Summary of the thesis for the degree of Master in Engineering

Power consumption evaluation of mobile radio access networks using a bottom-up approach: Modeling 4G networks and prospections of 5G deployment in Belgium

Author: Louis Golard
Supervisors: David Bol and Jérôme Louveaux
ICTEAM/ELEN, Université catholique de Louvain, Louvain-la-Neuve, Belgium
{louis.golard,david.bol,jerome.louveaux}@uclouvain.be

Abstract Despite the need to mitigate global warming, environmental impacts of mobile communication technologies tend to grow with the dramatic increase of mobile data traffic. Among all components of mobile networks, a significant part of power consumption is generated by base stations constituting the radio access networks (RANs). For the last few years, it is expected that 5G can reduce the energy consumption of RANs while supporting the expected growth of data traffic. However, mainstream research in this field focuses more on energy efficiency than on total power consumption, and the few existing RAN models become outdated. Therefore, this work aims to estimate the total energy consumption of broadband RANs in Belgium in 2020, and to forecast it by 2025 using six scenarios of 5G deployment. Models of current 4G networks and base station power consumption are determined based on data and field measurements from a Belgian mobile operator. The obtained relationship between the base station power consumption and the average traffic is linear with a power offset. This work also shows that 4G base stations are lightly loaded on average and that static energy consumption accounts for more than 80% of total RAN energy consumption in 2020. With a full 5G deployment, the annual mobile data traffic could reach over 1 400 PB in 2025 which is more than three times greater than in 2020. The related increase in total RAN energy consumption is more than 80%, but it can be limited to 27% if a sleep mode is implemented on 5G base stations. Indeed, the sleep mode feature could reduce the total power consumption of 5G base stations by about 60%, achieving a 10 times more energy efficient RAN in 5G than in 4G. Without 5G deployment, the total energy consumption of the 4G RAN only increases by 17% with a two-fold mobile data traffic growth. Excluding 4G decommissioning, scenario of no 5G deployment is the best in terms of RAN energy consumption, in spite of a lower overall energy efficiency. Using a rough estimate of base station embodied greenhouse gas emissions, there is also a clear upward trend of carbon footprint when the deployment of 5G is extensive. Hence, the main challenge for today's mobile networks is to reduce their total power consumption and carbon footprint in order to achieve climate targets.

Keywords *Sustainability, Climate Change, Information and Communication Technologies, 4G, 5G, Carbon Footprint, Energy Consumption, Radio Access Network, Base Station, Power Model, Sleep Mode.*

1 Introduction

Since the beginning of the industrialization era, greenhouse gas (GHG) emissions from human activities have caused global warming of approximately 1.0°C. If it continues to increase at the current rate, global warming is likely to reach and exceed 1.5°C between 2030 and 2050 [1]. The Paris Agreement adopted in 2015 aims to limit global warming to well below 2.0°C compared to pre-industrial levels, and preferably to 1.5°C [2]. Achieving this target requires to decrease the global net GHG emissions and to become net zero by 2050 [1]. Recently, the European Union have strengthened its climate ambitions and adopted the European Green Deal which sets a net GHG emissions reduction target of 55% by 2030 compared to 1990 levels [3].

Nowadays, the information and communication technologies (ICT) are responsible for an annual electricity consumption of approximately 1 400 TWh (6-7% of the annual global electricity consumption) and a carbon footprint of about 1.4 $GtCO_2e$ (3.5-4% of the annual global GHG emissions) [4]. Among ICT, the mobile wireless communications (better known as 2G, 3G, 4G and now 5G) show a clear carbon footprint growth over the last decade, while the footprint trend is less clear for the whole ICT sector [5, 6, 7, 8]. The infrastructure of mobile technologies can be divided in four segments: (i) the end-user terminals, (ii) the radio access network (RAN), (iii) the core network, and (iv) the data centers.The RAN and the core network are specific to mobile technologies and belongs to mobile operators, but the end-user terminals also embed other communication protocols (e.g. WiFi and Bluetooth), and the data centers communicate with the Internet also via fixed networks. The RAN is responsible for two third of the electricity consumption and one third of the carbon footprint of this sector. These figures could most likely continue to grow in the coming years [7, 9]. Accordingly, initiatives aimed at energy and carbon footprint reduction of mobile communications should focus on the RAN.

Mainstream research in mobile communications is mainly focused on improving the energy efficiency (in J/GB or Wh/GB) without deeply analyzing the absolute energy consumption (in J or Wh). Therefore, the carbon intensity of mobile communications (in gCO_2e/GB) has dramatically decreased during the last decade but not their absolute carbon footprint (in gCO_2e) which has increased, due to the exponential increase of data traffic (in GB) [5, 7]. This trend could continue in the future given the growing use of mobile broadband Internet by human consumers, the advent of Internet of Things (IoT) and the development of new connected applications [10, 11]. Most of the mobile data traffic is currently supported by 4G networks, but the increase in throughput and the new constraints of real-time communications could saturate the existing networks and may require deployment of new base stations (BSs) and/or a new technology such as 5G [12, 13].

The main objective of this work is to evaluate the energy footprint of mobile communication networks currently in use, and to forecast a trend for the coming years. More precisely, this work focuses on the total electricity consumption of the broadband RAN in Belgium in 2020 and 2025 under different scenarios of 5G deployment and mobile data traffic evolution. The assessment of the RAN uses a bottom-up approach, dividing the large-scale deployment over the country into several simpler small-scale systems. This methodology is adapted from an existing framework proposed by the EARTH project [14, 15]. The main contribution of our work is the modeling of BS power consumption as a function of the average data traffic, without going through signal propagation details. This allows to directly use an hourly data traffic profile as input of the RAN model. Power models of 4G BSs are determined using real measurements provided by a Belgian operator.

This work focuses only on the RAN and does not consider the impacts of the Internet backbone, the data centers, the core network nor the end-user devices. Moreover, this work does not perform a complete life cycle assessment (LCA) of BSs and the indicators of environmental impacts are limited to the electricity consumption and the carbon footprint including both the production and use phases. Also, this work only focuses on direct impacts of the RAN and does not consider the indirect effects of mobile technologies that can be positive (e.g. reducing transport emissions through connected vehicles) or negative (e.g. premature renewal of smartphones) but that are subject to many highly uncertain assumptions [16, 17] and are out of the scope of the study.

The following of this document is organised in three main sections. Section 2 explains the general modeling methodology used to model a countrywide RAN. Section 3 investigates the practical implementation of 4G RAN in Belgium today. Based on information from a Belgian operator, different types of BSs are analyzed in different deployment areas, and measurements are used to model the power consumption of 4G BSs as a function of the average data traffic. A prospective power model for 5G BSs is then proposed using a scaling of 4G power models. Section 4 defines different scenarios of macroscopic 5G deployment and mobile data traffic evolution by 2025. Accordingly, the total energy consumption and the carbon footprint of the RAN is evaluated in 2020 and for each scenario of 2025.

2 General methodology

The countrywide RAN model must be flexible and configurable such that it can be used to analyze both the current 4G RAN and different possible future 5G RANs. While a top-down approach would be used for 4G networks, it is not yet the case for 5G ones which are not already deployed on a large scale. Therefore, bottom-up modeling is more appropriate for a prospective evaluation. This section first introduces a framework classically used for RAN power consumption evaluation and then explains how it is adapted in this work.

2.1 The energy efficiency evaluation framework

To assess the RAN power consumption, the EARTH project has defined the energy efficiency evaluation framework (E^3F) [14, 15] that comprises two essential steps. The first step is to evaluate different *small-scale short-term* systems using statistical traffic models and specific small-scale deployment scenarios. The second step is to use *long-term* traffic models (over a full day) and *large-scale* deployment maps (over a whole country), to combine the small-scale short-term systems together and to obtain overall metrics. In the original paper, the E^3F was applied to a LTE network deployed within a representative European country.

2.1.1 Small-scale short-term evaluations

Traffic fluctuations in small-scale areas and short-term periods are modeled statistically, reflecting the packet transfer of different type of devices and applications (e.g. file downloading or video streaming). Classically, homogeneous space characteristics are assumed with regular hexagonal cells and uniformly distributed users. Then, the corresponding transmitted power of the BS is calculated and a power consumption model is applied

to compute the instantaneous power consumption. The E^{3}F considers different power models for various BS types.[15]. Hence, it requires to estimate the transmitted power of the BS according to the traffic load.

In wireless communications, the propagation of a signal from a BS to a user is subject to channel deterioration and the achievable throughput is constrained by interference from neighboring BSs. The power received by the device depends on the distance between the BS and the user, the carrier frequency, the path loss characteristics, the transmitting and receiving antennas gains, the shadowing, etc. To accurately model the signal propagation and the corresponding link budget, complex path loss models must be used, using parameters such as the type of urbanization, the BS height, the terminal speed, etc.

Small-scale short-term evaluations are conducted for all studied deployment scenarios (i.e. cities, suburbs and villages environments), and for a representative set of traffic load. These system-level evaluations provide power consumption and other performance metrics (e.g. the throughput) for each small-scale deployment and short-term traffic loads. These computations can be carried out by a system-level simulation platform.

2.1.2 Large-scale deployment and long-term traffic

The large-scale deployment modeling analyzes the RAN structure over a country according to the population density. With this approach, the main assumption is that the potential to generate traffic in an area is directly related to the population density of that area. E^{3}F divides the country into six small-scale deployment areas: super dense urban (SDU), dense urban (DU), urban, suburban and rural areas, as well as wilderness [18]. SDU reflects city centers with extreme traffic peaks while wilderness corresponds to areas that are sparsely populated.

The long-term traffic modeling establishes the profile of the traffic demand over time in the different deployment areas identified here above. In [15] and [18], a typical traffic profile for Europe is given with a peak in the evening and a valley during the night. If we consider that the data volume generated per subscriber does not depend on the area category, the areal data traffic demand is directly proportional to the population density.

For each small-scale deployment, long-term power consumption and other performance metrics are computed by weighted summing of the short-term metrics, according to the daily traffic profile. After, a weighted sum over the considered mix of deployment areas yields to the overall metrics for the large-scale deployment.

2.2 Evaluation methodology used in this work

The modeling of signal propagation from BSs to users is a complex and multi-variable issue where performances are sensitive among other things to the scheduling algorithms and to the modulation/demodulation techniques. To avoid carrying out such simulations, the small-scale short-term evaluations of this work use empirical models that directly link the BS power consumption to the data traffic. The complete model of the system is given by Figure 1.

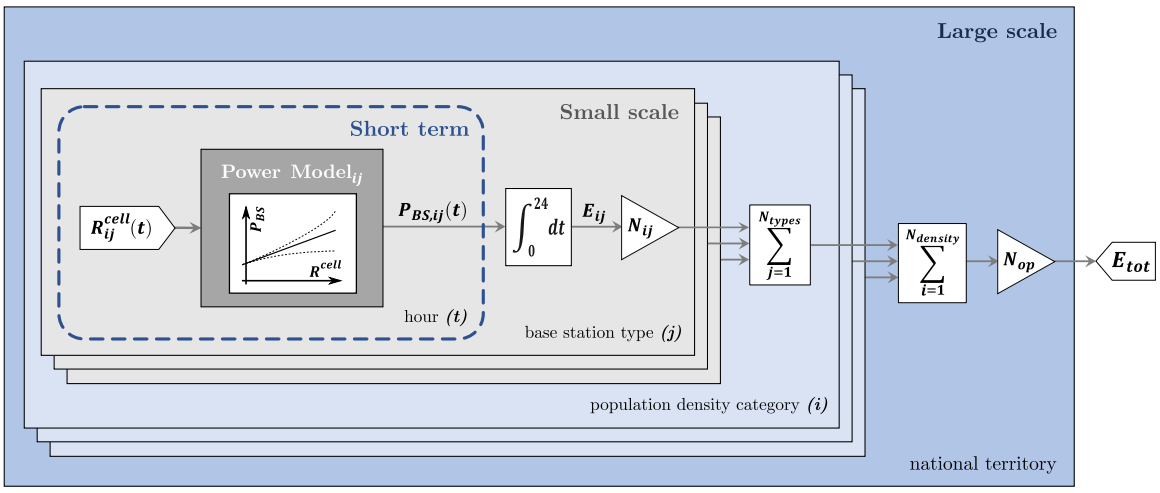


Figure 1: System model used in this work to evaluate the total energy consumption of several operators' RANs (4G or 5G) for a long-term (one day) and a large-scale (territory of a country) deployment scenario.

The input of the system is the provided data traffic in the cells. The traffic served by a BS is considered equivalent to the sum of the traffic generated by the users connected to that BS [19]. Using the instantaneous

traffic transmitted by a BS to determine its instantaneous power consumption is certainly not correct but the influence of each individual instantaneous channel on the average power consumption is reduced by averaging the traffic over several users. This effect is even more pronounced when the average is taken over a longer period of time. It is therefore assumed that the relationship between the average traffic in the cell during one hour and the average power consumption of the BS can be directly modeled with a limited error. The traffic is expressed by the throughput $R_{ij}^{cell}(t)$ (in Mbps) and $P_{BS,ij}(t)$ (in W) is the average BS power for time intervals t of one hour. The relationship linking throughput and power consumption is not known a priori and is modeled in Section 3.2.

The rest of the methodology is more similar to the E^{3}F. $R_{ij}^{cell}(t)$ varies every hour according to a daily profile and depends on the population covered by the cell. From the short-term to the long-term level, the average BS power for each hour of the day is integrated to obtain the daily BS energy consumption E_{ij} (in kWh). From the small-scale evaluations to the large-scale one, the large-scale characteristics of the RAN are taken into account. First, there is a number N_{types} of BS types with different features and thus different power versus traffic models. Second, the territory is divided into $N_{density}$ categories of deployment areas depending on the population density. The subscript i signifies the population density category and the subscript j represents the BS type. For each combination (i, j), there is a BS configuration that is present a number of times N_{ij} in the national RAN. As a result, considering a number of operators N_{op} with identical network deployment, the daily energy consumption of a BS and of the national RAN are respectively:

$$E_{ij} = \int_0^{24} P_{BS,ij}(t)\,\mathrm{d}t \tag{1}$$

$$E_{tot} = N_{op} \sum_{i=1}^{N_{density}} \sum_{j=1}^{N_{types}} N_{ij} E_{ij} \tag{2}$$

It is considered that the daily traffic profile remains unchanged whatever the day of the week and the period of the year. Therefore, the annual energy consumption is equal to E_{tot} multiplied by 365 days.

3 Modeling of 4G and 5G RANs

This section builds models of current 4G and future 5G RANs in Belgium according to the methodology presented in Section 2. For this purpose, the deployment of a Belgian operator's network is analyzed, mainly with respect to the population distribution over the country. Then, empirical power models of 4G BSs are established using field measurements and prospective models of 5G BSs are proposed by scaling the 4G models.

3.1 Large-scale deployment of 4G networks

The large-scale deployment of Belgian RANs must be divided into a limited but representative number of simpler small-scale systems. Hence, the general characteristics of Belgian mobile networks are analyzed in this section and the main features of the 4G RAN of one particular operator are extracted.

3.1.1 Overview of 4G deployment in Belgium and related assumptions

In Belgium, there are three mobile network operators (MNOs) that have their own mobile network infrastructures: Proximus, Orange and Base/Telenet. There are also mobile virtual network operators (MVNOs) that are virtual operators that do not own physical mobile networks and have to operate on the infrastructure of MNOs: e.g. LycaMobile, Mobile Vikings or JIM mobile. MNOs are dominant compared to MVNOs in terms of number of SIM cards. Proximus is the market leader with around 40% of active SIM cards while Orange and Base/Telenet represents between 25% to 30% of the market each. MVNOs cover the almost 10% remaining [20].

In this work, a *site* refers to a mobile network facility owned by an operator, and the term *base station* refers to the primary node of the RAN specific to a mobile generation for an operator on a physical site. Thus, 4G and 5G BSs are studied independently, but a single BS may support one or more frequency bands of the same mobile generation. Using databases for Wallonia [21], Brussels [22] and Flanders [23], the distribution of sites appears to be relatively similar with almost the same number of sites for each MNO, even if all they do not have the same market share. One reason for this is that a mobile network primarily provides coverage over the whole territory before focusing on the traffic load induced by the number of subscribers [15]. A strong assumption made here is that the deployment strategies of all operators are identical, regardless of their market share. Therefore, it is assumed that each operator has the same number of BSs N_{ij} for each configuration (i, j). Furthermore, only MNOs have BSs that consume energy and thus $N_{op} = 3$. As a result, to determine the N_{ij} coefficients, it is sufficient to analyse only one of the three networks of MNOs.

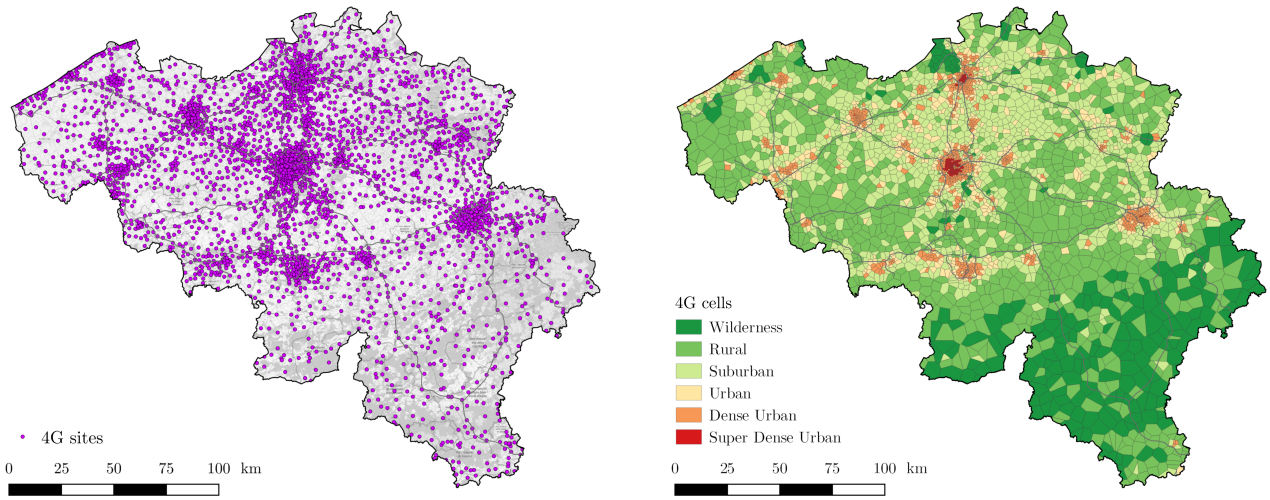
The Belgian Institute for Postal Services and Telecommunications (BIPT) also provides maps of 4G coverage over the whole territory [24]. The coverage is characterized by the signal strength categorised into 3 levels: satisfactory, good or very good. A very large part of the territory is covered by a satisfactory level of 4G signal for the three MNOs and the few areas that are not covered by 4G are mainly the large forests of the Ardennes region, which are unpopulated. Thus, in this work, full 4G coverage is assumed for the three RANs of MNOs.

Finally, most of the installed 4G BSs are of macro size supporting several sectors, usually 3. A few micro BSs are listed in the databases, deployed in subway stations of Brussels and in the Heysel exhibition halls [22] but their number is marginal. Hence, they are excluded from the modeling and only macro BSs are studied.

3.1.2 Analysis of one particular 4G RAN

Thanks to the assumption that all operators have identical 4G networks, only the RAN of one operator is analyzed to determine the types of 4G BSs deployed in Belgium, their characteristics and their respective number (i.e. the coefficients N_{ij}^{4G}). Each site of this operator is listed in a database which contains their GPS coordinates, their available carrier frequencies and their multiple-input multiple-output (MIMO) configuration.

Geographical distribution of 4G sites The geographical distribution of the 4G RAN is analyzed using the GPS coordinates of each site. Figure 2(a) shows the deployment of 4G sites of one operator in Belgium. Knowing the geography of Belgium, we observe that more 4G sites are located in large cities and along major roads (e.g. along highways) to provide service despite the higher speed of vehicles (high speed induces signal deterioration). The country is then divided into cells that are Voronoi zones[1] with 4G sites as centroids. The borders of Belgium are the system boundaries. This approach relies on the assumption of full 4G coverage and on the additional assumption that the size of one cell does not depend on the type of site deployed on it. For example, a site equipped with 3 bands does not have a larger (or smaller) cell than if it had been equipped with 1 or 2 bands only. Next, the population covered by each site is estimated by integrating the number of inhabitants included in its respective cell, based on demographic data reported for a grid of squares of 1 km^2 over the whole Belgian territory [25]. The population density of each cell is also calculated by dividing its estimated population by the area of its Voronoi zone. Finally, each 4G site is classified according to the population density of its cell. Figure 2(b) depicts the division of Belgium into Voronoi zones and the classification of the cells according to their population density. The same categories as E^{3}F (see Section 2.1) are considered such that $N_{density} = 6$. The classification intervals and average densities for each population density category are given in Table 1.



(a) 4G sites (b) 4G cells

Figure 2: Geographical distribution of 4G sites in Belgium and division of the territory into Voronoi zones with 4G sites as centroids. Each site is classified based on the population density computed over its whole cell.

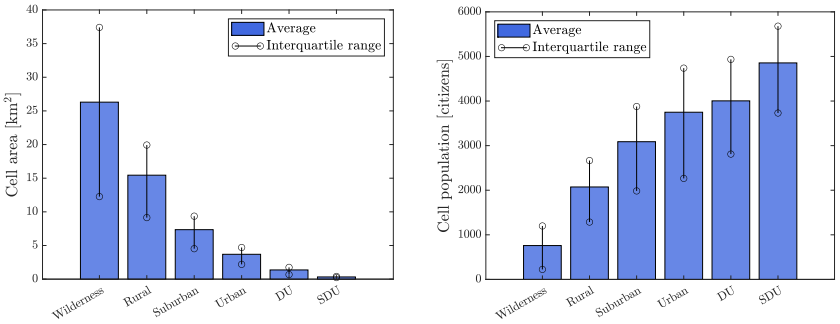
The average cell area and cell population are given in Figure 3(a) and in Figure 3(b) respectively, for each population density category. A relationship between these parameters and the population density clearly exists which confirms the assumption that large-scale modeling relates to the population structure of the country. The size of the cells decreases with increasing population density: from more than 25 km^2 in the wilderness

[1] A Voronoi zone around a centroid is the subset of points in space that lie closer to that centroid than any other. The boundaries of each Voronoi zone are the median lines between each pair of neighboring centroids.

to less than 1 km^2 in DU and SDU. In correspondence, the population covered by a site goes from about 750 in the wilderness to more than 4850 inhabitants in SDU. Yet, this total population does not directly represent the potential users. Indeed, not all inhabitants are users of mobile broadband networks (4G penetration ratio estimated to be about 82% in 2020 [20]), and not all users use the same operator. Assuming three identical operators, only one third of the population using 4G is potentially a user of one operator.

Cell category	**Population density [citizens/km^2]**	
	Interval	Average
Super dense urban	$\geq 10\,000$	14 828
Dense urban	$[1\,500; 10\,000[$	2 925
Urban	$[750; 1\,500[$	1 015
Suburban	$[250; 750[$	420
Rural	$[50; 250[$	134
Wilderness	< 50	29

Table 1: Classification of cells into six categories of population density using non-uniform length intervals.



(a) Cell area (b) Cell population

Figure 3: Average and interquartile range of 4G cell characteristics

Configuration of 4G sites Among all 4G sites, 27% are equipped with 1 frequency band, 51% with 2 bands and 21% with 3 bands. The 4-band sites currently represent less than 1% of 4G sites and this type is ignored for the modeling. To still count their power consumption, these sites are classified as 3-band sites. All the frequency bands use a frequency division duplex (FDD) paired spectrum. This means that for each carrier, one part of the spectrum is allocated to downlink (DL) and another part, of the same bandwidth, to uplink (UL) [26]. The most common 4G frequency carriers are 800 MHz with 2×10 MHz (10 MHz for DL and 10 MHz for UL), 1800 MHz with 2×20 MHz and 2600 MHz with 2×20 MHz. The 800 MHz band is said to be the *coverage band*, with better signal propagation capability. The additional bands are the *capacity bands* that increase the overall available bandwidth on the site, and thus, the maximum available throughput for devices capable of carrier aggregation. The 800 MHz band is the first to be used when only one band is deployed on a site, then the 1800 MHz band is added for the 2-band sites and finally the 2600 MHz band for the 3-band sites.

In cellular networks, each cell of a macro site is divided into one or more sectors. This means that the same bandwidth can be reused in each sector by using directional antennas (as opposed to omni-directional antennas which transmit in all directions). In the analyzed RAN, more than 92% of the macro-sites use 3 sectors, as it is classically the case. Some sites with 2 or 4 sectors also exist but are more uncommon. For the modeling, it is assumed that all sites use 3 sectors.

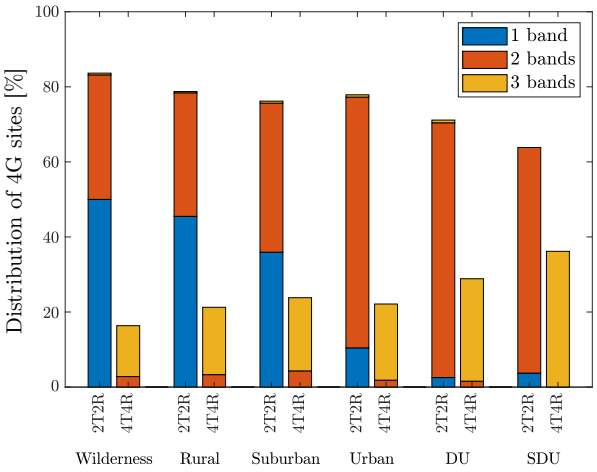


Figure 4: Normalized distribution of 4G sites according to their number of frequency bands, population density category and MIMO configuration.

There are also two main MIMO configurations for the BS antennas: 2T2R and 4T4R. A 2T2R antenna has one large beam comprising two cross-polarised channels for transmit and receive diversity, resulting in a total of 2 streams. With 4T4R antennas, there are two smaller beams with two cross-polarised channels per beam, making 4 streams overall [27]. Figure 4 shows the normalized distribution of 4G sites according to their number of bands and MIMO configuration for each population density category. Two important observations can be made. The first one is that the more densely populated the area, the fewer 1-band sites and the more 3-band sites are deployed (offering more overall bandwidth and thus more capacity). The second one is that a large majority of the 3-band sites use 4T4R (98%) while the two other types mainly use 2T2R. This is due to the more recent deployment of 3-band sites with newer and more advanced antenna technologies. For the modeling, all 1-band and 2-band sites use 2T2R and all 3-band sites use 4T4R.

Despite different number of transmitting antennas, the total radiated power does not change between 2T2R and 4T4R antennas according to information from the operator. The power is just distributed over more streams. The maximum power of 2T2R antennas is 80 W per sector for the 20 MHz bands, with 40 W

transmitted per stream. With 4T4R, there is also 80 W per sector but in that case there is only 20 W per stream. It is also assumed that the same power per Hertz (equivalent to 4 W/MHz) is used at 800 MHz which gives a maximum emitted power of 40 W over the 10 MHz bandwidth. These values are in line with the ITU 4G RAN evaluation guidelines [28]. However, these maximum radiated power may be adapted depending on the regulations. For instance in cities, electromagnetic radiation limits impose further restrictions on the maximum emitted power. Yet, these regulations are not taken into account in the model.

Simplified characterisation of 4G base stations Considering that for each 4G site only one 4G BS is deployed with 1, 2 or 3 bands, three main BS types are defined such that $N_{types}^{4G} = 3$. The features of these BSs are provided in Table 2. The large-scale system is thus composed of $N_{density} \times N_{types}^{4G} = 18$ small-scale configurations (i, j). The cell area and cell population of BSs are the average values given in Figure 3 for each of the six population density categories. The actual N_{ij}^{4G} coefficients are kept confidential.

Base station type	Number of bands [#]	Frequency carrier [MHz]	Duplexing mode	Downlink bandwidth [MHz]	MIMO configuration	Maximum TX power [W]	Number of sectors [#]
1 band	1	800	FDD	10	2T2R	40	3
2 bands	2	800/1800	FDD	10/20	2T2R	40/80	3
3 bands	3	800/1800/2600	FDD	10/20/20	4T4R	40/80/80	3

Table 2: Main features of the three 4G base station types considered in this work.

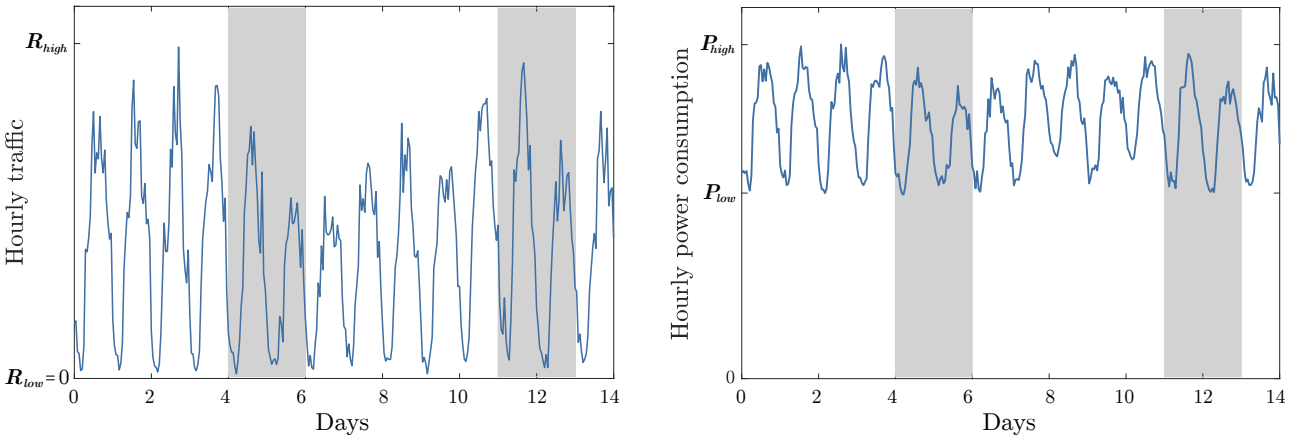
3.2 Power consumption models of 4G base stations

The power models of 4G BSs are first derived empirically based on field measurements. Then, these models are refined by considering the average cell capacity and extracting the main scaling parameters.

3.2.1 Empirical models using field measurements

For each small-scale configuration (i, j), a representative sample of 4G BSs is carried out. In particular, five BSs per configuration were randomly selected among all the 4G BSs of one operator, resulting in a sample size of 90 BSs (5 BSs $\times$ 18 configurations). For each of these BSs, measurements from the operator have been provided for the data traffic volume (in MBytes) and the energy consumption (in kWh) hour by hour during 14 days. These measurements are aggregated for all installed sectors of the 4G site. The recorded traffic is the sum of downlink and uplink data volumes on all 4G carriers, and the power consumption includes the digital baseband unit, the radio frequency transceivers and the power amplifiers, but excludes the active cooling system and the losses due to the AC/DC main power supply and DC/DC converters.

An example of such measurements for one 4G BS is shown in Figure 5. A periodic pattern is clearly visible and follows the day and night alternation. The hourly average data rate varies between a zero value and a high value, while the hourly average power consumption oscillates between a non-zero low value and a high value, which indicates a static power offset. The oscillation amplitudes of these two metrics are correlated: e.g. the peak traffic on Sundays is lower than on other days which is reflected in the power consumption.

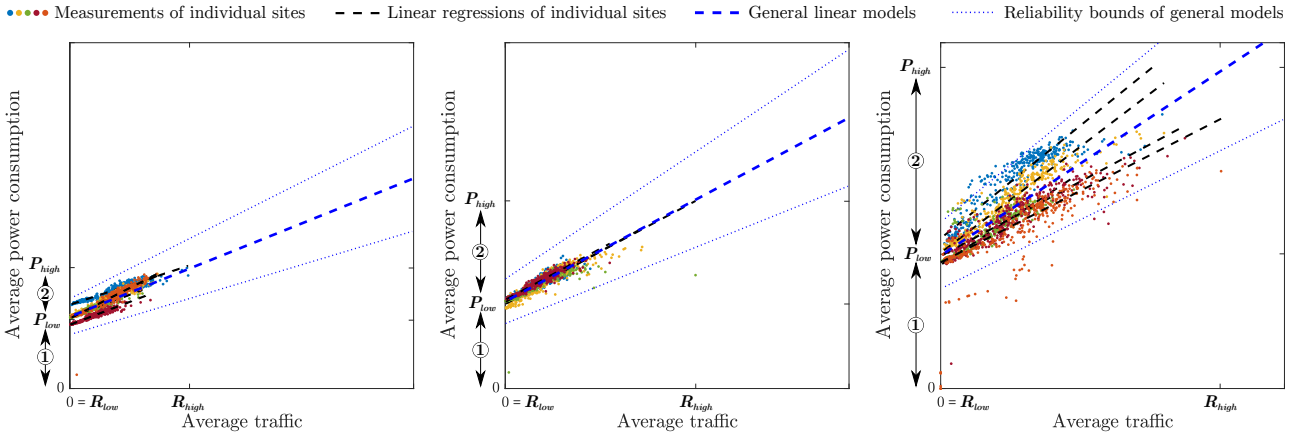


(a) Hourly mobile data traffic

(b) Hourly base station power consumption

Figure 5: Example of hourly measurements for one site during 14 days. Shaded areas highlight weekend days. Hourly traffic varies between $R_{low} = 0$ and R_{high}, and hourly power consumption between $P_{low} \neq 0$ and P_{high}.

For each site of the sample, a linear regression is computed between the power consumption and the traffic using a robust bisquare method to reduce the impact of outliers. Examples of linear regressions are provided in Figure 6 for sites belonging to the same population density category and grouped by their number of bands. These linear models consists of a constant term and a variable term that reflect the *static* power consumption (traffic-independent) and the *dynamic* power consumption (traffic-dependent), respectively. The resulting R^2 metric of all individual linear regressions is always between 0.85 and 0.95, meaning a good predictivity of the models and a nice linear relationship over the measured range[2]. The normalized root mean square error (NRMSE) of linear regressions is 2.5% on average, ranging from 0.4% to 5%, which is good enough to rely on the linear model. Moreover, no discontinuity is observed for the very low traffic hours, which suggests that there is no *sleep* mode implemented in these 4G BSs in no-load conditions (i.e. when there is no traffic).



(a) 1-band base stations (b) 2-band base stations (c) 3-band base stations

Figure 6: Measurements and linear power models of individual sites for each type of base station in suburban areas. The general model comprises a *static* power component ① and a linear *variable* power component ②. Reliability boundaries of ±25% surrounds the general model. The highest measured hourly traffic is R_{high}.

Despite the large variability in site characteristics of the sample, we can observe in Figure 6 that sites with the same number of bands belonging to the same category of population density behave generally in a similar way. Thus, a single model per configuration (i, j) can be carried out from models of individual sites. The model for each configuration (i, j) is defined by:

$$P_{BS,ij}\left(R_{ij}^{cell}\right) = P_{0,ij} + \Delta_{P,ij} \cdot R_{ij}^{cell} \tag{3}$$

where $P_{0,ij}$ is the power consumption (in W) at zero traffic load and $\Delta_{P,ij}$ is the variation of power consumption with traffic (in W/Mbps). The subscripts ij refer to the corresponding configuration (i, j). The model parameters of each configuration are estimated by averaging the corresponding parameters of all models of individual sites associated to this configuration.

The estimated model parameters and the associated fitting accuracy are kept confidential. In Figure 6, the resulting estimated models are plotted in blue. Maximum and minimum reliability bounds are arbitrarily set to ±25% of $P_{BS}\left(R^{cell}\right)$ and surround all site's models very well. The parameters P_0 are very similar for the same types of sites and whatever the deployment area. This is meaningful because the no-load power is not related to the signal propagation nor to path loss but depends on the BS design and its complexity. Therefore, the more bands installed on the site, the higher P_0 is. Moreover, the slope Δ_P is roughly comparable for deployment areas from wilderness to urban, but are lower in DU and SDU. It means that BSs consume less power for the same throughput in these areas, probably due to the significantly smaller cells.

3.2.2 Model extrapolation up to the maximum traffic load

In the literature, power models of BSs are usually expressed by a linear model as a function of the traffic load. This is consistent with the shape of our empirical models. In addition, the typical load region of the BS is generally far from the maximum load (i.e. $< 30\%$) [9, 29]. Therefore, in empirical models, only the typical traffic region is modelled through the measurements, because the operator makes sure that his BSs operate far from the saturation region (i.e. near the maximum traffic). In order to complete the models, the upper bound of the traffic load must be determined. In this work, this maximum operation point is called the *average capacity* and is defined as the maximum average throughput achievable by the BS during one hour when users

[2] R^2 is always between 0 and 1, the higher the better.

are uniformly distributed in the cell. In this section, the average spectral efficiencies (SEs) of each 4G frequency band are estimated and are then used to estimate the average capacities of each type of BSs.

To obtain the cell average SE of a site, throughputs of all users present in the cell would be known for many different positions. However, this type of data was not available for this work and another approach is proposed: every year, BIPT measures the customer experience of 4G networks via a *Drive Test Campaign* [30]. Using the raw measurements from this campaign, the average SEs are estimated for each carrier frequency and each population density category. Actually, the cell average SEs are rather stable whatever the category of population density: around 2 bps/Hz at 800 MHz and between 3 and 4 bps/Hz for the two other bands. These average SEs respect the ITU minimum requirement of 2.2 bps/Hz for LTE-Advanced macro-BSs [31].

Since $SE^{DL}_{avg,i,f}$ is the cell average downlink spectral efficiency (in bps/Hz) of population density category i and of frequency f, and considering a ratio α_{DL} of downlink traffic compared to the total traffic (i.e. DL + UL), the average capacity of a site of configuration (i, j) is given (in Mbps) by:

$$C^{cell}_{avg,ij} = \frac{N_{sect}}{\alpha_{DL}} \sum_{f=1}^{N_{bands,j}} BW_f \cdot SE^{DL}_{avg,i,f} \tag{4}$$

where N_{sect} is the number of sectors of the site, BW_f is the bandwith of frequency f and $N_{bands,j}$ is the number of bands of a site of type j. In 4G, the entire bandwidth is reused in each sector such that the frequency reuse factor is one. The actual bandwidth is 9 MHz at 800 MHz and 18 MHz at 1800 MHz and 2600 MHz.

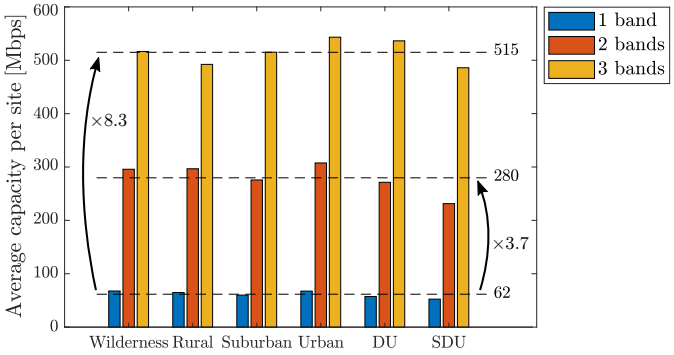


Figure 7: Average site capacities for each type of base station and each deployment area.

The average capacity of each type of site is given in Figure 7. For the same type of sites, the average capacity appears to be similar whatever the population density. This result may seem counter-intuitive as it could be thought that the average site capacity varies with the size of the cell. To explain this observation, we have to consider the effects related to the interference from neighbouring BSs and the variation of signal propagation due to urbanization. In any case, the 2-band sites have roughly 3.7 times the capacity of the 1-band sites, and the 3-band sites have around 8.3 times the capacity of 1-band sites.

The empirical 4G BS models can then be linearly extrapolated up to to their maximum load corresponding to the average capacity given by Figure 7. As expected, the average load during the hour of maximum traffic is well below the average capacity of sites. For 1-band sites, it corresponds to about 45% of the full load, for 2-band sites it is about 30% and for 3-band sites it is less than 20%. However, the projection of linear models beyond the measured range of traffic is prospective and uncertain. Indeed, there is no guarantee that the linear model continues up to the maximum load because non-linear saturation could occur, e.g. due to the saturation of power amplifiers, or a supra-linear effects could also appear, e.g. due to leakage currents in digital circuits.

3.2.3 Identification of the main scaling parameters

In Section 3.2.1, it is shown that P_0 and Δ_P change depending on the type of site, and that the model slopes are less steep in DU areas and even less steep in SDU areas. Based on these findings, the scaling parameters can be derived according to the characteristics of each type of BS. In [32], the main scaling parameters are the bandwidth, the spectral efficiency, the number of antennas and the quantization factor. Using a similar approach and limiting the number of degrees of freedom, the two main scaling parameters are:

- **Bandwidth (BW)**: This crucial parameter in wireless systems reflects the complexity of the digital baseband unit and the number of bands installed on the site. By assuming a uniform transmitted power density over the whole spectrum, it also reflects the dynamic part of the power amplifier consumption. In [32], it is also the most influential parameter.
- **Total number of streams ($Streams$)**: This parameter is the product of the number of streams per carrier by the number of carriers installed on the site. For example, for a 2-band site with 2T2R antennas, $Streams = 2 \times 2 = 4$, and for a 3-band site with 4T4R antennas $Streams = 3 \times 4 = 12$. Thereby, this parameter takes into account the number of carriers in the site, the number of transmission chains and the number of power amplifiers influencing the static power consumption.

The scaling equation is the following:

$$P_{scaled} = P_{ref} \left(\frac{BW_{scaled}}{BW_{ref}} \right)^{s_{BW}} \left(\frac{Streams_{scaled}}{Streams_{ref}} \right)^{s_{streams}} \tag{5}$$

where P_{scaled} is the scaled power having properties BW_{scaled}, and $Streams_{scaled}$ and P_{ref} is the reference power with properties BW_{ref} and $Streams_{ref}$. The scaling exponents are s_{BW} and $s_{streams}$.

The scaling exponents are determined using different pairs of BSs with different combinations of parameters. The range considered for scaling exponents is between 0 (independence) and 1 (linear dependency). The resulting scaling exponents are given in Table 3. On one hand, the scaling is not very pronounced for the no-load power because P_0 scales according to the third root of the BW ratio ($s_{BW} = 0.36 \approx 1/3$) and only according to the tenth root of the $Streams$ ratio ($s_{Streams} = 0.1 = 1/10$). This means that when the bandwidth of a BS doubles (or when an additional band of the same bandwidth is added), P_0 does not double but only increases by a factor of $2^{0.36} = 1.28$. On the other hand, the maximum power evolves almost linearly with the bandwidth ($s_{BW} = 0.98 \approx 1$) and with the third root of the $Streams$ ratio ($s_{Streams} = 0.36 \approx 1/3$).

The scaling exponents relative to the maximum power P_{max} are defined using models for wilderness to urban areas ($P_{scaled,max}$ is assumed to be the same in these areas). For the DU and SDU areas, a correction factor must be applied to obtain the maximum powers as follows:

$$P_{max,ij} = \xi_i \cdot P_{scaled,max,j} \tag{6}$$

where ξ_i is the correction factor for deployment area i. The correction factors are given in Table 4. Thus, P_{max} in DU is 25% lower than in areas between wilderness and urban, and almost 50% lower in SDU.

Power load	s_{BW}	$s_{Streams}$
No-load power P_0	0.36	0.10
Maximum power P_{max}	0.98	0.36

Table 3: Scaling exponents for no-load and maximum power.

Area type	ξ_i
Wilderness → Urban	1
Dense Urban	0.75
Super Dense Urban	0.52

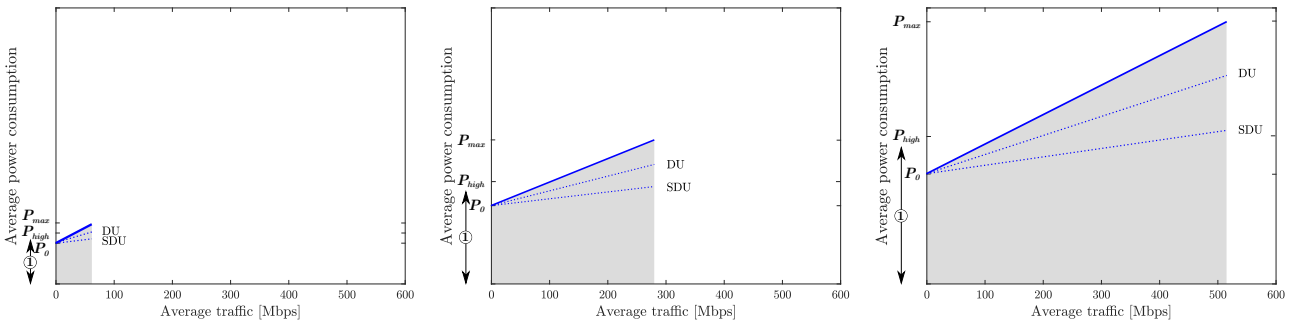
Table 4: Correction factor for maximum power.

3.2.4 Complete models of 4G base stations

Finally, the overhead due to the AC/DC main supply, the DC/DC converters and the active cooling is modelled using three efficiency coefficients, respectively η_{MS}, η_{DC} and η_{cool}, as in [32]. Moreover, the average capacity of BSs of the same type is considered constant across all deployment areas and hence only the type of the BS affects it. The same applies to the no-load power, which explains the drop of the subscript i for these two parameters. The complete parametric power model of 4G BSs as a function of the throughput is given by:

$$P_{BS,ij}\left(R_{ij}^{cell}\right) = \frac{1}{\eta_{MS} \cdot \eta_{DC} \cdot \eta_{cool}} \cdot \left(P_{0,j} + \frac{P_{max,ij} - P_{0,j}}{C_{avg,j}^{cell}} \cdot R_{ij}^{cell} \right) \tag{7}$$

For the modeling, the three efficiency coefficients are considered to be the same and are equal to 92% [32, 33]. The power models for the three types of 4G BSs considered in Belgium are shown in Figure 8.



(a) 1-band base stations (b) 2-band base stations (c) 3-band base stations

Figure 8: Complete power consumption models for each type of 4G base station as a function of the average traffic. Maximum ranges of power and traffic are very different in each case. The measurement range ① covers only a part of the maximum operating range and the linear projection beyond P_{high} and up to P_{max} is uncertain.

3.3 Prospective modeling of 5G base stations

Power models for 5G BSs are built prospectively under several assumptions. The same parametric model as for 4G BSs is used by scaling the parameters and considering technological improvements.

3.3.1 Assumptions about 5G base stations in the near future

In Belgium by 2025, 5G will be deployed but the detailed characteristics of 5G RANs are not yet known. To apply the evaluation methodology of this work, this section investigates the different possibilities of implementation for the 5G and define the most likely characteristics of 5G BSs.

Frequency carrier and bandwidth In 5G NR, two frequency ranges (FRs) can be used: FR1 (below 6 GHz) and FR2 (millimeter waves) [13]. Currently, BIPT has no plans to allow the use of FR2 in a short term in Belgium and this range is excluded from the following considerations [34]. In FR1, operators have typically the choice to use the bands of 700 MHz, 2100 MHz or 3500 MHz. The choice made in this work is to study only the 3.5 GHz band for 5G because it is the first band that allows for "*true*" 5G. Moreover, BIPT has not currently granted the 700 MHz band to operators while it has authorised the 3.5 GHz band [20]. At that frequency, 400 MHz of bandwidth is available but the three MNOs have only access to 50 MHz each. Hence, the bandwidth considered in this work is 50 MHz per operator. At 3.5 GHz, time division duplex (TDD) is more likely to be used with a division between UL and DL according to predefined frame structures. Thus, the spectrum is unpaired, unlike the 2×10 MHz or 2×20 MHz FDD paired spectrum in 4G. With this configuration, DL and UL can be asymmetric, which is useful with the asymmetric use of networks by mobile users. In general, the system configuration gives about 75% of time for DL and the remaining 25% of the time for UL [35, 36, 37].

Coverage The challenge of 5G is to provide a good coverage at a higher frequencies than 4G. These frequencies suffer from signal attenuation, mainly due to the smaller effective receiving antenna area, greater outdoor path loss and higher indoor penetration loss. Hopefully, it can be mitigated by using directive antennas on both transmitters and receivers, forming beams with strong gain in certain directions and limiting the emitted power in other directions. Considering an increase in antenna gain at the BS and at the mobile terminal, the resulting beamforming gain could compensate for the propagation losses at 3.5 GHz and 5G should be able to provide a coverage at least as good as 4G [36, 37]. On the field, Ericsson has conducted tests which have shown that the beamforming gain is lower than expected in reality, but still significant. They also show that massive-MIMO at 3.5 GHz provides a similar or better downlink coverage than at 1.8 GHz, which prove the feasibility of deploying 5G at existing 4G sites [35]. As a result, Nokia and Ericsson claim that the need for site densification with 5G is not as important as initially foreseen. However, even with massive-MIMO, they pointed out that uplink coverage at 3.5 GHz is worse than at 1.8 GHz, meaning that, in worst cases, mobile devices could require the support of a lower 4G frequency band for uplink transmissions [35, 36, 37]. For operators, the reuse of existing 4G sites to deploy 5G BSs would also drastically reduce costs and time for the roll-out. In this work, it is therefore assumed that there will be no need to increase the density of 5G sites at 3.5 GHz.

MIMO configuration At 3.5 GHz, BSs are still of macro size with 3 sectors, using active antennas. According to Ericsson, a good trade-off between complexity and performance is to use antennas with 16 to 32 radio chains in urban areas and between 8 and 16 radio chains in rural and suburban areas [38]. CommScope also claims that 16T16R could be the most feasible MIMO configuration for active antennas at frequencies below 6 GHz. In addition, when antennas are upgraded from 16T16R to 64T64R, the capacity increases by 50% but the costs increase by more than 150%. Hence, it does not seem attractive for operators to start their 5G roll-out with antennas more complex than 16T16R [27]. The MIMO configuration considered here is therefore 16T16R.

Table 5 summarises all the characteristics of the 5G NR BS studied in this work for a deployment in Belgium by 2025. This single type of base station is used for all population density category.

Base station type	**Number of bands [#]**	**Frequency carrier [GHz]**	**Duplexing mode**	**Overall bandwidth [MHz]**	**MIMO configuration**	**Maximum TX power [W]**	**Number of sectors [#]**
1 band	1	3.5	TDD 3:1	50	16T16R	200	3

Table 5: Main features of the single 5G base station type considered in this work.

3.3.2 Scaling of 4G power consumption models to 5G base stations

The scaling method defined for 4G BSs in Section 3.2.3 is used to build a power consumption model for 5G BSs. Of course, this scaling remains prospective and is not yet validated by on-site measurements, contrary to 4G models. The results obtained may therefore be open to discussion and are subject to uncertainty.

Scaling of no-load and maximum power The parameters BW and $Streams$ are used to scale the model from 4G to 5G, because the bandwidth and the number of streams defined for the 5G BS are not too far from those of the 4G 3-band BS. Hence, it is reasonable to assume that the scaling exponents of Table 3 remain valid in these conditions. For the DU and SDU deployment areas, the power correction factor ξ_i given by Table 4 is applied to maximum power, meaning that the behavior of the model with respect to population density remain the same for 4G and 5G. However, this may not be the case because the MIMO technique used in 5G is much more spatially selective, but without anything better, these factors are kept.

Technological evolutions A new parameter is used to address technological evolutions between 4G and 5G BSs. According to the trends presented in [32], a technological step induces a power consumption reduction of $\sqrt{2}$ for the analog circuitry and of 2 for the digital components, as well as a 4% decrease of power amplifier losses. Since the power breakdown of BSs is unknown, it is assumed that half of P_0 comes from analog components at no-load and the other half to digital components. The technological evolution factor is thus: $T^{5G}_{techno,0} = \frac{1}{\sqrt{2}} \cdot 0.5 + \frac{1}{2} \cdot 0.5 = 0.6$. At full load, the assumption is that one third of P_{max} is digital, one third is analog and the last third is for the power amplifiers. Therefore, the technological evolution factor becomes: $T^{5G}_{techno,max} = \frac{1}{\sqrt{2}} \cdot 0.33 + \frac{1}{2} \cdot 0.33 + 0.96 \cdot 0.33 = 0.72$. Finally, efficiency improvements are expected for the overhead in 5G. In [32], they consider that these efficiencies do not evolve and stay at 92%, but in [39] they assume a losses reduction of 1-2% between two generations. Huawei also claims that the power supply efficiency could reach 98% in its new BSs [33]. In this work, it is assumed that all efficiencies evolve from 92% to 95%, leading to an improvement of more than 9% (i.e. $(0.92/0.95)^3$) for the overhead. Altogether, the technology evolution achieved in 5G compared to 4G is 55% at no-load and 35% at full load.

By combining (5) and (7), the scaling equation from 4G to 5G becomes:

$$P^{5G}_{BS,i} = T^{5G}_{techno} \cdot \xi_i \frac{P^{4G}_{ref} \left(\frac{BW_{5G}}{BW_{4G}}\right)^{s_{BW}} \left(\frac{Streams_{5G}}{Streams_{4G}}\right)^{s_{streams}}}{\eta_{MS} \cdot \eta_{DC} \cdot \eta_{cool}} \tag{8}$$

where the $P^{5G}_{BS,i}$ is the power consumption of the entire 5G BS at either no-load or full load in deployment area i, T^{5G}_{techno} is the technology trend factor corresponding to the no- or full load, and P^{4G}_{ref} is the corresponding power consumption of the 4G BS. BW_{5G}/BW_{4G} and $Streams_{5G}/Streams_{4G}$ are the ratios of scaling parameters between 5G and 4G. For example, the scaling of P_0 from a 4G 2-band BS to a 5G BS involves: $(50/30)^{0.36}\,(16/4)^{0.1}$.

3.3.3 Average capacity of the 5G site

The adaptation of the model described by (7) requires to determine the average capacity of the 5G site. In 5G, it is often claimed that there are significant throughput gains due to higher bandwidth and improved SE. Unlike 4G, it is not possible here to rely on field measurements to estimate the cell average SE in 5G. As an alternative, simulations made by manufacturers [40, 41] are used, and the cell average SE in 5G is set to $SE^{5G}_{avg,DL} = 10$ bps/Hz and $SE^{5G}_{avg,UL} = 6$ bps/Hz in downlink and uplink respectively, regardless of population density. Remembering the TDD scheme of 75% DL and 25% UL, the average capacity of a 5G BS is:

$$C^{cell,5G}_{avg} = N_{sect} \cdot BW_{5G} \cdot \left(0.75 \cdot SE^{5G}_{avg,DL} + 0.25 \cdot SE^{5G}_{avg,UL}\right) \tag{9}$$

With a bandwidth of 50 MHz, the average capacity of a 3-sectors 5G BS is 1350 Mbps. This represents an increase of the average capacity of 160% compared to a 4G 3-band BS which has the same total bandwidth.

3.3.4 Sleep mode feature

A very interesting feature of 5G BSs is the sleep mode which could reduce the idle-state power consumption. When there is no traffic, the sleep mode sequentially disables components of the BS with deeper sleep-depths over time. This process is highly non-linear [32, 42]. Several papers investigate the achievable power savings with sleep mode for different traffic loads and reveal decreasing benefits with higher load [42, 43, 44]. In this work, sleep mode savings are modelled by a decaying exponential function of the form $0.75\exp(-10 \cdot load)$, resulting in a concave non-linear model that starts at the deepest sleep state at no-load and that tends to the linear model at high load. This means that during non-traffic hours, the use of a sleep mode can reduce power consumption by 75%. By applying the sleep mode to the linear model of 5G BSs, it becomes:

$$P^{5G,SM}_{BS,i}\left(R^{cell}_i\right) = \left(P^{5G}_{BS,0} + \frac{P^{5G}_{BS,max} - P^{5G}_{BS,0}}{C^{cell,5G}_{avg}} \cdot R^{cell}_i\right) \cdot \left(1 - 0.75\exp\left(-10 \cdot \frac{R^{cell}_i}{C^{cell,5G}_{avg}}\right)\right) \tag{10}$$

where $P^{5G,SM}_{BS,i}\left(R^{cell}_i\right)$ is the power consumption of a 5G BS with sleep mode as a function of the average traffic, and $P^{5G}_{BS,0}$ and $P^{5G}_{BS,max}$ are respectively the power consumption at no-load and at full load without sleep mode.

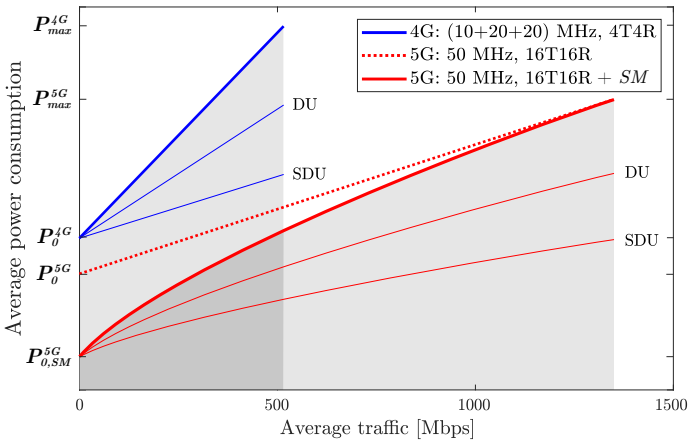


Figure 9: Power models of the 5G base station with and without sleep mode compared to the 4G 3-band base station.

Figure 9 shows the prospective model of 5G BS with and without sleep mode compared to the model of the 4G 3-band BS, which is the most advanced 4G BS. It illustrates the increase in capacity along the horizontal axis, and the power savings at zero and maximum load. The non-linear effect induced by the sleep mode is also visible and is stronger at low load where $P^{5G}_{0,SM} = 0.75 \cdot P^{5G}_{0}$. At high load, the model with sleep mode tends to the linear model without sleep mode and P^{5G}_{max} is the same with and without sleep mode. The values of the model parameters used in this work for 5G BSs are kept confidential.

3.4 Small-scale traffic profile

The small-scale traffic is the traffic transferred by a single BS. In our model, this traffic is constant in the short term, i.e. during intervals of one hour. In the long term, the traffic profile $R^{cell}_{ij}(t)$ characterizes the variations of the small-scale traffic during the 24 hours of a day. Once again, the measurements provided by the operator for the sample of 4G sites (see Section 3.2.1) are used.

To compare the traffic profiles of all the different sites, an approach is to normalize the traffic volume of each site by the number of its potential users. This is consistent with the initial assumption that traffic is generated directly by the residential population. As the traffic load of the studied operator is not strictly the same as the two others, a correction factor is applied to homogenise the traffic profile between operators, based on the total volume of mobile data transferred in Belgium. The obtained normalized profiles per inhabitant are in the same order of magnitude for all the different BS of the sample, and their shapes are similar whatever the type of BS and the deployment area. Nevertheless, a non-negligible variance is still observed due to the large variability of the sites studied, both in terms of cell size, number of users and consumer behaviour. In addition, weekends and weekdays do not usually have the same traffic because of changes in consumer behavior. The amplitude of the daily average profile also increases with the number of bands installed on the site and hence, the traffic profiles $R^{cell}_{ij}(t)$ must be different for each configuration (i, j). Besides, a general pattern of traffic can be extracted for the three types of 4G BSs. This profile is considered to be the same over the year, without any distinction between weekdays and weekends, or holiday periods.

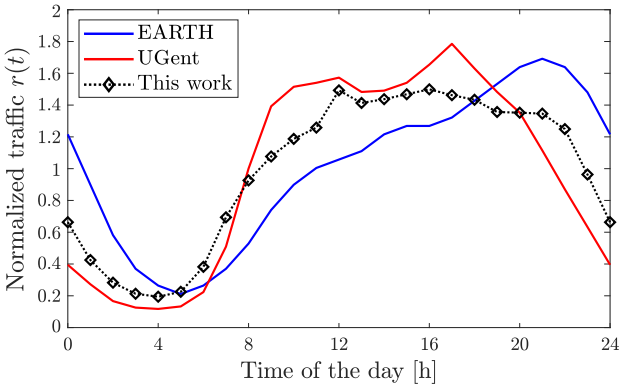


Figure 10: General traffic profile $r(t)$ using data from the Belgian operator, compared to two other profiles.

In Figure 10, the resulting traffic profile is compared with two other normalized profiles taken from the literature [15, 45]. These traffic profiles are further normalized to have a unit average throughput over the whole day. All the three profiles indicate the same trend with a valley during the night (because people are sleeping) and a level during the day (when people are awake and outside their homes). The small difference is that the peak in traffic occurs at lunchtime according to the measurements of the Belgian operator, whereas it occurs in the evening in the other models. This traffic profile is a periodic phenomenon of 24 hours that repeats every day.

Including the traffic per user and the cell population, the small-scale traffic per BS of configuration (i, j) is:

$$R^{cell}_{ij}(t) = \frac{Pop^{cell}_{avg,i}}{N_{op}} \cdot \alpha^{4G}_{pop} \cdot R^{user,4G}_{avg,j} \cdot r(t) \tag{11}$$

where $Pop^{cell}_{avg,i}$ is the average population of a cell of density category i (and independent of the BS type j), α^{4G}_{pop} is the penetration ratio of 4G in Belgium, $R^{user,4G}_{avg,j}$ is the daily average throughput per user served by a BS of

type j, and $r(t)$ is the general normalized traffic profile. Here, the operators are considered to be identical and $r(t)$ is independent of the configuration (i, j). The same equation can be applied to the 5G case by adapting the penetration ratio and the daily average throughput per user.

The average throughput per 4G user in 2020 is estimated to be 0.0084 Mbps for the 1-band sites, 0.0104 Mbps for the 2-band sites and 0.0174 Mbps for the 3-band sites. These figures may seem very low but they include inactive periods. For example, a user consuming 3 GB of mobile data per month only have an average throughput of 0.0093 Mbps over the month. Clearly, the instantaneous throughput per user is much higher during data packet transfers. Applying (11) in the heavily loaded situation (i.e. SDU area at noon for a 3-band sites), the throughput is 317 Mbps/km^2 or 35 Mbps/site in 2020, which corresponds more or less to the scenario studied by Ericsson for 2017 in [18]. However, this approach considers only the residential population covered by the BS but does not include other sources of data traffic such as tourist attractions or commercial centres. Hence, the normalisation per residential user does not always consider all users actually present in the cell.

4 Power consumption and carbon footprint of broadband RANs

The power consumption of the broadband RAN is estimated in 2020 considering only the deployment of 4G BSs. This situation is the reference case. As the 5G deployment strategies of the operators are still uncertain, six different scenarios are then studied to estimate the evolution of the RAN power consumption by 2025. Based on these six scenarios, the carbon footprint of the RAN is also estimated using a life-cycle approach.

4.1 Scenarios of 5G deployment

Regarding the 4G RAN, it is assumed that it is fixed for the next few years and that no more sites will be brought into service or removed by 2025. The type of BSs installed (1, 2 or 3 bands) does not change neither. Therefore, coefficients N_{op} and N_{ij}^{4G} are fixed and known for 4G. For the 5G RAN, one type of BS is considered which could be installed on existing 4G sites. Thus, the modeling of this RAN depends only on the population density category i and the number of BSs that would be deployed in each category is $N_i^{5G} = \sum_{j=1}^{N_{types}^{4G}} N_{ij}^{4G}$. The number of operators likely to deploy 5G is still $N_{op} = 3$ and their networks are supposed to be identical. Finally, the throughput for each hour, each mobile generation and each deployment category, namely $R_{ij}^{cell,4G}(t)$ and $R_i^{cell,5G}(t)$, are calculated according to (11). The shape of the traffic profile hour by hour during the day, $r(t)$, remains the same regardless of the mobile technology and the year analyzed. The only parameters that change for each scenario and over the years are the amplitudes of the 4G and 5G traffic profiles: $R_{avg,j}^{user,4G}$ and $R_{avg}^{user,5G}$, and the penetration ratio of 4G and 5G in Belgium: α_{pop}^{4G} and α_{pop}^{5G}.

The definition of each scenario is based on the division of the territory into categories of population density. For the sake of simplicity, when a scenario considers that a category is covered in 5G, it means that all the 4G sites in this category are also equipped with 5G BSs. This implies that all the population in that category is covered by 5G and is potentially a 5G user. Accordingly, the six scenarios studied by 2025 are:

(1) **No 5G deployment** (5G population coverage = **0%**): It is the reference scenario used for 2020 and 2025. Only 4G absorbs the increase in traffic volumes. This scenario is not possible for 2025 given the already started deployment of 5G, but it is useful to compare with other scenarios.

(2) **DU and SDU deployment** (5G population coverage = **37%**): Only densely populated areas are covered by 5G to increase the capacity of mobile broadband networks in city centres and their traffic hotspots.

(3) **Urban, DU and SDU deployment** (5G population coverage = **53%**): Wider coverage is offered by operators in the cities which allows to cover more than half of the Belgian population.

(4) **Suburban, urban, DU and SDU deployment** (5G population coverage = **83%**): The outskirts of cities and villages are equipped with 5G and a large majority of the population is covered with 5G.

(5) **Full 5G deployment** (5G population coverage = **100%**): The entire population is covered by 5G, even isolated homes in the wilderness. 5G coverage is identical to 4G coverage.

(6) **Full 5G deployment and decommissioning of 4G** (5G population coverage = **100%**): The entire population is covered by 5G and 4G BSs are removed. All mobile broadband users must therefore use 5G. This scenario corresponds to scenario (5) in terms of total traffic volume but here all the volume transferred using 4G has to move to 5G. This scenario is highly unlikely as it would mean that all mobile data users would have to change their smartphones for 5G-ready ones before 2025. Moreover, it would require a fully stand-alone 5G deployment as non-stand-alone 5G still requires 4G to operate.

4.1.1 Traffic forecasts based on historical data

Most of the studies of technological changes are based on the application of logistic models or S-curves. These curves are typically divided in three phases: (i) a short-term phase of fast growth (almost exponential), (ii) a mid-term phase of stable increase (almost linear near the inflection point), and (iii) a long-term phase of slower growth and saturation. The rationale behind this is that the growth slows down as the studied variable approaches its uppermost limit, essentially due to limits on the system [46, 47, 48]. Hence, these models are applicable to estimate the total number of mobile users in a bounded population and the associated traffic using a finite mobile network. In particular, the simple logistic curve is symmetric between two asymptotes and its inflection point lies mid-way between the asymptotes. The equation of a logistic curve is the following:

$$F(y) = \frac{K}{1 + \exp(-a(y-b))} \tag{12}$$

where $F(y)$ represents the cumulative adoption until year y, K is the ultimate market potential or saturation level, $a > 0$ indicates the intrinsic growth rate and b is the time scale offset.

The total population of Belgium is the limit used to determine the upper asymptote of the number of mobile data users. Based on data from StatBel [49], the mid-term evolution of the Belgian population is linear as:

$$Pop_{tot}^{\text{BEL}}(y) = -101.7 \times 10^6 + 56.1 \times 10^3 y \qquad \text{for } y \geq 2012 \tag{13}$$

With BIPT statistical data [50], the total number of SIM cards can be divided into human subscribers and machine-to-machine (M2M) devices. Here, only human consumers are considered (about 12 million SIMs in Belgium in 2020) and among them only a part has a subscription giving access to mobile data. Considering that each user has one and only one SIM card, the number of mobile data users is equal to the number of mobile data SIMs $SIM_{data}(y)$ (all generations included), which will saturate at $Pop_{tot}^{\text{BEL}}(y)$ as:

$$SIM_{data}(y) = \frac{Pop_{tot}^{\text{BEL}}(y)}{1 + \exp(-0.3(y - 2013.5))} \qquad \text{for } y \geq 2012 \tag{14}$$

Among these mobile data SIM cards, one part, $SIM_{3G}^{(s)}(y)$, uses only 3G; another part, $SIM_{4G}^{(s)}(y)$, has a 4G subscription; and a last part, $SIM_{5G}^{(s)}(y)$, could soon use 5G. Obviously:

$$SIM_{data}(y) = SIM_{3G}^{(s)}(y) + SIM_{4G}^{(s)}(y) + SIM_{5G}^{(s)}(y) \tag{15}$$

where the exponent (s) means that the variable depends on the scenario.

Depending on the scenario, more or less customers would use 5G, and consequently the numbers of 4G and 3G users would be the other ones. For each scenario, it is assumed that the potential 5G users are only those covered by 5G (representing a certain percentage of the population). Among them, only a fraction would indeed move to 5G between 2020 and 2025. The estimated penetration ratio of 5G, $\alpha_{5G}^{user,(s)}(y)$ hence follows a growing adoption rate up to a upper limit which is the percentage of population covered by 5G, $cov_{5G}^{(s)}$, yielding:

$$\alpha_{5G}^{user,(s)}(y) = \frac{cov_{5G}^{(s)}}{1 + \exp(-(y - 2022))} \qquad \text{for } y \geq 2021 \tag{16}$$

$$SIM_{5G}^{(s)}(y) = \alpha_{5G}^{user,(s)}(y) \cdot \frac{SIM_{data}(y)}{1 + \exp(-0.9(y - 2024))} \qquad \text{for } y \geq 2012 \tag{17}$$

$$SIM_{4G}^{(s)}(y) = \frac{SIM_{data}(y)}{1 + \exp(-0.6(y - 2015.5))} - SIM_{5G}^{(s)}(y) \qquad \text{for } y \geq 2012 \tag{18}$$

$$SIM_{3G}^{(s)}(y) = SIM_{data}(y) - SIM_{4G}^{(s)}(y) - SIM_{5G}^{(s)}(y) \qquad \text{for } y \geq 2012 \tag{19}$$

The forecasts here are purely prospective, which explains why different scenarios (s) are studied. With (16), more than 95% users covered by 5G actually use 5G in 2025. The remaining 5% of users still use 3G or 4G despite the availability of 5G. Moreover, the intrinsic growth parameter of 5G is defined greater than in 4G, itself greater than in 3G ($0.9 > 0.6 > 0.3$), meaning that the market penetration is faster every generation.

The average data volume per user of each generation is determined using [50], from which all mobile traffic reported per year is allocated to human consumers, including both MNOs, MVNOs and M2M traffic. The user

behaviour is considered to be independent of population density and depends only on the type of BS installed. In fact, this means that the BS performance (i.e. its capacity) drives the user behaviour and thus its average traffic. For example, a user that has access to a 5G BS with a 5G SIM will behave like a 5G customer, with much higher traffic than if he only had access to 4G. This is also true for the three different types of 4G BSs for which the average data rate per user is higher when there is more bands installed (see Section 3.4). A logistic model is used for the customer's data volumes with an upper limit for the monthly traffic per user. This is linked with network saturation which ultimately limits the maximum data traffic assuming a fixed network size. The monthly data volumes per 4G user, $V_{4G}^{user}(y)$, and per 5G user, $V_{5G}^{user}(y)$, are given (in GB) by:

$$V_{4G}^{user}(y) = \frac{7}{1 + \exp(-0.53(y - 2019.5))} \qquad \text{for } y \geq 2015 \tag{20}$$

$$V_{5G}^{user}(y) = \frac{42}{1 + \exp(-0.29(y - 2022))} \qquad \text{for } y \geq 2021 \tag{21}$$

The upper bound of monthly traffic per 4G user is set to 7 GB with the existing network unchanged. According to a report of Capgemini for BIPT [51], this volume of mobile data per customer could be reached in 2025. Following the associated growth trend, BIPT already predicts a high risk of 4G cell saturation in Brussels in 2022 [12]. This should be even more the case in 2025 with even more 4G traffic, which justifies this upper limit. For the monthly traffic per 5G user, the bound is set more arbitrarily to 42 GB. This is almost double of what is predicted by [51] for Belgium in 2040 (i.e. 22.27 GB/month/user). Since this limit is very high, only the first phase of the logistic curve is concerned by 2025, which roughly follows an exponential trend. Furthermore, in 2021, the traffic per 5G user is set to be the same as in 4G. The traffic per month and per user would reach 12.4 GB in 2025 using 5G, while it would be limited to 6.6 GB with 4G. Compared to the 3.9 GB in 2020, this corresponds to a compound annual growth rate (CAGR) of about 26%. This figure is in line with Ericsson's estimate of the mobile data traffic increase per smartphone for 2020-2026 in Western Europe [10].

The model parameters have been manually set to match the historical trends. Figure 11 illustrates the evolution of the number of users for each mobile generation compared to the Belgian population. The period of interest (i.e. 2020-2025) is shaded, and the six scenarios are indicated by dashed lines for 4G and 5G users. Figure 12 shows the mobile data traffic per month and per user in 4G and 5G. The traffic of each type of 4G BS is plotted in dotted lines while the weighted average traffic for all 4G BS types is shown with a dashed line. On both figures, the historical BIPT data are displayed as solid lines which shows a good match with the model.

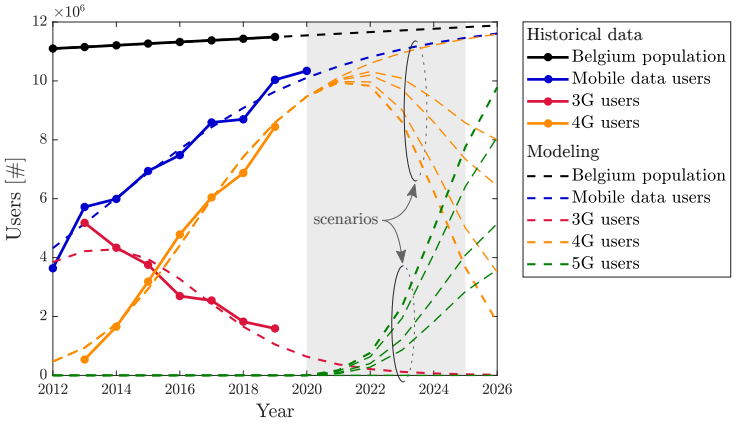


Figure 11: Historical data and modeling over time of the number of users per mobile generation compared to the Belgian population. From 2020, different scenarios are considered.

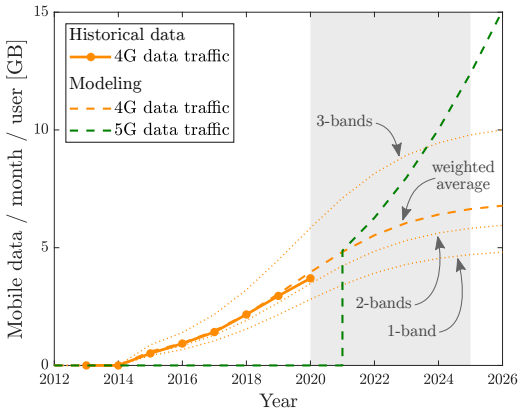


Figure 12: Evolution over time of 4G and 5G mobile data traffic per month and per user, for the different types of base stations.

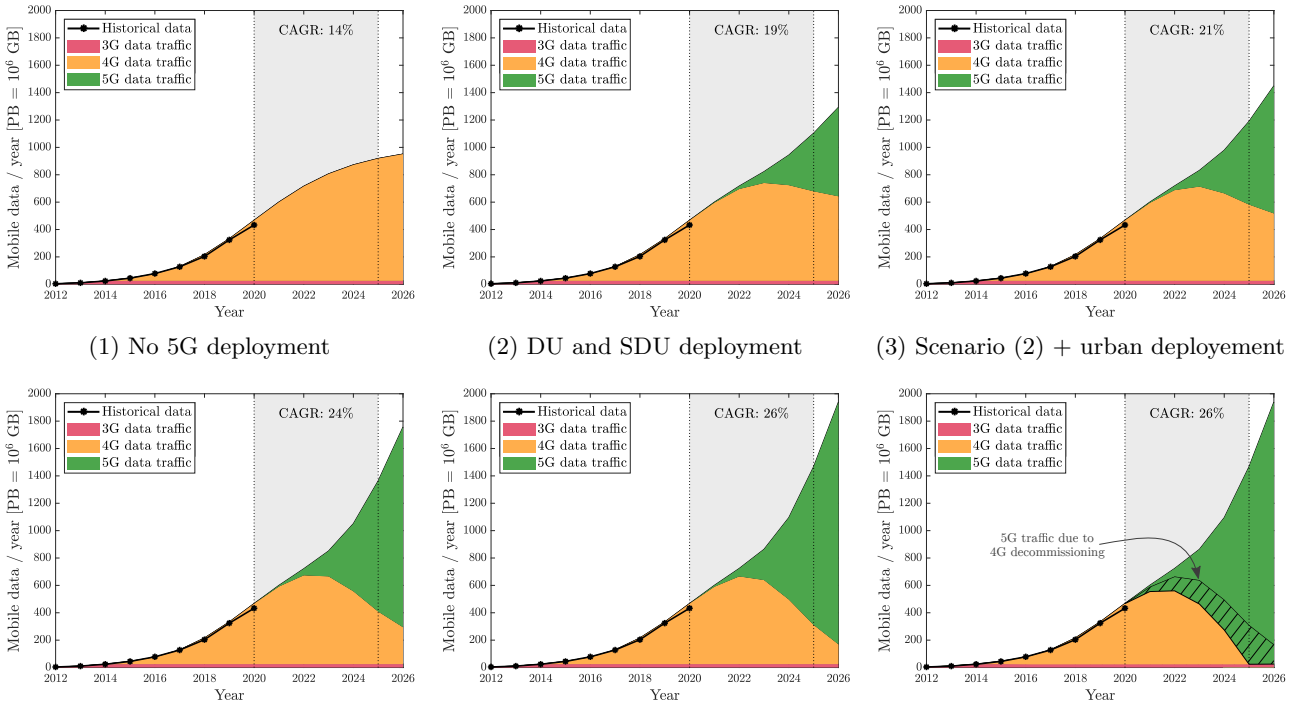
4.1.2 Total mobile data volumes of each scenario

By combining the two variables presented in Figure 11 and Figure 12, the total mobile data volume is estimated under the assumption that when a user has a subscription for one mobile generation, he only uses this generation. Thus, the total volume of mobile data transferred in Belgium during one year is given by:

$$V_{data}^{tot,(s)}(y) = V_{3G}^{tot} + 12 \cdot \left(SIM_{4G}^{(s)}(y) \cdot V_{4G}^{user,(s)}(y) + SIM_{5G}^{(s)}(y) \cdot V_{5G}^{user,(s)}(y) \right) \tag{22}$$

where the 3G volume V_{3G}^{tot} is considered constant over the years and independent of the scenario.

Volumes of mobile data per year for all scenarios are presented in Figure 13. These volumes are differentiated by generation and are given in petabytes (PB) = 10^6 GB. When there is no 5G deployment, 4G traffic saturates below 1 000 PB, which is twice the volume generated in 2020 (i.e. 433 PB [20]). When 5G is deployed, the overall traffic can continue to increase without short-term saturation. The rise in total data volumes depends directly on the deployment scenario to reach more than 1 400 PB in 2025 for scenario (5) and (6). This can then be seen as a *rebound effect* generated by the introduction of 5G [52]. Depending on the intensity of 5G deployment, part of the traffic shifts from 4G to 5G between 2020 and 2025, leading to a 4G traffic peak before a decline. From scenario (4) to (6), the total 4G volume becomes even lower in 2025 than in 2020. Finally, in scenario (6), 4G is decommissioned and the 4G traffic of scenario (5) is transformed into 5G traffic (represented by a hatched area). In that case, the 4G users being constrained to switch to 5G before 2025 do not change their consumption behaviour and the total data volume of these two scenarios does not change.



(1) No 5G deployment (2) DU and SDU deployment (3) Scenario (2) + urban deployement

(4) Scenario (3) + suburban deployment (5) Full 5G deployment (6) Scenario (5) + 4G decommissioning

Figure 13: Evolution over time of mobile data volumes per year and per mobile generation for all scenarios compared to historical trends. The CAGR is indicated for the period of interest, i.e. 2020-2025, which is shaded.

According to CISCO forecasts [11], the CAGR of mobile traffic in Western Europe for 2017-2022 would be 38%. In this work, the CAGR over the same period is around 41% in Belgium for each scenario because the impact of 5G deployment does not have much impact yet. For Ericsson [10], 60% of the population should be covered with 5G by 2026. This is between scenario (3) and (4). They also predict that 68% of mobile subscriptions would be in 5G, which is very close to scenario (4) where 70% of subscriptions are expected to use 5G by 2026. Finally, they forecast a CAGR of 25% for total mobile traffic in Western Europe for 2020-2026. The closest scenario is scenario (4) of widespread 5G deployment. On the other hand, according to Capgemini [51], the CAGR of total mobile traffic in Belgium for the next decade would be only 19%, reaching 1 100 PB in 2025. This corresponds rather to scenario (2) of targeted 5G deployment in city centres. BIPT also notes that due to the COVID crisis in 2020, the generalisation of home working has slowed down the growth of mobile data traffic, with consumers opting instead for Wi-Fi networks. Indeed, in 2020, mobile data traffic grew more slowly than in 2019, rising by 109 PB (+34%) over the year, while in 2019 the growth was 120 PB (+59%) [20].

4.2 Total power consumption of each scenario

In this section, the full model described in Section 2 is applied for the RAN in 2020 and 2025 according to the six scenarios of 5G deployment. 2020 is considered as the reference year and the transition period between 2020 and 2025 is not evaluated because the practical deployment strategies of the operators are too uncertain during this time frame. Moreover, only the impacts of 4G and 5G BSs (when they are deployed) are taken into account and the power consumption of 2G and 3G BSs are not included in the results.

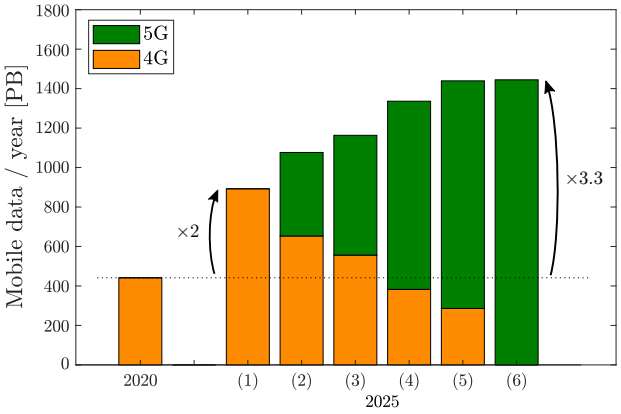
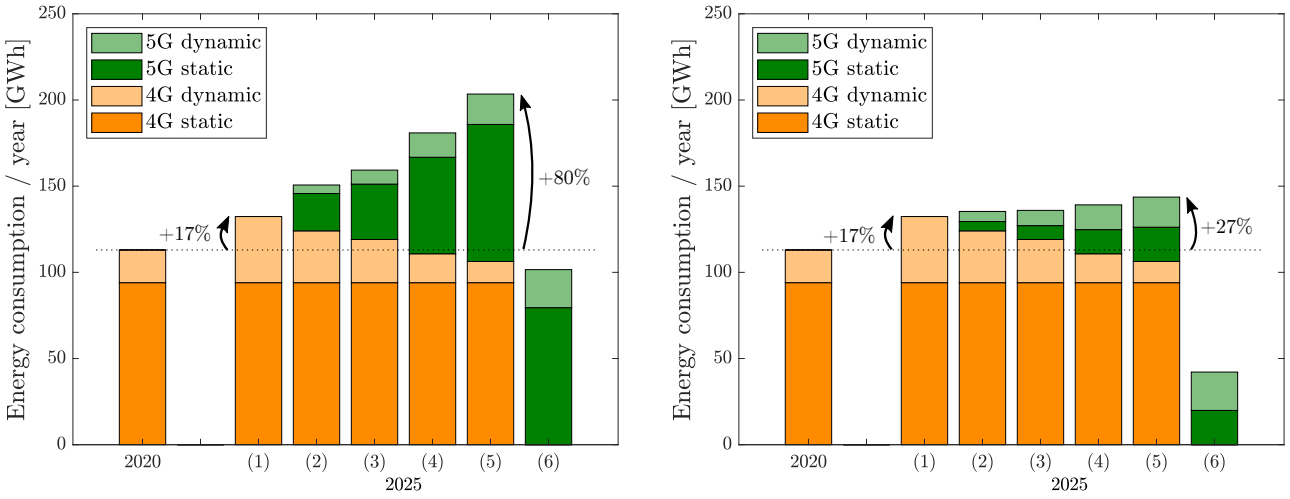


Figure 14: Comparison of mobile data volumes per year in 2020 and for each scenario in 2025.

Figure 14 presents a comparison of mobile data volumes per year in 2020 and for each scenario in 2025, split between 4G and 5G. In fact, this is a summary of Figure 13 for the two years of interest. Figure 15 shows the total annual energy consumption of the 4G and 5G RAN for each scenario. Each technology has a *static* (traffic independent) and a *dynamic* (traffic dependent) energy consumption part. Figure 15(a) presents the situation where 5G BSs do not use sleep mode and whose power model is simply linear with traffic. Figure 15(b) presents the situation where the sleep mode is active and for which a reduction of static power is possible. In practice, actual power consumptions should fall between these two cases.



(a) 5G without sleep mode

(b) 5G with sleep mode

Figure 15: Total electricity consumption of 4G and 5G RANs in 2020 and for each scenario in 2025. Static and dynamic power components are distinguished by different colors. Total power consumption is higher in 2025 than in 2020, except for scenario (6). Energy savings thanks to sleep mode are visible in (b) compared to (a).

The energy consumption corresponds only to the electricity consumption of the RAN (in GWh). As electricity is a secondary energy, a primary energy factor (PEF) must be applied to convert it into primary energy (for example, in tonnes of oil equivalent). The PEF is usually between 2 and 3 to take into account the efficiency of the electricity generation plants and the losses on the power grid [6, 53].

Figure 15 reveals that the total energy consumption increases in 2025 compared to 2020 for the first five scenarios, with or without sleep mode. When 5G is not deployed, the energy consumption only increases by 17% even if the total traffic doubles. With full 5G deployment, the increase in energy consumption is almost 80% without sleep mode and is limited to 27% using sleep mode. At the same time, the total traffic increases by 227%, showing a *relative decoupling* between data traffic and energy consumption, but not an *absolute decoupling* [4, 52]. Only scenario (6) with 4G decommissioning is energy beneficial in 2025 compared to 2020. In this very hypothetical case, the expected energy savings are 10% without sleep mode and 63% with sleep mode, and the decoupling is absolute because the total energy consumption decreases while the total traffic increases.

In 2025, the total energy consumption increases as the intensity of 5G deployment increases, and the same is true for the total data traffic. This highlights the effect of operating different mobile generations in parallel that stacks the respective power consumptions. This is mainly due to the large share of static power consumption, especially in 4G. Indeed, between 70% to 90% of the power consumption of 4G BSs is independent of traffic, and the same ratio is observed for 5G BSs. On another hand, by implementing an efficient sleep mode, it is possible to reduce the static consumption to about half of the BS consumption. This limits the increase in power consumption of the 5G RAN. The implementation of techniques to reduce the static power consumption appears to be crucial for the development of future generation of mobile communications.

4.2.1 Variability and uncertainty of results

The results of this work are sensitive to assumptions, simplifications and extrapolations, leading to uncertainty. In particular, there is a great degree of uncertainty about BS power models at high-load, especially for 5G. Fortunately, the traffic loads are low on average compared to the average site capacities and remain in the linear region of models. Whatever the scenario, the 4G traffic does not fall outside the measurement range of the 4G models. This may seem surprising since the mobile data traffic in 4G is supposed to double in the more extreme scenario compared to the time of the measurements. However, the available measurements already comprise a large variability of hourly traffic load depending on the days and the sites studied, giving several very high load measurements, very high above the average load. By contrast, the traffic profile used for the forecasts by 2025 does not take into account these possible variations between days and only considers the average hourly traffic that stays in the measurement range. As 4G traffic projections remain within the measurement range, it can be considered that these models are reliable and do not induce major variability in the results.

Besides, the scaling of 5G models induces more uncertainties. The traffic volume forecast by 2025 is also another uncertain aspect. One way to assess these uncertainties is to vary parameters to perform a sensitivity analysis. Using this method, the RAN power consumption can double in the worst case scenario with 5G deployment or it can decrease in the best case even without 4G decommissioning. This outlines that there is a very large degree of variation in the conclusions that can be drawn from this work. Depending on the deployment scenarios considered, the power performance of 5G BSs and the evolution of traffic, 5G deployment may have a positive or negative impact on the RAN power consumption in 2025 compared to the no deployment scenario. However, most of the scenarios rather predict a negative impact of 5G deployment on total network consumption compared to the case without 5G. The reason for this is among other things the induced *rebound effect* of 5G deployment that allow users to still generate more and more traffic. Even if this phenomenon is very present and observable in the ICT sector, it remains difficult to quantify and to model.

4.2.2 Comparison of results with other studies

In this section, results of this work are compared with those already obtained by other studies. Table 6 lists the annual RAN electricity consumption for different scopes and years, and the relative energy footprint with respect to the total electricity consumption of the analyzed country. The energy footprint of the 4G RAN estimated in this work corresponds to 0.13% of the electricity consumption in Belgium in 2020. This percentage is of the same order of magnitude than in other studies for 2015-2020. When a wider scope is analyzed, taking into account 2G, 3G and other operator activities (e.g. maintenance and offices) the percentage increases. Furthermore, the characteristics of 5G BSs may greatly vary between studies, considering or not BS densification, the use of millimeter waves or the access to very wide bandwidths. In this work, 5G BSs have rather simple characteristics, that are close to those of the current 4G+ BSs, and far from the more advanced 5G BSs.

Source	**Reference year**	**Annual electricity consumption of the RAN [TWh]**	**Annual electricity consumption of the country [TWh]**	**Relative electricity footprint [%]**
This work	2020	0.113 (4G RAN)	∼ 87 [54]	∼ 0.13
→ Belgium	2025	0.132 - 0.203 (4G+5G RANs)	-	-
Fehske et al. (2011) [5] → World	2020	48 - 109 (all RANs)	∼ 26 340 [54]	∼ 0.2 - 0.4
Andrae & Edler (2015) [55]	2020	10 - 120 (4G RAN)	25 000 - 28 000	0.05 - 0.5
→ World	2030	27 - 2 565 (4G+5G RAN)	35 000 - 61 000	0.08 - 4.2
Malmodin & Lundén (2016) [6] → Sweden	2015	0.7 (all RANs)	132 [54]	0.5
Malmodin & Lundén (2018) [56]	2015	25 (4G RAN)	22 470	0.1
→ World	2015	92 (all RANs)	22 470	0.4
Pihkola et al. (2018) [8] → Finland	2016	0.6 (all operator' activities)	85 [54]	0.7
Citizing - HCC (2020) [57]	2020	4.6 (all RANs+fixed)	∼ 470 [54]	∼ 1
→ France	2030	9.9 - 16.6 (all RANs+fixed)	-	-

Table 6: Benchmarking of the estimated annual RAN electricity consumption with other studies. When no electricity consumption of the country is given by the study, the relative footprint is based on IEA data [54]. The tilde means a projection in 2020 based on the latest available data (2018 or 2019). When no estimate of electricity consumption is given in the study for 2025 or 2030, the relative footprint is not calculated.

4.3 Energy efficiency of each scenario

For a same amount of energy, it is expected that a 5G BS transfers more data than a 4G BS [58, 59]. By dividing the data traffic volumes given by Figure 14 by the RAN energy consumption of Figure 15, a simple energy efficiency metric (in kbits/J) can be calculated. The obtained energy efficiencies are shown in Figure 16 for each generation independently and for the two mobile networks together (4G RAN + 5G RAN).

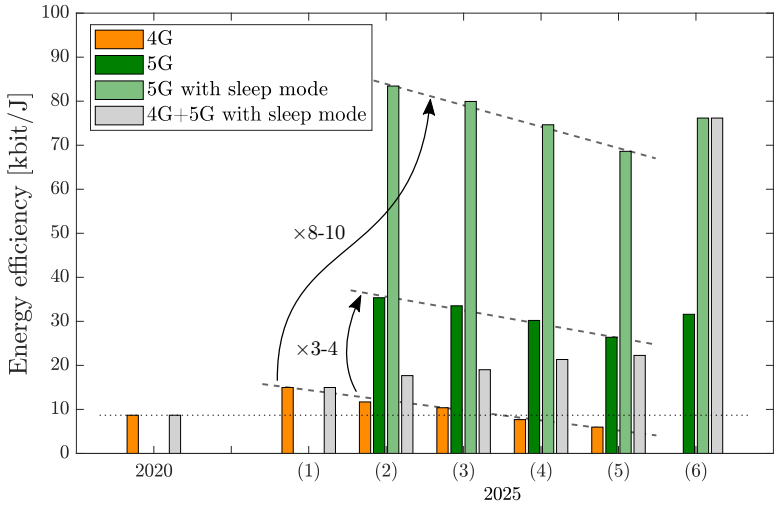


Figure 16: Energy efficiency of the 4G and 5G RAN without and with sleep mode in 2020 and for each scenario in 2025. Obviously, some situations do not exist (e.g. there is no 5G in 2020). The grey bar depicts the overall energy efficiency of the 4G RAN plus the 5G RAN with sleep mode. In 2025, the 5G RAN is 3-4 times more efficient than the 4G RAN, and even 8-10 times more efficient when the sleep mode is used.

Clearly, there is an improvement in energy efficiency between the 4G and the 5G RAN. Without sleep mode, the energy efficiency of the 5G RAN is 3 to 4 times better than the 4G RAN, and this gain ranges from 8 to 10 with sleep mode. Without 5G deployment, the energy efficiency of the 4G RAN still increases by 75% between 2020 and 2025 because more traffic is transferred without requiring much additional dynamic power (the static power is dominant and remains constant). However, the energy efficiency of the 5G RAN is maximal when 5G BSs are deployed only in SDU, and thus when the BSs are more heavily loaded (i.e. scenario (2)). The more massive the 5G deployment, the less loaded the BSs on average, as they occupy less densely populated areas, thus reducing their efficiency. Overall, the energy efficiency for the whole network (4G + 5G) increases with a more intense 5G deployment even if the gain is minor between scenario (4) and scenario (5). Of course, in scenario (6) the energy efficiency of the 5G RAN is higher than in scenario (5) despite the same deployment because 5G BSs are more loaded due to 4G decommissioning. Moreover, in scenario (6), the efficiency of the whole network corresponds exactly to the efficiency of the 5G RAN alone and is much higher than in the other scenarios because the energy consumption of the 4G RAN has disappeared.

To summarize, the increase in energy efficiency does not necessary mean a decrease in total energy consumption. If the increase in traffic is greater than the improvement in energy efficiency, energy savings per bit of data are counterbalanced by the traffic growth. This phenomenon is the result of the *rebound effect* induced on user behavior by the 5G deployment [52]. As an example, in scenario (4) (the most likely scenario according to Ericsson's forecasts), the total energy efficiency gain for the whole network with sleep mode is by a factor of 2.46. But at the same time, the increase in traffic is by a factor of 3.03. Thus, the overall energy consumption increases by a factor of $3.03/2.46 = 1.23$. This highlights the fact that thinking about impacts of a technology should not only be done in terms of energy efficiency but also in terms of total energy consumption.

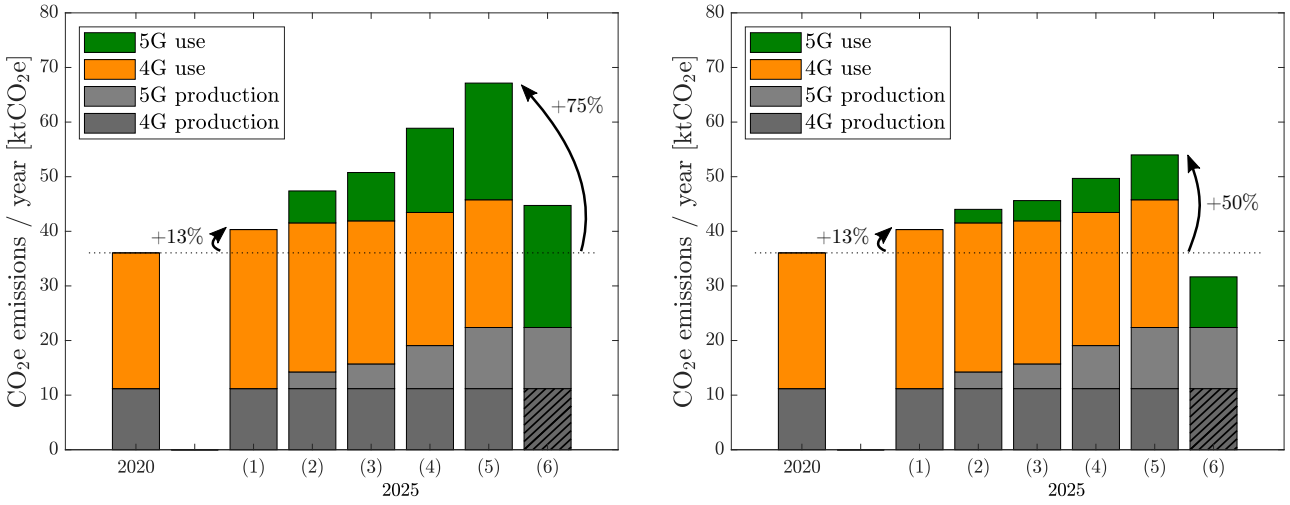
4.4 Total carbon footprint of each scenario

The carbon footprint of the mobile network is mainly composed of GHG emissions from the use and the production phases. The aim of this section is not to provide exact carbon footprint figures, but rather to introduce a life-cycle approach to approximate the carbon footprint of the RAN by 2025. Indeed, to obtain more accurate estimates, it is required to perform a full LCA, which is not done here.

The carbon footprint of the use phase is estimated using a simple CO_2e emission factor per kWh of electricity. Such a factor for Belgium is estimated to be 0.220 kgCO_2e/kWh [60]. The carbon footprint of the production phase is much more uncertain as there are very few sources that assess this parameter. Thus, an

unsure emission factor of 12 tCO_2e/BS is used. This figure comes from a study of Ericsson for LTE BSs [61]. This figure is also in line with [62] that estimates the embodied energy of a BS to 20 MWh. Applying the world average emission factor of 0.6 $kgCO_2e$/kWh [6], this gives also 12 tCO_2e/BS. The emission factor for the production phase is supposed to be identical for 4G and 5G BSs (for lack of a better estimate). Then, the production phase emissions have to be spread over the whole lifetime of the BS. In this work, the BS lifetime is set to be 12 years, which gives 1 tCO_2e/BS/year. Furthermore, this allows to cover the period 2014-2026, from the beginning of the 4G deployment, until at least 2025, our latest year of interest. This means that the embodied carbon impact of 4G BSs can be held constant for the entire duration of the analysis, whatever the scenario studied for 2025. In other studies, an average BS lifetime of 10 years is rather used [61, 62].

Figure 17 presents the obtained carbon footprint for 4G and 5G mobile networks in 2020 and for each scenario in 2025. For 4G, the findings seem to be consistent with other studies which estimate that the production phase of BSs typically accounts for a quarter of the total footprint [5, 6, 17]. In 2025, the total carbon footprint increases with more massive deployments of 5G because more BSs are produced and deployed to cover additional territories. In scenario (5), the annual carbon footprint of the production phase is equivalent in 4G and 5G as there is the same number of BSs deployed for each generation. In scenario (6), the carbon footprint of the production of 4G BSs is still present despite their decommissioning (shown with hatching) to allow a fair comparison with the other five scenarios. The reason for this is that the allocation of the production footprint is spread over 12 years and at least until 2026. If it were decided that 4G would be removed earlier, the production footprint should have been amortized over fewer years, resulting in more emissions per year.



(a) 5G without sleep mode

(b) 5G with sleep mode

Figure 17: Total carbon footprint of 4G and 5G RANs in 2020 and for each scenario in 2025 considering the production and the use phases. Total carbon footprint is higher in 2025 than in 2020, except for scenario (6) with sleep mode. GHG emission savings thanks to sleep mode are visible in (b) compared to (a).

When the production phase is considered, the trend of impacts due to 5G deployment is even more visible in some cases than when focusing only on the operating electricity consumption. For instance, the total annual GHG emissions of full 5G deployment with sleep mode while maintaining 4G increase by 50% compared to 2020, whereas this increase is 27% in the same case when analyzing operating electricity only. Overall, the only scenario that reduces the annual carbon footprint is the full 5G deployment with sleep mode and with 4G decommissioning. However, this scenario is highly unlikely in practice due to technical and market constraints. On average, the carbon intensity of mobile data transfers is 0.08 $kgCO_2$/GB in 2020 and decreases to 0.05 $kgCO_2$/GB in 2025 for the five scenarios keeping 4G active. The decommissioning of 4G may decrease it further to 0.03 $kgCO_2$/GB. Furthermore, when a sleep mode is implemented, the share of operational carbon footprint of 5G BSs decreases to reach about one half of the total carbon footprint. This means that thinking about 5G deployment solely in terms of use phase impacts actually only takes one part of the overall impacts into account.

Once again, the figures mentioned here are very uncertain and should be analyzed with caution. Only the general approach and the main trends should be considered. To estimate the global climate impacts of 5G deployment, an analysis should also be made including all the impacts related to the advent of new 5G devices, the densification of IoT, the massive smartphone replacement, the increased use of the core network and data centres, etc. Therefore, there is a real need to define more accurate carbon impact models of mobile communication technologies to guide operators' deployment choices in order to reach climate targets on schedule.

5 Conclusion

For several decades, global warming is accelerating at an alarming rate due to GHG emissions from human activities. As a result, all economic sectors have to address this issue and rapidly reduce their carbon footprint. More specifically, the environmental impact of mobile communications tends to grow due to the huge mobile data traffic increase. In this sector, a significant part of the power consumption is due to base stations constituting the RAN. In recent years, 5G has been heralded as a new technology that could reduce the power consumption of the RAN through much more energy efficient base stations. At the same time, this technology has been developed to support the expected growth of data traffic. However, if the increase in data traffic is faster than the improvements in energy efficiency, the absolute power consumption will continue to grow. This phenomenon is directly related to the so-called rebound effect. Hence, to assess the energy footprint of a new mobile generation deployment, for instance 5G, it is more appropriate to focus on the absolute energy consumption than on the energy efficiency per unit of data transferred.

Therefore, this work aims to quantify the overall energy consumption of mobile broadband RANs in Belgium and to evaluate the impact of 5G deployment. Models of current 4G networks and 4G base station power consumption are determined using data and field measurements from a Belgian operator. Models of 5G base station power consumption are defined by scaling the 4G models to reflect technology improvements and changes in key parameters. Then, the total energy consumption of mobile networks is estimated in 2020 and forecasted in 2025 for different 5G deployment scenarios.

Field measurements show that the relationship between the base station power consumption and the average traffic is linear with a static power consumption offset. Since mobile networks are dimensioned to cope with traffic peaks and to cover the whole territory, base stations are lightly loaded on average. In 2020, static energy consumption accounts for more than 80% of the total RAN power consumption and only the remaining part depends dynamically on the data traffic. To reduce the static power consumption, sleep modes can be used. This feature implemented in 5G base stations could reduce their total power consumption by about 60%. The 5G base station could then be up to 10 times more energy efficient than 4G base stations.

However, in a full 5G deployment scenario, the volume of annual mobile data traffic could reach over 1 400 PB in 2025 which is 3.3 times more than in 2020. In this scenario, the total energy consumption of broadband RANs increases by 80%, but is limited to 27% if 5G base stations use the sleep mode feature. On the contrary, in the scenario of no 5G deployment, the total energy consumption only increases by 17% with a two-fold growth of mobile data traffic. Without considering 4G decommissioning, it appears that the scenario of no 5G deployment is the best in terms of energy consumption. This finding reveals that operating two mobile generations in parallel increases the total energy footprint, mainly due to the large proportion of static power consumption. On another hand, the deployment of 5G base stations greatly enhances the mobile network capacity and thus avoids the saturation of 4G networks in densely populated city centres. To limit the energy footprint increase while preventing 4G saturation, a good compromise would be to restrict the deployment of 5G to densely populated areas and mobile traffic hotspots. Indeed, it is useless to cover the rural and wilderness areas with 5G because the capacity of 4G networks can be sufficient there.

Moreover, estimating impacts of the whole life cycle of base stations requires data and specific information from raw materials extraction to production, use and end-of-life. Unfortunately, there is currently a lack of information from manufacturers to accurately assess the embodied energy consumption and GHG emissions of these equipments. Using a rough estimate of base station embodied greenhouse gas emissions, there is a clear upward trend of carbon footprint when the deployment of 5G is extensive. This is a consequence of the large number of newer base stations that need to be produced.

Future work is needed to corroborate the bottom-up modeling with a top-down approach using overall power consumption of Belgian operators. It would also be interesting to validate the power models of 5G base stations with field measurements. Another improvement could be to refine the base station power models for shorter time intervals than one hour, in order to analyze more precisely the effects of instantaneous traffic fluctuations. Lastly, to assess the total environmental impacts of mobile communication technologies, impacts of terminals, data centres and core networks should also be taken into account.

References

[1] IPCC. "Summary for Policymakers. Global Warming of 1.5°C, an IPCC Special Report." *Press* (2018).

[2] UNFCCC. *The Paris Agreement.* (2015). URL: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement (visited 2021-08-18).

[3] European Comission. *A European Green Deal.* (2019). URL: https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en (visited 2021-08-18).

[4] D. Bol et al. "Moore's Law and ICT Innovation in the Anthropocene". *IEEE DATE* (2021).

[5] A. Fehske et al. "The global footprint of mobile communications: The ecological and economic perspective". *IEEE Commun. Mag.* (2011).

[6] J. Malmodin and D. Lundén. "The energy and carbon footprint of the ICT and E&M sector in Sweden 1990-2015 and beyond". *ICT4S* (2016).

[7] C. Freitag et al. "The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations". *Patterns* (2021).

[8] H. Pihkola et al. "Evaluating the Energy Consumption of Mobile Data Transfer—From Technology Development to Consumer Behaviour and Life Cycle Thinking". *Sustainability* (2018).

[9] J. Malmodin. "The power consumption of mobile and fixed network data services - The case of streaming video and downloading large files". *EGG+* (2020).

[10] Ericsson. *Ericsson Mobility Report 2020.* 2020.

[11] CISCO. *Visual Networking Index: Global Mobile Data Traffic Forecast, 2017–2022.* (2019).

[12] BIPT. *Communication concernant les risques de saturation des réseaux mobiles.* (2021).

[13] N. Al-Falahy and O. Y. Alani. "Technologies for 5G Networks: Challenges and Opportunities". *IT Pro.* (2017).

[14] G. Auer et al. "Cellular Energy Efficiency Evaluation Framework". *IEEE VTC* (2011).

[15] G. Auer et al. "How much energy is needed to run a wireless network?" *IEEE Wireless Comm.* (2011).

[16] Global e-Sustainability Initiative (GeSI). *SMARTer2030: ICT Solutions for 21st Century Challenges.* (2015).

[17] J. Bieser et al. *Next generation mobile networks: Problem or opportunity for climate protection?* (2020).

[18] M. Olsson et al. "Energy performance evaluation revisited: Methodology, models and results". *IEEE WiMob* (2016).

[19] Y. Chen et al. "Evaluation of Potential Energy Efficiency Gain of 5G Wireless Networks" (2015).

[20] BIPT. *Rapport annuel 2020.* (2021).

[21] Geoportail de Wallonie. *Cadastre des antennes émettrices stationnaires de Wallonie.* (2020). URL: https://geoportail.wallonie.be/walonmap#SHARE=C300838ACD2B2802E053D0AFA49D9390 (visited 2020-12-05).

[22] Bruxelles Environnement. *Carte des antennes émettrices à Bruxelles.* (2020). URL: https://geodata.environnement.brussels/client/view/3a33e35f-6b64-4b28-bb50-5b4c6b7cb29c (visited 2021-03-09).

[23] Geopunt Vlaanderen. *Zendantennes.* (2020). URL: http://www.geopunt.be/catalogus/datasetfolder/00b8c6b6-437d-4d35-9b39-7ad300f11b94 (visited 2021-03-11).

[24] BIPT. *Atlas mobile.* (2020). URL: https://www.bipt-data.be/fr/projects/atlas/mobile (visited 2021-02-14).

[25] IWEPS. *Degré de densité de la population des communes belges.* (2011). URL: https://www.iweps.be/indicateur-statistique/degre-de-densite-de-population-communes-belges-methode-dg-regio/ (visited 2020-11-19).

[26] S. Parkvall et al. "Evolution of LTE toward IMT-advanced". *IEEE Commun. Mag.* (2011).

[27] A. Mandhyan. *4G and 5G Capacity solutions - comparative study.* CommScope White Paper. (2019).

[28] ITU-R M.2135-1. *Guidelines for evaluation of radio interface technologies for IMT-Advanced.* (2009).

[29] P. Frenger and R. Tano. "More Capacity and Less Power: How 5G NR Can Reduce Network Energy Consumption". *IEEE VTC* (2019).

[30] BIPT and CommSquare. *Drive test campaign 2020: results.* (2020).

[31] ITU-R M.2134. *Requirements related to technical performance for IMT-Advanced radio interface(s).* (2008).

[32] B. Debaillie et al. "A Flexible and Future-Proof Power Model for Cellular Base Stations". *IEEE VTC* (2015).

[33] Huawei. *5G Power.* White Paper. (2019).

[34] Bruxelles Environnement. *La 5G : principes et enjeux.* (2021). URL: https://environnement.brussels/thematiques/ondes-et-antennes/comment-ca-marche/la-5g-principes-et-enjeux (visited 2021-04-19).

[35] B. Halvarsson et al. "5G NR Testbed 3.5 GHz Coverage Results". *IEEE VTC* (2018).

[36] F. Kronestedt et al. *The advantages of combining 5G NR with LTE at existing sites.* Ericsson White Paper. (2018).

[37] Nokia. *Optimizing 5G Coverage: Beamforming, Low-band and Dual Connectivity.* White Paper. (2020).

[38] P. von Butovitsch et al. *Advanced antenna systems for 5G networks.* Ericsson White Paper. (2018).

[39] G. Auer et al. *Energy efficiency analysis of the reference systems, areas of improvements and target breakdown.* EARTH Project D2.3. (2010).

[40] S.-J. Oh. *IMT-2020 Evaluation report.* TTA SPG33 Technical report. (2019).

[41] 3GPP. *TDocs at Gothenburg meeting.* (2018). URL: https://www.3gpp.org/DynaReport/TDocExMtg--R1-94--18796.htm (visited 2021-05-29).

[42] P. Lähdekorpi et al. "Energy efficiency of 5G mobile networks with base station sleep modes". *IEEE CSCN* (2017).

[43] F. Salem et al. "Advanced Sleep Modes and Their Impact on Flow-Level Performance of 5G Networks". *IEEE VTC* (2017).

[44] F. Salem et al. "Traffic-aware Advanced Sleep Modes management in 5G networks". *IEEE WCNC* (2019).

[45] M. Deruyck et al. "Reducing the power consumption in LTE-Advanced wireless access networks by a capacity based deployment tool". *Radio Science* (2014).

[46] D. Kucharavy and R. De Guio. "Logistic substitution model and technological forecasting". *Procedia Eng.* (2011).

[47] A. Jha and D. Saha. ""Forecasting and analysing the characteristics of 3G and 4G mobile broadband diffusion in India: A comparative evaluation of Bass, Norton- Bass, Gompertz, and logistic growth models"". *Technological Forecasting and Social Change* 152 (Dec. 2019).

[48] H. Shin et al. "Forecasting the video data traffic of 5G services in South Korea". *Technological Forecasting and Social Change* (2020).

[49] StatBel. *Structure de la population.* (2020). URL: `https://statbel.fgov.be/fr/themes/population/structure-de-la-population#figures` (visited 2021-04-11).

[50] BIPT. *Situation du secteur des communications électroniques 2019: données.* (2020). URL: `https://www.ibpt.be/operateurs/publication/situation-du-secteur-des-communications-electroniques-2019-donnees` (visited 2021-02-14).

[51] BIPT. *Rapport de Capgemini Invent concernant l'évolution des données mobiles liées au spectre sous licence en Belgique et l'impact sur la présence des médias.* (2020).

[52] L. Hilty and B. Aebischer. *ICT Innovations for Sustainability.* (2015).

[53] A. Esser and F. Sensfuss. *Evaluation of primary energy factor calculation options for electricity.* (2016).

[54] IEA Data and statistics. *Electricity consumption.* (2020). URL: `https://www.iea.org/data-and-statistics/data-browser?country=WORLD&fuel=Electricity%20and%20heat&indicator=TotElecCons` (visited 2021-06-03).

[55] A. Andrae and T. Edler. "On Global Electricity Usage of Communication Technology: Trends to 2030". *Challenges* (2015).

[56] J. Malmodin and D. Lundén. "The Energy and Carbon Footprint of the Global ICT and E&M Sectors 2010–2015". *Sustainability* (2018).

[57] Citizing pour le HCC. *Déploiement de la 5G en France : Quel impact sur la consommation d'énergie et l'empreinte carbone ?* (2020).

[58] S. Tombaz et al. "Energy Performance of 5G-NX Wireless Access Utilizing Massive Beamforming and an Ultra-Lean System Design". *IEEE GLOBECOM* (2015).

[59] P. Frenger and K. W. Helmersson. "Energy Efficient 5G NR Street-Macro Deployment in a Dense Urban Scenario". *IEEE GLOBECOM* (2019).

[60] ADEME. *Mix électrique autres pays.* (2013). URL: `https://www.bilans-ges.ademe.fr/documentation/UPLOAD_DOC_FR/index.htm?moyenne_par_pays.htm` (visited 2021-06-03).

[61] P. Bergmark. *Life cycle assessment of an LTE base station based on primary data.* Ericsson presentation. (2015).

[62] I. Humar et al. "Rethinking energy efficiency models of cellular networks with embodied energy". *IEEE Network* (2011).