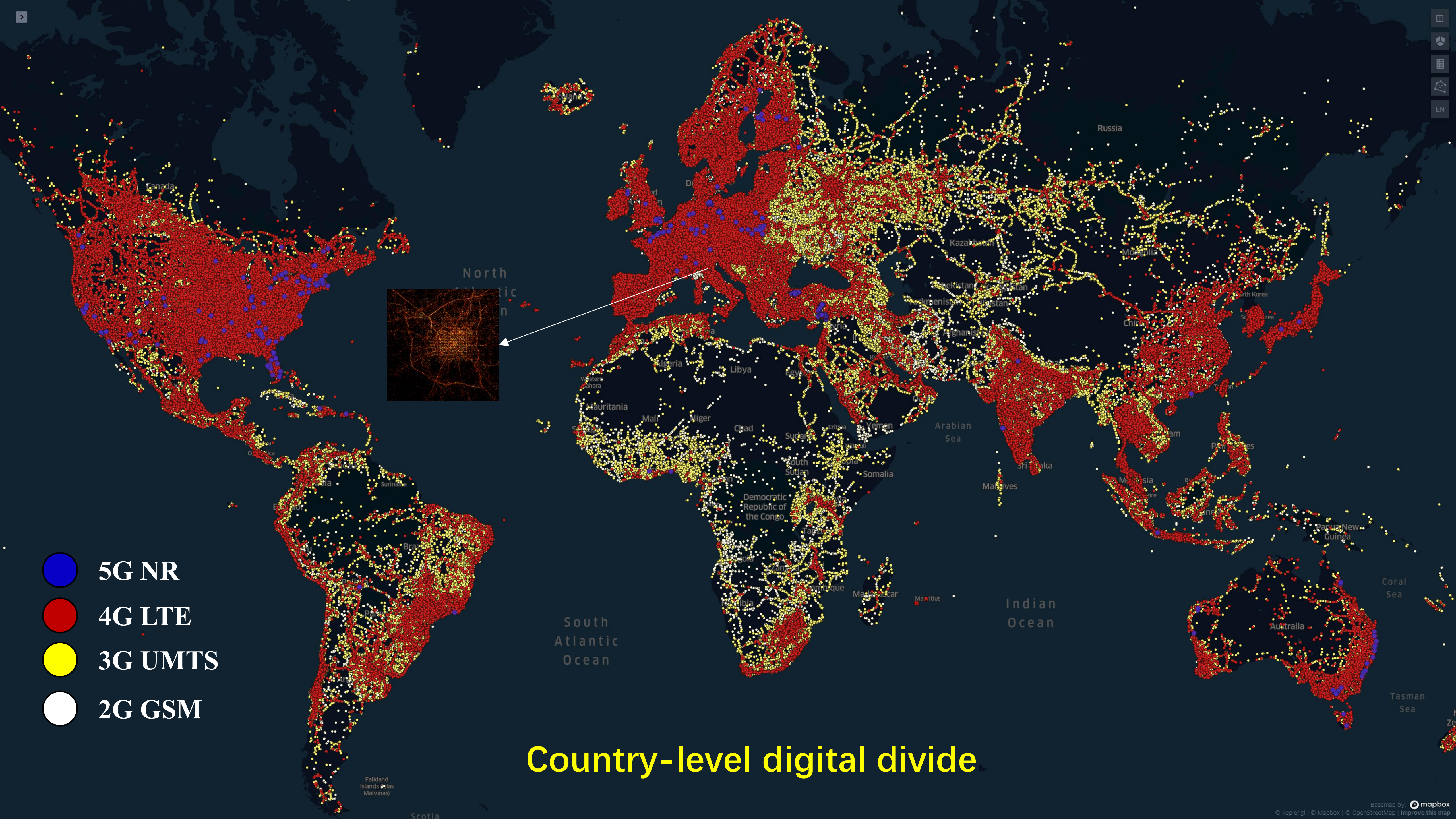
Population Data

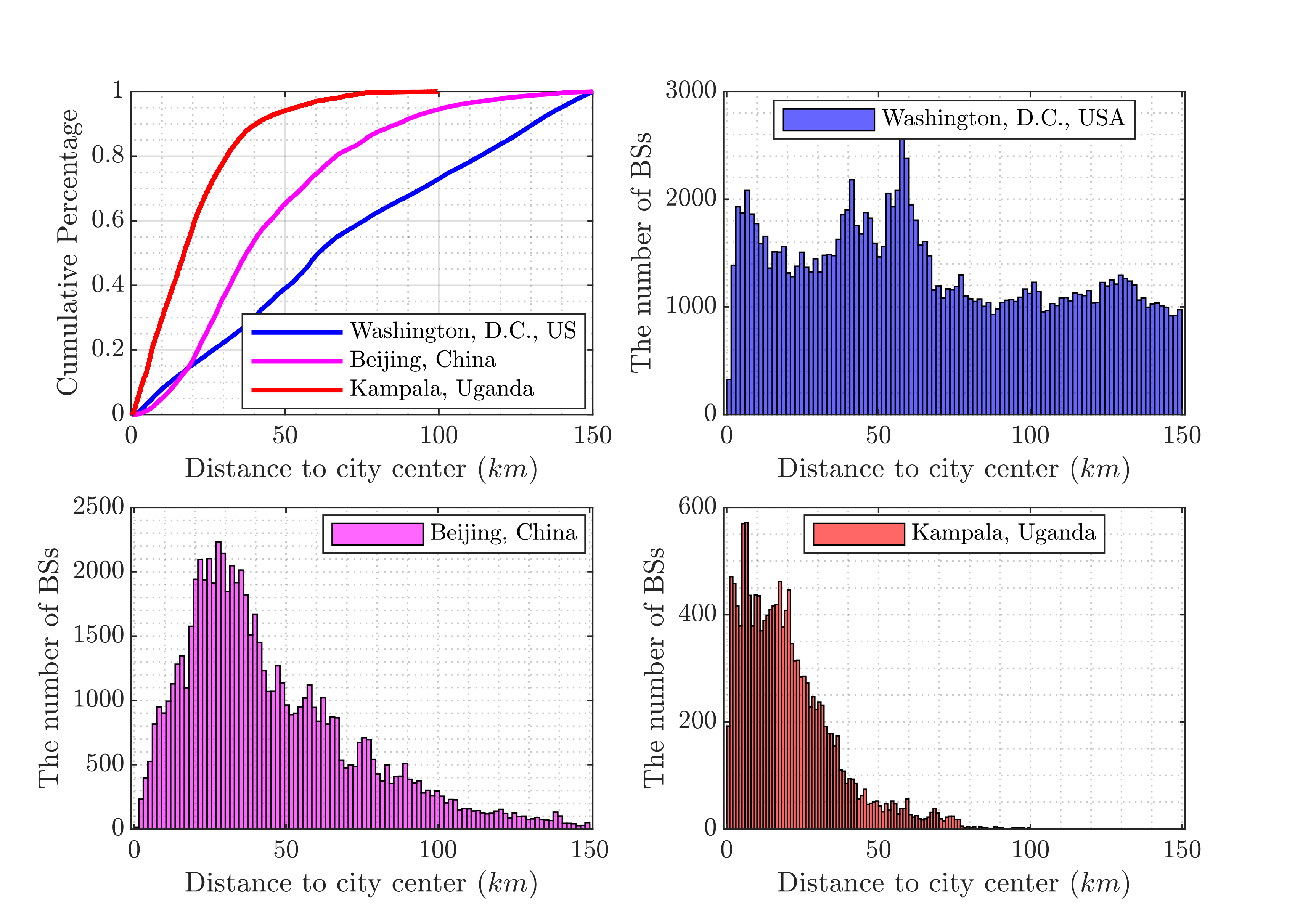
- CSV files are in the format *latitude*, *longitude*, *population*;
- The latitude and longitude are the coordinates of the center of the 1-arc-second-by-1-arc-second grid cell (*30m* * *30m*);
- The value is the (statistical) *number of people in that grid*;
- The values are obtained by using *computer vison techniques* to identify buildings from public map services and *adjusted* to match the census population with UN estimates.

Dataset > HighResolutionPopulation > KSA > pop.csv

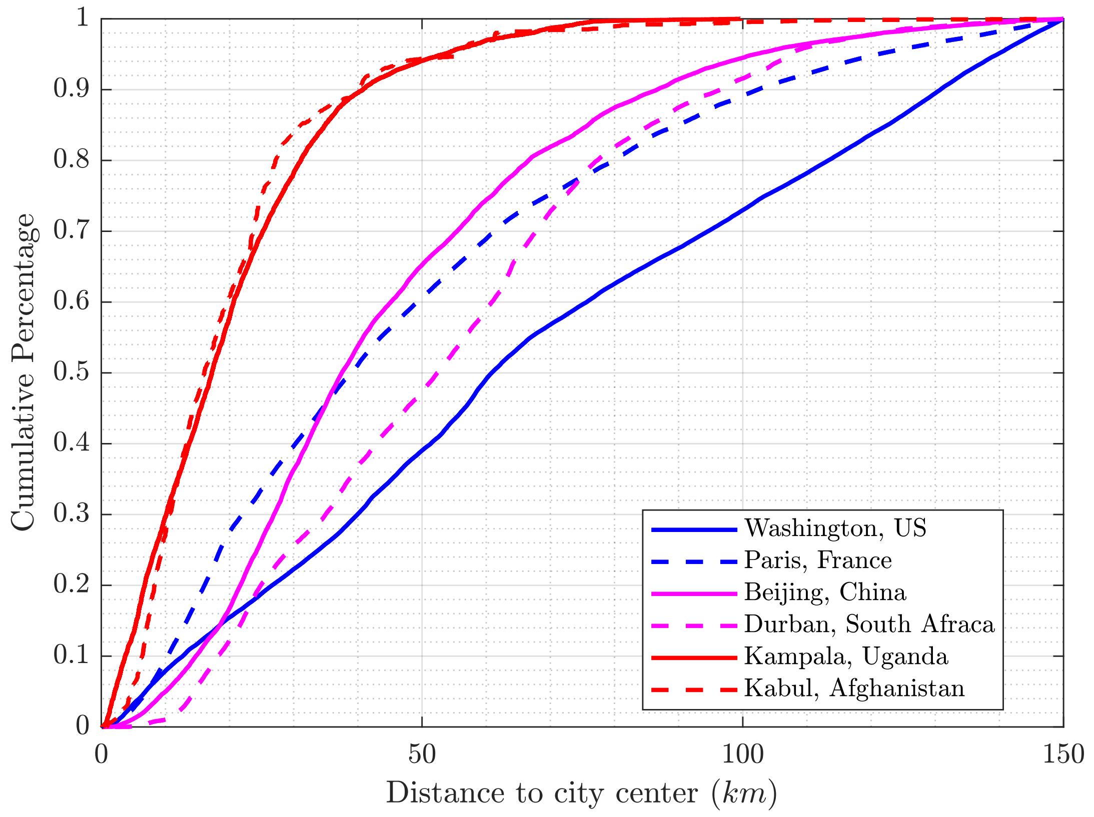
```
"Lat","Lon","Population"
"27.000138888910488","49.64791666670638","6.22612410417777"
"27.010416666688272","49.657361111150834","6.22612410417777"
```

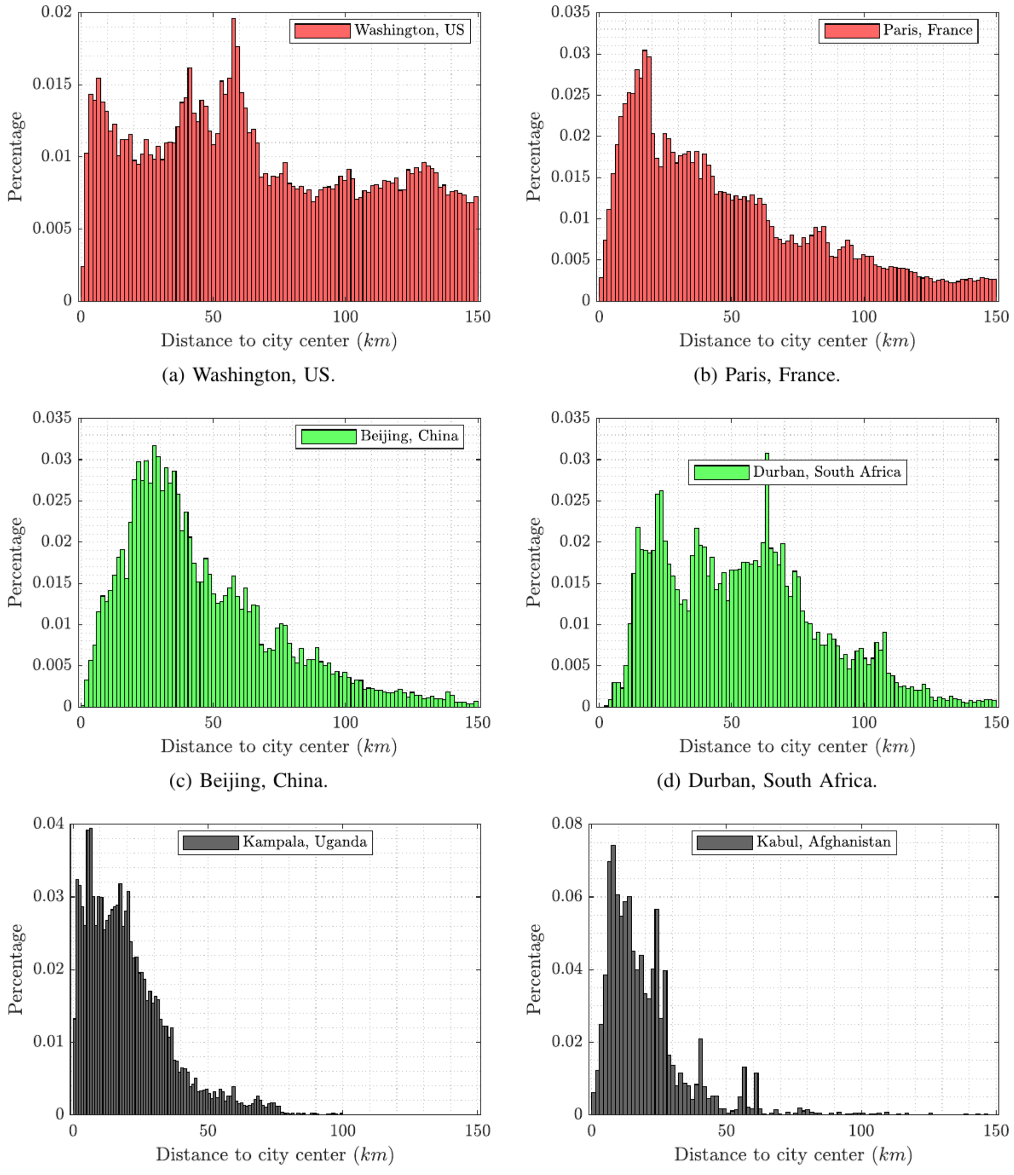
Facebook: https://dataforgood.fb.com/docs/methodology-high-resolution-population-density-maps-demographic-estimates/


5G NR
4G LTE
3G UMTS
2G GSM
Country-level digital divide
Russia
Libya
Chad
Mauritania
Somalia
Democratic Republic of the Congo
Arabian Sea
Maldives
Indian Ocean
South Atlantic Ocean
Australia
Coral Sea
Tasman Sea
Falkland Islands (Islas Malvinas)
Basemap by: mapbox
© kepler.gl | © Mapbox | © OpenStreetMap | Improve this map



City/area-level digital divide

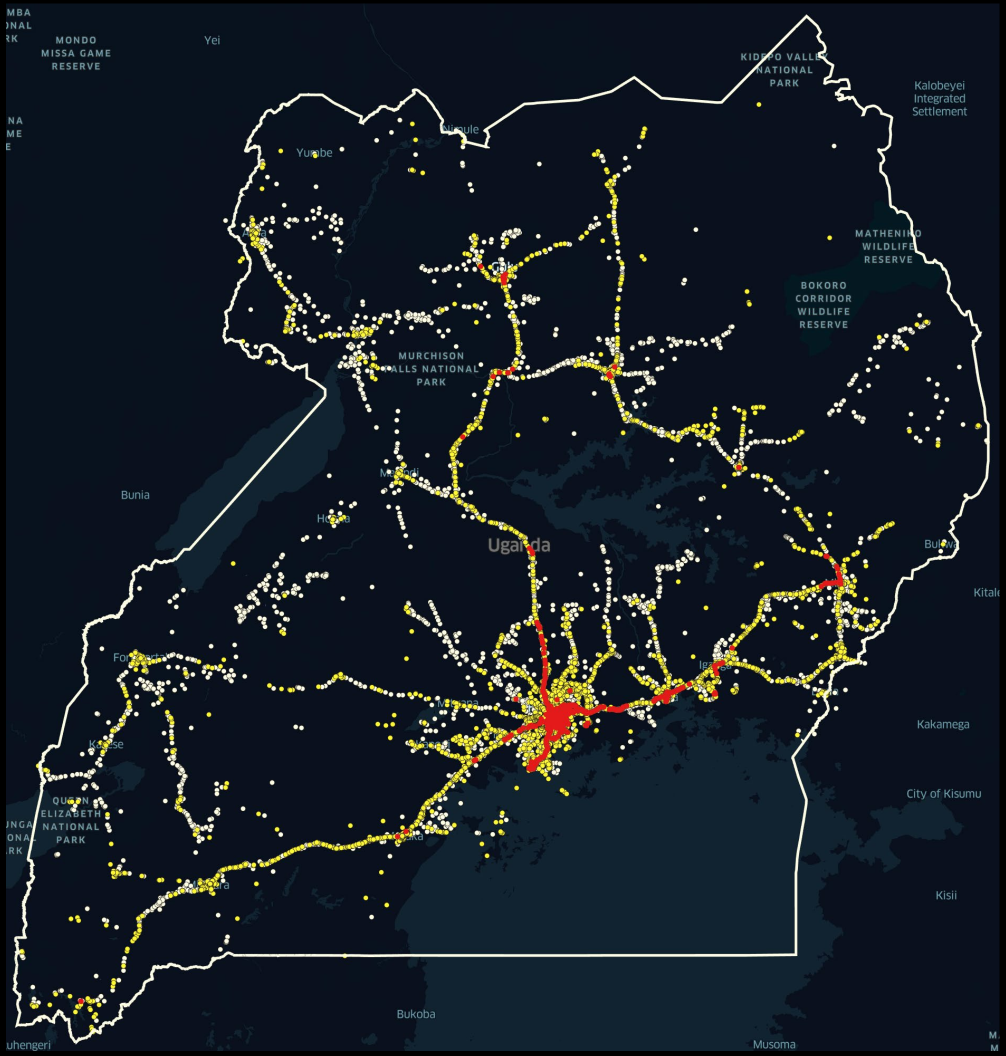




(a) Washington, US.
(b) Paris, France.
(c) Beijing, China.
(d) Durban, South Africa.
(e) Kampala, Uganda.
(f) Kabul, Afghanistan

Should We Fill All the Coverage Holes?

- MNO or governments need using AI and data mining techniques to find out which places needs to be covered accurately
  - ✓ Introduce *population data* into consideration, *define an imbalance indicator* for each place based on BS data and population data, and fill coverage holes with high imbalance
  - ✓ Similar ideas can be applied to city center (find out which places' QoS need to be improved), but more data and complicated algorithms are needed

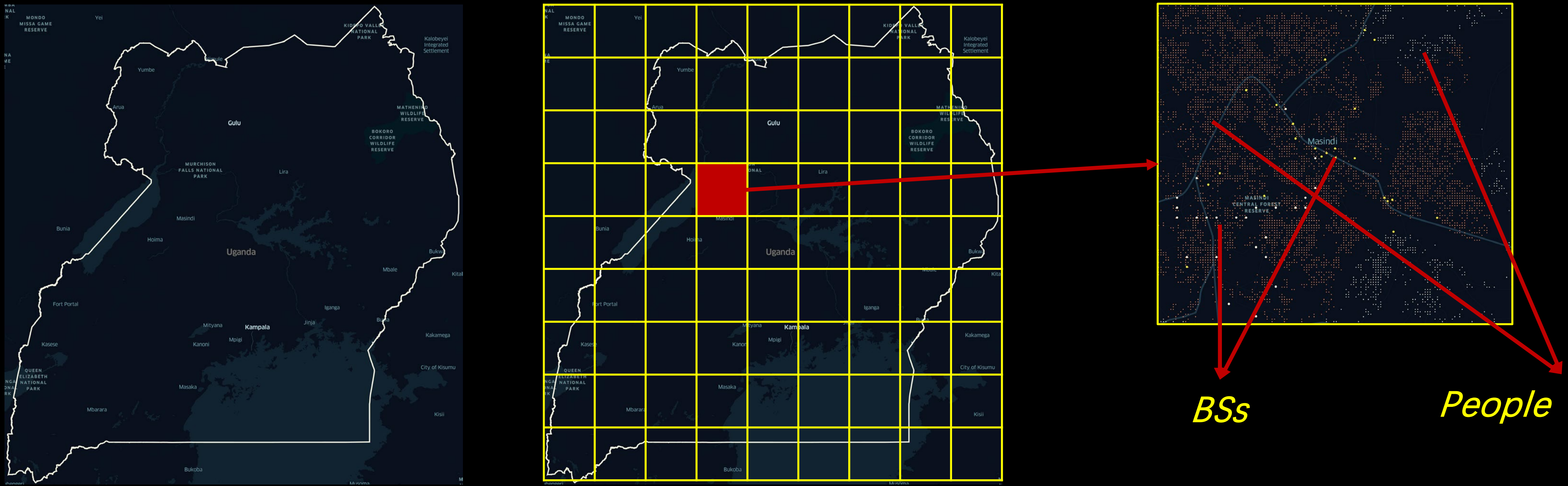


Methodology

a) Given a polygon of a country and calculate the max and min of longitude and latitude;

b) Split the square into I × J grids;

c) Calculate the number of BSs $B_{(i,j)}$ and the number of peoples $P_{(i,j)}$ in each grid (i, j);



Methodology

a) Given a polygon of a country and calculate the max and min of longitude and latitude;

b) Split the square into I × J grids;

c) Calculate the number of BSs $B_{(i,j)}$ and the number of peoples $P_{(i,j)}$ in each grid (i, j);

d) Obtain the imbalance index value as follows:

$$I_{(i,j)} = \begin{cases} 0, & P_{(i,j)} = 0 \ \& \ B_{(i,j)} = 0, \\ \gamma, & P_{(i,j)} > 0 \ \& \ B_{(i,j)} \geqslant \alpha, \\ P_{(i,j)}/\beta, & P_{(i,j)} > 0 \ \& \ B_{(i,j)} = 0, \\ P_{(i,j)}/\beta B_{(i,j)}, & otherwise. \end{cases}$$

No people and no BS in the grid

# of BS is larger than a threshold and we treat this grid low imbalance (gamma=0.01)

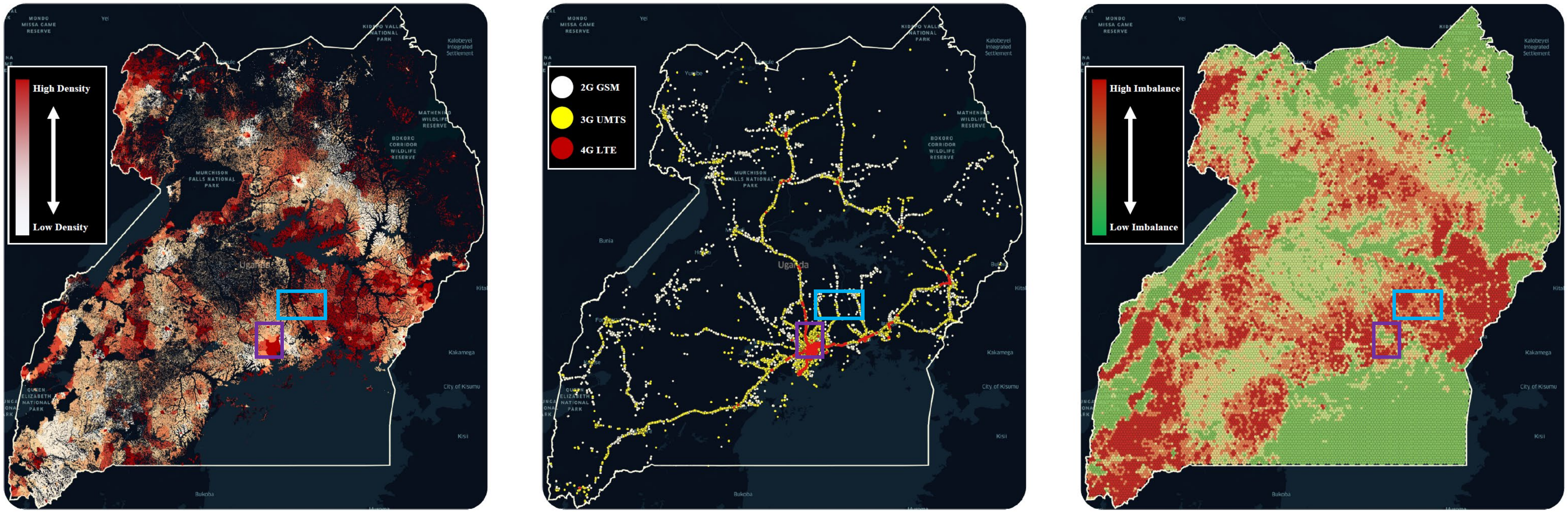
The grid has nonzero population but zero BS

# of users per BS can serve, and we set it to 100

BS threshold that indicates 'good' connectivity, we set it to 5

https://blog.wirelessmoves.com/2016/02/lte-and-the-number-of-simultaneously-connected-users.html

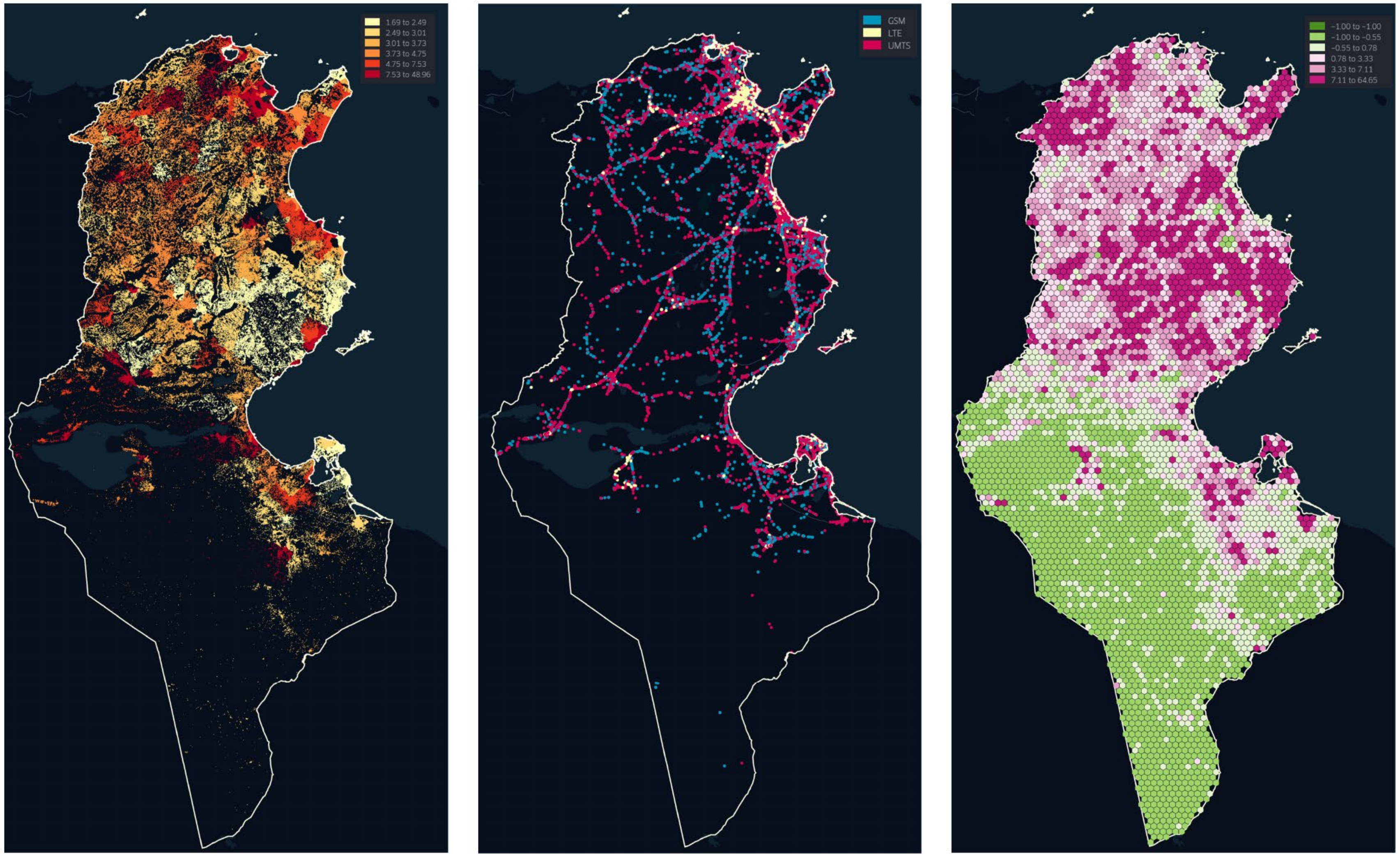
Imbalance Index Visualization: Uganda



(a) Population distribution. (b) BSs distribution. (c) Imbalance index.

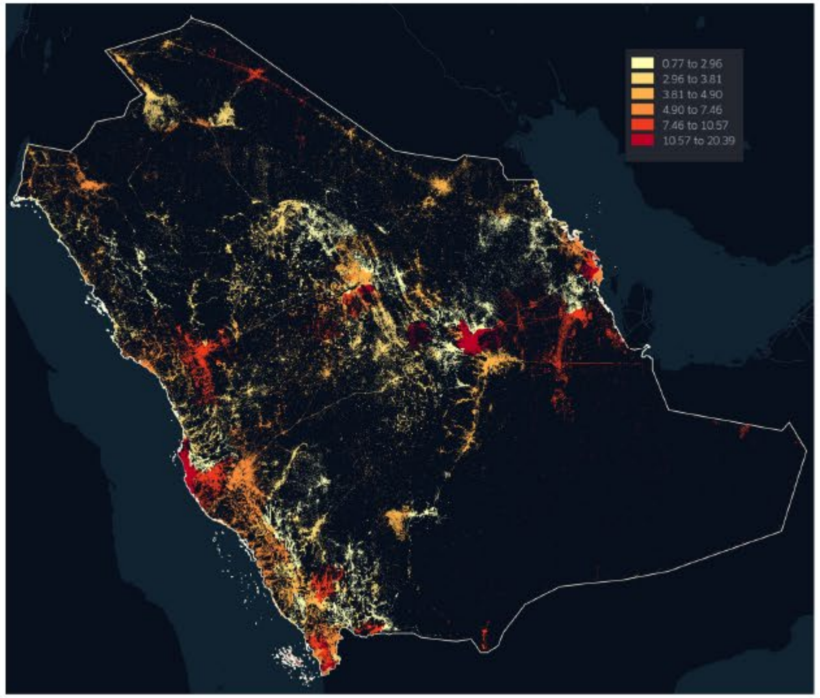
City center has relatively low imbalance index, *rural areas* have high imbalance index values.

Imbalance Index Visualization: Tunisia

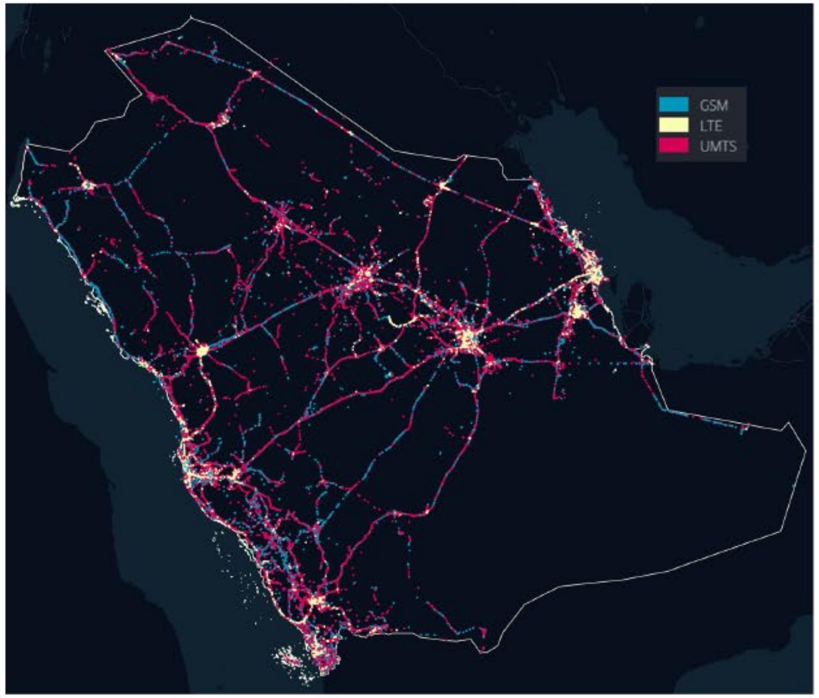


(a) Population distribution. (b) BS distribution. (c) Imbalance index visualization.

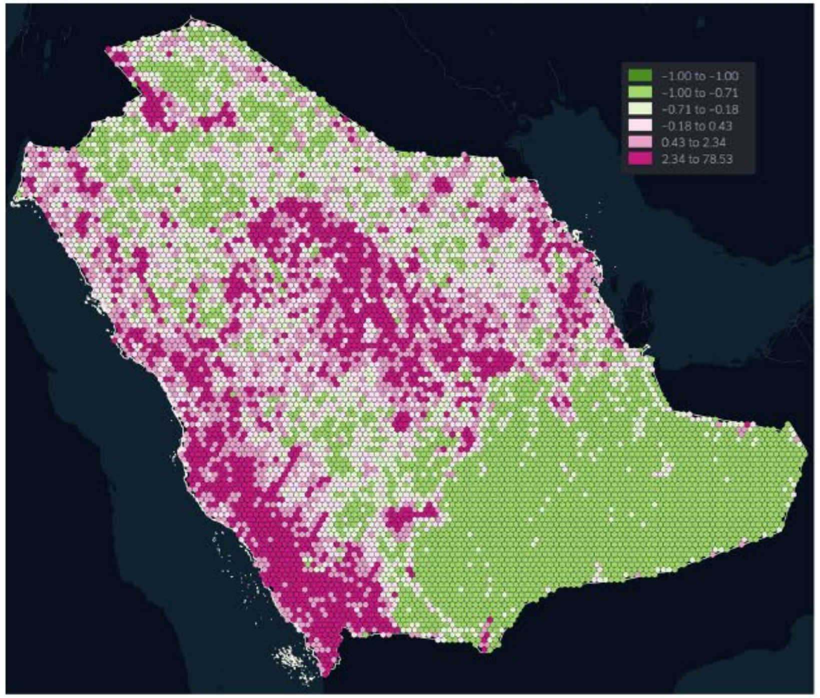
Imbalance Index: Saudi Arabia



(d) Population distribution.

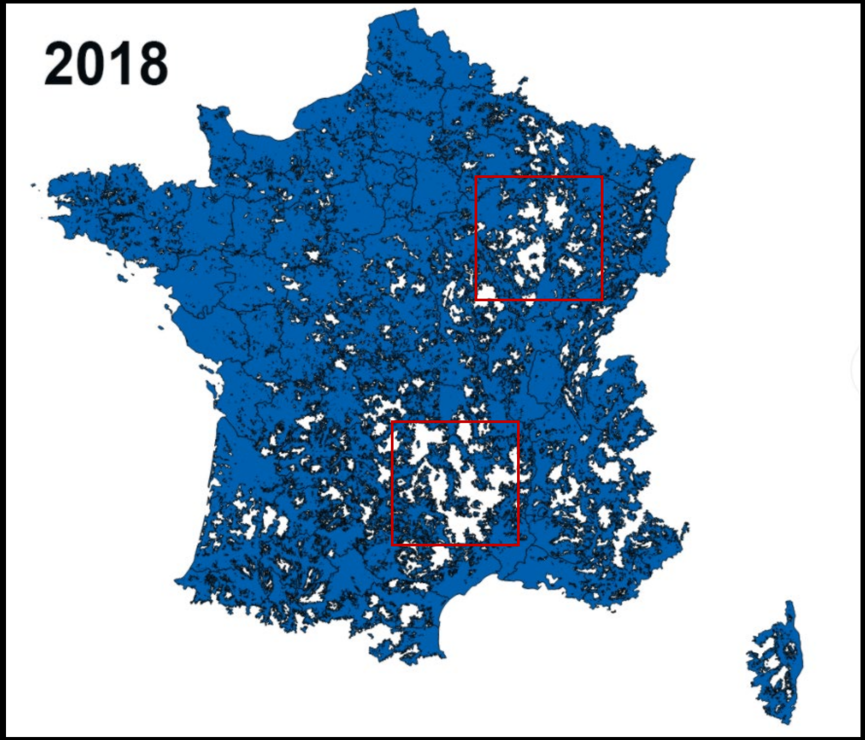


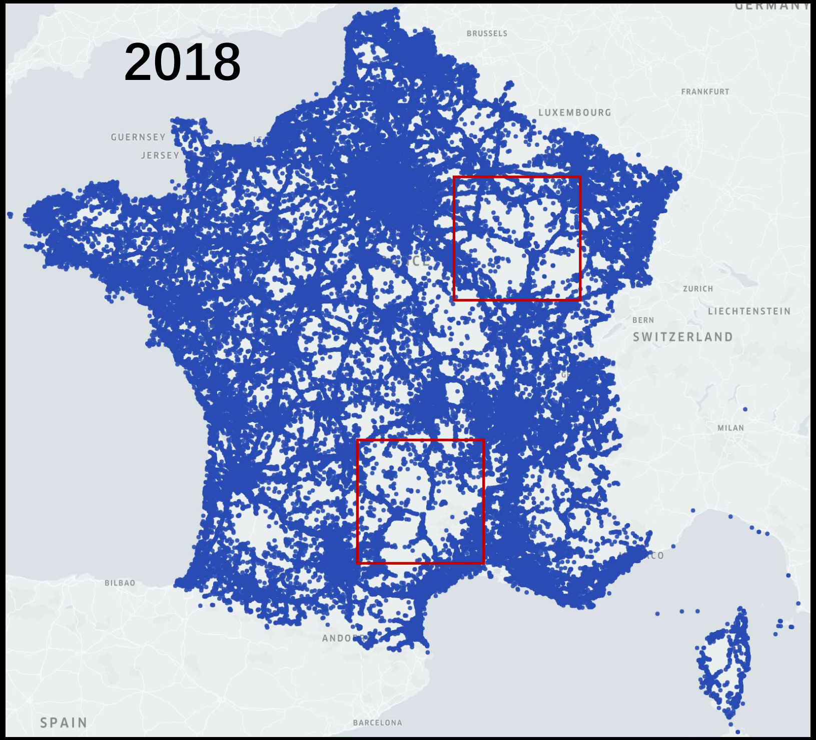
(e) BS distribution.



(f) Imbalance index visualization.

Validation on Digital Divide





- **This map is taken from here**: https://www.lagazettedescommunes.com/621283/en-cartes-et-en-graphiques-comment-avance-la-couverture-mobile-en-4g-en-france-22/
- This map is made base on real coverage data, which is available here: https://www.data.gouv.fr/fr/datasets/5b2b6715c751df6acaf0c2ee/#_

- **4G LTE BS (Cell ID) distribution till the end of 2018, using data from OpenCellId.**
- **Results are similar but the map looks sparser as we do not have the precise location and coverage information.**

Thanks!

Code: https://github.com/chuanting/misc

Population Data

Methodology: High Resolution Population Density Maps + Demographic Estimates

In partnership with the Center for International Earth Science Information Network (CIESEN) at Columbia University, we use state-of-the-art computer vision techniques to identify buildings from publicly accessible mapping services to create the world's most accurate population datasets. Our maps are available at 30-meter resolution - much more accurate than existing high-resolution maps, which are only available at a resolution of 100 meters. These high-resolution maps estimate the number of people living within 30-meter grid tiles in nearly every country around the world. Additionally, our datasets provide insights on the distribution of certain populations within each country, including the number of children under five, the number of women of reproductive age, as well as young and elderly populations, at unprecedentedly high resolutions. These maps aren't built using Facebook data and instead rely on combining the power of machine vision AI with satellite imagery and census information. One use case for these maps are disease prevention efforts - gender and age are crucial indicators for the transmission and control of diseases. These high-resolution maps can provide the necessary insights for health organizations to allocate resources and control outbreaks.

Adjustments to match the census population with the UN estimates are applied at the national level. The UN estimate for a given country (or state/territory) is divided by the total census estimate of population for the given country. The resulting adjustment factor is multiplied by each administrative unit census value for the target year.