# README

**About the data set**

>> 5G High Density Demand (HDD) Dataset in Liverpool City Region, UK

>> Mukesh Kumar Maheshwari1(ORCID: [0000-0002-2843-9758](https://orcid.org/0000-0002-2843-9758)), Alessandro Raschella1 (ORCID: [0000-0002-1626-8947](https://orcid.org/0000-0002-1626-8947)), Michael Mackay1 (ORCID: [0000-0001-9013-7884](https://orcid.org/0000-0001-9013-7884)), Max Hashem Eiza1 (ORCID: [0000-0001-9114-8577](https://orcid.org/0000-0001-9114-8577)), Jon Wetherall2 and Jen Laing2

1School of Computer Science and Mathematics, Liverpool John Moores University, Liverpool, L3 3AF, UK.

2 CGA Simulation, L8 1XE, UK.

**Organisations**

Liverpool John Moores University, Liverpool, UK

CGA Simulation, Liverpool, UK.

Rights holder(s) for the dataset

Alessandro Raschella

Year of publication

The article is in process.

**Description**

Provide a brief abstract or description of the dataset

*>> The wireless network data are a feasible way to understand the user behavior in a given environment and may be utilized for analysis, prediction and optimization. On the other hand, datasets from wireless service providers are not publicly available, and obtaining a dataset in real time is challenging. In this work, we present a 5G dense deployment dataset obtained from the Liverpool City Region High Density Demand (LCR HDD) project. The project involves network deployment and assessment at Salt Tar and the ACC Arena event venues located in the city of Liverpool. Digital twin technology is considered to generate the dataset, which is inputted to a system level simulator for data modeling and analysis. The data set consists of 3,000 users in the Salt Tar venue and 12,000 users in the ACC Arena venue with features including users' position, traffic type, Radio Unit (RU) association, Signal to Interference and Noise Ratio (SINR), Physical Resource Blocks (PRB), throughput, Block Error Rate (BLER), and a total length of 10,000 samples. The dataset is validated through experimental measurements and is released in a simple format for easy access.*

* Highlight key information such as data collection or generation methods, important characteristics, etc.

>> Digital twin model and Matlab based simulator

How the dataset should be cited <https://doi.org/10.24377/LJMU.d.00000236>

M. K. Maheshwari, et al., “5G High Density Demand Dataset in Liverpool City Region, UK”, *Sci Data.* **(**Article in progress)

**Contact details**

* Provide email for the PI

[m.i.mackay@ljmu.ac.uk](mailto:m.i.mackay@ljmu.ac.uk)

**Terms of use**

Creative Commons licence **Attribution 4.0 International**

**Project and funding information**

Title: Liverpool City Region High Density Demand (LCR HDD) project.

The project start and end dates - September 2023–March 2025

**Funding organisation name**

Department for Digital, Culture, Media and Sport, Innovate UK

**Contents**

The files are arranged in folder as shown in following figure, and they are name according to data feature type and UE Id. e.g. The file SINRDL\_UE\_ID\_0\_499 in folder SINR\_ACC\_Arena, present the DL SINR for user 0 to 499 in ACC Arena event venue.

**A diagram of a computer network

AI-generated content may be incorrect.**

File formats

>> .CSV files.

**Methods**

The data set is generated through a Liverpool 5G (L5G) simulator. The L5G simulator consists of two main components: a digital twin-based model of the venues and a Matlab-based simulator of the deployments. The digital twin component is developed by CGA simulation, a virtual simulation capable of solving real-world problems based on realistic user behaviour in terms of mobility and data traffic, achieved from real events. The Matlab-based component is developed from the Vienna 5G System Level (SL) Simulator,

which has the capability to mimic buildings, streets, radio environment conditions, such as path loss and shadowing, developed according to real-world scenarios based on data available in the OpenStreetMap (OSM) database. The MATLAB-based component can gather information from the radio environments and compute key performance parameters based on the input generated by the digital twin-based component, including downlink (DL) and uplink (UL) SINR (dB), user throughput (Mbps), PRB utilization, and BLER at every timestamp for each user, considering their current position and application.

* If detailed methods are available in a published or forthcoming article, you can refer to it here and provide a brief citation.

>> M. K. Maheshwari, et al., “5G High Density Demand Dataset in Liverpool City Region, UK”, *Sci Data.* **(**Article in progress).