Final Report on WP2: Simulation and Modelling for Spectrum Sharing

Spectrum Sandbox ITT Project

Queen Mary University of London, Telet Research, Aetha Consulting, Federated Wireless

Date: 27.03.2025

Table of Contents

1. Introduction 4
1.1 Project Overview 4
1.2 Work Package Structure..4
WP1: Spectrum Data Collection and Measurement (Led by Telet Research) 4
WP2: Simulation and Modelling of Dynamic Spectrum Sharing (Led by QMUL)..................5
WP3: Economic and Regulatory Impact Assessment (Led by Aetha Consulting) 6
1.3 Focus of this report: WP2 – Simulation and Modelling...6
1.4 Expected Outcomes of WP2 7
2. Literature Review: Key Technologies and Strategies in Spectrum Sharing 8
2.1 Ray Tracing: A High-Precision Approach to Signal Propagation 8
2.2 Communication Channels: Mapping Performance Across Spectrum Scenarios..............15
2.3 The Evolution of Spectrum Sharing: Approaches and Challenges...................................19
2.3.1 CR-Based Spectrum Sharing..19
2.3.2 D2D-Based Spectrum Sharing..20
2.3.3 IBFD-Based Spectrum Sensing 22
2.3.4 NOMA-Based Spectrum Sharing 23
2.4 Machine Learning for Smarter Spectrum Access and Interference Control 24
3. Simulation and Modelling Framework..25
3.1 Simulation Approach and Tools Used..25
3.1.1 Choice of MATLAB as the Simulation Platform..26
3.1.2. Ray Tracing for Realistic Environment Modeling 26
3.1.3 Machine Learning for Interference Prediction and Path Loss Estimation 27
3.2 Ray Tracing for Signal Propagation Analysis..27
3.2.1 Frequency Bands and Deployment Scenarios..27
3.2.2 Key Propagation Effects Considered 28
3.3 Communication Channel Model 28
4. Insights from WP2: Validating and Benchmarking Performance..29
4.1 Benchmarking Accuracy: Comparing Simulations with Real-World Data 29

4.1.1 Testing in Rural and Urban Environments: Case Studies from Chalke Valley & Liverpool 36
4.2 Bridging the Gap: Understanding Discrepancies and Material Effects 38
4.3 Interference Dynamics and Performance Evaluation For 5G Network in Band 3 and n77 40
4.3.1 Band n77 gNodeB threshold simulation 40
4.3.2 Band n77 TDD synchronisation simulation 43
4.3.3 Band 3 cross Bandwidth simulation 45
4.3.4 Throughput coverage 47
4.4 Challenges, Limitations and Considerations 49
5. Innovations and Future Directions 51
5.1 AI-Powered Path Loss Prediction: A Smarter Alternative to Traditional Models 51
5.1.1 Data Preparation and Feature Extraction 51
5.1.2 Model Architecture 52
5.1.3 Fine-Tuning for Accuracy: Optimising the Learning Process 53
5.1.4 Measuring Success: Evaluating the AI Model's Performance 54
5.1.5 Key Observations: Strengths and Weaknesses of AI-Based Predictions 56
5.1.6 Refining the Model: Future Enhancements for Better Accuracy 57
5.1.7 Use Case: Base Station Optimisation Methodology 58
5.2 Enhancing Simulations with LiDAR: Improving Environmental Accuracy 66
5.2.1 Data Source 68
5.2.2 Methodology 69
6. Conclusion 74
References 77

1. Introduction

1.1 Project Overview

The increasing demand for wireless communication services has led to severe spectrum congestion [1, 2, 3], particularly in urban areas where the radio frequency (RF) spectrum is often inefficiently allocated. For this reason, the QMUL Spectrum Sandbox initiative has been taken, which is a collaborative research effort aimed at advancing dynamic spectrum-sharing solutions to improve wireless network efficiency, particularly in underutilised spectrum bands. In 2019, Ofcom [4] set out two license products:

- Shared Access License (SAL), which gives access to four spectrum bands (1800 MHz, 2300 MHz, $3800-4200$ MHz, and $2425-2650$ MHz) [5] assuming rule compliance and non-interference with protected systems.
- Local Access License (LAL), which provides a way for 3rd parties to access spectrum licensed to the UK's Mobile Network Operators (MNOs) that is not currently in use nor planned for use in the near future [6].

Current spectrum licensing procedures are often time-consuming and administratively complex, where receiving a SAL can take about 6 to 12 months for approval, and LALs have an inherent high degree of uncertainty for access and continued operation due to the requirement to protect future incumbent MNO operations. This greatly limits their feasibility for commercial use.

This project proposes a Dynamic Spectrum Access (DSA) framework, allowing automated, real-time spectrum allocation in under two minutes, ensuring efficient usage while preventing harmful interference with primary mobile operators. By leveraging advanced modelling, machine learning, and real-world testing, this research aims to develop a scalable dynamic spectrum-sharing system. This project is led by a consortium comprising **Queen Mary University of London (QMUL)**, Telet Research, Aetha Consulting, and Federated Wireless, each bringing specialised expertise in wireless communications, spectrum policy, and economic assessment.

The project is structured into three key Work Packages (WPs) to comprehensively address spectrum-sharing feasibility, technical implementation, and economic impact. This is also shown in the form of a workflow diagram in Figure 1.

1.2 Work Package Structure

WP1: Spectrum Data Collection and Measurement (Led by Telet Research)

WP1 focuses on real-world spectrum measurement and data collection to establish a foundation for dynamic spectrum assignment. The objectives include:

- Deploying 5G radio equipment (e.g., cellXica M5Q, Picocom PC802) to scan and collect real-time spectrum usage data.
- Assessing interference conditions in urban and rural environments through practical field tests.
- Validating dynamic assignment feasibility using Federated Wireless's Spectrum Access System (SAS).
- Conducting long-term studies on spectrum availability and impact.

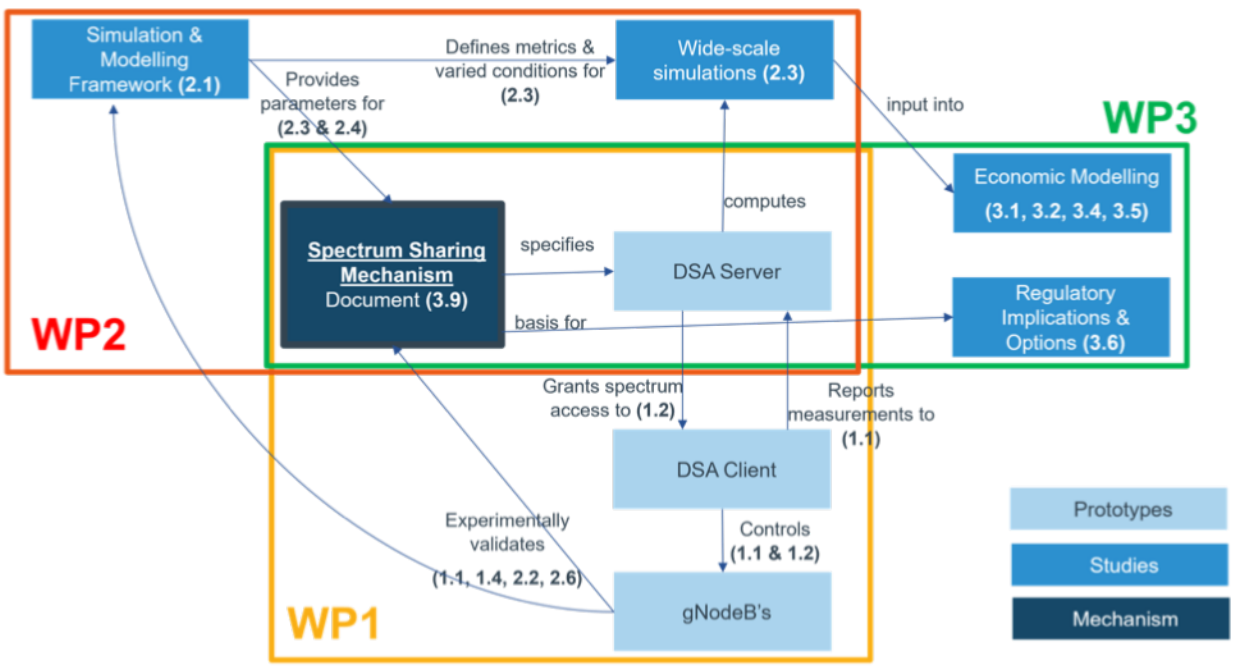


Figure 1: Integration of spectrum sharing mechanism within project workflow (see, Q5 (Methodology) of Ref. [7]).

WP2: Simulation and Modelling of Dynamic Spectrum Sharing (Led by QMUL)

WP2, which is the focus of this report, utilises simulation and modelling techniques to assess the scalability and impact of dynamic spectrum-sharing solutions. It builds upon the real-world data from WP1 to:

- Develop an accurate simulation framework for spectrum sharing.
- Analyse the performance of trade-offs under different deployment scenarios.
- Investigate machine learning (ML)-based optimisation to enhance decision-making.

WP3: Economic and Regulatory Impact Assessment (Led by Aetha Consulting)

WP3 evaluates the economic feasibility and regulatory implications of dynamic spectrum sharing. Key objectives include:

- Quantifying economic benefits for different stakeholders.
- Assessing the regulatory framework required for implementation.
- Engaging with industry stakeholders through workshops to validate findings.
- Developing a roadmap for large-scale adoption of dynamic spectrum licensing.

1.3 Focus of this report: WP2 – Simulation and Modelling

This report primarily covers WP2, which focuses on assessing the feasibility of dynamic spectrum assignment through advanced simulations and modelling techniques. The key aspects of WP2 include:

1. Integration of Ray Tracing for Realistic Propagation Modelling:
   Unlike conventional statistical path loss models, which rely on empirical or stochastic formulations, this study employs ray tracing-based simulations to incorporate critical multipath effects, diffraction, and reflection mechanisms. This methodology provides a granular and physics-based representation of signal propagation, enhancing the fidelity of spectrum-sharing assessments in complex environments. Currently, Ofcom's coverage maps for base stations do not account for environmental factors, often assuming uniform circular coverage, which does not reflect actual propagation characteristics. In contrast, ray tracing provides a more detailed and realistic representation of coverage, typically revealing smaller coverage areas than those depicted by Ofcom. Based on the findings in this report, we recommend that Ofcom consider using more accurate models including ray tracing and machine learning models to maximise the usage of spectrum, mitigate coverage gaps (not-spots), and ensure more reliable mobile connectivity. Most importantly, it would be useful to revisit and modify legacy deterministic models based on accurate numerical and ML approaches.

2. Validation Against Real-World Data through WP1 Integration:
   A key aspect of this work is the validation of simulation results against real-world measurement data from WP1. By systematically comparing simulated path loss and interference patterns with field measurements, this approach ensures regulatory compliance and industrial applicability while improving model reliability for real-world deployments.

3. Dynamic spectrum sharing at 1800 MHz in Band 3: This study explores the feasibility of dynamic spectrum sharing in the 1800 MHz Band 3, a spectrum range traditionally allocated for mobile communication. The purpose is to eliminate existing mobile "not-spots" by providing mobile coverage in many of those areas through smaller local mobile operators using Local Access Licensing in a more cost-effective manner than the national mobile operators such as EE, O2, Vodafone and Three. The investigation assesses interference management strategies, coexistence with incumbent users, and real-time spectrum access techniques, contributing to future spectrum efficiency enhancements.

4. Optimising Base Station Deployment for Targeted Local Coverage in the 3800-4200 MHz Band: This analysis identifies optimal configurations and minimum base station deployments required to deliver effective local-area coverage within the 3800-4200 MHz band under Shared Access Licensing conditions, ensuring efficient spectrum use and interference avoidance with incumbent operators. Our findings aim to support cost-effective infrastructure planning specifically tailored to private network deployments.

5. Machine Learning (ML) for Path Loss and Interference Prediction: Traditional deterministic and empirical models for path loss estimation are often computationally intensive or lack adaptability to dynamic environments. This work leverages machine learning techniques to enhance propagation modelling accuracy, enabling the real-time prediction of interference levels. By integrating ML-based models, the study reduces computational overhead while improving spectrum allocation decisions.

6. Enhancing environmental models using LiDAR data: OpenStreetMap (OSM) data, frequently used for urban environment modelling, often lacks fine-grained building details required for precise propagation simulations. To address this, an ML-driven approach is proposed to integrate LiDAR data, allowing for the automatic extraction of structural information and material properties. This fusion enhances environmental model accuracy, leading to more reliable spectrum-sharing simulations.

1.4 Expected Outcomes of WP2

The findings from WP2 will:

- Provide a validated, data-driven framework for dynamic spectrum sharing.
- Offer insights into AI-driven spectrum management and efficient interference mitigation.
- Support WP3 by quantifying the technical and economic viability of dynamic spectrum licensing.
- Contribute to regulatory discussions on scalable and secure spectrum access frameworks.

2. Literature Review: Key Technologies and Strategies in Spectrum Sharing

2.1 Ray Tracing: A High-Precision Approach to Signal Propagation

Wireless communication systems rely on radio waves to transmit signals, but the propagation of these waves is influenced by complex interactions with the environment. Factors such as reflection, diffraction, refraction, and scattering significantly impact signal strength, coverage, and quality of service. Accurate propagation modelling is essential for predicting signal behaviour, estimating interference, and optimising spectrum-sharing strategies.

Traditional empirical models, such as the Hata Model [8] and the Okumura Model [9], approximate path loss based on system parameters like frequency, antenna height, and terrain properties. While these models provide quick estimates of signal attenuation over large areas, they assume simplified propagation conditions and are valid only for environments that resemble their original calibration scenarios. Consequently, they fail to capture fine-grained spatial and temporal variations in urban and dense environments, where obstacles like buildings, trees, and vehicles create complex multipath effects.

Unlike these empirical models, ray tracing is a deterministic approach that models electromagnetic waves as individual rays, tracing their interactions with physical surfaces in a given three-dimensional (3D) environment. Ray tracing accurately simulates multipath propagation by accounting for the laws of optics and wave physics, making it particularly suitable for urban and heterogeneous environments.

Ray tracing treats radio waves as narrow beams of energy that travel in straight lines through a homogeneous medium. As these rays propagate, they interact with the environment in multiple ways, which influence signal behaviour [10]:

Interaction	**Description**
Line of sight (LOS)	The ray travels directly from the transmitter to the receiver.
Reflection	The ray reflects off the surface according to the law of reflection.
Refraction (transmission)	The ray refracts as it moves into a new medium, according to the law of refraction.

Diffraction	The ray diffracts off the surface according to the law of diffraction. One ray can spawn many diffracted rays.
Diffuse scattering	The ray interacts with a rough surface such as the ocean or a building facade.

Ray tracing models [11] perform numerical simulations to:

- Predict the paths of rays from transmitters to receivers, considering multiple propagation effects.
- Estimate signal strength, phase changes, and path loss for each ray based on environmental interactions.

By leveraging high-fidelity 3D maps, ray tracing enables the precise modelling of radio environments under different spectrum-sharing scenarios. For example, in an urban 5G network, ray tracing can simulate reflections of skyscrapers, diffraction around street corners, and penetration losses through windows, ensuring accurate coverage predictions.

<u>Propagation Loss</u>

Effect of surface materials:

An important factor in ray tracing-based propagation modelling is the impact of surface materials on reflection losses. Different building materials affect the absorption and reflection of electromagnetic waves, which influences overall signal propagation. The ray tracing model incorporates surface material properties into path loss calculations by considering their complex relative permittivity (ε_r), which describes how the material interacts with electric fields.

The complex permittivity is given by:

$$\varepsilon_r = \varepsilon_r' + j\varepsilon_r''$$

where: ε_r' is the real part of the relative permittivity, determining the ability of the material to store electrical energy, ε_r'' is the imaginary part, related to the conductivity σ of the material, which contributes to signal attenuation, ε_0 is the permittivity of free space, f is the frequency in Hz.

The imaginary part of permittivity is calculated as:

$$\varepsilon_r'' = \frac{\sigma}{2\pi\varepsilon_0 f}.$$

For different building materials, ε_r' and σ can be estimated using empirical formulas:

$$\varepsilon_r' = af^b$$
$$\sigma = cf^d$$

where a, b, c, d are material-specific constants derived from experimental data.

These calculations are guided by ITU-R recommendations, including ITU-R P.2040-3 [12] and ITU-R P.527-5 through ITU-R P.527-6 [13], which provide methods and reference values for estimating permittivity and conductivity across different frequencies.

This material-aware ray tracing approach enables more realistic simulations by accounting for frequency-dependent reflection, refraction, and absorption effects. For example:

- Glass windows cause partial transmission and reflection, affecting indoor penetration losses.
- Concrete walls have high permittivity, leading to strong reflections and significant attenuation.
- Metallic surfaces act as perfect reflectors, contributing to signal multipath effects.

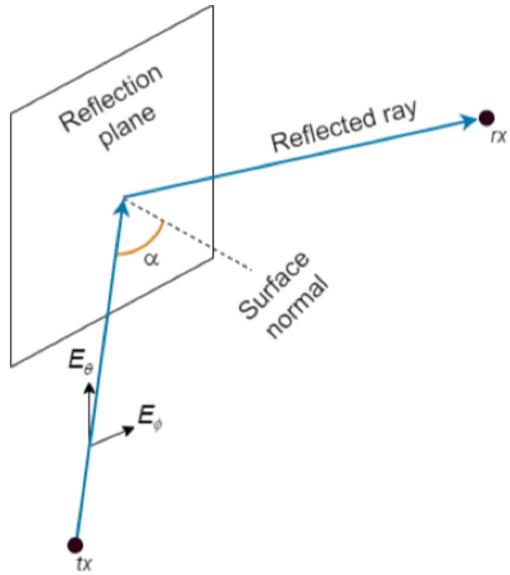


Figure 2: Reflection of a Ray [14].

Reflection loss:

Figure 2 shows a reflection path from a transmitter site *tx* to a receiver site *rx*.

The model determines polarisation and reflection loss using these steps.

1. Track the propagation of the ray in 3-D space by calculating the propagation matrix P. The matrix is a repeating product, where i is the number of reflection points.

$$P = \prod_i P_i \, .$$

For each reflection, calculate P_i by transforming the global coordinates of the incident electromagnetic field into the local coordinates of the reflection plane, multiplying the result by a reflection coefficient matrix, and transforming the coordinates back into the original global coordinate system [15]. The equations for P_i and P_0 are:

$$P_i = \begin{bmatrix} s_{out} & p_{out} & k_{out} \end{bmatrix} \begin{bmatrix} R_V(\alpha) & 0 & 0 \\ 0 & R_H(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_{in} & p_{in} & k_{in} \end{bmatrix}_i^{-1}$$

$$P_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where:

- s, p, and k form a basis for the plane of incidence (the plane created by the incident ray and the surface normal of the reflection plane). s and p are perpendicular and parallel, respectively, to the plane of incidence.
- k_{in} and k_{out} are the directions (in global coordinates) of the incident and exiting rays, respectively.
- s_{in} and s_{out} are the directions (in global coordinates) of the horizontal polarisations for the incident and exiting rays, respectively.
- p_{in} and p_{out} are the directions (in global coordinates) of the vertical polarisations for the incident and exiting rays, respectively.
- R_H and R_V are the Fresnel reflection coefficients for the horizontal and vertical polarisations, respectively. α is the incident angle of the ray and ε_r is the complex relative permittivity of the material.

$$R_H(\alpha) = \frac{cos(\alpha) - \sqrt{\left(\varepsilon_r - sin^2(\alpha)\right)/\varepsilon_r^2}}{cos(\alpha) + \sqrt{\left(\varepsilon_r - sin^2(\alpha)\right)/\varepsilon_r^2}}$$

$$R_V(\alpha) = \frac{cos(\alpha) - \sqrt{\varepsilon_r - sin^2(\alpha)}}{cos(\alpha) + \sqrt{\varepsilon_r - sin^2(\alpha)}}$$

2. Project the propagation matrix P into a 2-by-2 polarisation matrix R. The model rotates the coordinate systems for the transmitter and receiver so that they are in global coordinates.

$$R = \begin{bmatrix} H_{in} \bullet H_{rx} & V_{in} \bullet H_{rx} \\ H_{in} \bullet V_{rx} & V_{in} \bullet V_{rx} \end{bmatrix}$$

$$H_{in} = P\left(V_{tx} \times k_{tx}\right)$$

$$V_{in} = PV_{tx}$$

where:

- H_{rx} and V_{rx} are the directions (in global coordinates) of the horizontal $\left(E_{\theta}\right)$ and vertical $\left(E_{\varnothing}\right)$ polarisations, respectively, for the receiver.
- H_{in} and V_{in} are the directions (in global coordinates) of the propagated horizontal and vertical polarisations, respectively.
- V_{tx} is the direction (in global coordinates) of the nominal vertical polarisation for the ray departing the transmitter.
- k_{tx} is the direction (in global coordinates) of the ray departing the transmitter.

3. Specify the normalised horizontal and vertical polarisations of the electric field at the transmitter and receiver by using the 2-by-1 Jones polarisation vectors J_{tx} and J_{rx}, respectively. If either the transmitter or receiver are unpolarised, then the model assumes $J_{tx} = J_{rx} = \frac{\sqrt{2}}{2}\begin{bmatrix}1\\1\end{bmatrix}$.
4. Calculate the polarisation and reflection loss IL by combining R, J_{tx} and J_{rx}.

$$IL = -20log_{10}\left|J_{rx}^{-1}RJ_{tx}\right|$$

Diffraction loss:

The model calculates diffraction loss by using computations based on the Uniform Theory of Diffraction (UTD) [16].

For a first-order signal diffraction, the equation for path loss, PL_D, is:

$$PL_D = JV'_{rx}H_{diffl}JV_{tx},$$

where:

- JV_{rx} and JV_{tx} are polarisation vectors for the receiver and transmitter, respectively, specified as Jones vectors.
- H_{diffl} is the diffraction matrix.

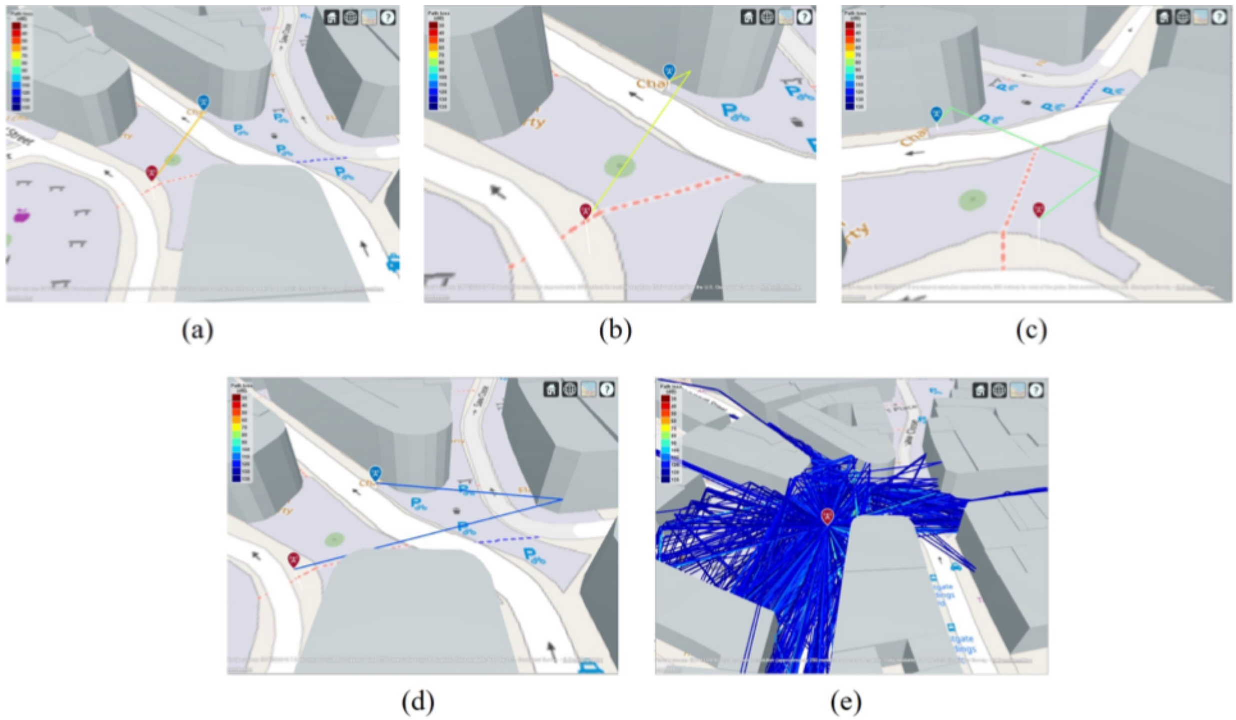


Figure 3: Ray tracing visualisation of path loss in an urban environment. The transmitter (Tx) is marked in a red marker pin, receiver (Rx) in a blue marker pin. (a) Line of sight, (b) non-line of sight having a single reflection, (c) non-line of sight having a double reflection, (d) non-line of sight having single diffraction, (e) all possible combinations of rays.

The equation for the diffraction matrix contains three terms.

- The first term is a geometric coupling matrix that rotates the polarisation vector from the basis of the ray coordinates to the basis of the edge-fixed incidence plane. The edge-fixed incidence plane contains the ray and the edge.
- The second term is a polarisation matrix containing diffraction coefficients for the local horizontal and vertical polarisations, $D_\perp$ and $D_\|$, and an amplitude scaling factor. For more information about the diffraction coefficients and amplitude scaling factor, see [16] and [17].
- The third term is a geometric coupling matrix that rotates the polarisation vector from the basis of the edge-fixed incidence plane to the basis of the edge-fixed diffraction plane. The edge-fixed diffraction plane contains the diffracted ray and the edge.

High-fidelity ray tracing models, such as those implemented in Remcom Wireless InSite provide the necessary computational power to simulate realistic propagation conditions. Figure 3 illustrates an example of a ray tracing simulation for an urban environment, highlighting multipath propagation effects.

2.2 Communication Channels: Mapping Performance Across Spectrum Scenarios

In addition to ray tracing, communication channel models provide a mathematical representation of the transmission medium, enabling the evaluation of network performance under different spectrum-sharing scenarios. An illustration of spectrum sharing and mechanism to allocate unused spectrum to the secondary base station, in our case a gNodeB (gNB) is shown in Figures 4 and 5, respectively. These models incorporate key parameters tabulated in Table 1 according to the 3GPP standards for 5G NR [18].

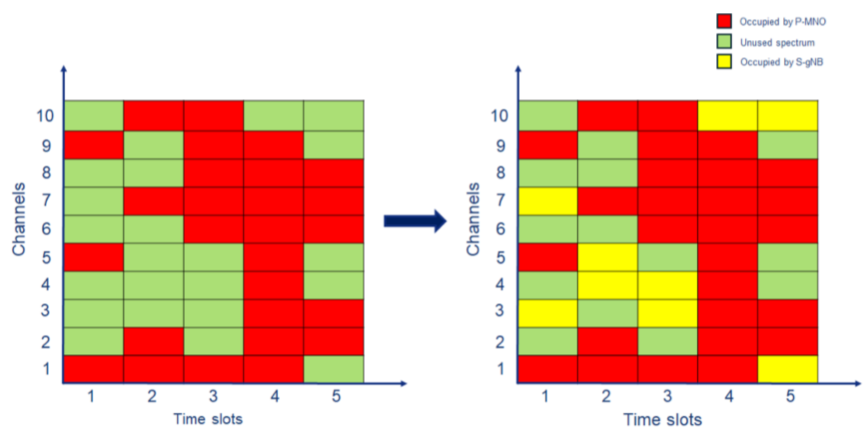


Figure 4: Illustration of a change in spectrum allocation over time, showing areas occupied by primary mobile network operator (P-MNO), secondary gNodeB (S-gNB), and unused spectrum.

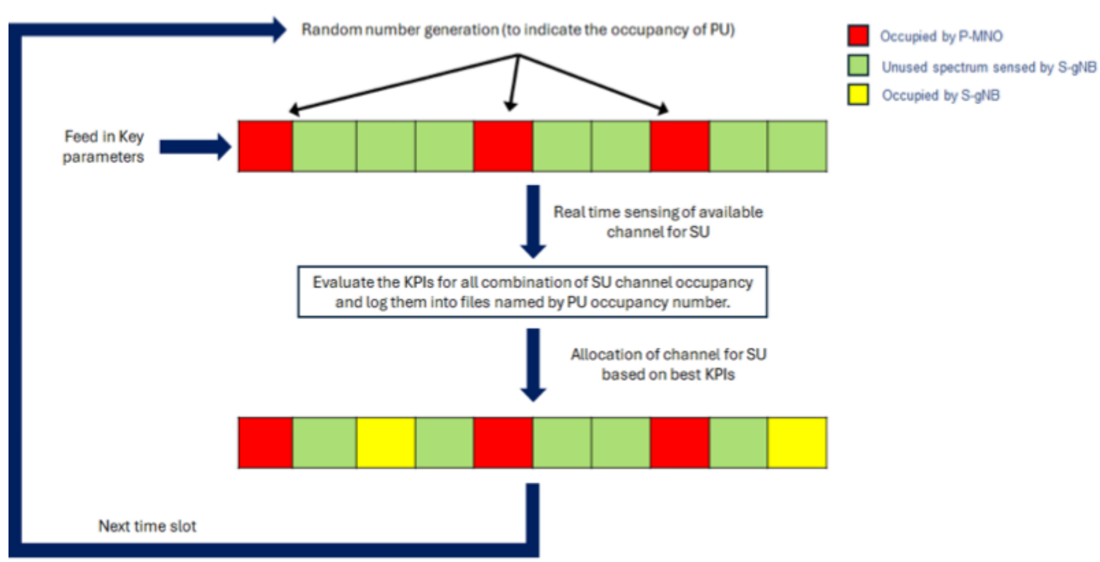


Figure 5: A visual representation of the steps involved in dynamic spectrum access, from random PU occupancy generation to SU channel allocation based on key performance indicators (KPIs).

Table 1: Key parameters for 5G against 3GPP standards

Parameter	**Band 3**	**Band n77**	**Description**
Carrier Frequency	1805 MHz to 1880 MHz (uplink) 1710 MHz to 1785 MHz (downlink) [19]	3700 MHz to 4200 MHz (uplink and downlink) [19]	The frequency at which PDSCH operates.
Bandwidth	Varies based on Resource Block and Subcarrier Spacing [20]	Varies based on Resource Block and Subcarrier Spacing [20]	Overall bandwidth allocated for PDSCH
Subcarrier Spacing (SCS)	15 kHz 30 kHz and 60 kHz (optional for specific scenarios) [21]	15 kHz 30 kHz (standard for n77) 60 kHz (optional for specific scenarios) [21]	Common spacings: 15 kHz, 30 kHz, 60 kHz, or 120 kHz.
Modulation Scheme	QPSK, 16QAM, 64QAM, 256QAM [22]	QPSK, 16QAM, 64QAM, 256QAM [22]	Types: QPSK, 16QAM, 64QAM, or 256QAM.
Modulation and Coding Scheme (MCS)	Table 5.1.3.1-2: MCS index table 2 for PDSCH [23]	Table 5.1.3.1-2: MCS index table 2 for PDSCH [23]	MCS index determines modulation type and coding rate.
Transport Block Size	Varies based on MCS; can be significant depending on configuration	Varies based on MCS; can be substantial depending on configuration	Size of the data being transmitted, influenced by MCS.
Coding Rate	Varies based on MCS; typically ranges from 1/3 to 1 [24]	Varies based on MCS; typically ranges from 1/3 to 1 [24]	Rate of error correction coding applied to the data.

Resource Block (RB) Allocation	Varies based on Reference channel [24]	Varies based on Reference channel [24]	Number of resource blocks assigned to the PDSCH.
Time Allocation	For transmission scheme 1 of the PDSCH, the UE may assume that a gNB transmission on the PDSCH would be performed with up to 8 transmission layers on antenna ports 1000-1011 as defined in Clause 7.3.1.4 of [4, TS 38.211], subject to the DM-RS reception procedures in Clause 5.1.6.2. [27]		Duration in time domain (slots or symbols) for PDSCH.
Slot Configuration	subcarrier spacing configuration µ varying along with Nsymb which also effect on the cyclic prefix [26]		Specific slots assigned for PDSCH transmission.
Reference Signal Configuration	Varies based on Reference channel [24]	Varies based on Reference channel [24]	DMRS configuration, including type and ports.
Power Control [25]	Wide Area BS (Note) Medium Range BS ≤ 38 dBm Local Area BS ≤ 24 dBm	Wide Area BS (Note) Medium Range BS ≤ 38 dBm Local Area BS ≤ 24 dBm	Settings for downlink power control and UE power control.

WP2 employs a MATLAB-based channel model, parameterised according to empirical data from Work Package 1 (WP1), which captures real-world spectrum usage, interference conditions, and network load variations. This approach ensures that the simulated scenarios reflect actual deployment environments, enhancing the reliability and applicability of the findings. Moreover, the integration of bit-level simulations allows for the analysis of key performance indicators (KPIs), such as bit error rate (BER), throughput, and latency, under different spectrum-sharing configurations. Some of the KPIs are listed in Table 2.

Table 2: KPIs in wireless communication channel model

KPIs	**Definition**	**Importance**
Signal-to-Noise Ratio (SNR)	Ratio of signal power to noise power.	Higher SNR indicates better signal quality and fewer errors.
Signal-to-Interference-plus-Noise Ratio (SINR)	The ratio of signal power to the sum of interference and noise power.	Accounts for interference from other signals, crucial for multi-user communication.
Bit Error Rate (BER)	The fraction of bits received in error over the total transmitted bits.	A lower BER indicates higher communication reliability.
Frame Error Rate (FER)	The ratio of incorrectly received frames to the total transmitted frames.	Used in packet-switched networks to measure reliability.
Channel Capacity	Maximum achievable data rate of the channel under given conditions.	Determines the theoretical limit for data transmission.
Throughput	The actual amount of data successfully transmitted per second.	A real-world performance measure, lower than channel capacity due to losses.
Spectral Efficiency	The number of bits transmitted per second per Hz of bandwidth.	Measures efficient utilisation of frequency spectrum.
Path Loss	Reduction in signal strength as it propagates through space.	Affects coverage and power requirements.
Delay Spread	The difference between the arrival times of the earliest and latest significant multipath components.	Affects inter-symbol interference (ISI) and equalisation requirements.
Latency	The time delay in data transmission from source to destination.	Critical for real-time applications like VoIP and gaming.

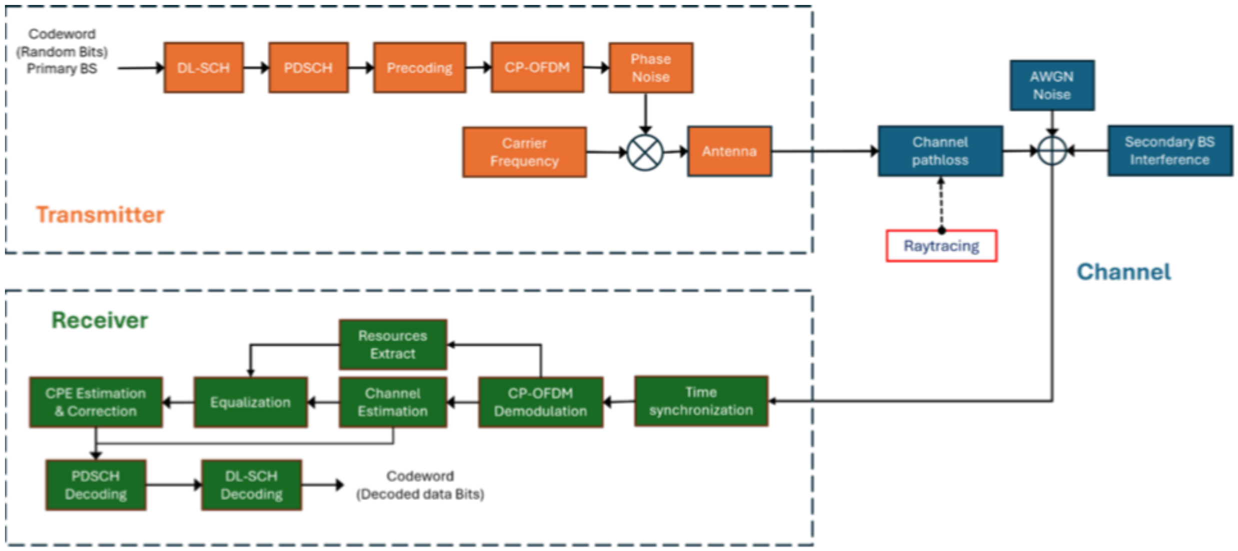


Figure 6: 5G NR downlink physical layer processing chain, showing the transmitter and receiver functionalities, including channel estimation, equalization, and demodulation.

While deterministic channel models offer a high degree of accuracy, they are computationally intensive, particularly when simulating large-scale networks. To address this challenge, hybrid approaches that combine ray tracing-based deterministic models with statistical approximations are being developed. This ensures a balance between simulation accuracy and computational efficiency, allowing for extensive parameter sweeps across different deployment scenarios. An example of a communication wireless channel model is shown in Figure 6.

2.3 The Evolution of Spectrum Sharing: Approaches and Challenges

The basic idea of spectrum sharing is to allow transceivers to use idle or underutilised spectrum bands temporally and geographically. Indeed, spectrum sharing is a promising strategy to tackle the imbalance between limited spectrum resources and unprecedented traffic demands [28]. We provide a survey of some of the techniques for spectrum sharing.

2.3.1 CR-Based Spectrum Sharing

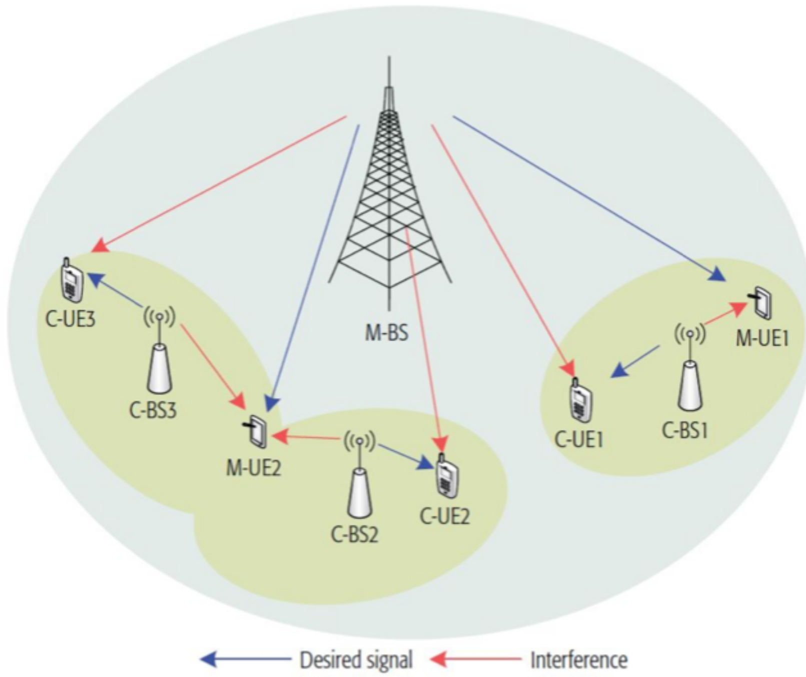


Figure 7: A typical scenario of CR-based spectrum sharing. C-BS: Cognitive base station, C-UE: Cognitive user equipment [1].

Cognitive radio (CR) is an effective technique to enhance spectrum efficiency and potentially ameliorate the spectrum scarcity problem [29]. By sensing the radio environment, cognitive users can adaptively configure transmitters and protect incumbent users. In general, a cognitive period consists of two phases: spectrum sensing and cognitive transmission. In the spectrum sensing phase, cognitive users sense the radio environment and collect spectrum information (e.g., occupation status, traffic, energy, channel gain). In the cognitive transmission phase, cognitive users select the best spectrum bands and adapt transmissions according to the collected spectrum information. A typical scenario of underlay spectrum sharing in a microcell is provided in Figure 7, where the macro-base station (M-BS) and macro user types of equipment (M-UEs) denote incumbent macro nodes.

2.3.2 D2D-Based Spectrum Sharing

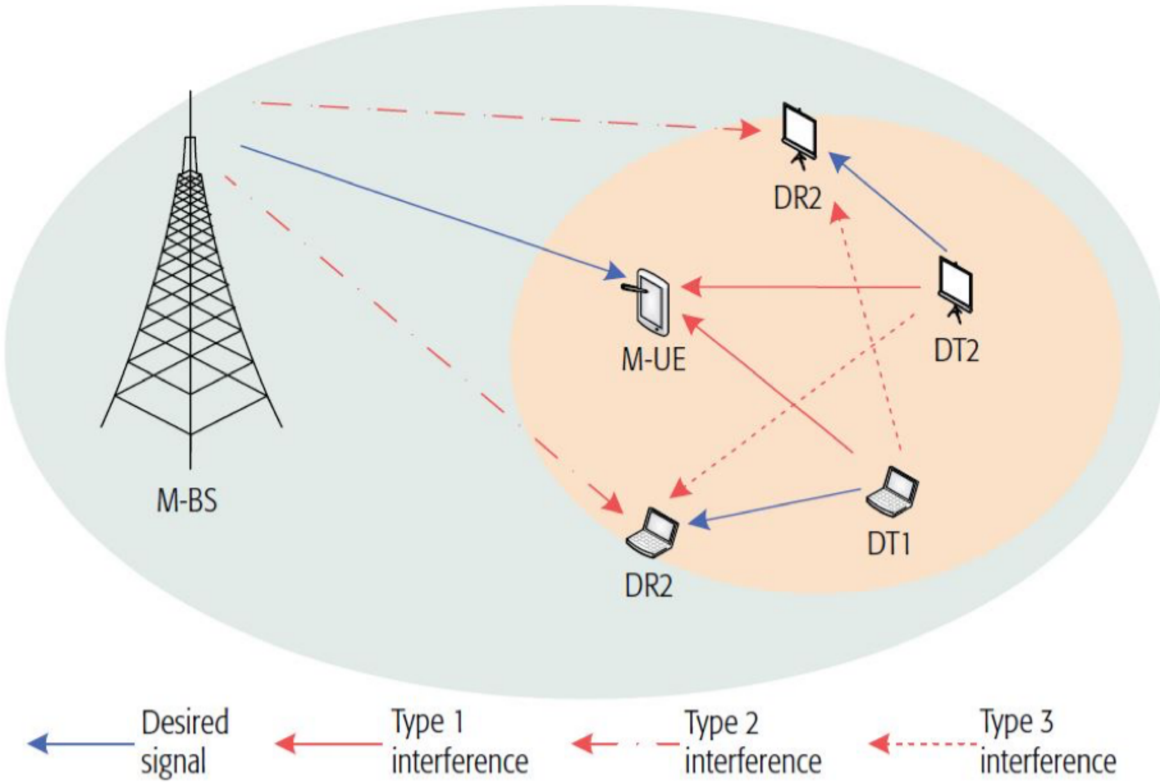


Figure 8: Three types of interference in an underlay D2D communication network. DT: D2D transmitter; DR: D2D receiver [1].

Two users are allowed to communicate via a device-to-device (D2D) connection when they are close to each other and far away from the M-BS [30]. By reducing the communication distance, D2D communication can improve throughput, reduce energy consumption, and decrease latency. To enhance spectrum efficiency, it is suggested that D2D communications share the spectrum bands of M-UEs under the central management of the M-BS, namely, underlay D2D communication. There are usually three types of interference in an underlay D2D communication network, as shown in Figure 8:

- Type I: interference from D2D communications to M-UE communications
- Type II: interference from M-UE communications to D2D communications
- Type III: interference from D2D communications to D2D communications

The interference management between M-UE communications and D2D communications is the key issue in underlay D2D communications.

2.3.3 IBFD-Based Spectrum Sensing

In traditional wireless communications, a node is allowed to receive-transmit signals in orthogonal time slots (i.e., half-duplex mode) or simultaneously receive-transmit signals in orthogonal frequency bands (i.e., out-of-band full-duplex mode) due to implementation simplicity. Hence, two-time slots or two frequency bands are needed to accomplish a reception and a transmission. Both half-duplex mode and out-of-band full-duplex mode are inefficient in terms of spectrum efficiency. To deal with the issue, in-band full-duplex (IBFD) communication is proposed [31].

The basic idea of IBFD communication is to enable a node to receive-transmit signals on the same spectrum band simultaneously. As shown in Figure 9, a node may receive-transmit the same signal when the node acts as an IBFD relay, and a node may also receive-transmit different signals when the node is an IBFD transceiver (both the M-BS and M-UE are IBFD transceivers in Figure 9b). Thus, an IBFD node needs only one timeslot or frequency band to accomplish a reception and a transmission. Compared to the half-duplex and out-of-band full-duplex communications, IBFD communication theoretically doubles the spectrum efficiency. However, the transmission of the IBFD node may cause severe self-interference in its reception. In particular, self-interference is generated at the IBFD relay in Figure 9a and is generated at each IBFD transceiver in Figure 9b. This brings about theoretical improvement of spectrum efficiency. Thus, the bottleneck of the IBFD communication is the self-interference cancellation technique.

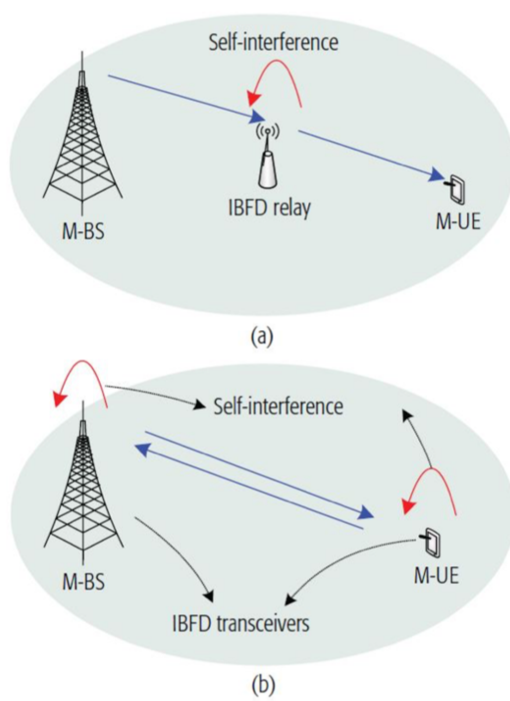


Figure 9: Two typical scenarios of IBFD communications: a) receive and transmit the same signal simultaneously; b) receive and transmit different signals simultaneously [1].

2.3.4 NOMA-Based Spectrum Sharing

In the previous generations of mobile communications, time/frequency/code domains were divided into orthogonal channels for orthogonal multiple access (OMA). OMA can avoid co-channel interference of multiple transmissions and reduce the processing complexity at receivers. However, OMA degrades the spectrum efficiency due to the exclusive channel occupation. To deal with the issue, NOMA is proposed and has been extensively studied in recent years. Different from OMA, NOMA allows the BS to schedule multiple users on a single channel at the same time and enhances the spectrum efficiency [32]. Although NOMA is generally a multiple-access technique, we can also regard it as a spectrum-sharing technique from the aspect that NOMA allows multiple transmissions on the same spectrum band at the same time and provides spectrum-sharing gains.

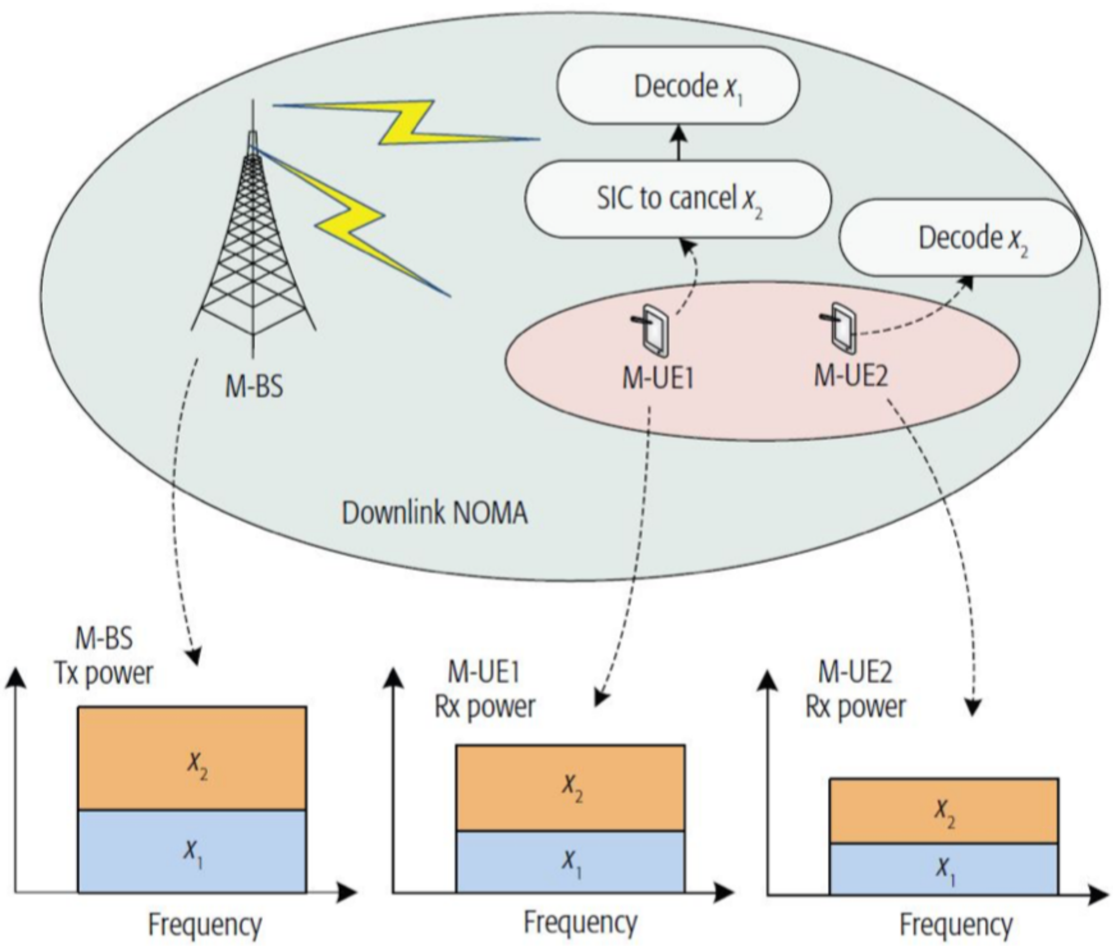


Figure 10: The principle of a two-user downlink basic NOMA transmission [1].

The key idea of the basic NOMA scheme is to exploit the power domain instead of time/frequency/code domains. In particular, the transmissions in the basic NOMA scheme are allowed to share the same time/frequency/code but with different power levels. Although different transmission signals may cause co-channel interference with each other at receivers in the basic NOMA system, the receivers can mitigate the co-channel interference and extract desired components from the received signals with the successive interference cancellation (SIC) technique. In this way, the basic NOMA scheme enhances the spectrum efficiency compared with the OMA. The principle of a two-user downlink basic NOMA transmission is shown in Figure 10.

2.4 Machine Learning for Smarter Spectrum Access and Interference Control

Machine Learning (ML) is a subset of artificial intelligence (AI) that enables systems to learn patterns from data and make decisions without explicit programming. ML models improve their accuracy over time as they are exposed to more data, making them particularly useful for complex, data-driven applications. In the field of wireless communications, ML has emerged as a powerful tool for solving problems related to spectrum sharing, interference management, and dynamic resource allocation. Traditional spectrum management approaches rely on predefined models and static policies, which often fail to adapt to the ever-changing wireless environment. ML, by contrast, enables adaptive decision-making, allowing networks to optimise spectrum usage efficiently while minimising interference.

ML has become a pivotal tool in enhancing spectrum sharing in wireless communications. By enabling systems to learn from data and make adaptive decisions, ML addresses the complexities of dynamic spectrum environments, leading to more efficient utilisation of available frequencies. In the realm of spectrum sensing, ML algorithms have been employed to improve the detection of available channels. For instance, deep learning models can analyse spectral data to identify underutilised frequencies, thereby facilitating more efficient spectrum sharing among users [33]. Regarding dynamic spectrum access, reinforcement learning techniques have been applied to enable systems to make real-time decisions about spectrum utilisation. These methods allow devices to adapt their transmission strategies based on the current spectrum environment, leading to improved coexistence and reduced interference [34]. In terms of interference management, ML has been utilised to predict and mitigate interference among users. By analysing patterns in spectrum usage, ML models can forecast potential interference scenarios and adjust transmission parameters proactively to maintain communication quality [35]. Furthermore, ML has been explored in the context of federated learning for spectrum sharing, where multiple devices collaboratively learn a shared model while keeping their data localised. This approach enhances privacy and reduces the need for centralised data collection, which is beneficial in dynamic spectrum environments [36]. The incorporation of ML into WP2 not only reduces computational overhead but also enables real-time adaptability, ensuring that spectrum resources are allocated dynamically based on current demand and environmental conditions.

3. Simulation and Modelling Framework

3.1 Simulation Approach and Tools Used

Work Package 2 (WP2) of the Spectrum Sandbox ITT Project employs a MATLAB-based simulation framework to analyse and optimise dynamic spectrum-sharing solutions. The simulation framework integrates three core components: communication channel models, ray tracing techniques, and machine learning-driven interference prediction as shown in Figure 11. These elements work together to provide a realistic and comprehensive evaluation of wireless network performance under different deployment scenarios. Though machine learning is part of the next step moving forward beyond the project, it has been used preliminary (the prototype and internal working highlighted in §5) to showcase how it can be integrated into the system to predict realistic path loss under different deployment scenarios.

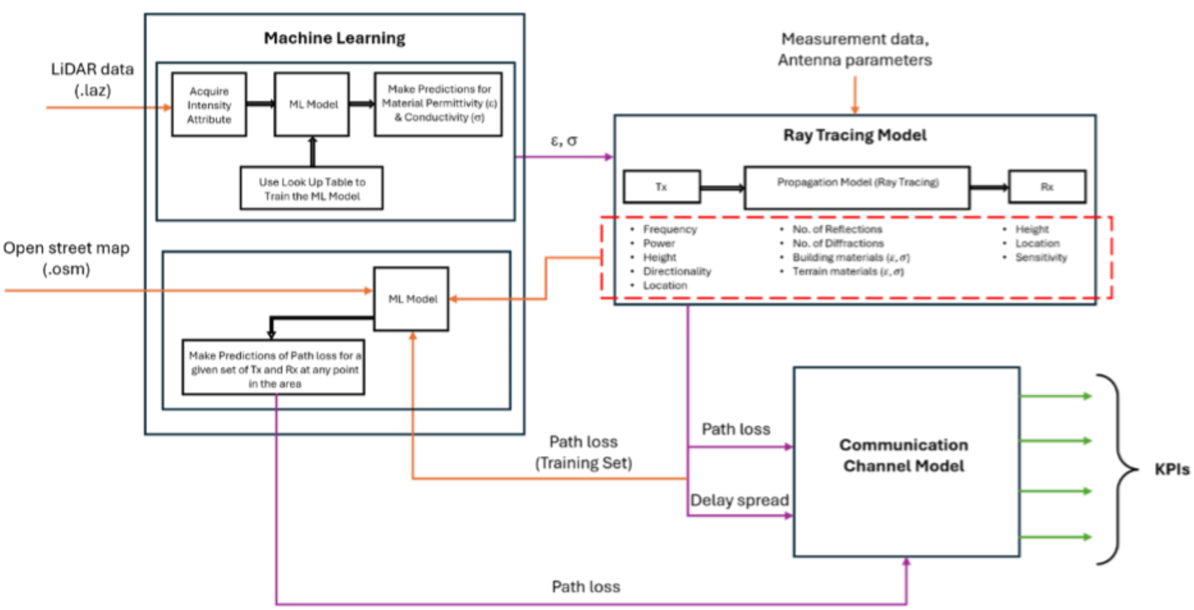


Figure 11: Flow diagram between different components namely machine learning, ray tracing and communication channel model.

3.1.1 Choice of MATLAB as the Simulation Platform

MATLAB is selected as the primary simulation tool due to its robust computational capabilities and seamless integration with specialised toolboxes such as:

- Antenna Toolbox, which facilitates the design, analysis, and visualisation of antennas, ensuring accurate radiation pattern modelling.
- 5G Toolbox, which enables simulation of 5G New Radio (NR) waveforms, resource allocation, and MIMO antenna configurations.
- RF Propagation Toolbox, which supports ray tracing-based path loss estimation, enabling detailed urban and rural propagation modelling.

MATLAB is also chosen for the following practical advantages:

1. Cost-effectiveness and Licensing: MATLAB is cost-effective for academic researchers and universities that have campus-wide licenses. In contrast, commercial software such as Remcom Wireless InSite, which specialises in ray tracing simulations, can be prohibitively expensive for non-industrial users.
2. Customisation and Flexibility: MATLAB allows users to define custom antenna patterns and propagation models, offering greater flexibility in wireless communication research. Many proprietary tools, including Remcom Wireless InSite, restrict customisation, making it difficult to tailor simulations to specific research requirements.

By leveraging MATLAB's built-in customisation, extensibility, and affordability, WP2 ensures that the simulation results are aligned with real-world network behaviour, improving the predictive accuracy of spectrum-sharing strategies.

3.1.2. Ray Tracing for Realistic Environment Modeling

The MATLAB ray tracing engine is employed to simulate:

- Urban, suburban, and rural settings, ensuring that spectrum-sharing solutions are tested across a variety of topographies.
- Fine-grained propagation effects, including multipath reflections, diffractions, and scattering, are essential for accurate interference prediction.

Ray tracing ensures that WP2's simulation framework closely replicates actual radio environments, thereby improving the reliability of spectrum-sharing strategies before field deployment.

3.1.3 Machine Learning for Interference Prediction and Path Loss Estimation

Machine learning (ML) technique integration into the simulation framework is part of the innovation work. ML primarily serves two purposes:

- Improve path loss prediction accuracy, reducing the reliance on predefined empirical models.
- Enhances building modelling and predicts material properties for more accurate environmental modelling.

The concept of ML integration in our simulation framework is further detailed in Section 5. The integration of ML can enhance the efficiency and scalability of WP2's modelling efforts, ensuring that simulations remain computationally feasible even for large-scale network deployments.

3.2 Ray Tracing for Signal Propagation Analysis

Ray tracing is a fundamental component of WP2's simulation approach, providing realistic modelling of wireless signal propagation. The ray tracing methodology aligns with 3GPP standards, ensuring compliance with industry benchmarks.

3.2.1 Frequency Bands and Deployment Scenarios

WP2 evaluates spectrum-sharing solutions across two critical frequency bands:

- Band 3 (1800 MHz): A widely used LTE frequency, essential for macrocell coverage and indoor penetration.
- Band n77 (3800–4200 MHz): A mid-band frequency, widely utilised for 5G deployments, offering a balance between capacity and coverage.

By analysing these bands, WP2 ensures that both existing and future spectrum-sharing solutions are assessed under realistic deployment conditions.

3.2.2 Key Propagation Effects Considered

Ray tracing models account for various real-world propagation effects, including:

- Path Loss: The reduction in signal strength as the wave propagates through space.
- Shadowing: Obstructions, such as buildings and trees, causing signal attenuation.
- Multipath Fading: The interaction of signals travelling along different paths, leading to constructive or destructive interference.

3.3 Communication Channel Model

The communication channel model incorporated in the simulation and modelling plan is designed to adhere to 3GPP industry standards, as shown in Figure 12, for spectrum sharing, ensuring an accurate representation of real-world signal propagation and interference dynamics. Figure 12 shows an illustration of a 5G network setup with potential interference abiding by the 3GPP industry standards for spectrum sharing.

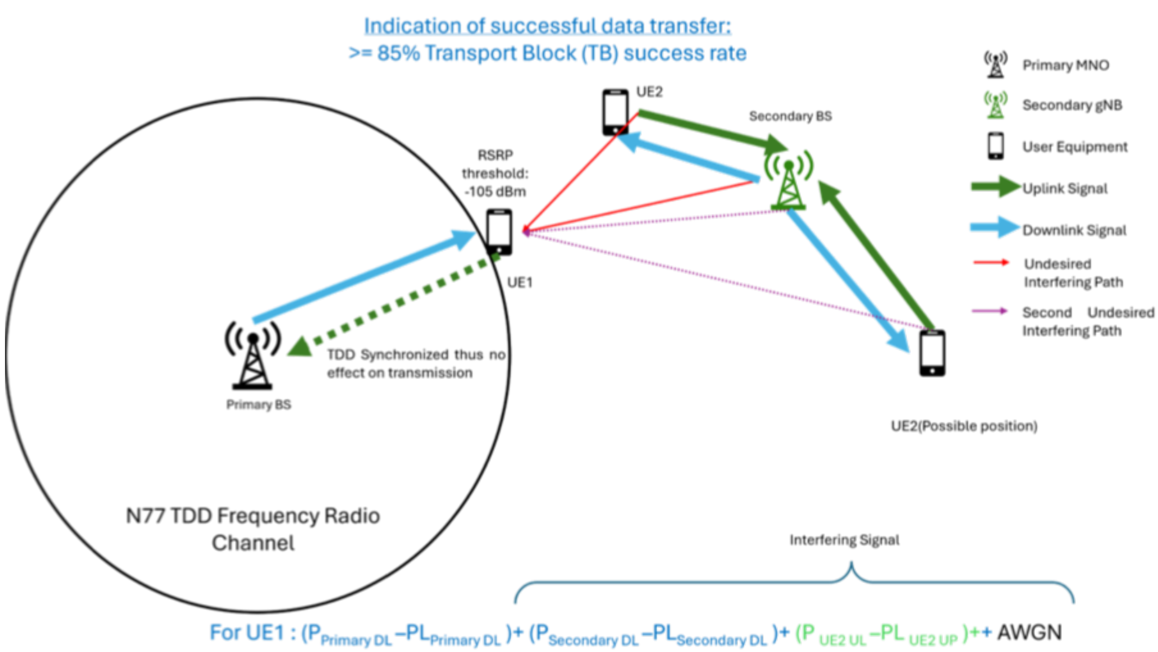


Figure 12: Illustration of a 5G network setup with potential interference abiding by the 3GPP industry standards for spectrum sharing.

4. Insights from WP2: Validating and Benchmarking Performance

4.1 Benchmarking Accuracy: Comparing Simulations with Real-World Data

To ensure the accuracy and reliability of the ray tracing-based simulation used in WP2, it is essential to compare its performance with findings from established studies in the literature. A study by Zakaria et al. presented real-world measurements in urban and suburban environments at 3.5 GHz as shown in Figure 13 [37]. Comparison with this study indicates that the WP2 ray tracing simulation model demonstrates similar performance trends with that of the real-world measurements. This comparison also highlights areas where refinements are needed. For example, while the simulation effectively models diffraction and reflection effects,

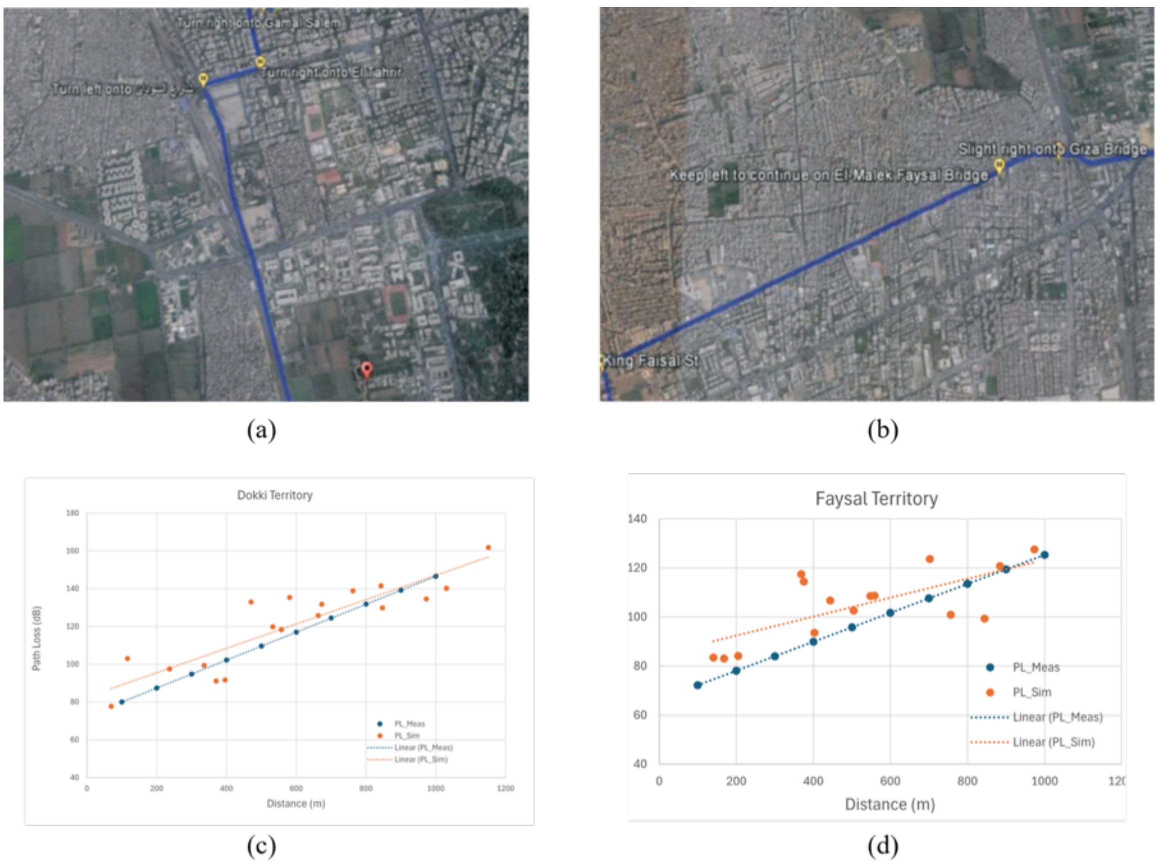


Figure 13: Validation of path loss simulation models against real-world measurements at 3.5 GHz [37]. (a) Indication of the route used for data collection in Dokki territory. (b) Indication of the route used for data collection in Faysal territory. (c) Comparison of measured and simulated path loss in Dokki territory as an example of an urban scenario. (d) Comparison of measured and simulated path loss in Faysal territory as an example of a suburban scenario.

differences in building material characterisation and transient obstacles such as moving vehicles and temporary structures contribute to observed deviations from real-world measurements.

Beyond regulatory benchmarking, WP2 simulations are cross-validated against real-world data collected in WP1, which includes field measurements of received signal strength and path loss from Bath test sites. By aligning the simulated propagation characteristics with empirical observations, discrepancies can be identified and corrected, thereby refining the predictive capabilities of the ray tracing and machine learning-driven models. This validation is particularly crucial for ensuring that the simulated interference predictions and spectrum-sharing strategies are representative of actual network conditions. The simulation setup configuration is tabulated in Table 3.

Table 3: MATLAB® simulation configuration

Transmitter Information	
Frequency	4000 MHz (n77 Band)
Transmitting power	5.01 W (37 dBm)
Antenna height	5 m
Directionality	Isotropic
Ray Tracing Propagation Model	
Number of reflections	2
Number of diffractions	1
Building material	Concrete
Terrain material	Concrete
Receiver Information	
Antenna height	1 m
Directionality	Isotropic

The comparison with WP1 data, as shown in Figure 14, provides insights into how accurately the model captures real-world attenuation, multipath effects, and shadowing in diverse environments. Figure 14 takes into account one gNB situated on Green Park Road (Latitude: 51.37933323, Longitude: -2.36481137) operating at 4080 MHz frequency (n77 Band). The data points shown in the Figure are the path loss obtained from User Equipment with the Network Signal Guru software installed and latched to this particular gNB.

The cumulative distribution function (cdf) of the path loss for the WP1 measurement data (black solid line) and WP2 Ray Tracing method (blue solid line) are shown in Figure 15. Initially, we utilised the in-built isotropic antenna for the Ray Tracing method in MATLAB. However, this is different from the antenna used in the Bath test bed. The actual antenna used in the test bed, its specification and the radiation pattern are shown in Figure 16. Since we were unable to obtain the precise radiation pattern of the antennas deployed in the Bath city centre from Telet, we simulated the antenna in CST such that it approximately matches the radiation pattern of the actual antenna. Figure 17 shows our simulated radiation pattern of the 4-port omnidirectional antenna. As can be seen between Figures 16 and 17, the simulated radiation may not be an exact match but is approximately comparable to the one used for measurements. This is then incorporated into the Ray Tracing MATLAB to obtain the respective path loss. This is shown as a blue dotted line in Figure 14. We find from Figure 15 that the Ray Tracing method applied on the OSM file downloaded directly from the OpenStreetMap [38] website has a large portion of samples with optimistic values of the path loss. To account for more comparable results, we have incorporated OSM files obtained after processing LiDAR data. The LiDAR processing is explained more in detail under §5.2. The results of the path loss obtained by operating the Ray Tracing method on the LiDAR-based OSM file lead to a more reasonable agreement between the simulation and the measurement with about 40% data on the pessimistic side and 60% on the optimistic side. The discrepancy between the simulation and the measurement can be due to several factors: (a) dynamic real-world environment including moving objects such as vehicles, pedestrians causing shadowing, diffraction and time-varying multipath effects, (b) weather and atmospheric effects, (c) Tx power and hardware variability, (d) incomplete or simplified building and terrain representation.

Figure 14: Path loss data around the Bath area using Telet gNBs. (a) WP1: field measurement, (b) WP2: Ray Tracing using OSM downloaded directly from OpenStreetMap [38] with in-built MATLAB isotropic antenna, (c) WP2: Ray Tracing using OSM downloaded directly from OpenStreetMap [38] with CST based 4-port omnidirectional antenna, (d) WP2: Ray Tracing using OSM obtained after processing on LiDAR data with in-built MATLAB isotropic antenna, (e) WP2: Ray Tracing using OSM obtained after ML processing on LiDAR data with CST based 4-port omnidirectional antenna.

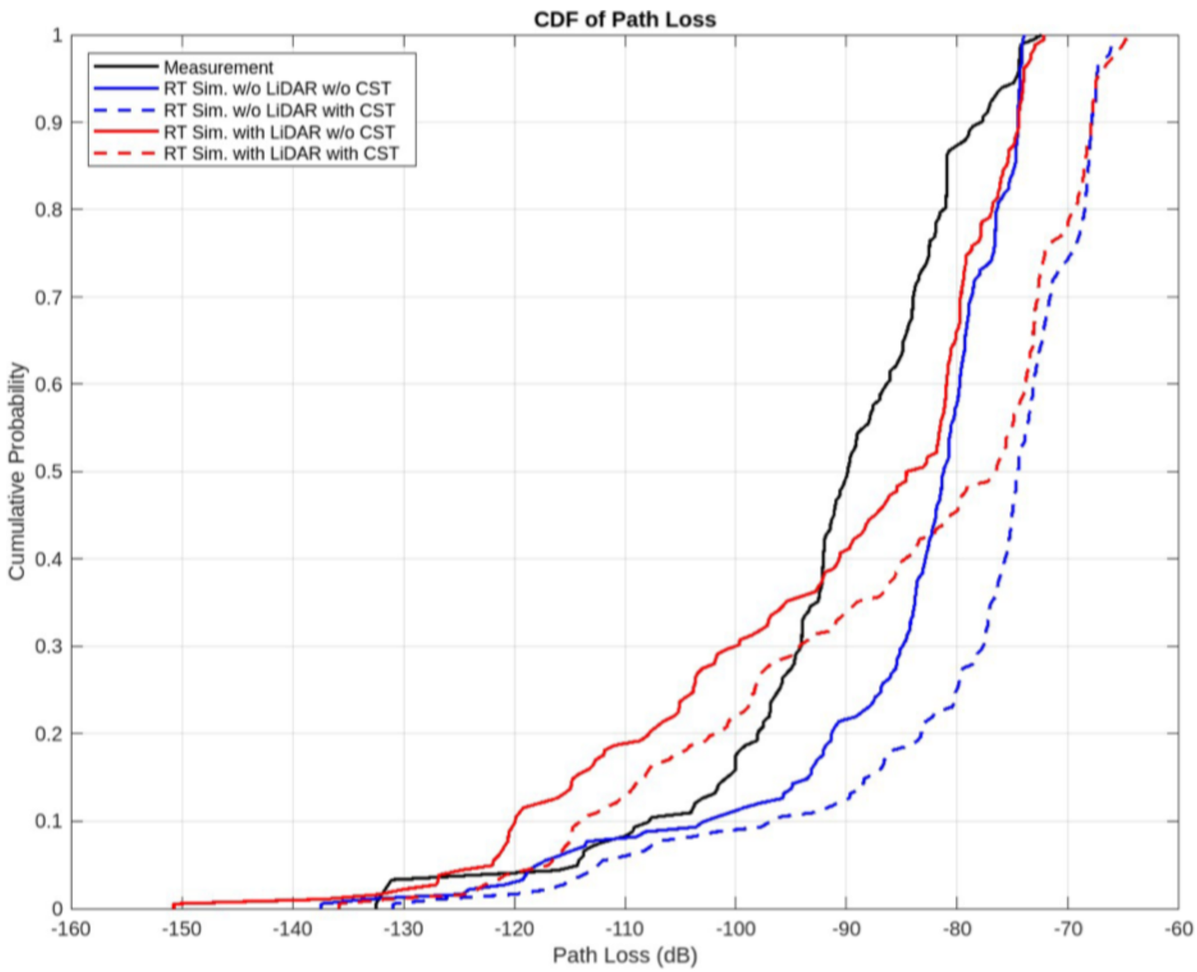


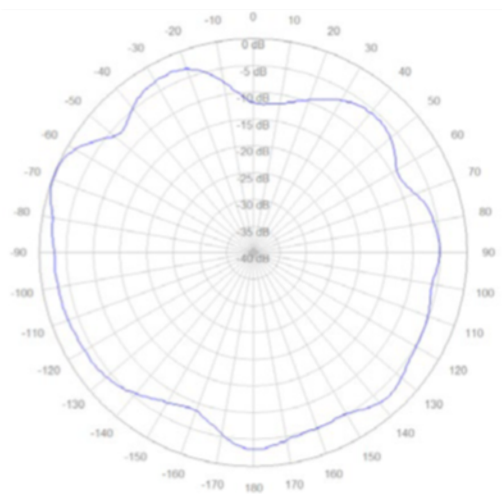
Figure 15: Evaluating the performance of Ray Tracing simulations for path loss prediction in Bath City Center by comparing CDF curves with measured data, with and without the inclusion of LiDAR data, and antenna from CST Studio Suite.

(a)

(b)

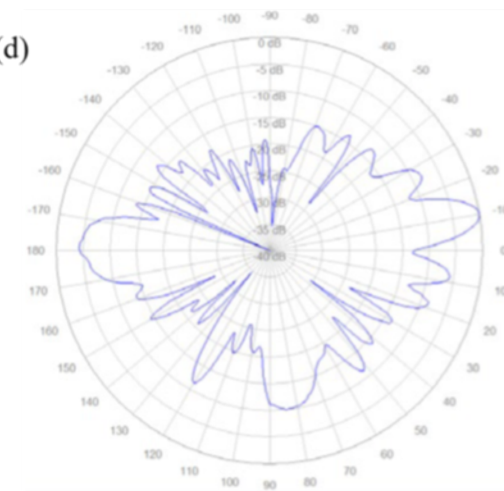
Electrical Specifications	
Frequency Range(MHz)	3800-4200
Gain(dBi)	10.0
VSWR	≤1.8
Polarization	Vertical
Horizontal Beam width	360°
Vertical Beam width	9±1°
Isolation(dB)	≥20
Input Impedance(Ω)	50
Max. Input Power(W)	100
Lightning Protection	DC Ground
Mechanical Specifications	
Connector	4xN Female
Dimensions(mm)	Φ165x530
Raadome Material	UV-PVC
Operating Temperature(°C)	-40 to +60
Mounting	Pole Mount

(c)



H

(d)



V

Figure 16: (a) 4-port omnidirectional antenna used in Bath City Center test bed, (b) specifications of the antenna, (c) H-plane radiation pattern, (d) V-plane radiation pattern.

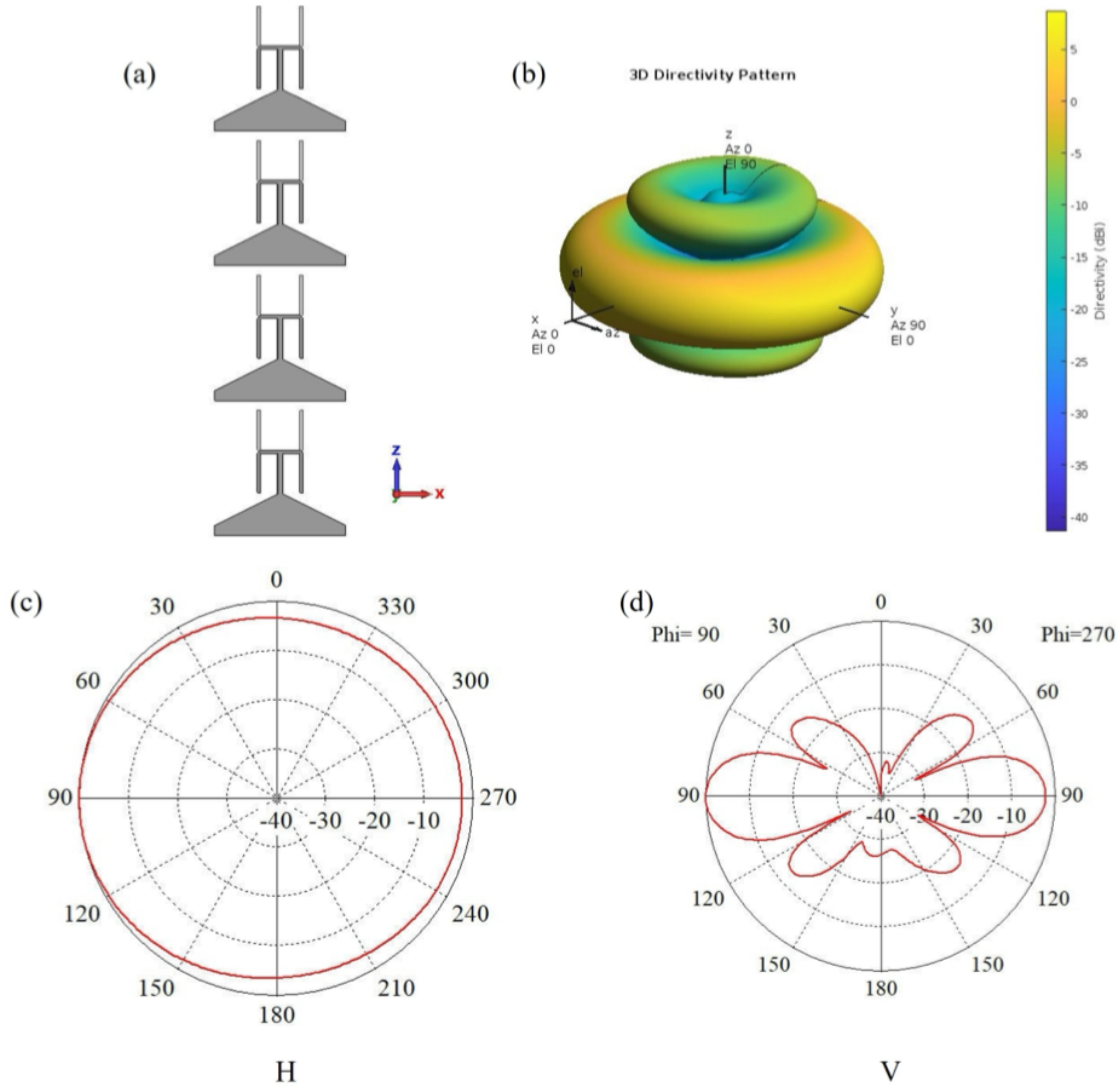


Figure 17: (a) 4-port omnidirectional antenna design in CST, (b) Simulated 3D radiation pattern of the 4-port omnidirectional antenna in CST Studio Suite, (c) H-plane radiation pattern of the simulated antenna, (d) V-plane radiation pattern of the simulated antenna.

4.1.1 Testing in Rural and Urban Environments: Case Studies from Chalke Valley & Liverpool

The Ray Tracing method was further applied to additional locations, including Chalke Valley and Liverpool, to represent rural and urban environments, respectively, at Band 3 (1.8 GHz) and n77 (4 GHz) frequencies.

Figure 18: Coverage plot in the region of Chalke Valley (a rural scenario) for an area of 200 m radius.

Chalke Valley

The Chalke Valley is a 13-mile valley in Wiltshire and Dorset, England. It stretches out along twisting lanes from Salisbury to Shaftesbury through quintessentially English villages and hamlets, against a backdrop of the rich, lush landscapes of the rivers Ebble and Chalke [39]. This is a classic example of a rural scenario. As part of the Cranborne Chase Area of Outstanding Natural Beauty, the valley boasts of patchwork farmland, chalk downland, rolling hills, rich green water meadows and sprawling bluebell woods. The location features low-rise buildings spread all over the place in small clusters with the majority being green coverage. For modelling purposes, we set the parameters the same as shown in Table 3 but with the exception that the "Terrain Material" is now chosen as "vegetation". The coverage map for Band 3 and n77 can be found in Figure 18. Additionally, the cumulative distribution function (CDF) of path loss for this area is also illustrated in Figure 20.

Liverpool

Liverpool is a Cathedral and port city in Merseyside, England, situated on the eastern side of the Mersey Estuary, near the Irish Sea. Liverpool is the fifth largest city in the United Kingdom with one of the most densely populated areas in England with a population of over 1.5 million. The city is heavily developed with residential, commercial, and industrial buildings, extensive transport networks, and public amenities. All these contribute to Liverpool being classified as an urban area. As a result, the MATLAB simulation setup has been kept identical to the one used in Bath and shown in Table 3.

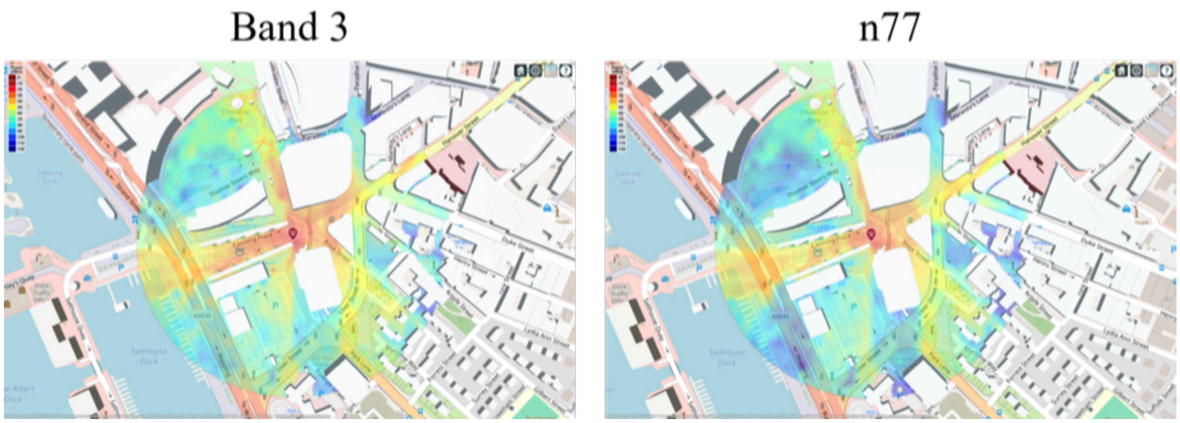


Figure 19: Coverage plot in the region of Liverpool (an urban scenario) for an area of 200 m radius.

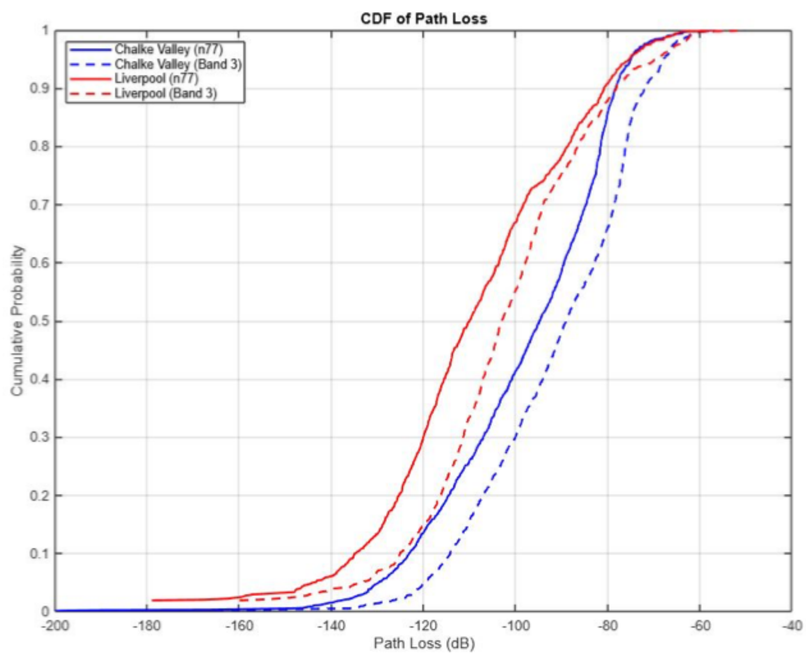


Figure 20: Coverage plot in the region of Chalke Valley (a rural scenario) and Liverpool (an urban scenario) for an area of 200 m radius.

The resulting coverage map in this area is presented in Figure 19. Again, the cumulative distribution function (CDF) of path loss for this area is shown in Figure 20. As observed between Figure 18 and Figure 19, path loss is notably higher in Liverpool compared to Chalke Valley, which aligns with the well-established trend of increased path loss in urban environments due to higher building density and multipath propagation. Furthermore, the CDF curve for the n77 frequency (4 GHz) is consistently shifted to the left compared to Band 3 (1.8 GHz), confirming the expected increase in path loss at higher frequencies.

4.2 Bridging the Gap: Understanding Discrepancies and Material Effects

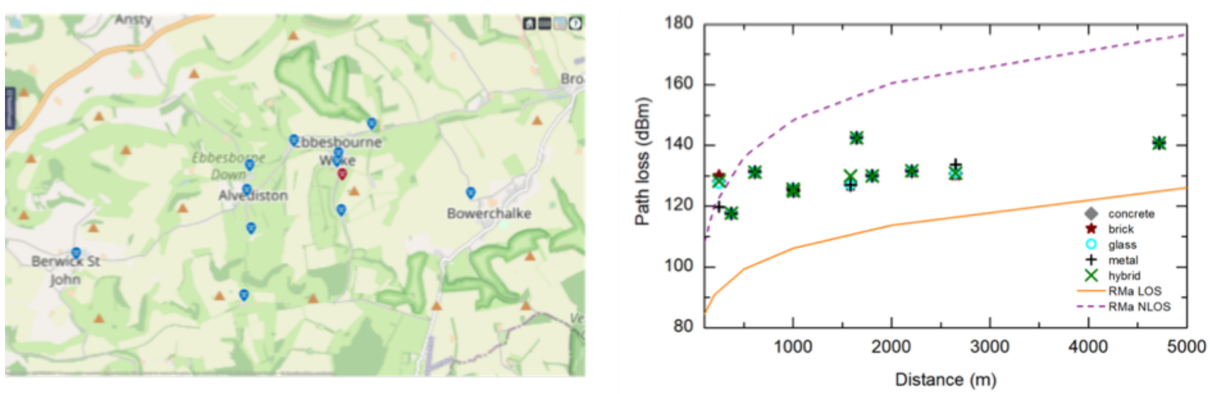
To ensure a fair and meaningful comparison, the simulation settings were carefully chosen to align with the real-world measurement setup, maintaining consistency with actual deployment conditions. The simulation was conducted for Band n77 (3.8–4.2 GHz), which is one of the key frequency bands designated for testing within the QMUL Spectrum Sandbox. The environmental models used in the simulation were designed to replicate urban landscapes, incorporating structural and terrain materials predominantly composed of concrete, a common material in densely built environments. To enhance the realism of the propagation model, key effects such as diffraction, reflection, and shadowing were accounted for, ensuring that the impact of different surface materials on signal attenuation was accurately represented.

Despite these efforts, discrepancies between the simulated results and real-world WP1 measurement data were observed. These differences arise due to several factors, primarily environmental uncertainties, material property assumptions, and dynamic interference effects. Real-world measurement campaigns inherently include transient obstacles, such as moving vehicles, pedestrians, and temporary structures, which introduce signal variations that are difficult to fully capture in static 3D simulation models. As a result, localised differences in received signal strength between measured and simulated data were observed, particularly in areas where frequent obstructions were present.

Another potential source of discrepancy is the generalisation of material properties within the ray tracing model. While the simulation employs ITU-R standardised material coefficients for buildings and surfaces, real-world structures exhibit variations in thickness, surface roughness, and material composition, leading to deviations in how signals interact with them. For example, real-world buildings may consist of a mix of glass, metal reinforcements, and varying concrete densities, affecting signal propagation in ways that standardised models may not fully capture. These material-dependent losses are particularly evident at higher frequencies, where reflections, diffractions, and absorptions are more sensitive to the specific characteristics of the medium.

(a) (b)

Figure 21: Customising building materials using MATLAB in Ray Tracing. (a) The region of interest (ROI) is indicated in Pink. The inset shows the buildings within the ROI on a new map. (b) Buildings identified as 1, 2, and 3 are customised with materials glass (indicated by blue), brick (indicated by red) and metal (indicated by black), respectively. All the other buildings are customised with concrete material (indicated by light grey).



(a) (b)

Figure 22: Ray Tracing path loss predictions for various building material (concrete, brick, glass, metal, hybrid) types in a rural scenario (Chalke Valley) along with Rural Macro (RMa) line of sight (LOS) model and RMa non-line of sight (NLOS) model.

Moreover, limitations in antenna configuration and transmitter power settings contribute to differences between the simulated and measured plots. In the real-world WP1 measurements, details regarding the dynamic transmitter power were unknown, requiring assumptions to be made in the simulation. Consequently, a static antenna transmitter power and an omnidirectional CW antenna were used in the simulation setup to approximate the measurement conditions. Additionally, unknown factors such as interference from neighbouring wireless systems, atmospheric effects, and hardware variations in the measurement equipment may have further influenced the real-world results, leading to deviations from the simulated predictions.

With a more precise understanding of the environment and building materials, more accurate ray tracing-based propagation models can be developed. How this improved knowledge can be obtained is further discussed in §6: Innovations and Future Directions as a flow diagram utilising more accurate LiDAR data to extract building and terrain material information. A small example of how this refined knowledge can be applied in ray tracing simulations is provided in Figure 21, illustrating how enhanced modelling techniques can reduce discrepancies between simulation and reality. Additionally, the effect of different building materials on path loss has been depicted in Figure 22, highlighting the significant role material properties play in propagation modelling.

4.3 Interference Dynamics and Performance Evaluation For 5G Network in Band 3 and n77

4.3.1 Band n77 gNodeB threshold simulation

Based on the 5G NR downlink physical layer processing chain shown in Figure 6, we explored the relationship between the transmitting power of a 5G gNodeB (gNB) and the resulting Transport Success Rate (TBSR) between the UE and the MNO (Mobile Network Operator) with a specific reference point of -70 dBm received power at the UE in each of the modulation and coding schemes (MCS). This setup is illustrated in Figure 23 and the results are presented in Figure 24. The different lines illustrate the impact of the MCS on the TBSR performance.

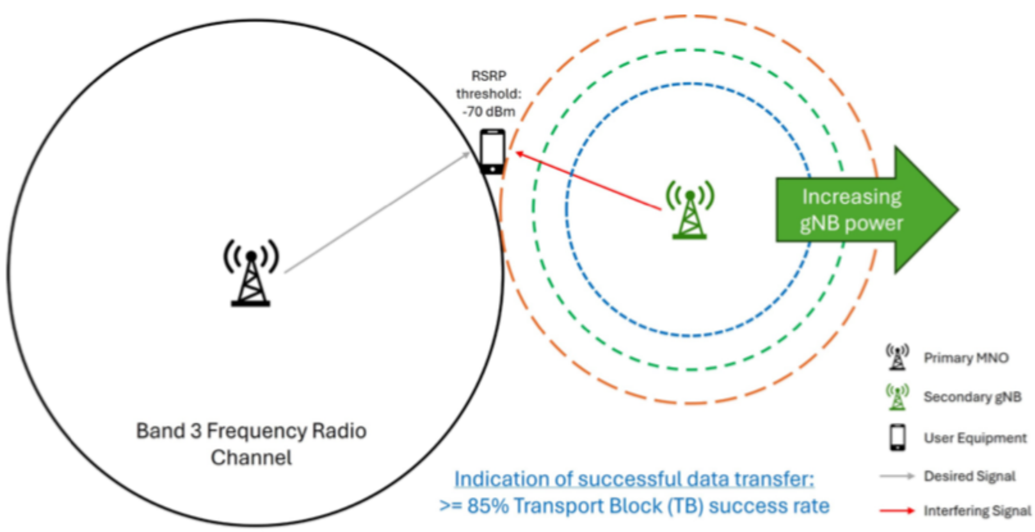


Figure 23: Illustration of how increasing secondary gNB transmit power affects the TBSR from primary MNO to the UE in a 5G network.

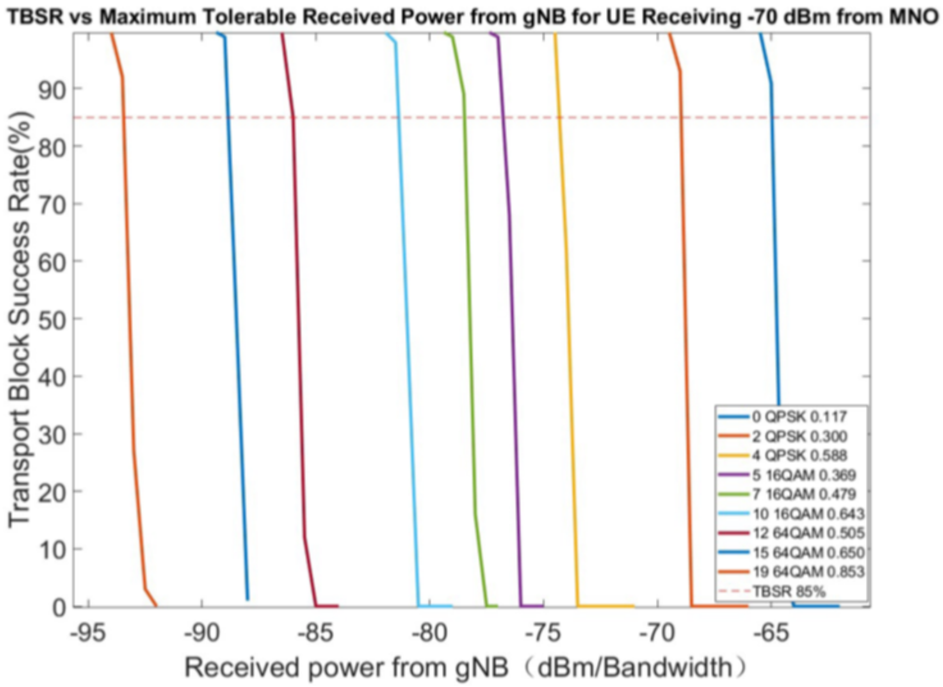


Figure 24: Transport Block Success Rate (TBSR) vs. maximum tolerable received power from a gNB for a user equipment (UE) receiving -70 dBm from an MNO.

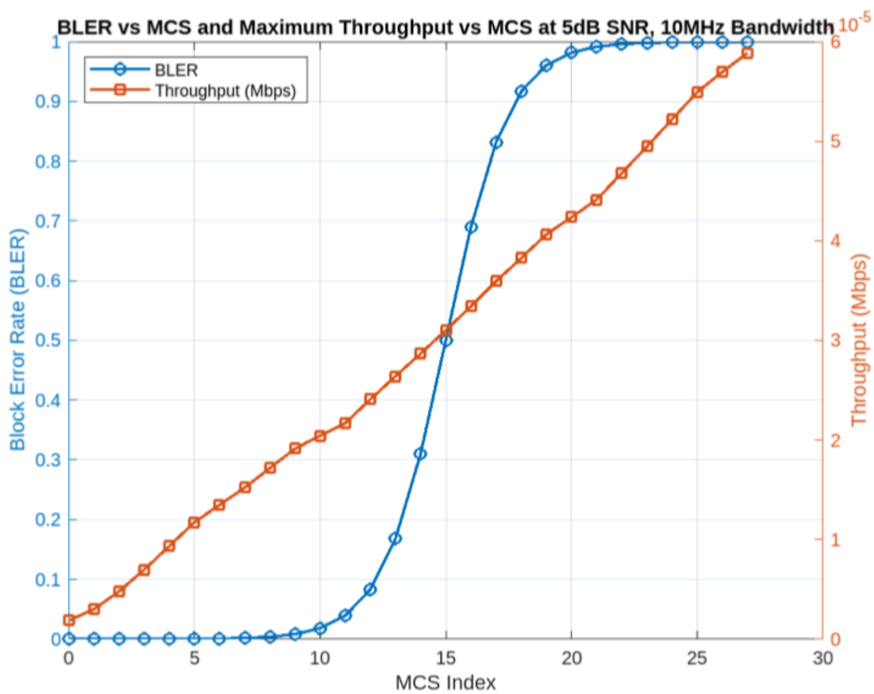


Figure 25: BLER and Throughput vs. MCS at SINR = 5 dB, 10 MHz BW

Table 4: Impact of gNB received power on 5 MHz and 10 MHz with MNO received power of -100 dBm per BW.

MCS Index	Code Rate	Modulation Method	SNR (dB) on 5 MHz	SNR (dB) on 10 MHz
1	193/1024	QPSK	0.7	-0.49
4	602/1024	QPSK	0.73	0.4
7	490/1024	16QAM	5.09	2.95
10	658/1024	16QAM	6.08	3.55
13	567/1024	64QAM	6.88	4.02
			gNB Maximum power (dBm)	gNB Maximum power (dBm)
1	193/1024	QPSK	-101.72	-101.27
4	602/1024	QPSK	-101.77	-102.71
7	490/1024	16QAM	-108.84	-108.67
10	658/1024	16QAM	-111.69	-111.76
13	567/1024	64QAM	-115.84	-116.96

TBSR and the MCS are closely related performance metrics in wireless communication systems, particularly in 5G and LTE networks. The TBSR represents the percentage of successfully received transport blocks (TBs) without errors after decoding, whereas the MCS defines the modulation order and coding rate, determining how much data is transmitted within a transport block. MCS is a key parameter in link adaptation, dynamically adjusting the modulation and coding based on channel conditions to optimise data transmission. The relationship between MCS with Block Errors Rate (BLER) and throughput can be described as follows in Figure 25:

i. Higher MCS values (higher-order modulation and lower coding redundancy): In good channel conditions, where the SNR and channel quality indicator (CQI) are high, the system selects a higher MCS (e.g., 256-QAM with a high coding rate). This increases spectral efficiency, allowing more bits per symbol, but also makes the transmission more vulnerable to errors. If channel conditions degrade due to interference or fading, TBSR drops as BLER increases, leading to retransmissions and lower throughput.

ii. Lower MCS values (lower-order modulation and higher coding redundancy): In poor channel conditions, a lower MCS (e.g., QPSK or 16-QAM with a lower coding rate) is selected to improve transmission reliability. The lower MCS increases coding redundancy, making the transmission more robust against noise and interference. As a result, TBSR remains high since BLER is low, but at the cost of reduced throughput due to a lower bit-per-symbol transmission rate.

In the context of cell edge simulations, the scenario where the received power from the MNO is -100 dBm must be considered. Typically, -105 dBm is used as the cell coverage limit, but with an Additive White Gaussian Noise (AWGN) spectral density of -174 dBm/Hz, the total noise power increases from -107 dBm to -104 dBm. Given that this noise level may exceed the received MNO power of -105 dBm, the simulation assumes a received MNO power of -100 dBm per unit bandwidth to ensure a more realistic assessment of performance at the cell edge.

The results presented in Table 4 demonstrate that increasing the bandwidth from 5 MHz to 10 MHz results in a reduction in SNR across all MCS levels. This degradation is particularly pronounced for higher-order modulation schemes, such as 16-QAM and 64-QAM. This trend aligns with theoretical expectations, as in a scenario with fixed total transmission power, doubling the bandwidth from 5 MHz to 10 MHz leads to a 3 dB increase in noise power, thereby reducing the available SNR.

Consequently, as the bandwidth increases, the receiver experiences a higher AWGN power, leading to an overall reduction in SNR. The impact of this SNR degradation varies across different MCS levels. For lower MCS levels (QPSK at MCS 1 and 4), the degradation is relatively minor because the gNBs signal remains the dominant factor influencing noise levels, allowing the receiver to maintain a more stable SNR. In contrast, for higher MCS levels (16-QAM and 64-QAM at MCS 7, 10, and 13), the drop in SNR is more significant, closely aligning with the theoretical 3 dB prediction. This is because, at higher MCS levels, AWGN becomes the dominant noise factor, making the signal more susceptible to degradation as bandwidth increases.

This trend highlights the increased sensitivity of higher-order modulations to noise power, reinforcing the importance of power control, error correction, and adaptive equalisation techniques in mitigating SNR losses in wideband 5G transmissions.

Overall, these results highlight the critical interplay between bandwidth, modulation scheme, and power control in 5G NR networks. The observed SNR degradation at higher bandwidths suggests that receivers must be equipped with advanced equalisation and interference mitigation techniques to compensate for the power-spreading effect. Additionally, the adaptive power control observed in the gNB transmission suggests that 5G systems intelligently regulate power levels to optimise performance across different bandwidth configurations. These findings reinforce the importance of dynamic modulation and power adaptation strategies to ensure consistent network reliability and spectral efficiency in varying bandwidth scenarios.

4.3.2 Band n77 TDD synchronisation simulation

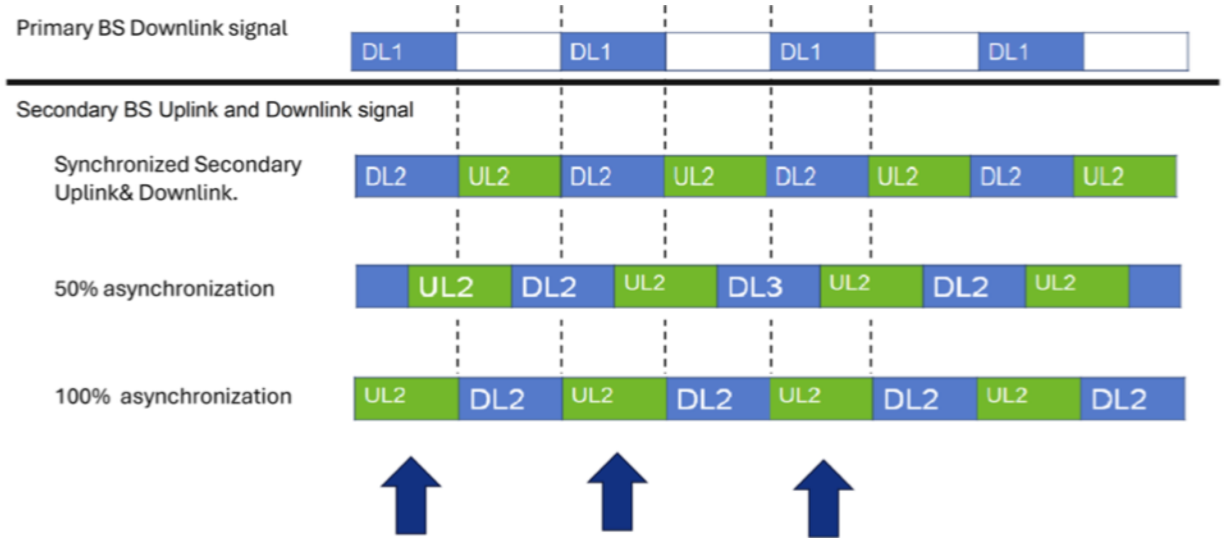


Figure 26: Illustration of Primary BS Downlink (DL1) and Secondary BS Uplink/Downlink (DL2/UL2) signal synchronisation, showing scenarios with perfect synchronisation, 50% asynchronisation, and 100% synchronisation.

We also looked at the scenarios where the different TDD synchronisation and asynchronisation of downlink (and uplink) from (and to) the secondary gNB can affect the performance of a primary MNO interacting with a UE. The scenario is illustrated in Figure 26.

First, the impact of the uplink transmission from the secondary gNB on the downlink performance of the primary MNO is analysed. This scenario, depicted in Figure 11, involves the primary BS communicating with UE1, while UE2 interacts with the secondary BS. Various DL:UL ratios are considered, and the resulting impact is illustrated in Figure 27, with threshold gNB received power values listed in Table 5. As observed in Figure 27, the uplink interference is minimal due to the frequency separation between DL and UL, where the downlink operates at a higher frequency.

A specific MCS, code rate and modulation scheme is considered, and the observed trends remain consistent across different configurations of these parameters.

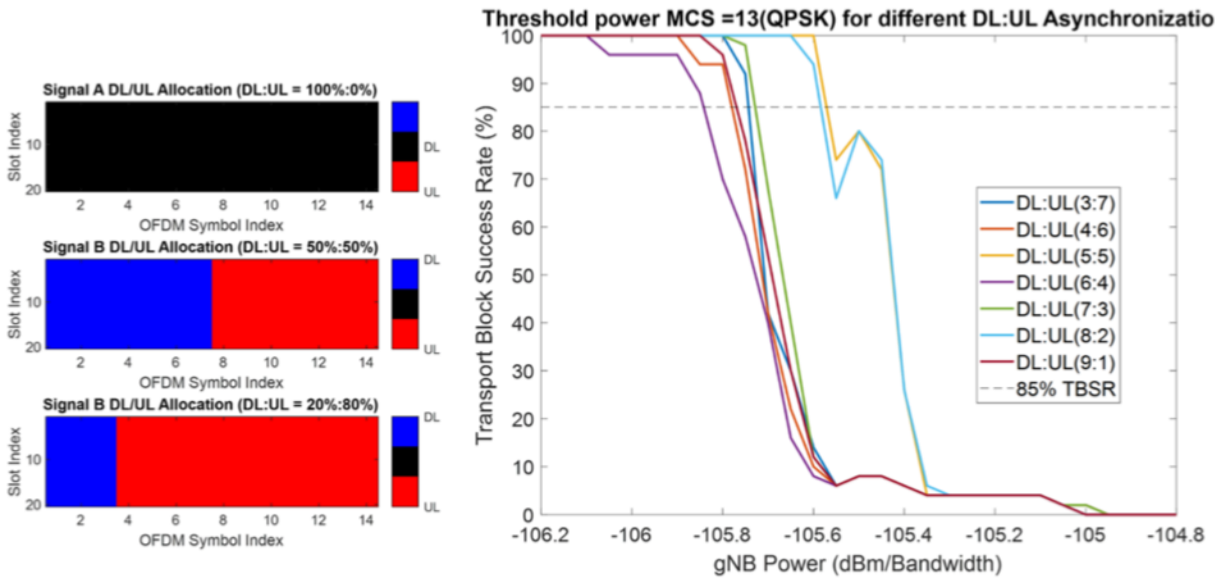


Figure 27: Impact of different UL/DL ratios on the TBSR performance as a function of gNB transmit power.

Table 5: Threshold gNB power in maintaining 85% TBSR from MNO for different downlink-uplink ratios.

DL: UL	MCS	Code Rate	Modulation	Max gNB Power (dBm)
3:7	13	526/1024	QPSK	-105.744
4:6	13	526/1024	QPSK	-105.78
5:5	13	526/1024	QPSK	-105.572
6:4	13	526/1024	QPSK	-105.842
7:3	13	526/1024	QPSK	-105.728
8:2	13	526/1024	QPSK	-105.585
9:1	13	526/1024	QPSK	-105.77

4.3.3 Band 3 cross Bandwidth simulation

Next, the impact of the downlink transmission from the secondary gNB on the downlink performance of the primary MNO is analysed. The centre frequency of operation is set at Band 3 1.8 GHz. We take the following two scenarios considering the subcarrier spacing (SCS) to be the same in both cases to understand the impact of the downlink:

- Same bandwidth (BW)
- Different BW

Table 6 presents a comparison of interference with different MNO and gNB bandwidth configurations. The evaluated scenarios include MNO bandwidths of 60 MHz alongside gNB bandwidths of 10 MHz, 20MHz, 30MHz, 40 MHz, 50 MHz, and 60 MHz. The increase in MNO bandwidth leads to a rise in AWGN power, which exceeds -105 dBm, making signal decoding infeasible in the simulation. To ensure that the gNB remains the dominant interference source, the MNO received power is fixed at -70 dBm, and the gNB received power is fixed at -81 dBm.

It can be observed that in each of the scenarios, the interference from gNBs decreases (i.e., MCS increases) as the gNB bandwidth increases. This is evident from the fact that MCS levels define the modulation order and coding rate, determining how much SNR is required for reliable decoding. Higher MCS values demand significantly higher SNR, meaning even small increases in interference from neighbouring transmitters disrupt the signal decoding process, leading to higher packet error rates. With increasing gNB bandwidth, the SNR for the MNO decreases, hence higher MCS can keep up with tolerable TBSR.

Table 6: Comparison of maximum tolerable gNB received power (dBm/BW) among three different combinations of MNO (40 MHz and 60MHz) and gNB (10 MHz- 60 MHz) bandwidth.

MNO 40MHz		MNO 60 MHz	
gNB Bandwidth	MCS	gNB Bandwidth	MCS
10 MHz	7	10 MHz	5
20MHz	7	20MHz	7
30MHz	8	30MHz	7
40MHz	9	40MHz	7
50MHz	10	50MHz	7
60MHz	10	60MHz	8

The results show that the more bandwidths overlap the less interference appears. The impact of interference is not only determined by total power but also by how that power is distributed across the spectrum. With same-bandwidth interference, the total interfering power is concentrated only within the same occupied frequency range as the MNO's signal, making it more manageable for equalisation and filtering techniques to compensate for. In contrast, when the interfering base station has a larger bandwidth, its power is spread across a wider frequency range, reducing the power spectral density (PSD) per Hz. The lower PSD per Hz reduces the total interference power within the overlapping range.

An interesting observation arises when the MNO and gNB operate the same bandwidth while operating at fixed power levels of -70 dBm for the MNO and -81 dBm for the gNB. The results indicate MCS is higher for a 40 MHz bandwidth than for a 60 MHz bandwidth. Specifically, when both the MNO and gNB operate at 40 MHz, a greater number of MCS levels are achievable compared to operations at 60 MHz.

This trend suggests that as the operation bandwidth increases, interference also increases, leading to higher SNR requirements to achieve the same MCS levels. Consequently, the ability to sustain high MCS levels is reduced at larger bandwidths, emphasising the trade-off between bandwidth expansion and interference management in co-channel operation.

4.3.4 Throughput coverage

A simulation of throughput coverage in Bath city centre was conducted, as illustrated in Figure 28. The figure provides a comparative visualisation of throughput (in Mbps) during a simulated handoff scenario, where a user transitions between the coverage areas of two base stations. The images labelled (a) and (b) correspond to different inter-site distances, offering insights into the optimal minimum separation required between the base stations to ensure a seamless handoff.

This analysis is crucial for optimising network planning and deployment in urban environments, where maintaining consistent connectivity is essential for user experience. A poorly optimised handoff region can lead to high packet loss, increased latency, and degraded Quality of Service (QoS), particularly for applications such as video streaming, online gaming, and real-time communications. By determining the ideal base station separation, network operators can minimise handoff failures, reduce call drops, and improve overall throughput, ultimately enhancing the efficiency and reliability of mobile communication networks.

By simulating the TBSR for each transmission, the throughput coverage map can be generated. The throughput coverage maps provide an insightful visualisation of the impact of interference and signal propagation on 5G NR network performance. The simulation results indicate that high throughput regions, represented by red and yellow areas, are primarily concentrated near the primary base station, where the received signal strength is high, and the impact of interference is minimal. These regions benefit from higher-order modulation schemes, such as 64-QAM or 256-QAM, which enable increased data rates. As the distance from the base station increases or interference levels rise, throughput degrades progressively, as seen in the transition from green to blue regions. The analysis highlights that interference from a secondary base station significantly influences throughput distribution. In scenarios where the interfering transmitter operates at a high-power level, a notable reduction in throughput is observed across the coverage area. The presence of dark blue and purple zones indicates areas with severe degradation in signal quality, likely due to high interference levels, substantial path loss, or the use of lower-order modulation schemes such as QPSK with a low coding rate. This suggests that interference mitigation strategies, such as power control and adaptive resource allocation, are crucial in maintaining network performance.

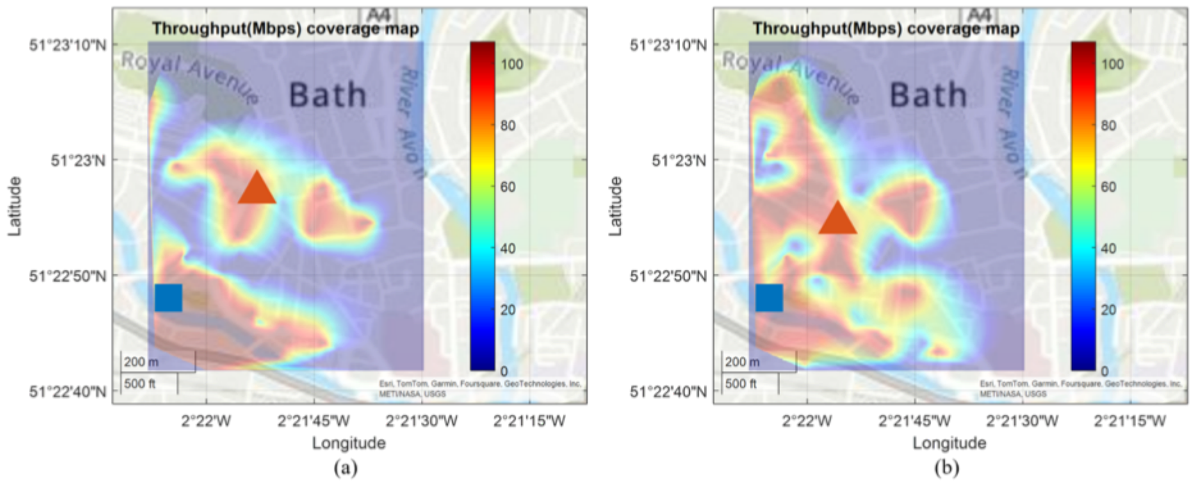


Figure 28: Visualisation of a change in throughput during the handoff scenario as a user moves from the coverage area of one base station (blue square) to another (orange triangle), when the distance between the two base stations is (a) 350 m and (b) 280 m.

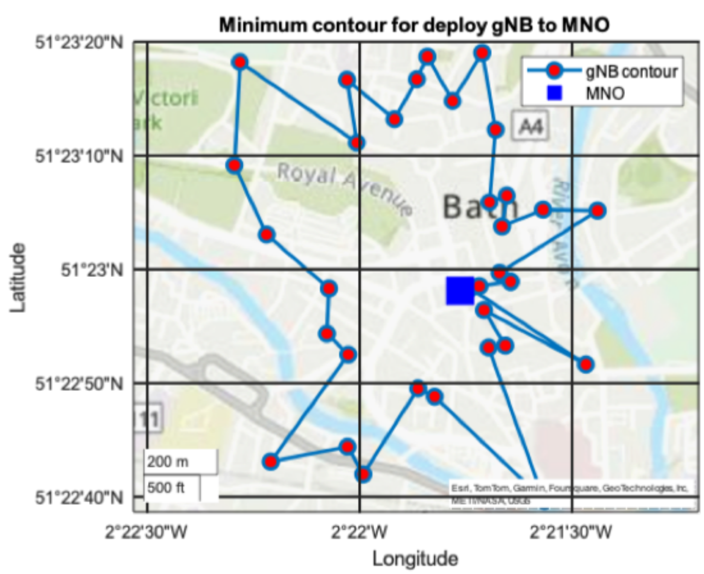


Figure 29: Visualisation of an area within which a second gNB cannot be deployed.

A key implication of Figure 29 is their relevance to network densification in 5G deployments. The results provide quantitative insights into how closely base stations can be placed while ensuring minimal interference. By adjusting transmission power and optimising resource allocation, spectral efficiency can be maximised while preserving an acceptable quality of service (QoS). The study demonstrates the necessity of balancing cell density with interference management to maintain high network throughput. Future work can extend this analysis by incorporating real-world propagation conditions and dynamic user mobility models to further refine network planning strategies.

4.4 Challenges, Limitations and Considerations

Despite the promising findings of WP2, several challenges must be addressed to further enhance the accuracy, efficiency, and applicability of the simulation and modelling framework. A key limitation arises from the computational complexity of high-fidelity ray tracing simulations, which require substantial processing power. The detailed modelling of electromagnetic wave propagation, including multipath effects, diffraction, scattering, and material interactions, significantly increases computational demands. Consequently, the number of simulated scenarios is inherently constrained by available hardware resources, necessitating a careful selection of representative environments to balance accuracy and feasibility. To overcome these computational limitations, we propose the integration of Machine Learning (ML) techniques as a complementary approach to traditional ray tracing. ML can be leveraged to optimise simulation efficiency, predict propagation characteristics, and enhance the fidelity of results without requiring exhaustive ray tracing calculations. The specific methodologies and implementation strategies for this integration are detailed in the next section, Innovations and Future Directions.

Another challenge stems from the limited material libraries available in MATLAB-based Ray Tracing. Currently, predefined databases for buildings, terrain, and surface materials are insufficient for capturing the full complexity of real-world environments. This limitation reduces the accuracy of simulated results, particularly in diverse urban and rural landscapes where material properties significantly impact signal propagation. Moreover, the lack of detailed environmental representation restricts the ability to model intricate electromagnetic interactions. To address this issue, Machine Learning can be employed to generate high-resolution environmental models, which can then be incorporated into the Ray Tracing framework. By utilising ML-based algorithms to infer and reconstruct missing material data, a more precise and realistic representation of physical environments can be achieved. This approach enables the creation of customised and highly detailed simulation scenarios, thereby improving the accuracy of signal propagation analysis and optimising network performance predictions.

Additionally, data integration from WP1 presents a significant challenge, as harmonising real-world measurement data with simulated models requires meticulous calibration and validation efforts. Field measurements collected in WP1 serve as a vital benchmark for assessing the accuracy of ray tracing and machine learning-based predictions. However, discrepancies between measured and simulated results often arise due to differences in environmental conditions, transient obstructions, and variations in hardware configurations. Addressing these inconsistencies requires iterative refinements, where simulation parameters such as material properties, antenna configurations, and interference sources are continuously adjusted to improve alignment with real-world observations. The process of fine-tuning these models demands a comprehensive approach, incorporating machine learning-based calibration techniques and enhanced geospatial data integration to bridge the gap between theory and practice.

5. Innovations and Future Directions

5.1 AI-Powered Path Loss Prediction: A Smarter Alternative to Traditional Models

The increasing complexity of urban, suburban, and rural environments necessitates a robust method for predicting wireless path loss. Given that ray tracing models take a long time to simulate, a robust and efficient solution is essential. In this work, we develop a machine learning pipeline that leverages simulation data generated by ray tracing models as well as real-world measurement data. The primary objective is to predict the path loss at any receiver point when the map of an area and the transmitter location are provided. Initially, the model is trained using simulated data from the Bath area and later validated on real measurement data. This pipeline is designed for scalability and can be extended to different geographies by adapting to various environmental characteristics.

5.1.1 Data Preparation and Feature Extraction

OpenStreetMap Data and Regional Division

The base map data is sourced from OpenStreetMap (OSM). The Bath area is divided into several key regions—Bath City Centre, Bath South, Bath North, Bath Southeast, and Bath West—to capture a diverse range of building geometries and environmental conditions. The input for our machine learning pipeline comprises both simulated data (from the ray tracing model) and features extracted from the OSM map.

Building Feature Extraction

For each building, the following features are derived:

- Height (h): Provided either directly in the map or set to a default value.
- Centroid Coordinates (C_{lat}, C_{long}): Calculated as the geometric centre of the building polygon.
- Area (A): The total area enclosed by the building footprint.
- Perimeter (P): The length of the building's boundary.
- Circularity (Φ): A shape descriptor.

These features are assembled into a feature vector for each building, which is normalised later.

Transmitter and Receiver Data: The ray tracing simulations provide pairs of transmitter and receiver coordinates along with the corresponding path loss measurements. The final data structure combines the Tx/Rx locations into an 8-dimensional input for subsequent regression tasks.

5.1.2 Model Architecture

Our approach employs a dual-model system that integrates a transformer-based module for encoding building features with a residual path loss encoder for regression as shown in Figure 30.

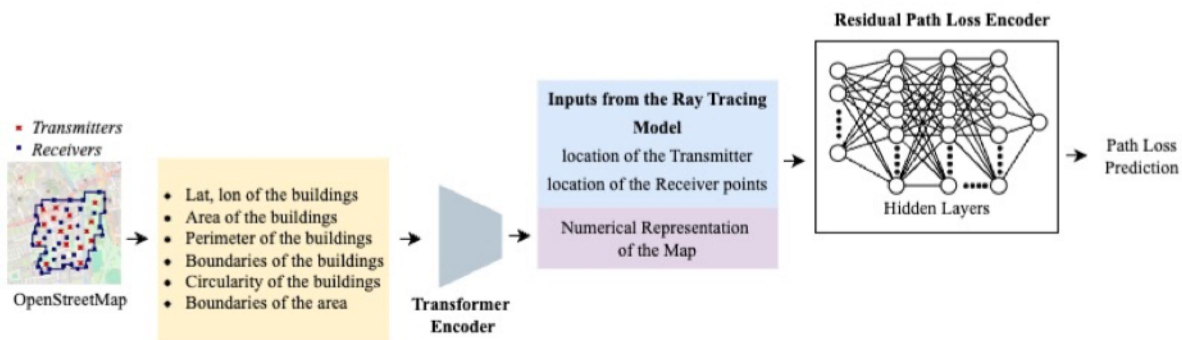


Figure 30: Introduced Machine Learning Based Dual Model System for predicting the Path Loss.

Transformer Pooling Model: **The Transformer Pooling Model** aggregates building features using a transformer encoder. Its key steps are as follows:

- Input Projection: The 6-dimensional building feature vector is linearly projected to a higher-dimensional hidden space.
- Layer Normalisation: Applied immediately after the projection to stabilise the training process.
- Transformer Encoder: Consists of multiple layers of self-attention and feed-forward networks. The attention mechanism is given by:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V$$

where $Q, K and V$ denote the query, key, and value matrices respectively, and d is the dimensionality of the key.

- Mean Pooling: After the transformer layers, the features are aggregated using a simple mean operation.
- Final Projection and Activation: The aggregated feature vector is passed through a fully connected layer and activated by ReLU.

Residual Path Loss Encoder: The second model in our pipeline is a Residual Path Loss Encoder. It accepts an 8-dimensional vector, which is a concatenation of the transformer-derived building embedding and the Tx/Rx location information, and outputs a scalar prediction for path loss. The residual blocks in this model follow the form:

$$y = ReLU\left(f(x)\right) + x$$

where $f(x)$ is a transformation performed by a fully connected layer (possibly with dropout) and the addition ensures that the original input x is retained through the block, enabling better gradient flow during training.

A series of such blocks are stacked with increasing and then decreasing dimensionality, culminating in a final linear layer that outputs the predicted path loss value. This structure helps the network learn complex, nonlinear relationships while preserving information from the original inputs.

Loss Function: To train the model, we use the Smooth L1 Loss defined as:

$$L\left(x, y\right) = \begin{cases} 0.5(x-y)^2, if\left|x-y\right| < 1 \\ \left|x-y\right| - 0.5, otherwise \end{cases}$$

This loss function is less sensitive to outliers than the mean squared error and provides a balance between L1 and L2 loss behaviours.

5.1.3 Fine-Tuning for Accuracy: Optimising the Learning Process

Data Loading and Batching

The training data is organised by map folders where each folder contains an OSM file and the corresponding simulation data in Excel format. A custom data loader is used to:

- Parse the OSM file to extract building features.
- Read and filter the simulation data.
- Concatenate and batch the data appropriately.

The data are grouped by map ID to compute building embeddings on the fly for each unique region.

Optimiser and Learning Rate Scheduling: The models are jointly trained using the AdamW optimiser. The learning rate is dynamically adjusted using an Exponential Learning Rate scheduler defined by: $\eta_{t+1} = \gamma \eta_t$

where η_t is the learning rate at iteration t and γ is a decay factor (typically close to 1, e.g., 0.999999). Gradient clipping is employed to prevent exploding gradients, with the norm of the gradients clipped to a maximum value.

5.1.4 Measuring Success: Evaluating the AI Model's Performance

For evaluating the model performance, one transmitter location in Bath City Centre is taken into consideration. These sites provide a controlled yet diverse set of conditions within a densely built urban environment, allowing for a thorough assessment of the model's prediction capabilities under real-world propagation scenarios. This evaluation framework not only facilitates direct comparison with simulated data but also helps in identifying specific areas for further model enhancement and calibration. The trained machine learning predicted the 182 receiver points path loss with an average of 8.745 dB accuracy. Furthermore, to illustrate this validation loss, a CDF plot of the path loss between the two models and the path loss field plot at the Bath City Centre are shown in Figure 31 and Figure 32, respectively. Table 7 shows specifically the superiority of ML over RT in terms of computational time, i.e., ML is 1000x faster than RT.

This is worth mentioning that the 1000x increase does not result from MATLAB inefficiency specifically (RT methods perform thousands of individual ray-path computations that can take several minutes or even hours for extensive regions, depending on their complexity. This is particularly true for upper 6GHz, where the wavelength is shorter, requiring higher resolution in the Ray-Tracing algorithm), but rather from the inherent computational complexity of the RT propagation method, which the ML model significantly simplifies once trained. In other words, the large speed increase comes from shifting away from physics-based calculations (traditionally slow and detailed) to predictive ML-based methods (fast and approximate). This advancement provides regulators like Ofcom and spectrum users with a powerful, rapid, yet still reliable way to manage spectrum sharing and interference scenarios, significantly enhancing efficiency and reducing the cost and time associated with spectrum management.

In addition to the substantial speed improvements demonstrated by our machine learning (ML) model, there is significant potential for extending this approach toward real-time spectrum management. Specifically, ML models could predict spectrum usage dynamically, adapting rapidly to changing network conditions and user demands. By accurately forecasting usage patterns and interference potential, ML would facilitate unprecedented levels of spectrum sharing, even among multiple MNOs and diverse spectrum users simultaneously.

Again, implementing ML-driven spectrum prediction can significantly reduce reliance on physical spectrum sensing devices. By accurately predicting spectrum occupancy based on historical usage, environmental conditions, and other context-specific data, ML models minimise the need for extensive real-time monitoring hardware, leading to lower infrastructure costs (through reducing hardware installation and operational costs), reduced operational complexity (through maintaining fewer devices), improved scalability and flexibility (through reducing the reliance on dense deployments of spectrum sensors), complementary sensing capability, enhanced responsiveness and reliability.

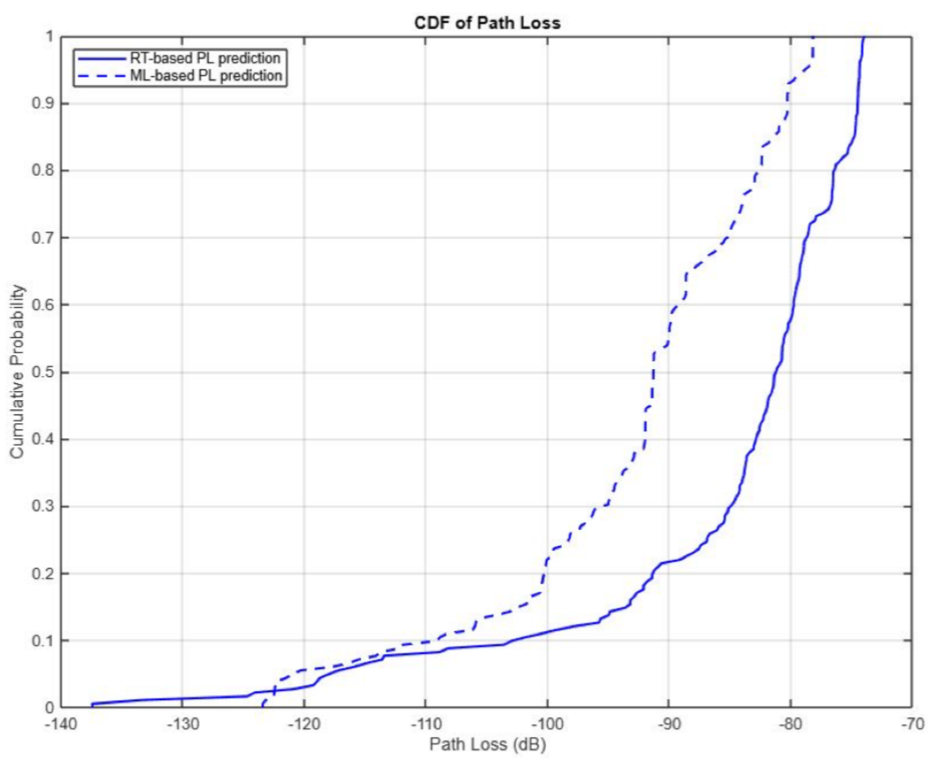
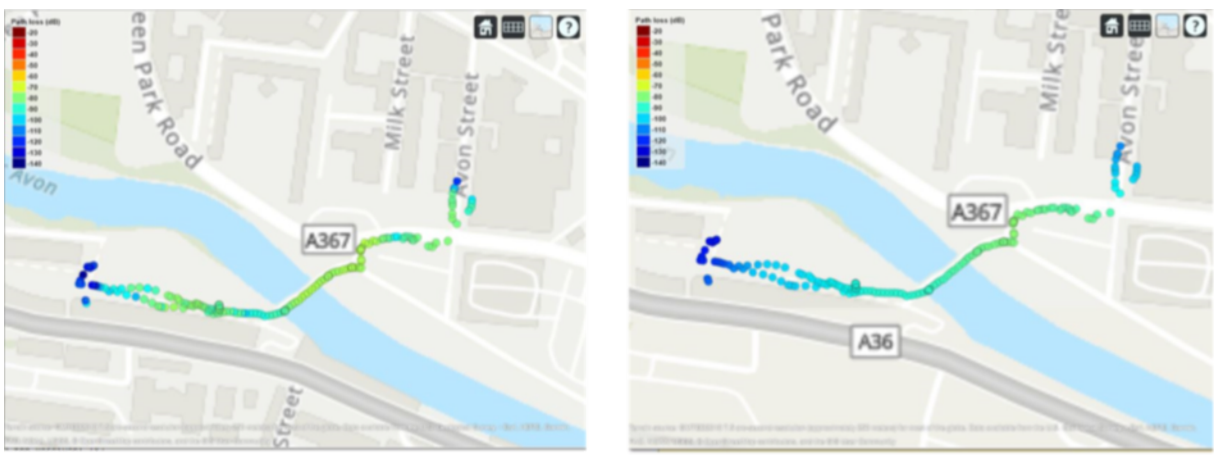


Figure 31: CDF comparison between Ray Tracing model and Machine Learning

Table 7: Comparison in the processing time between the MATLAB-based Ray Tracing model and Python-based Machine Learning model for predicting path loss between one Tx and one Rx.

	Programming Language	Hardware Processing Power (GHz)	Time for simulating one Tx and one Rx (sec)
Ray Tracing	MATLAB R2024b	2.90	~ 10
Machine Learning	Python 3.11.7	2.749	~ 0.01



(a) (b)

Figure 32: Path loss field plot in the Bath City Center using: (a) Ray Tracing method and (b) Machine Learning method.

5.1.5 Key Observations: Strengths and Weaknesses of AI-Based Predictions

Training Dynamics

During training, we observed that the integration of the transformer pooling model significantly improved the extraction of spatial features from the OSM data. The use of residual blocks in the path loss encoder allowed better information retention and gradient flow. In early epochs, the loss decreases rapidly as the model learns the basic mapping from inputs to outputs; however, fine-tuning in later epochs is crucial to handle the complex interactions between building geometry and signal propagation.

5.1.6 Refining the Model: Future Enhancements for Better Accuracy

Enhancements can be achieved by incorporating additional parameters from the OSM maps such as road networks, vegetation, and other urban features—to further bridge the gap between simulation data and real-world measurements. Moreover, increasing the diversity of the training dataset to include a wider range of scenarios from urban, suburban, and rural areas can significantly enhance the model's generalisability. Finally, integrating an active learning pipeline would enable the model to continuously update its parameters as new measurement data become available, thereby improving its accuracy over time.

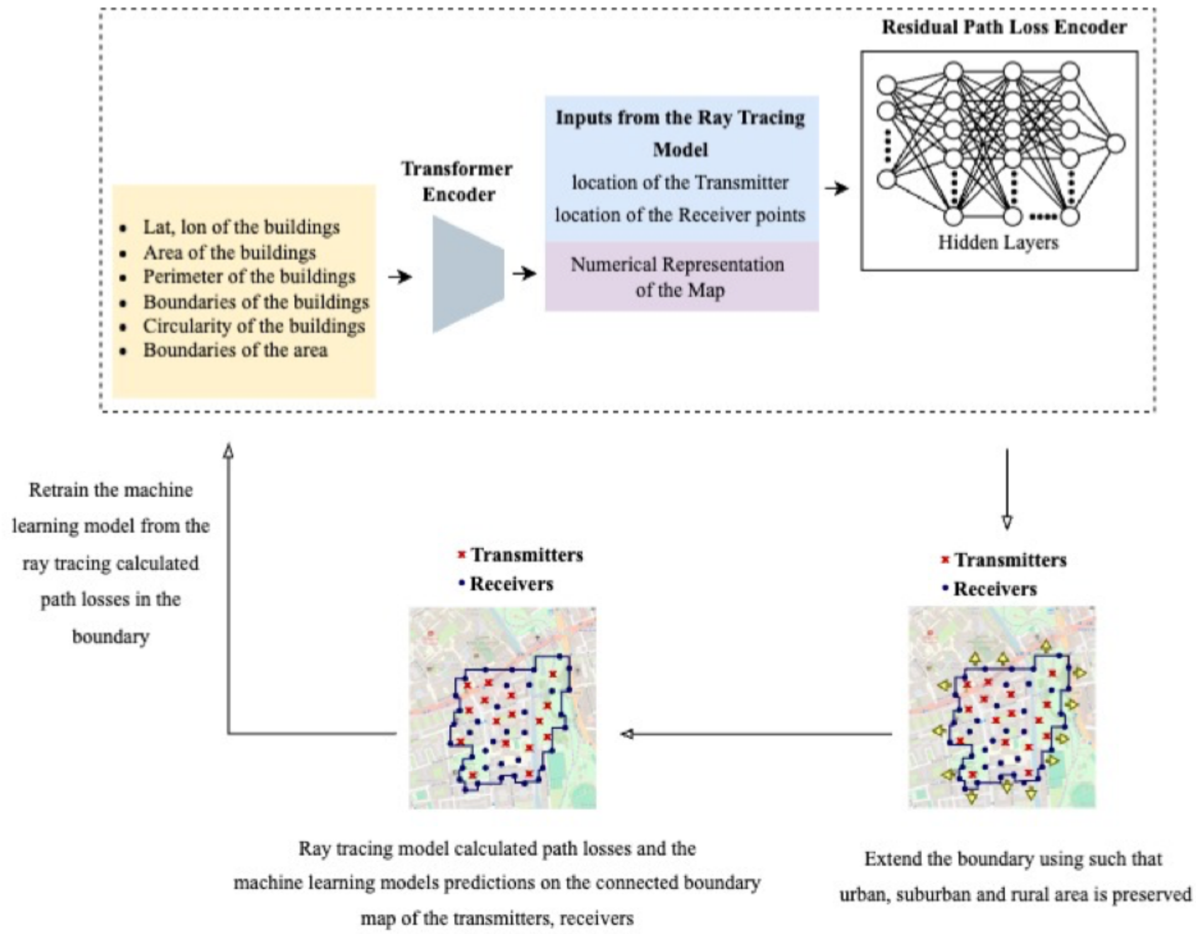


Figure 33: A scalable approach for path loss prediction in the UK by expanding the geographical boundary.

In addition to the potential enhancements mentioned earlier, further improvements can be made by expanding the model to cover the entire UK map. When extending the geographical boundaries, it is critical to ensure that the expansion preserves the distinct characteristics of urban, suburban, and rural areas (Figure 33). This involves the careful design of boundary expansion strategies that maintain the inherent diversity of the UK's landscape. By incorporating representative data from densely built urban centres, moderately developed suburban regions, and sparsely populated rural areas, the model can be trained to capture the unique propagation conditions and structural variations present across different environments. This approach not only enhances the model's generalisability but also improves prediction accuracy when applied to varied real-world scenarios throughout the UK.

The developed machine learning pipeline demonstrates the feasibility of predicting wireless path loss by combining simulated ray tracing data with real-world OSM features. By leveraging a transformer-based module for spatial feature extraction and a residual network for regression, the pipeline effectively learns complex relationships that govern signal propagation in diverse environments.

Future work will focus on scaling the model to cover larger geographic areas, such as the entire UK, and developing specialised sub-models for distinct environments. The ensemble approach—where urban, suburban, and rural predictions are combined—could provide a more robust and generalisable prediction system. In addition, the adoption of an active learning strategy promises continuous improvement as more data becomes available.

5.1.7 Use Case: Base Station Optimisation Methodology

In radio network planning, optimal transmitter placement is critical for ensuring reliable coverage across a target region. In urban environments, physical obstacles (e.g., buildings) can significantly affect signal propagation, making it challenging to achieve good coverage with minimal infrastructure. This section presents a gradient-based optimisation approach that integrates:

1. A Transformer-based model to summarise the built environment (buildings, geometry, and other features).
2. A Path Loss prediction model to forecast signal strength between a transmitter and arbitrary receiver points.
3. A cost function designed to balance coverage performance, coverage overlap, transmitter proximity constraints, and the overall number of transmitters.

By iteratively refining transmitter coordinates using gradient descent, the method converges on an effective layout under real-world constraints derived from OpenStreetMap (OSM) data.

Data Preparation

Data preparation ensures the optimisation routine has all necessary inputs: (1) Region boundaries, (2) building footprints and attributes, (3) receiver grid, (4) (optional) street nodes for projection.

OSM Data Extraction

1. Boundaries:
    - The .osm file includes bounding box tags (min_lat, min_lon, max_lat, max_lon) or sufficient node data to deduce these.
    - These limits define the rectangular region in which transmitters may be placed and across which coverage must be ensured.
2. Building data:
    - Each building footprint in OSM is stored as a collection of nodes (latitude–longitude points).
    - By iterating over these nodes and forming polygons, we can calculate: centroid (mean location of the building), perimeter (distance around the polygon), area (size of the building footprint), circularity (a shape descriptor relating area and perimeter), height.
    - These features capture how buildings might obstruct or reflect signals.
3. Street data (optional):
    - The .osm file also holds highway or road elements.
    - Extracting and storing the nodes (lat–lon pairs) of these roads allows an optional step later where each transmitter is snapped to the closest street node.

Building Vector Encoding

1. Rationale: Buildings significantly influence signal propagation. However, explicitly modelling each building for path loss can be cumbersome.
2. Transformer Pooling model:
    - Takes in a list of building feature vectors (height, area, etc.) for the region.
    - Processes them via a small Transformer encoder, aggregating the features into a single environment vector (bld_vec).
    - This vector effectively represents the overall urban environment's complexity.

Such an encoding step helps the subsequent path loss model adapt to different contexts (e.g., dense high-rise areas vs. suburban regions).

Receiver Grid Generation

1. Evenly spaced grid:
   - A grid of points is laid across the bounding box in latitude-longitude space. For example, a 50 × 50 grid yields 2500 receiver points.
   - At each of these points, we will predict coverage strength.
2. Purpose:
   - Ensures a comprehensive sampling of the region's signal quality.
   - The optimisation routine attempts to guarantee that each grid point meets a minimum path loss threshold (i.e., is "covered").

Path Loss Prediction Model

A neural network called PathLossEncoder predicts how strong (in dB) the signal is likely to be at any receiver location, given:

1. The building vector (bld_vec) – summarising the region.
2. Transmitter coordinates – latitude and longitude of the transmitter.
3. Receiver coordinates – latitude and longitude of the grid point.

The model outputs a numerical value interpreted as the predicted path loss (PL). Internal details often include:

- Fully Connected Layers: The network consists of linear transformations with non-linear activations (e.g., Leaky ReLU).
- Batch Normalisation / Dropout: Enhances stability and helps the model generalise.

Interpretation: In this particular code structure, a higher PL value corresponds to better signal strength The model is trained (before optimisation) on a dataset of known transmitter-receiver pairs and measured or simulated path loss.

Cost Function

The cost function encodes the network planner's objectives. Specifically:

1. Coverage (Uncovered Penalty)
   - Each grid point's effective coverage is computed by summing the coverage indicators from all transmitters.
   - If a point's coverage sum is below 1 (meaning no transmitter provides sufficient signal), we add a penalty.

2. Overlap Penalty
   - If the sum of indicators at a point exceeds 1, we interpret it as "redundant coverage".
   - While some overlap can be beneficial, excessive overlap is penalised to avoid resource waste.
3. Transmitter Separation Penalty
   - A minimum distance min_tx_sep\text{min_tx_sep}min_tx_sep is enforced between transmitters.
   - If any pair is too close, a penalty is added proportionally to how much they violate the separation constraint.
4. Transmitter Count Penalty
   - The total cost increases linearly or quadratically (depending on the chosen factor) with the number of transmitters.
   - This encourages solutions that use fewer transmitters, assuming coverage goals can still be met.

Formally, these penalty terms are summed into a single scalar, which the optimiser attempts to minimise. Crucially, all these terms are differentiable (thanks to the careful design of the coverage indicator via a sigmoid function), allowing gradient-based optimisation.

Optimisation Procedure

Initialisation

1. Random placement: We specify a target number of transmitters, say N. Each transmitter's latitude and longitude are initialised randomly within the bounding box.

2. Trainable parameters: These coordinates (lat–lon) become PyTorch parameters, meaning they will be updated automatically via gradient descent.

Gradient Descent

1. Compute predictions: For each transmitter-receiver pair (in the coverage grid), predict the path loss via PathLossEncoder. Convert these predictions into coverage indicators (using a threshold-based sigmoid).
2. Calculate cost: Sum penalties for coverage shortfalls, overlap, transmitter proximity violations, and total transmitter count.
3. Backpropagation: PyTorch calculates how sensitive each cost component is to the transmitter coordinates.
4. Parameter update: An optimiser (e.g., Adam) moves each transmitter coordinate in the direction that reduces the overall cost.

Additional Constraints & Techniques

1. Bounding box clamping: After each iteration, transmitter coordinates are “clamped” to remain within the bounding box, so they never drift outside.
2. Street projection: If realistic deployment requires placing transmitters on actual streets, we can snap each updated transmitter to the closest street node. The .osm file’s highway data helps identify these possible locations.

Early stopping

We track changes in the cost function. If cost improvement falls below a certain threshold for a predetermined number of iterations (“patience”), the process halts early. This avoids unnecessary computation once it appears we have plateaued near a good solution.

Finding the Optimal Number of Transmitters

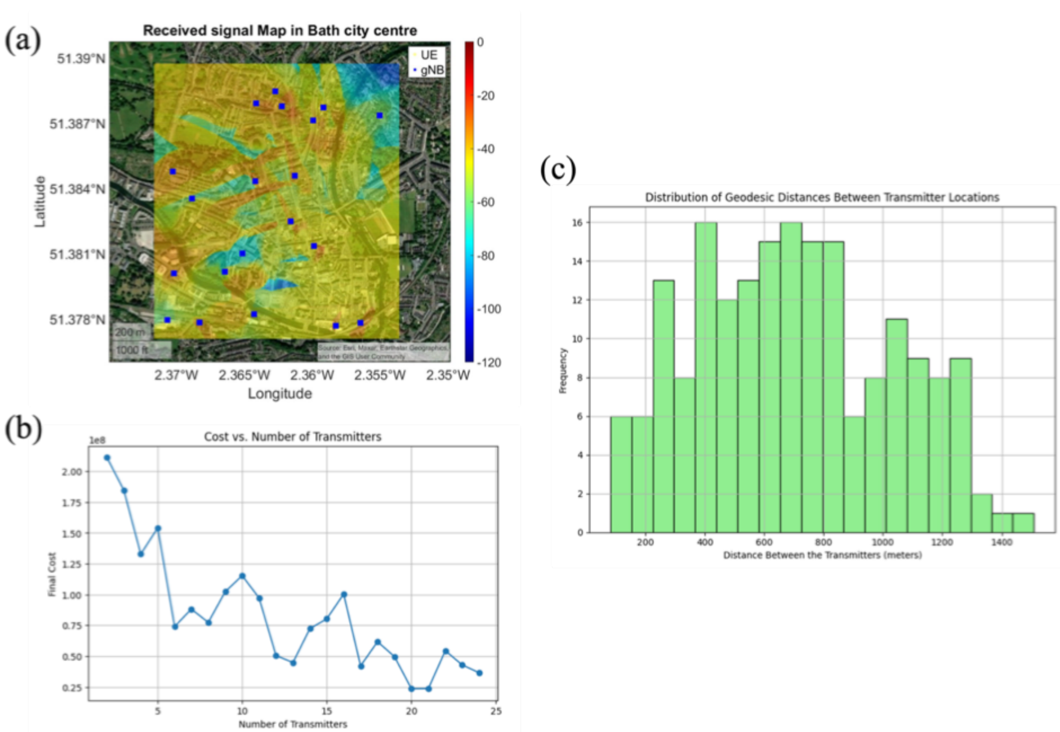
1. Candidate range: Rather than fixing the transmitter count, the procedure can be run multiple times for a range of values (e.g., 2 to 25).
2. Compare final costs: For each fixed number of transmitters, we note the minimised cost after the optimisation loop is completed.
3. Select best: We choose the number of transmitters that yields the lowest final cost. This typically reflects an ideal balance between coverage and infrastructure expense.

Coverage Identification

After obtaining an optimised solution, the final layout is assessed visually:

1. Radial Boundary Computation: For each transmitter, we sample multiple radial directions (e.g., 36 angles). We step outward from the transmitter, predicting path loss at small increments (e.g., 0.001° each). Once the path loss dips below the coverage threshold for too many consecutive steps, we mark that boundary point.
2. Polygon formation: Connecting these boundary points yields a "coverage polygon" around each transmitter. These polygons can be plotted to show how far coverage extends in each direction.

Through this data-driven, gradient-based approach, radio planners can identify transmitter placements that achieve desired coverage while minimising overlapping, meeting separation constraints, and limiting the total number of transmitters. The integration of real geographic data (OSM) and neural network-based path loss predictions provides a modern, flexible framework suited to a variety of urban environments.



Application of ML Path Loss Model in Bath and London for n77 Frequency Band

The locations of gNB are optimised to determine what number of gNBs. The cost function is based on uncovered penalty, overlap penalty, minimum separation penalty, transmitter count penalty.

<u>Bath</u>

The optimised coverage map for Bath city centre is shown in Figure 34(a), featuring 20 gNBs, which is concluded from cost function evaluation shown in Figure 34b). This demonstrates comprehensive coverage, ensuring that all areas receive signal strength of at least -105 dBm (no dark blue area). The geodesic distance between 2 gNBs, shown in Figure 34(c), is spread across 400 to 900 meters.

<u>London</u>

The optimised coverage map for London Blackfriars is shown in Figure 35(a), featuring 11 gNBs, which is concluded from the cost function evaluation shown in Figure 35(b). This demonstrates comprehensive coverage, ensuring that all areas receive signal strength of at least -105 dBm (no dark blue area). The geodesic distance between 2 gNBs, shown in Figure 35(c), appears mostly around 600 meters.

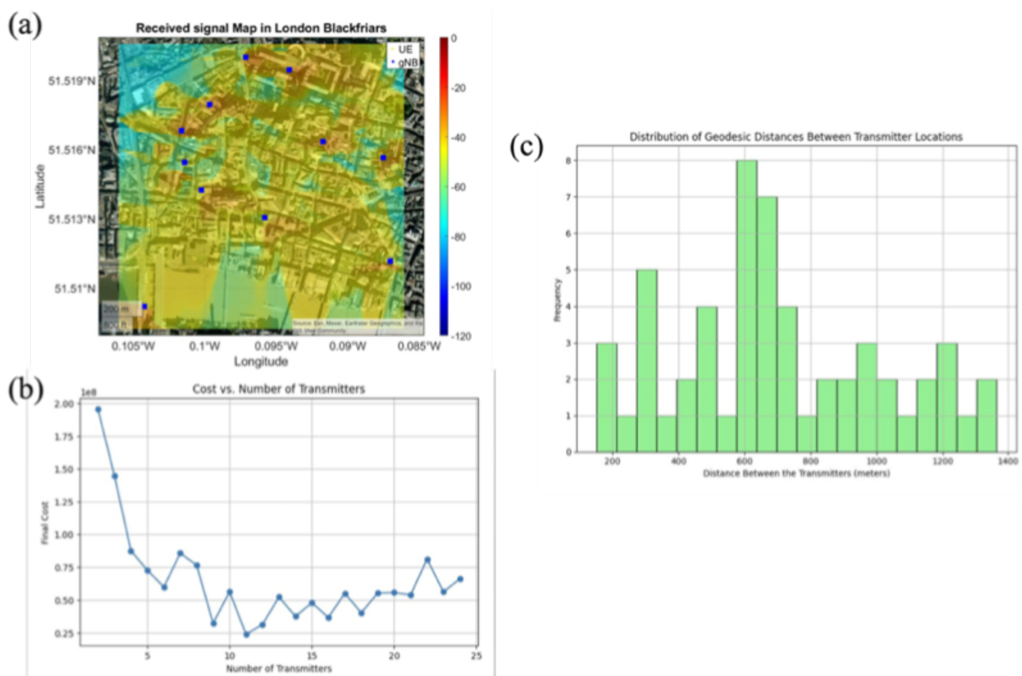


Figure 35: (a) Received coverage signal map for London Blackfriars at n77 frequency band, (b) cost function vs number of transmitters, (c) count of geodesic distances between each of 2 gNBs.

Application of ML Path Loss Model in Bath and London for Band 3

The location of gNBs is optimised to determine the number of gNBs using ML. The cost function is based on the Overlap with the MNO Penalty, Uncovered Penalty, Overlap with the gNB Penalty, Minimum Separation Penalty, and Transmitter Count Penalty.

Bath

The optimised coverage map Figure 36(a) for Bath city centre, featuring six gNBs. Figure 36(b) indicates that when more gNBs are included in the coverage area, the cost function varies for ML, suggesting there are multiple options for “maximum signal coverage” for minimal interference. We chose 10 transmitters in total, of which 6 are new gNBs. The distance between 2 gNBs appears to be mostly 500 meters (Figure 36c).

London

The optimised coverage map Figure 37(a) for London Blackfriars, featuring four gNBs, Figure 37b indicates that when more gNBs are included in the coverage area, the cost function increases for ML, suggesting only 8 (4 new ones) B3 transmitters are allowed for minimal interference. The average distance between 2 gNBs is 550 meters (Figure 37c).

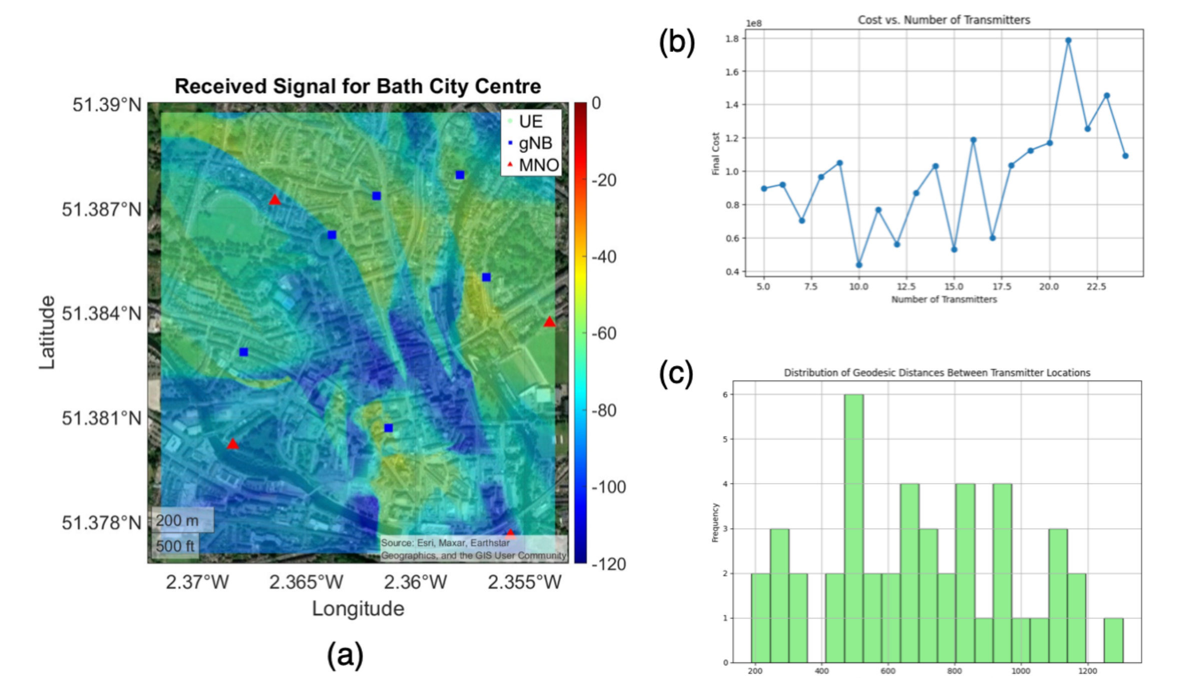


Figure 36: (a) Received coverage signal map for the Bath City Centre at Band 3, (b) cost function vs number of transmitters, (c) count of geodesic distances between each of 2 gNBs.

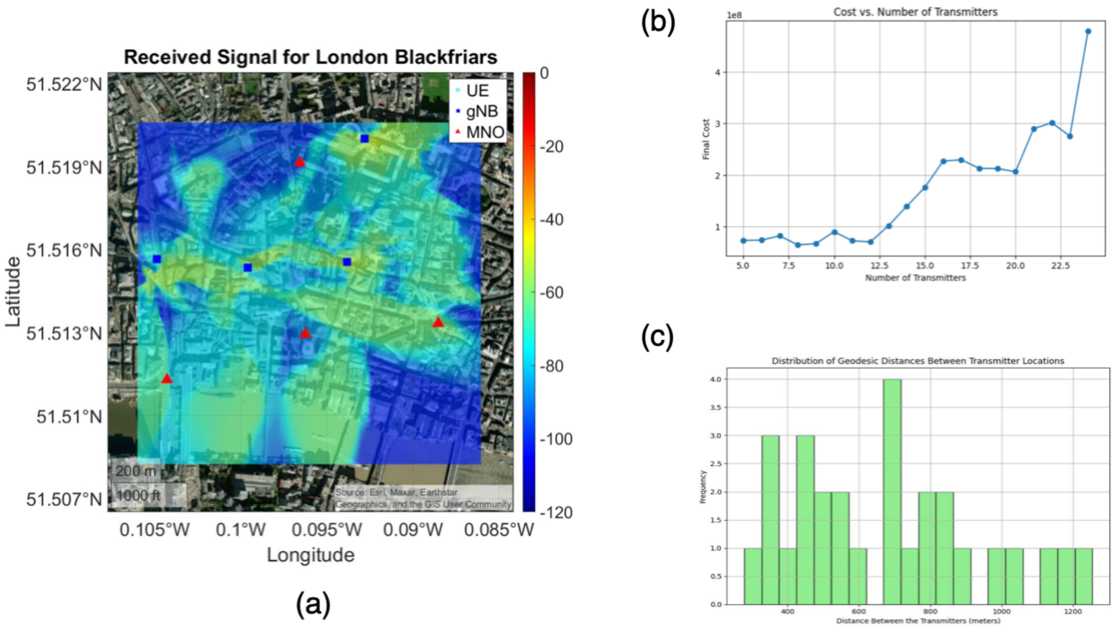


Figure 37: (a) Received coverage signal map for London Blackfriars at Band 3, (b) cost function vs number of transmitters, (c) count of geodesic distances between each of 2 gNBs.

5.2 Enhancing Simulations with LiDAR: Improving Environmental Accuracy

The advancement of wireless communication systems, particularly in the domain of spectrum sharing and urban propagation modelling, necessitates precise environmental characterisation. A critical innovation in this regard is the integration of deep learning methodologies with open-source geospatial databases, as illustrated in Figure 38. This approach significantly enhances the fidelity of radio wave propagation simulations by improving the accuracy of environmental representations.

The efficacy of ray-tracing models is inherently dependent on the granularity and precision of the simulated environment. OpenStreetMap (OSM), a widely utilized open-access geographic dataset, provides extensive spatial information, encompassing building footprints, vegetation coverage, fences, and utility infrastructure. Among these, the structural attributes of buildings exert the most profound influence on urban propagation modelling. However, the quality of building-related data retrieved from OSM—particularly building heights—exhibits substantial limitations. These inaccuracies stem from the constraints of community-driven data annotation and the inherent challenges of human perception in satellite image-based geospatial mapping.

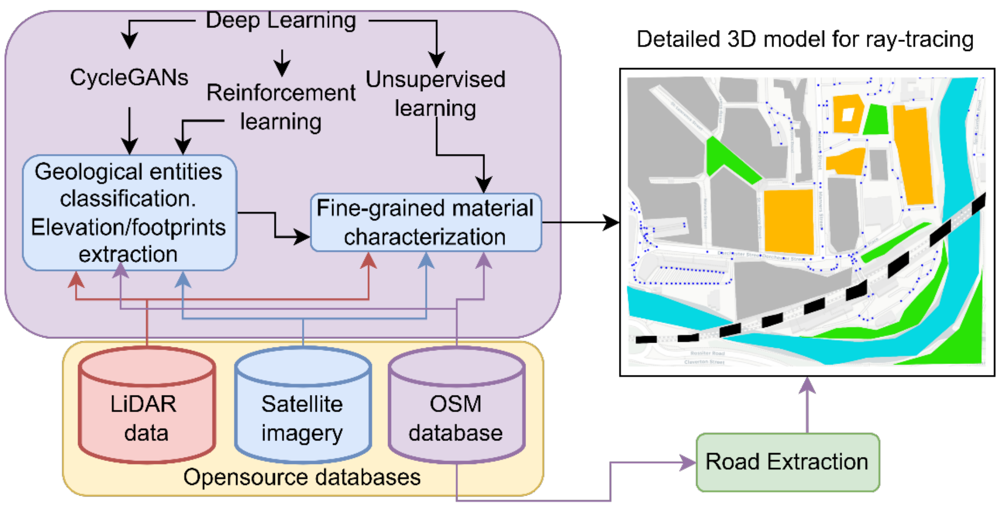


Figure 38: Utilising machine learning to extract geological information, classify materials, and determine material properties from LiDAR, satellite, and OSM data to construct a comprehensive 3D city model for ray-tracing applications.

Beyond geometric accuracy, the electromagnetic properties of urban entities play a pivotal role in ray-tracing simulations. Structural surfaces exhibit diverse material compositions, including concrete, brick, stone, and metal, each influencing electromagnetic wave interactions through distinct permittivity and conductivity properties. Unfortunately, such material-specific data is rarely available in OSM, further limiting the precision of conventional propagation models.

Deep learning offers a transformative potential for extracting critical environmental attributes from multiple heterogeneous data sources. The proposed ensemble deep learning framework, illustrated in Figure 36, integrates both structured and unstructured datasets to refine urban environment representation with superior accuracy. Key advancements include:

1. **Building and Vegetation Heights:** Aerial Light Detection and Ranging (LiDAR) data provides highly precise vertical measurements, offering a robust solution for the accurate estimation of entity heights. LiDAR's established efficacy in vertical imaging makes it indispensable for refining three-dimensional urban models.

2. **Entity Footprints and Boundary Delineation:** LiDAR point clouds, when processed using deep learning-based classification techniques, facilitate the extraction of fine-grained boundaries for various urban entities. Additionally, high-resolution satellite imagery can be incorporated into the ensemble model, leveraging its superior horizontal resolution to enhance planimetric accuracy.

3. **Material Characterisation for Electromagnetic Modeling:** The intensity of LiDAR returns is intrinsically linked to surface material properties, offering a promising avenue for

material classification. By analysing reflectance variations, deep learning algorithms can infer the composition of building facades, road surfaces, and other urban elements. This enables the estimation of key electromagnetic properties such as permittivity and conductivity, thereby significantly improving the realism of propagation simulations.

In summary, the proposed deep learning-driven approach for environmental modelling bridges critical gaps in conventional geospatial datasets. By integrating LiDAR, satellite imagery, and data-driven material classification, this methodology enables a highly precise and electromagnetically relevant representation of urban environments, thus advancing the reliability of next-generation wireless communication simulations.

To address these limitations, as Proof of Concept, we developed a methodology leveraging Lidar (Light Detection and Ranging) data to derive more precise building footprint and height information, which can be seamlessly integrated into ray tracing models for enhanced environmental simulation fidelity. The improved environmental model we derived provides a close resemblance to the experimental data.

5.2.1 Data Source

The Lidar data utilised in this study is sourced from the Environment Agency National Lidar Programme [40], which provides high-resolution elevation data at a 1-meter spatial resolution across England. Aerial Lidar data, a form of point cloud data, offers highly accurate topographic elevation measurements of the Earth's surface. The dataset follows the ASPRS LAS 1.4 classification standard, categorising elements such as ground, low, medium, and high vegetation, as well as buildings. This classification enables the precise extraction of building distribution, footprints, and heights within the study area.

Figure 39 illustrates the process of acquiring LiDAR data for the area surrounding Bath City Centre via the National LiDAR Programme website. To obtain the required dataset, the user first delineates the area of interest by defining a polygon. Subsequently, the desired data category can be specified and downloaded.

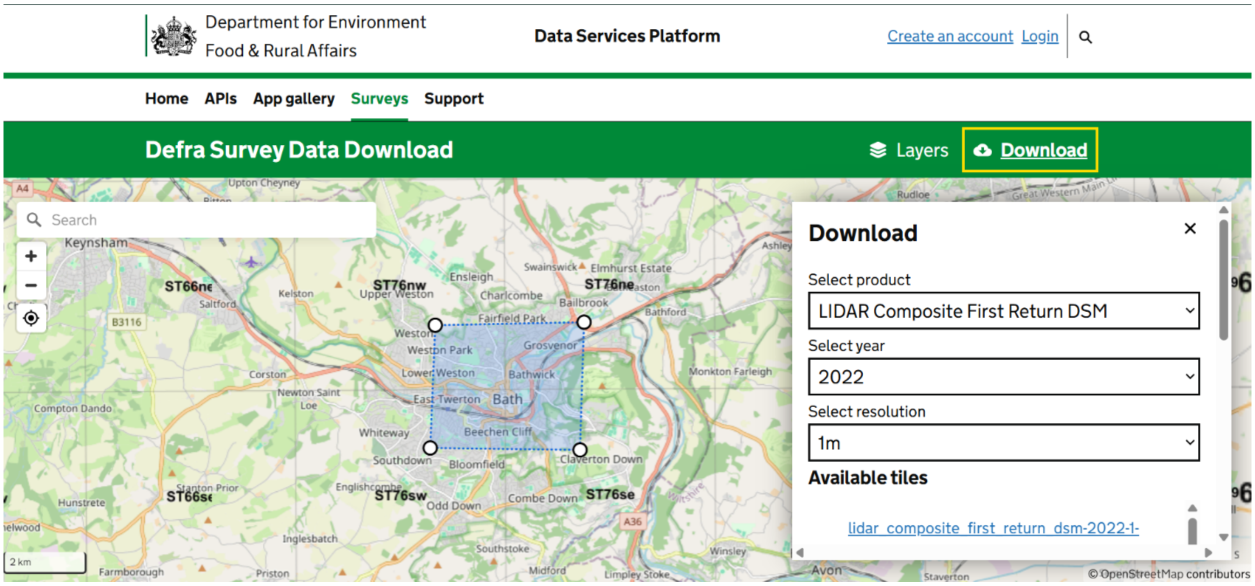


Figure 39: Downloading Lidar data from the Defra web interface. The tiles which are visible in the map can be selected by the polygon tool and downloaded.

Once the data is downloaded, we can visualise the classified Lidar data using Geographic Information System (GIS) software such as QGIS or Python language. A visualisation of classified LiDAR data of Bath City Centre is shown in Figure 40(a).

5.2.2 Methodology

Building Footprint Extraction

To delineate building footprints, we extracted Lidar data points classified as buildings and generated a raster layer, representing a two-dimensional projection of the point cloud onto the Earth's surface. Since only the building-classified points were retained, the resultant raster contained discrete structures corresponding to building locations. These structures were then vectorised through a polygonisation process to generate convex hulls around each detected building, effectively delineating their footprints. The vectorised data enables conversion into multiple GIS file formats, including OSM and GPKG, facilitating its integration into ray tracing simulations.

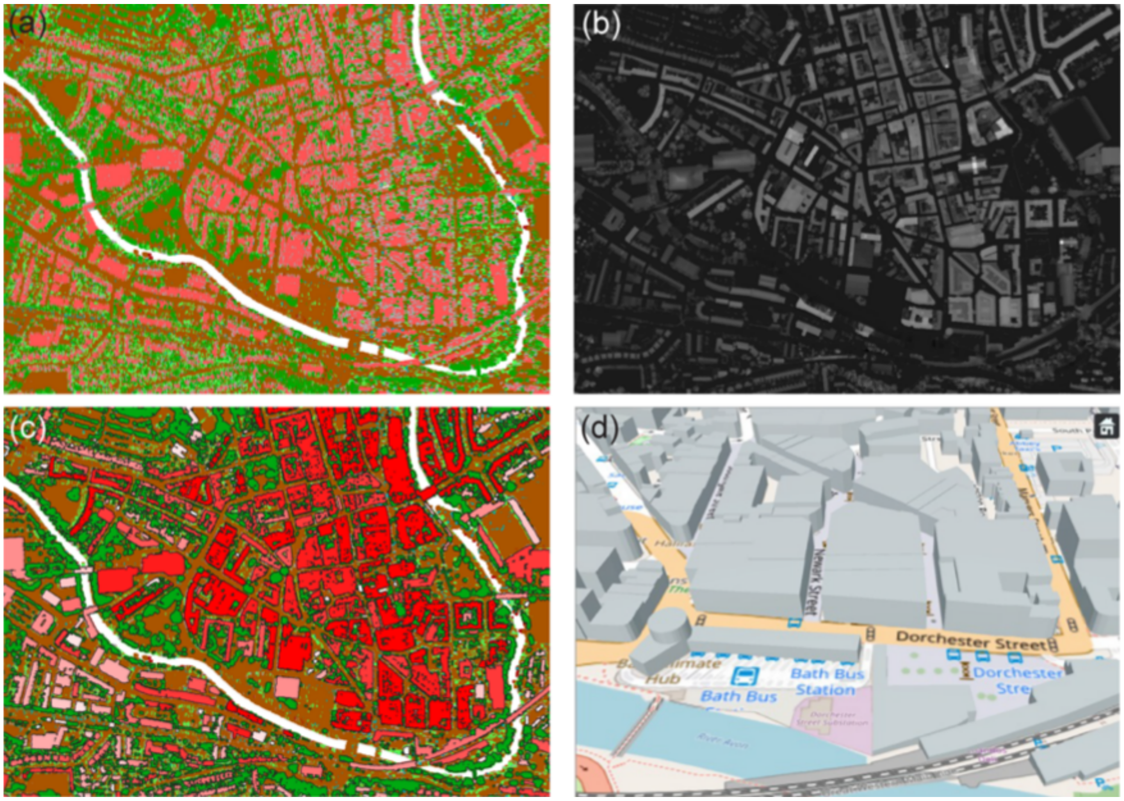


Figure 40: a) The Lidar data of Bath-CC, classified between buildings, vegetation and ground. b) The height variation on the earth's surface. c) The buildings, with height colour-coded. d) A 3D view of Bath-CC generated using the processed Lidar data.

Building Height Extraction

A critical prerequisite for accurate building height estimation is the establishment of a Digital Terrain Model (DTM), which represents the bare-earth elevation. To construct the DTM, we extracted ground-classified Lidar points and projected them onto a raster layer, effectively generating a continuous surface model. Data gaps in the rasterised terrain were interpolated to ensure a seamless and accurate DTM representation.

Following the DTM construction, we derived the Digital Surface Model (DSM), which encapsulates the elevation variations due to both natural and anthropogenic features. This was achieved by incorporating all non-noise Lidar classifications into a rasterised elevation model. The absolute height of objects on the Earth's surface was then computed as the difference between the DSM and DTM, expressed mathematically as:

$$height = DSM - DTM$$

Figure 40(b) shows the variation of height across Bath City Centre using the LiDAR data. To determine building heights, the previously extracted building footprints were overlaid onto the DSM-derived height variation data. The mean zonal height statistic of the height layer within each vectorised footprint was calculated, providing an estimate of the average building height. This derived metric serves as a reliable representation of building height information extracted from Lidar data.

Figure 40(c) shows the extracted building footprints, coloured based on the average height of the building. The accurately extracted height shows the reducing trend of building heights when moving away from the city centre. Figure 40(d) is a three-dimensional visualisation of the extracted buildings, overlayed on the OpenStreetMap data. A comparison between the LiDAR-processed OSM file and the OSM file obtained from the OpenStreetMap [38] is shown in Figure 41. It can be observed that LiDAR provides much better resolution of the building structures in a given area.

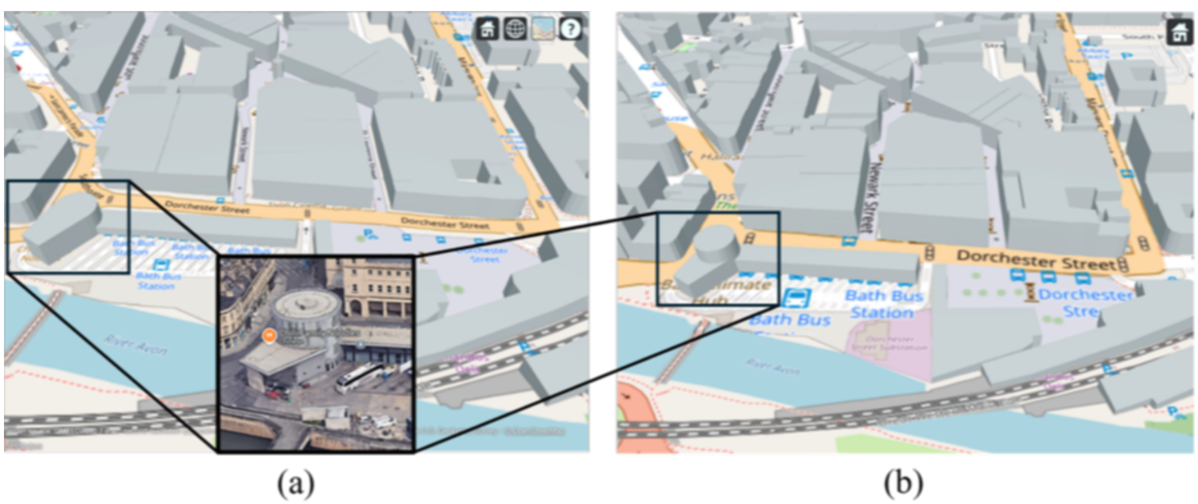


Figure 41: Bath City Center. (a) OSM from OpenStreetMap [38], (b) OSM from QGIS after LiDAR analysis. (Inset: Google Earth Map)

Software and Processing Tools

Lidar data processing and geospatial analysis were conducted using QGIS, an open-source GIS software widely employed for terrain and environmental modelling. Figure 42(a) shows the QGIS interface. The vectorised building footprints were converted to the OSM format using JOSM (Java OpenStreetMap Editor). To ensure adherence to OSM standards, a custom Python script was developed for post-processing, standardising building annotations and height attributes to facilitate seamless integration with MATLAB-based ray tracing simulations. Figure 42(b) shows the example data before processing, while Figure 42(c) shows the results after being processed from the Python code.

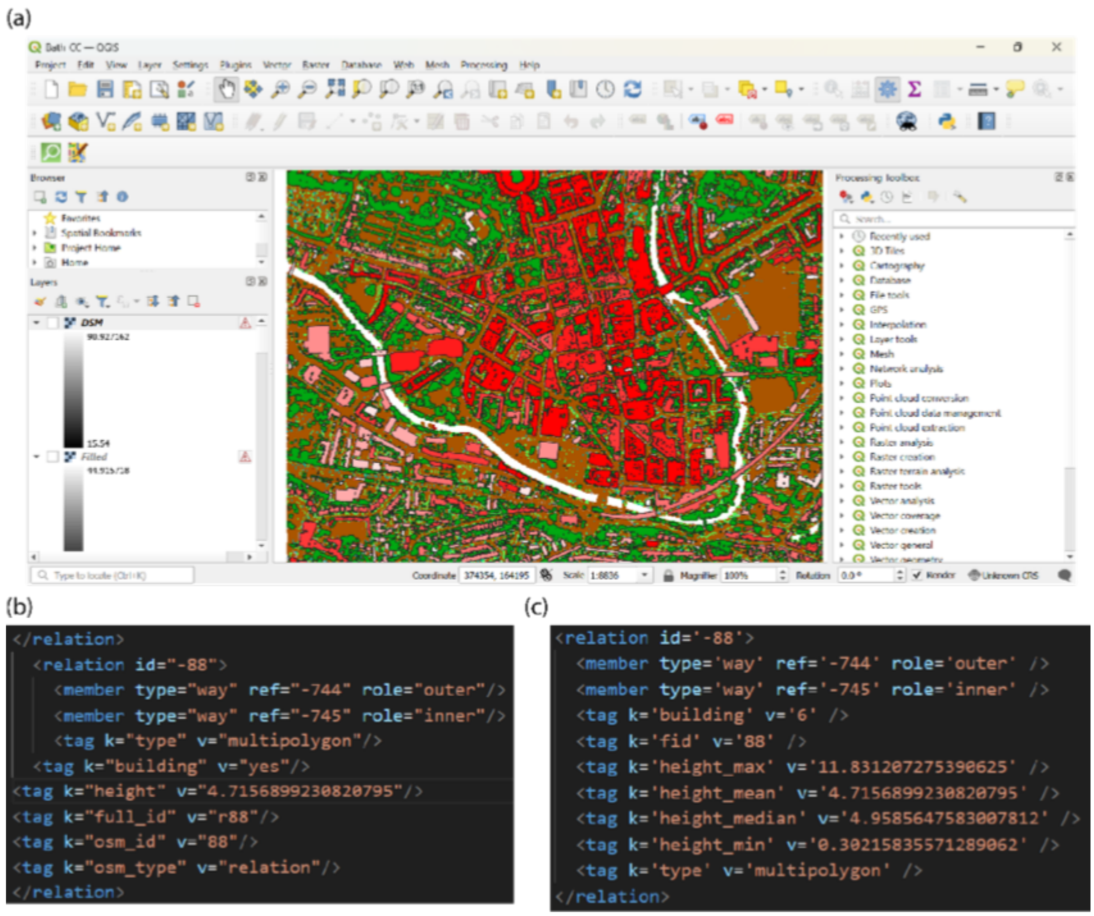


Figure 42: a) The QGIS interface to process the Lidar data to derive the building footprints and heights. b) The OSM data after post-processing using the python code. c)The un-processed OSM data produced using the Lidar data, with custom attributes.

The proposed Lidar-based methodology significantly enhances the accuracy of building footprint and height estimation, addressing the inherent limitations of community-annotated OSM data. By leveraging high-resolution Lidar elevation datasets, the resulting geospatial models offer improved precision for ray tracing simulations, as demonstrated earlier, ultimately contributing to more reliable urban propagation modelling and environmental analysis.

Thus, we observe that Ray tracing combined with LiDAR provides detailed and accurate representation of the physical environment (buildings, vegetation), significantly improving propagation predictions compared to traditional deterministic models. Deterministic methods, such as empirical or statistical models (e.g., ITU-R models, Hata-Okumura), typically provide generalised predictions that may not adequately capture local variations or complex urban/ suburban geometries. In contrast, RT with LiDAR captures precise geometries and material properties, resulting in more realistic and reliable predictions. More precise predictions directly enable more accurate interference modelling, facilitating more effective spectrum sharing and efficient spatial reuse.

Moving towards Ray Tracing, supported by high-resolution LiDAR data would enable Ofcom and other stakeholders to leverage significantly more precise propagation information. This approach can greatly improve spectrum allocation efficiency, interference management, and overall utilization of bands designated for shared or private access.

Typically, when compared to deterministic or empirical models, RT combined with LiDAR can yield differences of 5 to 15 dB or more in predicted propagation losses, depending on environment complexity and frequency. This improvement translates directly into more accurate and less conservative interference coordination and frequency reuse criteria, potentially freeing significant additional spectrum resources.

At upper 6 GHz, propagation conditions become increasingly sensitive to environment-specific clutter and multipath scenarios. RT and LiDAR significantly improve the predictive accuracy of link quality and interference levels in these environments. This is particularly important as at those higher frequencies there is greater path loss and increased sensitivity to physical obstructions, and thus, RT and LiDAR can allow tighter coordination and denser spatial reuse. Hence, the increased accuracy of RT and LiDAR allows less conservative interference management, enabling more effective sharing scenarios and a higher density of deployments without unacceptable interference. For Ofcom, embracing this modelling approach would support: (1) More precise licensing conditions, (2) better-informed policy decisions on coexistence criteria, including potentially reduced protection distances, and (3) enhanced flexibility and efficiency in shared band management. For SAL licensees, this would mean: (1) reduced deployment costs through optimized infrastructure planning, (2) increased certainty and reliability of service performance, and (3) improved spectrum efficiency, potentially allowing greater bandwidth and higher-quality services.

For the next phase of this research, we will develop the deep learning framework mentioned earlier, integrating Lidar data with satellite imagery, photogrammetry, and OpenStreetMap data to develop a more sophisticated and precise urban model. Deep learning techniques, particularly CNNs, graph neural networks transformer-based architectures, even multi-modal architectures exhibit exceptional promise in automated feature extraction, classification, and 3D reconstruction for combined geospatial data. By capitalising on our expertise in deep learning, we aim to develop data-driven urban models that significantly enhance the accuracy of electromagnetic wave propagation simulations and other computational models reliant on precise geospatial data.

6. Conclusion

The work presented in this report marks a significant advancement in the development of high-fidelity simulation frameworks for dynamic spectrum sharing. Within WP2 of the Spectrum Sandbox ITT Project, we have integrated ray tracing, machine learning, and real-world measurement data to enhance the accuracy of signal propagation modelling and interference prediction. These contributions are crucial for designing efficient spectrum-sharing strategies that align with real-world network conditions and regulatory frameworks.

A major achievement of this work is the integration of machine learning (ML) techniques which represent a significant innovation in propagation modeling and interference mitigation. ML-based models have been utilised to improve the accuracy of path loss estimation, material classification, and dynamic interference prediction. Unlike static propagation models, ML-driven approaches allow for adaptive learning, where the models refine their predictions based on continuous data acquisition. This capability is particularly beneficial for spectrum-sharing scenarios, where interference patterns fluctuate due to the presence of multiple network operators and dynamic spectrum access users. Not only that, but it also reduces the time of simulation for predicting path loss humongous by the order of 10^3. This can help with providing very high-resolution data for an area in comparatively much less time. The use of LiDAR data has further improved environmental modelling accuracy, filling gaps in traditional datasets such as OpenStreetMap (OSM).

This study also examines critical deployment considerations for spectrum sharing, particularly within the 1800 MHz (Band 3) and 3800–4200 MHz (n77) bands. In Band 3, we explore the feasibility of dynamic spectrum access, ensuring efficient coexistence with incumbent users while minimizing interference. Additionally, we address coverage gaps ("not-spots") by determining the minimum number of base stations required to eliminate these gaps while maintaining reliable connectivity. For the 3800–4200 MHz spectrum, we extend our analysis to optimize base station deployment for targeted local area coverage and robust connectivity. The key distinction between our approaches for Band 3 and n77 lies in their respective deployment strategies:

- Band 3 (1800 MHz): Since incumbent Mobile Network Operators (MNOs) are already present, our focus is on eliminating mobile not-spots. Here, smaller local mobile operators can cost-effectively extend coverage using dynamic spectrum sharing, facilitating Local Access Licensing (LAL).
- n77 (3800–4200 MHz): In contrast, this band primarily enables Shared Access Licensing (SAL). Base stations can be deployed to provide local coverage without interfering with existing incumbent operators, ensuring an optimal balance between spectrum efficiency and network expansion.

These insights are crucial for network planning and infrastructure optimisation, allowing operators to balance coverage, capacity, and cost-effectiveness in future spectrum-sharing frameworks.

Despite the promising results, several challenges remain. One of the primary limitations of ray tracing-based simulation is its high computational complexity. Simulating large-scale environments with full multipath analysis requires significant processing power, limiting the number of scenarios that can be tested within practical time constraints. Additionally, while ITU-R standards provide generalised material properties, real-world variations in dielectric constants, surface roughness, and construction materials introduce discrepancies between simulated and measured data. These uncertainties affect the precision of attenuation and reflection modelling, particularly in dense urban environments where multiple interactions influence signal propagation.

Another challenge lies in interference modelling, particularly in dynamic spectrum-sharing environments where neighbouring cells operate at varying power levels. The tolerance to inter-cell interference decreases with increasing Modulation and Coding Scheme (MCS), meaning that high-throughput transmissions become more susceptible to degradation in the presence of strong adjacent-channel signals. The static power assumptions used in the current simulation framework do not fully capture the adaptive nature of real-world power control mechanisms, leading to variations in observed vs. predicted performance. Future enhancements will incorporate real-time power adaptation models, improving the accuracy of interference mitigation strategies.

Looking ahead, several key areas of future research and development have been identified to enhance the scope and effectiveness of the simulation framework.

1. Computational optimisations, such as GPU-accelerated ray tracing and hybrid ML-physics models, will be explored to reduce simulation time without sacrificing accuracy.
2. Implement ML-driven approaches for improved extraction of building and terrain features using LiDAR data. The objective is to reduce discrepancies in path loss predictions between simulated results and empirical data.
3. Increase the accuracy of antenna radiation pattern modelling and material property databases through additional field measurements and lab validations.
4. Extend the geographic and environmental diversity of the tested scenarios, focusing on varied rural, suburban, and dense urban settings to further generalise findings.
5. Conduct comprehensive field trials to validate real-time spectrum access decision-making algorithms developed during WP1, assessing their performance under varied operational conditions.

6. Continue refining ML models for path loss and interference prediction, improving their adaptability and real-time applicability, and integrate these into dynamic spectrum-sharing decisions.
7. Deeper integration of AI-driven real-time spectrum allocation techniques will enable dynamic network adaptation, allowing systems to intelligently reconfigure frequency allocations based on evolving interference conditions.
8. Increase stakeholder workshops and industry consultations to ensure the technological, regulatory, and economic findings remain relevant and impactful for policy and practical deployments.

The findings of WP2 provide a strong foundation for future spectrum-sharing research, regulatory decision-making, and industrial deployments. By combining rigorous theoretical modelling, data-driven validation, and practical deployment insights, this work contributes to the advancement of next-generation wireless networks, ensuring efficient spectrum utilisation, improved connectivity, and robust interference management. As spectrum scarcity continues to challenge wireless communication, the innovations presented in this study pave the way for smarter, more adaptable, and more resilient wireless ecosystems.

References

[1] Zhang, L., Xiao, M., Wu, G., Alam, M., Liang, Y.C. and Li, S. "A survey of advanced techniques for spectrum sharing in 5G networks." IEEE Wireless Communications 24.5 (2017): 44-51.

[2] Parvini, M., Zarif, A.H., Nouruzi, A., Mokari, N., Javan, M.R., Abbasi, B., Ghasemi, A. and Yanikomeroglu, H. "Spectrum sharing schemes from 4G to 5G and beyond: Protocol flow, regulation, ecosystem, economic." IEEE Open Journal of the Communications Society 4 (2023): 464-517.

[3] Saha, R. K., and John M. C. "Dynamic spectrum sharing for 5G NR and 4G LTE coexistence-A comprehensive review." IEEE Open Journal of the Communications Society 5 (2024): 795-835.

[4] Ofcom, "Enabling wireless innovation through local licensing," 25 July 2019.

[5] Ofcom, "Shared Access Licence: Guidance Document," 20 September, 2022.

[6] Ofcom, "Local Access Licence: Guidance Document," July 2019.

[7] Joint Bid for DSIT Spectrum Sandbox Tender, February 2024.

[8] Hata, M. "Empirical formula for propagation loss in land mobile radio services." *IEEE Transactions on Vehicular Technology* 29.3 (2013): 317-325.

[9] Okumura, Y. "Field strength and its variability in VHF and UHF land-mobile radio service." *Review of the Electrical Communication Laboratory* 16.9 (1968).

[10] Schaubach, Kurt R., Nathaniel J. Davis, and Theodore S. Rappaport. "A ray tracing method for predicting path loss and delay spread in microcellular environments." *[1992 Proceedings] Vehicular Technology Society 42nd VTS Conference-Frontiers of Technology*. IEEE, 1992.

[11] Yun, Z., and Iskander, M. F. "Ray tracing for radio propagation modeling: Principles and applications." *IEEE Access* 3 (2015): 1089-1100.

[12] International Telecommunications Union Radiocommunication Sector. *Effects of Building Materials and Structures on Radiowave Propagation Above About 100MHz.* Recommendation P.2040. ITU-R, approved August 23, 2023. https://www.itu.int/rec/R-REC-P.2040/en.

[13] International Telecommunications Union Radiocommunication Sector. Electrical Characteristics of the Surface of the Earth. Recommendation P.527. ITU-R, approved September 27, 2021. https://www.itu.int/rec/R-REC-P.527/en.

[14] https://uk.mathworks.com/help/comm/ug/ray-tracing-for-wireless-communications.html.

[15] Chipman, R., Lam, W. S. T., and Young, G. *Polarized light and optical systems*. CRC press, 2018.

[16] McNamara, D. A., C. W. I. Pistorius, and J. A. G. Malherbe. *Introduction to the Uniform Geometrical Theory of Diffraction*. Boston: Artech House, 1990.

[17] International Telecommunications Union Radiocommunication Sector. *Propagation by diffraction*. Recommendation P.526-15. ITU-R, approved October 21, 2019. https://www.itu.int/rec/R-REC-P.526/en.

[18] 3GPP TS 38.101-1 V18.6.0 (2024). *NR; User Equipment (UE) radio transmission and reception*.

[19] 3GPP TS 38.101-1 V18.6.0 (2024-06) 5.2 *Operating bands.*

[20] 3GPP TS 38.104 Table 5.3.2-1: *Transmission bandwidth configuration NRB.*

[21] 3GPP TS 38.101-1 V18.6.0 (2024-06) Table 5.3.5-1 *Channel bandwidths for each NR band.*

[22] 3GPP TS 38.211 V18.3.0 (2024-06) 6.3.1.2 *Modulation*.

[23] 3GPP TS 38.214 Table 5.1.3.1-2: *MCS index table 2 for PDSCH*.

[24] 3GPP TS 38.104 Annex A (normative): *Reference measurement channels.*

[25] 3GPP TS 38.104 6.2 *Base station output power.*

[26] 3GPP TS 38.211 V18.3.0 4.3.2 *Slots*.

[27] 3GPP TS 38.214 5.1 *UE procedure for receiving the physical downlink shared channel.*

[28] Yang, C., Li, J., Guizani, M., Anpalagan, A. and Elkashlan, M. "Advanced spectrum sharing in 5G cognitive heterogeneous networks." IEEE Wireless Communications 23.2 (2016): 94-101.

[29] Ahmad, A., Ahmad, S., Rehmani, M.H. and Hassan, N.U. "A survey on radio resource allocation in cognitive radio sensor networks." IEEE Communications Surveys & Tutorials 17.2 (2015): 888-917.

[30] Liu, J., Kato, N., Ma, J. and Kadowaki, N. "Device-to-device communication in LTE-advanced networks: A survey." IEEE Communications Surveys & Tutorials 17.4 (2014): 1923-1940.

[31] Kim, D., Lee, H. and Hong, D. "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers." IEEE Communications Surveys & Tutorials 17.4 (2015): 2017-2046.

[32] Ding, Z., Liu, Y., Choi, J., Sun, Q., Elkashlan, M., Chih-Lin, I. and Poor, H.V. "Application of non-orthogonal multiple access in LTE and 5G networks." IEEE Communications Magazine 55.2 (2017): 185-191.

[33] Francisco R. V. G., José M. B. da S. Jr., Charles C. C., Gabor F., Mats B. and Carlo F. (2024), "Machine Learning for Spectrum Sharing: A Survey", Foundations and Trends® in Networking: Vol. 14: No. 1-2, pp 1-159.

[34] Cohen, K.. "Machine learning for spectrum access and sharing." *Machine Learning for Future Wireless Communications* (2020): 1-25.

[35] Sun, H., Rose Q. H., and Yi Q. "Secure Spectrum Sharing with Machine Learning: An Overview." (2024): 115-134.

[36] Vo, V., et al. "Security and Privacy of 6G Federated Learning-enabled Dynamic Spectrum Sharing." *arXiv preprint arXiv:2406.12330* (2024).

[37] Zakaria, Yahia A., et al. "Propagation measurements and calculation of path loss exponent for outdoor cellular communication systems at 3.5 GHz." Radioelectronics and Communications Systems 64 (2021): 247-254.

[38] OpenStreetMap (https://www.openstreetmap.org/).

[39] Chalke Valley – Whites (https://hwwhite.co.uk/the-chalke-valley/).

[40] National LIDAR Programme (https://www.data.gov.uk/dataset/f0db0249-f17b-4036-9e65-309148c97ce4/national-lidar-programme)