15   Performance, Energy-Efficiency and Techno-Economic Assessment

Michał Maternia[1], Jose F. Monserrat[2], David Martín-Sacristán[2], Yong Wu[3], Changqing Yang[3], Mauro Boldi[4], Yu Bao[5], Frederic Pujol[6], Giuseppe Piro[7], Gennaro Boggia[7], Alessandro Grassi[7], Hans-Otto Scheck[8], Ioannis-Prodromos Belikaidis[9], Andreas Georgakopoulos[9], Katerina Demesticha[9] and Panagiotis Demestichas[10]

[1]Nokia, Poland
[2]Universitat Politècnica de València, Spain
[3]Huawei, China.
[4]Telecom Italia, Italy
[5]Orange Labs, France
[6]IDATE, France
[7]Politecnico di Bari, Italy
[8]Nokia, Sweden
[9]WINGS ICT Solutions, Greece
[10]University of Piraeus, Greece

15.1   Introduction

In the race towards 5th generation (5G) cellular, many stakeholders are involved with their own proposals for technologies, algorithms and procedures. Especially in the current phase of definition of the new system, the right choice between these proposals is essential to secure the success and widespread adoption of 5G. In particular, it is expected that the new generation will enable novel business opportunities, as stressed in Chapter 2, and provide a performance leap that will justify expenses related to its development and deployment.

The introduction of the new cellular standard is a very long and expensive process and, additionally, major design agreements are extremely difficult to revert once the system is mature.

Hence, the need to quantify the performance of key design concepts long before any type of hardware implementation is available (e.g., prototypes or trial equipment) is extremely important. Such evaluation is often done by means of 'pen and paper' analysis or through various computer simulations. In order to provide an accurate and unbiased assessment, the right evaluation procedures, metrics and models need to be discussed and agreed among the stakeholders involved in the process. Additionally, an economic evaluation of the introduction of 5G needs to be taken into account, as also stressed in Chapter 2.

This chapter is organized as follows. Section 15.2 provides information on major performance evaluation frameworks that were developed for legacy and 5G cellular systems. Additionally, this section presents performance assessments for some of the representative 5G use cases, as introduced in Section 2.5. Section 15.3 focuses on network energy efficiency, i.e., the capability of the network to operate with certain (preferably low) electric power consumption. In recent years, optimization of this factor has become one of the most important criteria for mobile network operators (MNOs) and a major target for telecom equipment vendors. Section 15.4 then introduces a techno-economic analysis of potential 5G deployments, giving insight into costs and benefits of running the 5G network. Finally, Section 15.5 summarizes this chapter and highlights the key performance and economic evaluation results that 5G brings.

15.2   Performance evaluation framework

In recent years, wireless telecommunications have been at the forefront of the technological development of the modern society. As explained in Chapter 2, comparing to the legacy solutions, the 5G system is designed not only to push to the extreme the broadband access for humans, but it is also expected to focus on different kinds of machine-type communications and related services. This diversity of services motivates the evolution and extension of the existing evaluation methodology, mainly focused on human-based mobile broadband communication, to a comprehensive and unbiased **5G evaluation framework** allowing for a fair comparison of proposed concepts, which will give an insight into the achievable performance in most relevant 5G use cases.

15.2.1 IMT-A evaluation framework

Even if the 5G system will introduce a set of brand new functionalities, the evaluation framework does not need to be developed from scratch. There is a clear experience of success in the process followed for the evaluation of International Mobile Telecommunications Advanced (IMT-A). IMT-A is recognized as the 4th generation (4G) of cellular networks, the one generally recognized as the enabler of the widespread success of the mobile broadband applications. In the 2007-2008 timeframe, the radio sector of the International Telecommunication Union (ITU-R) Working Party (WP) 5D developed the evaluation guideline for IMT-A for its performance assessment, and defined performance metrics and technical performance requirements. A summary of these can be found in two ITU-R reports, i.e., the report on guidelines for the evaluation [1], and the report on radio interface requirements [2]. The former report contains the detailed simulation assumptions and the evaluation methodologies for IMT-A. This document represents a significant reference ensuring the proper harmonisation of the tools used by the

independent external evaluation groups for the performance evaluation of IMT-A technology candidates. It includes three components of the evaluation framework:

- *Test environments*, which consist of:
  - A *traffic model* based on the service to be evaluated;
  - A *deployment scenario*, which provides the geographical characteristics where the service is deployed (e.g., indoor hotspot, dense urban area);
  - An *evaluation configuration*, i.e., the assumed evaluation parameters applied to the selected traffic (service) and deployment scenario.
- *Evaluation methodology and procedures* for each key performance indicator (KPI):
  - High-level assessment method, e.g., inspection, analysis or simulation, as defined in more detail in Section 15.2.3;
  - Detailed evaluation method and procedure.
- *Evaluation models*, e.g., channel model, etc.

In 2009 and afterwards, the 3rd Generation Partnership Project (3GPP) has taken the IMT-A evaluation method (with minor enhancements) as the baseline for the assessment of LTE and LTE-A. The main assumptions for system level simulation are captured in [3], while link level simulation considerations (needed e.g., for performance evaluation of advanced receivers) can be found in [4].

On the other hand, the IEEE body has also elaborated a methodology to evaluate IEEE 802.16m, also referred to as Worldwide Interoperability for Microwave Access (WiMAX), which was its IMT-A proposal. This methodology is captured in [5], and has many commonalities with those of 3GPP, although it is more detailed in some parts, such as link-to-system mapping or traffic models.

15.2.2 IMT-2020 evaluation process and framework

The formal evaluation framework for 5G, known as IMT-2020 evaluation framework, is the official process of ITU-R, as also described in Section 17.2.2. The most obvious candidate to meet the IMT-2020 requirements is the New Radio (NR) standard developed by 3GPP.

ITU-R has developed reports on minimum technical requirements [6] and evaluation guidelines [7] for IMT-2020 technology proposals, which will be used in the IMT-2020 evaluation and submission process, also detailed in Section 17.2.2. In addition, Task Group 5/1 in ITU-R is tasked with conducting the sharing and compatibility studies for World Radio Conference 2019 (WRC-19), in order to secure 5G spectrum globally.

The main steps of the ITU-R process and alignment with 3GPP work are shown in Figure 15-1. As an initial step, a Circular Letter to invite technology proposals was released in March 2016. Further steps consist of the submission of technology proposals, followed by their official evaluation. It should be noted that ITU-R itself doesn't develop the technical specifications nor the evaluation of candidates. It instead announces a call for external evaluation bodies and requests the contribution from the scientific world to complete this task.

2015	2016	2017	2018	2019	2020	2021

3GPP: Rel. 13, Rel. 14, Rel. 15, Rel. 16, Rel. 17

ITU: 5G Vision, Requirements, WS, Proposals, IMT-2020 Specs

WRC'15   WRC'19

Figure 15-1. ITU-R process and its alignment with 3GPP specifications.

In [8], the detailed IMT-2020 submission and evaluation process is defined. Nine steps are described to approve a candidate radio interface technology (RIT) or a set of RITs (SRIT) as a part of the IMT-2020 specification. Here, step 2 and step 7 define the "entry criteria" and the "exit criteria" for the process, respectively. Any candidate RIT/SRIT needs to fulfil the requirements defined in step 2 according to the following submission process:

- A RIT needs to fulfil the minimum requirements for at least three test environments: two test environments related to enhanced mobile broadband (eMBB), and one test environment for massive machine-type communications (mMTC) or ultra-reliable low-latency communications (URLLC), representing the 3 main 5G service types defined in Section 2.2;
- A SRIT consists of a number of component RITs complementing each other, with each component RIT fulfilling the minimum requirements of at least two test environments and together as a SRIT fulfilling the minimum requirements of at least four test environments comprising the three service types.

If it fails, the proposed RIT/SRIT cannot enter the ITU-R submission process for IMT-2020. In step 7, a candidate RIT/SRIT that successfully passed the step 2 needs to further fulfil the following requirement to be approved as (part of) IMT-2020: the RIT/SRIT must meet the requirements for all five test environments defined in [7], comprising the three main service types eMBB, mMTC, and URLLC.

This process allows the proposals to target an initial capability as defined in step 2, and then later achieve the full capability with respect to all 5G usage scenarios, for instance through a further development or consensus building. By these means, a powerful and unified 5G standard that gains wide industry and regional support is expected.

Just like in the case of IMT-A evaluation, 3GPP has committed to submit its proposal to the IMT-2020 process in the beginning of 2020. First considerations on scenarios, requirements and models for (but not limited to) this process can be found in [9].

15.2.3 5G PPP evaluation framework

Before the official IMT-2020 evaluation framework was established, several 5G Public-Private-Partnership (5G PPP) projects have investigated the topic of 5G evaluation to satisfy the needs and challenges set for 5G in different fora, all identifying the need to update the evaluation methodologies considered in previous standards. Firstly, because of the novel requirements posed

by new services, scenarios or system configurations, and secondly because of the new technologies that are foreseen to fulfil these requirements, as for instance outlined in [10].

Concerning the 5G services, as explained in Section 2.2, the focus of 5G is not only on mobile broadband, but also on massive connectivity and reliable and low-latency communication. This implies new devices for machine-type communication as well as new traffic models and KPIs (e.g., reliability or connection density). Additionally, the scenarios investigated for 5G are not just based on regular homogeneous base station (BS) placements, but are increasingly closer to realistic heterogeneous deployments. Moreover, the 5G system is assumed to be based on the integration of multiple radio access technologies (RATs), such as evolved LTE, NR, Wi-Fi etc., and involve multi-connectivity among these or among multiple transmission points, as covered in Section 6.5. Finally, 5G will also involve operations at higher frequencies with large carrier bandwidths, and the usage of massive MIMO, as detailed in Section 11.5. These aspects also require extensions of the channel models to properly capture radio propagation at higher frequencies and the correlation of channel characteristics at different carrier frequencies, as described in Section 4.3.4.

With regard to the new technologies that are being proposed to fulfil the 5G requirements, the use of new waveforms, as detailed in Section 11.3, has gained quite some attention. There is a clear impact of this change on the 5G evaluation in the link-level modelling and link-to-system mapping, requiring the development of new models. 3D beamforming is another technology that influences the simulation models and especially the channel models, as 3D extension becomes a must for an accurate evaluation. In addition, device-to-device (D2D) communication capabilities, or the formation of moving networks, should be integrated in the 5G evaluation methodology and also have implications on the channel modelling needs, as detailed in Section 4.3.7.

In response to these factors, a new evaluation framework has been detailed by 5G PPP in [12]. This document analyses 5G use cases investigated in several 5G PPP projects, defines the appropriate KPIs, and proposes a set of performance evaluation models.

The first set of KPIs are the so-called **inspection KPIs**, whose evaluation is based on the examination of statements from each specific 5G proposal. The inspection KPIs are basically questions that can be answered with a yes or no. 5G PPP considers six of these:

- **Bandwidth and channel bandwidth scalability**, referring to the ability of the system to operate with different bandwidths (at least supporting 1 GHz) and carrier frequencies;
- **Deployment in IMT bands**, i.e. allowing to deploy 5G in at least one identified IMT band;
- **Operation above 6 GHz**;
- **Spectrum flexibility**, i.e. the capability to accommodate different downlink (DL) and uplink (UL) transmission patterns in both paired and unpaired frequency bands;
- The support of **inter-system handover** between 5G and at least one legacy system;
- Efficient **support of a wide range of services** over a continuous single block of spectrum.

All inspection KPIs were evaluated positively in [13] for the 5G system proposed in METIS-II. For **analytical KPIs,** namely those KPIs that are evaluated through calculations based on available technical information, the 5G PPP framework assumes the following:

- **Mobility interruption time**: a time span during which a user equipment (UE) cannot exchange user plane packets with any BS during transitions between the cells. This KPI relates to the capability of 5G to provide a continuous connectivity for devices on the move. It was shown that 0 ms interruption time is possible, if multi-connectivity solutions are employed, as discussed in Section 6.5. This is in line with ITU-R requirements.

- **Peak data rate**: the highest theoretical single user data rate, assuming error-free transmission and utilization of all radio resources for the corresponding link direction. This value is linked to the maximum supported number of MIMO streams, modulation order, coding scheme, and transmission bandwidth. Although peak data rates are unlikely to be experienced in realistic operations (a KPI of *experienced* user data rate is a far more accurate approximation, as explained later on), but they show a potential of the cellular system to cater for the needs of broadband services. For peak data rates, values of 21.7 Gbps and 12.4 Gbps were assessed for DL and UP, respectively, which is above the ITU-R target.
- **mMTC device energy consumption**: reflected through the device battery lifetime without recharging and using a single 5 Wh battery, under the assumption that the device is stationary and the energy consumption is related only to communication aspects. This KPI reflects the ability of the 5G system to provide an energy efficient procedure for emerging Internet of Things (IoT) services. Assuming a sporadic data transfer of low payloads, a lifetime of 10 years or more can be reached.
- **Control plane latency**: represents the transition time from an inactive and energy efficient mode (e.g., when devices do not exchange any user data with the network) to an active mode. Low values are necessary for energy efficiency reasons and to provide an always-connected experience. For the new generation of cellular devices, when using the newly introduced radio resource control (RRC) connected inactive state as detailed in Section 13.3, control plane (CP) latency can be as low as 7.125 ms, i.e. far below the target of 20 ms set by ITU-R.
- **User plane latency**: defined as the one-way transmission time of a packet between the transmitter and the receiver. This KPI not only relates to the efficiency of the radio interface, but also, e.g., to the handling of data buffers at the devices side. User plane latency is assumed to comprise the following steps: (1) transmitter processing delay at the BS, (2) frame alignment, (3) synchronization, (4) transmission of a packet over a number of transmit time intervals (TTIs), (5) Hybrid Automatic Repeat reQuest (HARQ) retransmission probability, and (6) receiver processing delay in the UE. Taking into account all these factors and assuming a 0.125 ms TTI for a single packet transmission, a value of 0.763 ms can be obtained, which is below the 1 ms target of ITU-R.

In order to complete the evaluation of 5G system concepts, link-level and system-level simulations are required to assess the KPIs that depend on the actual propagation conditions or system load. The **simulation KPIs** considered in the 5G PPP framework are:

- **Experienced user throughput**: the instantaneous data rate measured separately for DL and UL;
- **Traffic volume density**: the total number of bits correctly received by the infrastructure (in UL) or UE (in DL), measured over a certain geographical area and a period of time divided by the considered area and period;
- **End-to-end (E2E) latency**: one trip time or round-trip time in a packet transmission. In each case, the time is measured at the interfaces between layers 2 and 3;
- **Reliability**: the percentage of packets successfully received in a system within the maximum E2E latency. In this context, **availability** is typically defined both as the percentage of locations where the user gets the quality of experienced desired, and the probability of a service not being blocked;

- **Retainability**: the percentage of time when transmissions fulfil the experienced user throughput or reliability requirements;
- **mMTC device density**: the maximum density of users supported in a spatial area with a minimum percentage of messages correctly received;
- **RAN energy efficiency**: defined as the overall energy consumption in the 5G RAN compared to the energy consumption in the RAN of legacy systems;
- **Supported velocity**: an estimate of the maximum velocity for which a certain data rate can be achieved'
- **Complexity**: a KPI that may refer to the size or volume of an analogue component, the number of operations of a digital process, or the cost of a certain implementation.
- **Coverage**: While different definitions exist, one common definition in the context of broadband services [11] is that this is calculated as the experienced user throughput over the target value and expressed as a percentage. More precisely, a user is assigned a value of 100 % if its throughput is equal or higher than the target, and a proportionally lower percentage otherwise. It is averaged over all realizations.

The novelty in the 5G PPP framework does not only come from the definition of new KPIs, but also in the proposal of new simulation deployment scenarios and models. These models are covered in [12], and detailed in several deliverables from METIS-II [13], FANTASTIC-5G [11], mmMAGIC [14] and SPEED-5G [15]. In addition, some characteristics of the 3GPP evaluation framework for 5G covered by [9] have been incorporated in the 5G PPP framework.

On one hand, proposed models and configurations consider synthetic deployment scenarios, namely indoor hotspot, urban macro, outdoor small cells, and rural macro or long-range communications. Several possible configurations are provided for each one. On the other hand, the major novelty of the deployment scenarios in the 5G PPP framework comes from the definition of realistic deployment scenarios, such as an indoor office scenario or the so-called *Madrid grid* [13]. The definition of these aims at providing more realistic conditions for the evaluation of 5G.

The deployment scenarios definition is complemented with the specification of the user, traffic, channel and mobility models for individual use cases. Concerning the channel modelling, the basic references are the IMT-A channel models [1][3], extended to support higher frequencies, bandwidths, numbers of antennas and 3D models [16]. Moreover, additional aspects have been included for 3D modelling [17], the support of high speeds [18], for propagation in small cells [19], and for vehicular communication links including direct communication between vehicles [20], as also covered in Section 4.3.7. With regard to traffic models, both full buffer and bursty File Transfer Protocol (FTP)-like traffic is considered. The latter have different parameterizations ranging from the simplest case with fixed packet sizes and packet inter-arrival times to complex random values of these parameters generated according to exponential or Poisson distributions.

Compared to analytical KPIs, evaluation results for simulation KPIs vary strongly between use cases, deployment scenarios and traffic models, therefore it is impractical to discuss them without a wider context. The following section gives exemplary evaluations of simulation KPIs and a basic background information derived from selected 5G PPP projects.

15.2.3.1 FANTASTIC-5G

The goal of FANTASTIC-5G was to define a flexible multi-service air interface for the 5G system [21]. It focused on 7 different use cases (i.e., 50 Mbps everywhere, high speed train, sensor networks, tactile internet, automatic traffic control/driving, broadcast like services, and dense urban society, cf. Section 2.5) and developed a number of technical solutions allowing the 5G air interface to significantly improve the performance of the baseline 4G technology, as well as to satisfy the respective 5G KPIs.

Wide area coverage scenario.
Mobile users require high-speed Internet connection for advanced interactive services. To provide a satisfactory experience, a minimum data rate should be consistently provided to all the users, even at the cell-edge. Most studies set this limit to at least 50 Mbps in DL and 25 Mpbs in UL [22]. To reach this goal, FANTASTIC-5G enhanced the typical macro-cell deployment with the usage of massive MIMO, based on a 2-stage precoding strategy [23]. The first stage uses a grid-of-beams configuration, which uses a beamforming matrix to create a regular grid of highly directional signals in the angular domain. The second stage is a regularized zero-forcing precoder working on top of the first stage, which can be seen as a virtual array of high-gain antennas. Additionally, FANTASTIC-5G also formulated a coordinated beamforming method for inter-cell interference reduction, derived from [24]. It involves a sub-sectorization of the cells and the activation of specific patterns of sub-sectors, which reduces the average interference among adjacent cells. To demonstrate the performance gain provided by developed technical components, various system level simulations were conducted. First of all, three main scenarios were taken into account: (1) rural, with 100 users/km² and the inter-site distance (ISD) set to 1000 m, suburban with 400 users/km² and an ISD of 600 m, and urban with 2500 users/km² and an ISD of 200 m. At the physical layer, an array of 16x8x2 antennas with half-wavelength spacing over a bandwidth of 100 MHz was used [25].

Evaluation results for a rural scenario are depicted in Figure 15-2 and demonstrate the traffic density going from 1.38 Gbps in legacy LTE-A up to 12.4 - 14.7 Gbps in the new solutions proposed for 5G. For the other scenarios, the absolute values change depending on the ISD, but the gain is similar. Figure 15-3 shows the service coverage, defined here as the ratio of the experienced user throughput over the target throughput (50 Mbps in DL), limited to 100 % when the throughput is exceeded. Moreover, in this case it is possible to observe that the new techniques can greatly outperform LTE-A, as the coverage is increased from 25-50 % to more than 95 %.
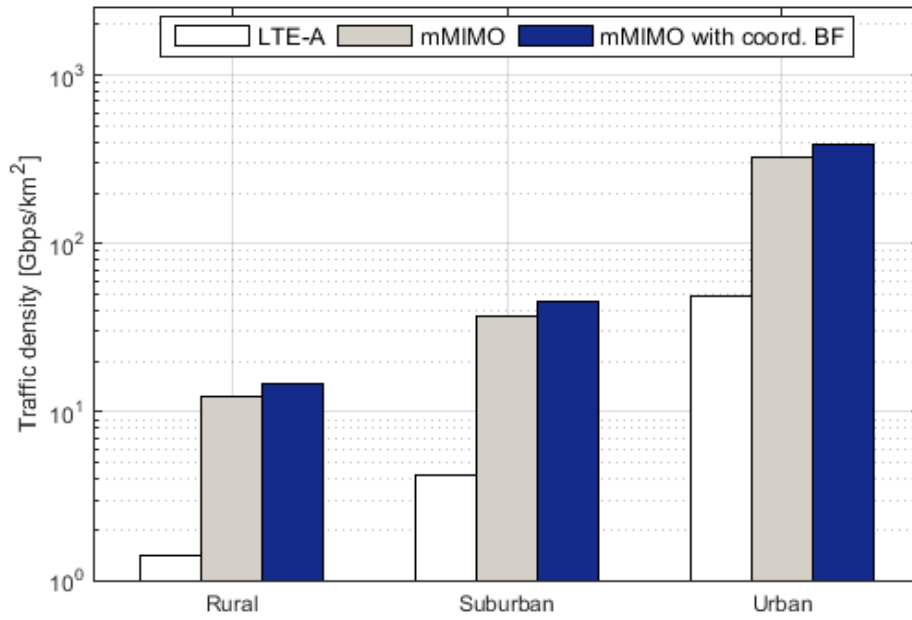
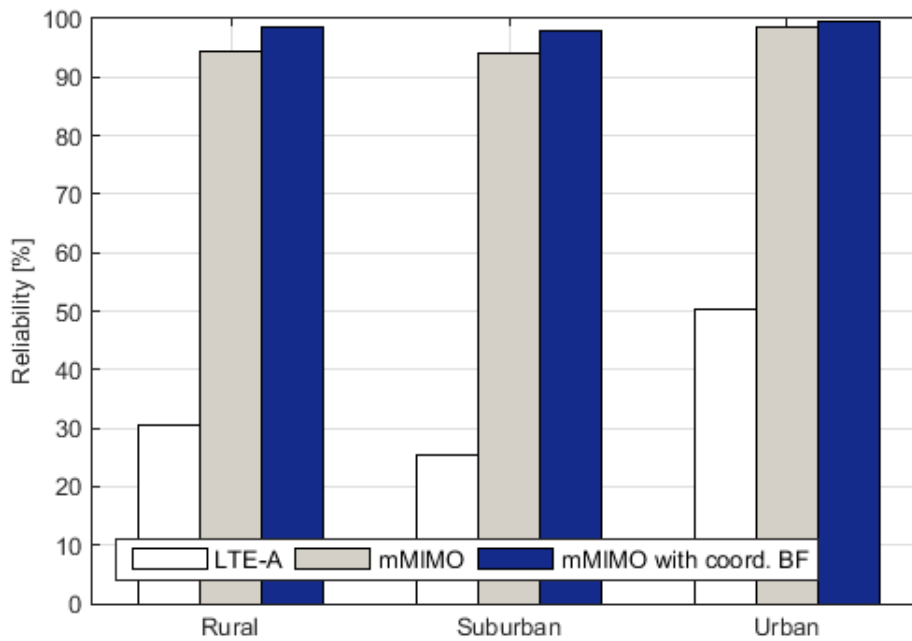Figure 15-2. Traffic density for the "wide area coverage" deployment scenario.

Figure 15-3. Coverage obtained for a 50 Mbps target user throughput in DL in the "wide area coverage" deployment scenario.

Video broadcasting scenario

In this scenario, the aim is to investigate the efficiency of one-to-many transmission techniques able to allow a large number of mobile users to receive the same real-time video stream. As a baseline approach, the 4G technology proposes the multicast-broadcast single frequency network (MBSFN) technique [3], where multiple BSs transmit the exactly same signal under tight synchronization, and all the replicas add up in power at the mobile users. To improve efficiency and reliability, FANTASTIC-5G enhanced the simple MBSFN approach with two new technical components, namely an adaptive selection of the modulation and coding scheme (MCS) and a HARQ retransmission of broadcast packets [26]. Specifically, the MCS is chosen based on the users' channel quality indication (CQI) feedback, similarly to non-MBSFN operation, and broadcast packets that are received with errors by some users are retransmitted only to such users, over dedicated unicast bearers using HARQ. This concept is evaluated using a transmission of a high definition video encoded at 17 Mbps, assuming a bandwidth of 20 MHz and 6 dedicated sub-frames for every radio frame made up by 10 sub-frames [25]. Without loss of generality, it is assumed that the single frequency network (SFN) is extended on a large scale. However, for each cell, only the closest rings of adjacent cells are able to boost the channel quality through a constructive signal. The others, instead, become progressively more interfering as the propagation delay exceeds the duration of the cyclic prefix. Results from system-level simulations are reported in Figure 15-4 and 15-5. They show that there is an optimal value for the MCS, because for lower values the throughput is reduced, and for higher values too much reception errors occur. The most advanced approach automatically selects the best MCS value without static or manual configuration. In addition, the packet loss rate is reduced compared to the baseline approach (for any MCS value) thanks to the HARQ retransmissions.
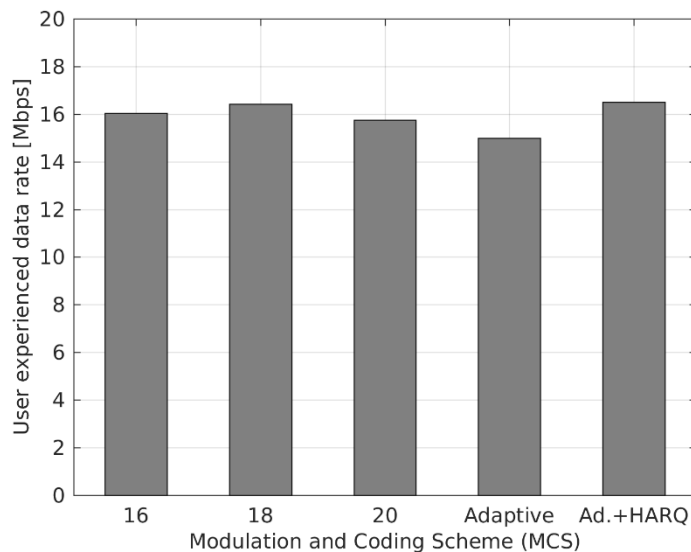

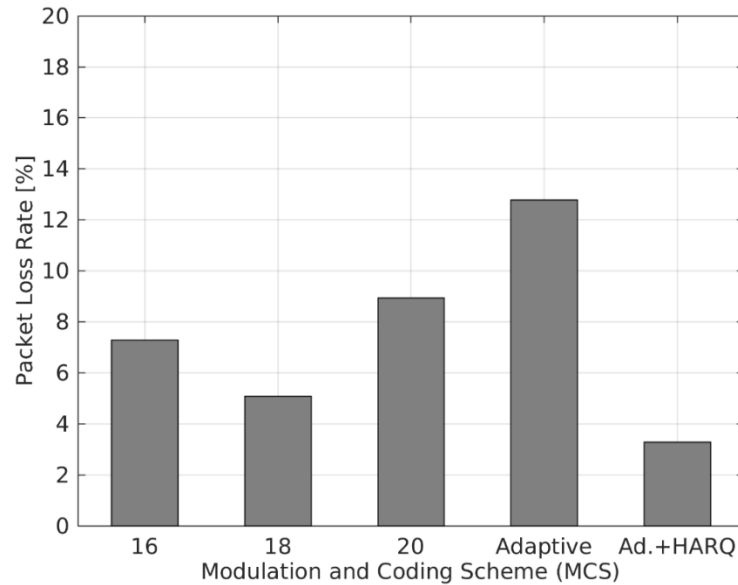Figure 15-4. User experienced data rate for the "video broadcasting" scenario.

Figure 15-5. Packet loss rate for the "video broadcasting" scenario.

15.2.3.2   METIS-II

METIS-II has developed the overall 5G RAN design and provided the technical enablers needed for an efficient integration and use of the various 5G technologies and components. Additionally, METIS-II has provided the 5G collaboration framework within 5G PPP for a common evaluation of 5G RAN concepts, and prepared concerted action towards regulatory and standardisation bodies. A summary of the simulation KPIs is captured in Table 15-1 and followed by two exemplary evaluations of the METIS-II use cases.

Table 15-1. Summary of simulation performance evaluation results from METIS-II [13].

Use case	KPI	Expected performance	Evaluated performance
Dense urban information society	Experienced user throughput	300 Mbps	> 1 Gbps
Virtual reality office		Up to 5 Gbps	7.85 Gbps
Broadband access everywhere		50/25 Mbps for DL/UL	50/25 Mbps for DL/UL
Massive deployment of sensors and actuators	mMTC device density	> 1 million/km^2	4 mln/km^2

Connected cars	Reliability	99.999 % at 50/1000 m for urban/highway	99.999 % at 45/150 m for urban/highway with 20 MHz Requirement fulfilled with 40/100 MHz for urban/highway

Dense urban information society use case

In this use case, users exchange information with cloud servers and with other users, devices or sensors located in close vicinity, see also Section 2.5.1. A heterogeneous network of BSs deployed in a dense urban environment caters for connectivity requirements. Two important KPIs that are used to quantify 5G performance are network energy efficiency, covered in more detail in Section 15.3.4, and traffic volume density. The latter reflects the capability of the 5G system to handle massive traffic, which is expected especially in cities due to the large concentration of users in limited areas. Results presented in Figure 15-6 show the DL traffic volume density that may be supported by 5G for a given traffic load, represented by the packet arrival rate for individual users. It should be highlighted that these results were obtained for simulations limited to operations using 100 MHz only, mainly for simulation complexity reasons. In practice, bandwidths as high as 1 GHz can be expected in such deployments, leading to a potential improvement factor of at least 10, given that beyond a linear scaling of capacity with system bandwidth also pooling and multiplexing gains would be expected.
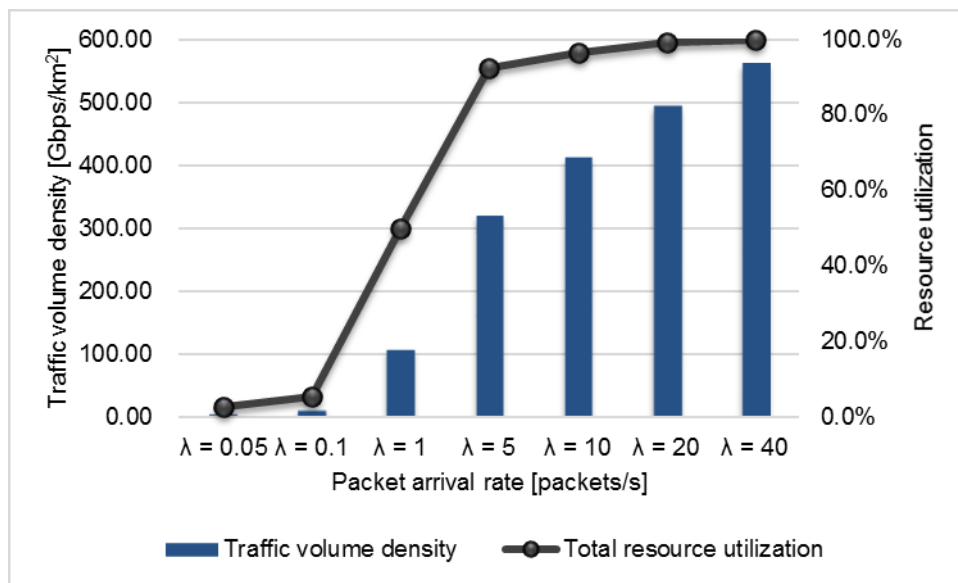

Figure 15-6. Traffic volume density vs. packet arrival rate for dense urban information society.

Connected cars use case

This use case addresses the information exchange among vehicles and with the infrastructure to enable a safer and more efficient transportation and real-time remote computing for mobile terminals. The evaluations conducted in METIS-II focused on traffic safety and efficiency, where a reliability of 99.999 % for the transmission of packets is required with a maximum E2E delay of

5 ms, considering certain communication ranges that depend on the mobility scenario. To assess this requirement, it is assumed that each vehicle periodically broadcasts packets of at least 1600 payload bytes with a repetition frequency of at least 5-10 Hz. Three main mobility environments are considered, namely an urban, rural and highway environment, with a maximum speed of 60 km/h, 120 km/h and 250 km/h, along with a required coverage range of 50 m, 500 m and 1 km, respectively. Further, three relevant scenarios are envisioned, namely a realistic urban scenario, earlier on already introduced as the so-called Madrid grid, a synthetic urban scenario, and a highway scenario [27]. The last two are based on scenarios defined by 3GPP in [20]. Concerning the density of vehicles, the specific values considered in the evaluation were 1000 vehicles/km^2 in the urban realistic scenario, 595 vehicles/km^2 in the urban synthetic scenario, and 10.25 vehicles per lane and km in the highway scenario. METIS-II has assessed the ability of a preliminary 5G system with direct vehicle-to-vehicle (V2V) communication and a centralized resource allocation over the 5.9 GHz band to fulfil the requirements of the connected cars use case. From the technical enablers proposed in METIS-II, this system considers the availability of large bandwidths for V2V communications, but no multi-antenna transmission scheme is used. Therefore, the reader should consider these results as a baseline evaluation that could be further improved by a final 5G solution. The latest results of this evaluation at the time of writing this book can be found in [28], with more details on evaluation models and results. As an example, Figure 15-7 shows the packet reception ratio for different distances and system bandwidths in the urban realistic scenario. The reliability requirement is fulfilled using a bandwidth of 40 MHz, while the system is already close to the target performance with a bandwidth of 20 MHz. Results are similar in the urban synthetic scenario. On the contrary, in the highway scenario, a large bandwidth of 100 MHz is needed to fulfil the reliability level required for a range of 1 km, while the range would be 600 m with 50 MHz. In conclusion, the results have shown that the requirements can be fulfilled with system bandwidths between 30 and 100 MHz, depending on the scenario. In addition, it seems feasible to reduce the needed bandwidth down to around 20 MHz in urban scenarios and 50 MHz in the highway scenario with a more advanced 5G system.
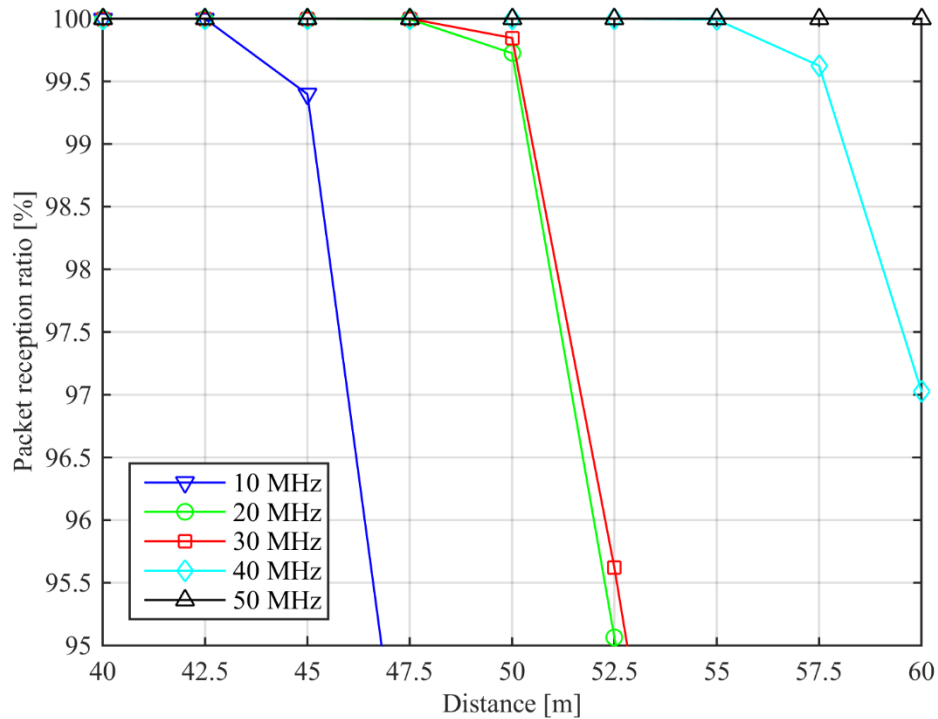
Figure 15-7. Packet reception ratio vs. distance in the Madrid Grid urban scenario.

15.2.3.3   SPEED-5G

SPEED-5G's main objective was to investigate and develop technologies that address the well-known challenges w.r.t. the predicted growth in mobile connections and traffic volume by successfully addressing the lack of dynamic control across wireless network resources, so far leading to unbalanced spectrum loads and a perceived capacity bottleneck. Consequently, SPEED-5G has focused on resource management with the three degrees of freedom of densification, rationalized traffic allocation over heterogeneous wireless technologies, and a better load balancing across available spectrum.

In particular, SPEED-5G has focused on 4 different use-cases: Massive IoT, broadband wireless, ultra-reliable communications, and high-speed mobility. The project has investigated various scenarios where capacity demands are the highest, but also where extended dynamic spectrum access (eDSA) [29] is expected to be most effective for exploiting co-operation across technologies and bands.

Future dense urban use case
One of the main solutions that have been investigated is in the context of broadband wireless with hierarchical management capabilities; that is, blending distributed and centralized management of ultra-dense multi-RAT and multi-band networks. In SPEED-5G, centralized management is used as a baseline and can be expanded with distributed management by moving management decisions related to RAT, spectrum or channel selection closer to the node level.

In order to obtain QoS and capacity expansion, operators and regulatory bodies are increasingly pursuing policy innovations based on the paradigm of shared spectrum, which allows spectrum bands that are under-utilized by primary owners to be exploited opportunistically by secondary users. Specifically, the solution captured below focuses on the Spectrum Access System (SAS) in the 3.5 GHz band, which consists of a hierarchical three-tier model: an incumbent with high priority, Priority Access Licenses (PAL) with medium priority, and General Authorized Access (GAA) with the lowest usage priority, see also Section 3.2. The higher priority users have a better utilization of channels compared to those at lowest priority. PAL and GAA users are controlled by the SAS and thus must register and check all of their operations in order to provide an interference-free environment to higher-tier users (i.e., incumbents). SAS is a monitoring system that checks whether a given user category can transmit over a specific channel, or whether the user should change the channel in order to avoid any interference to the higher priority users. In this way, priorities are formed between the users, namely high, medium and low priority for the incumbent, PAL and GAA users, respectively.

Performance evaluation was carried out in a heterogeneous network deployment with 19 macro BSs complemented with 285 small cells, and with 8000 UEs, each one downloading 2 MB packets with a different packet arrival rate, according to an FTP traffic model. Figure 15-8 illustrates the relative average DL throughput of a UE belonging to different access priorities. Specifically, it is shown that users (especially with higher priority, such as incumbents) can experience higher throughputs as the packet arrival rate increases, but after a certain point the PAL and GAA (i.e., the lower priority users) start to compete for radio resources and their throughput drops. Moreover, the relative packet transmission latency is better for higher priority users as the packet arrival rate increases, as depicted in Figure 15-9.
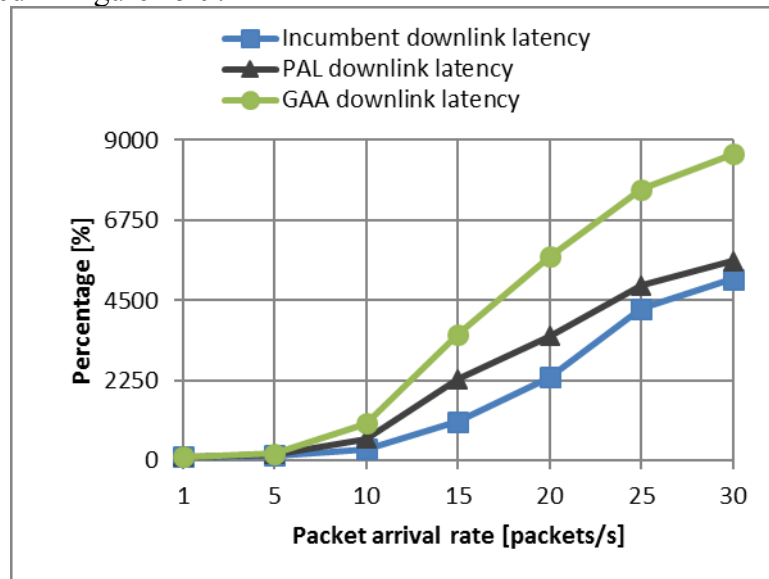

Figure 15-8. Relative increase of average DL throughput and latency for different access priorities. Performance achieved at packet arrival rate of 1 packet/s (low load) is the baseline (100 %).
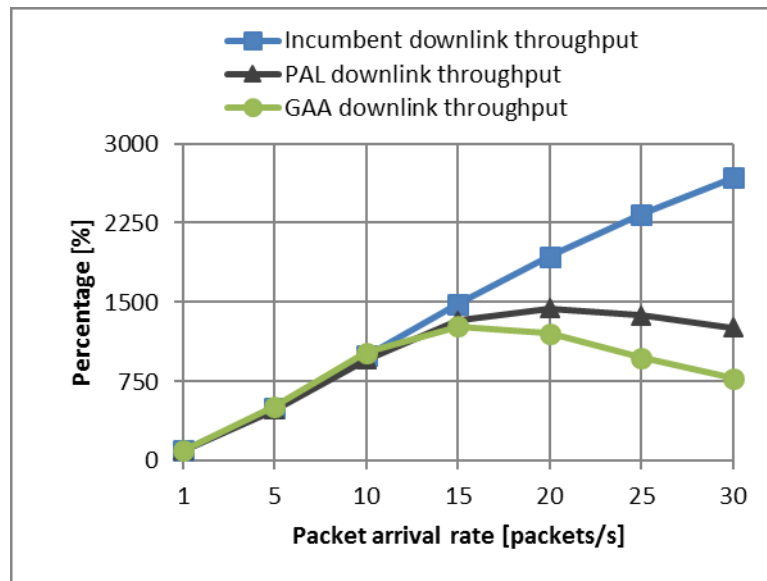
Figure 15-9. Relative increase of average packet transmission latency for different access priorities. Performance achieved at packet arrival rate of 1 packet/s (low load) is the baseline (100 %).

15.3 Network energy efficiency

As stated previously, network energy efficiency is an important indicator for the deployed 5G networks, considering the growth of traffic and the number of users. In this section, we elaborate on why the energy efficiency has become so relevant recently, and we introduce energy efficiency metrics and methods of measurements. Finally, we propose a preliminary evaluation, based on simulations, of the energy efficiency of 5G versus the legacy radio systems.

15.3.1 Why is network energy efficiency important?

Network energy consumption translates to a substantial cost for operators. In mature markets, energy costs account for 10-15 % of the total network operating expenditure (OPEX) and can reach up to 50 % in developing markets with a high number of off-grid sites, or where only a poor-quality electricity grid is available [30].

Because of the rapidly increasing usage of mobile broadband connectivity, the largest network operators have recently reported a growth of 15-35 % in their network energy consumption, the main reason being an increasing demand for mobile network coverage and capacity. Part of this growth is driven by the rise of the global mobile broadband subscriber base, which is expected to grow beyond six billion subscriptions globally in 2017, increasing by 10 % per year, with a mobile broadband penetration likely hitting 100 % by 2020. On top of this, each mobile broadband subscriber will use an average of 25-50 % more data per year, resulting in an expected sevenfold increase of mobile data traffic between 2016 and 2021 [31].

In mature markets, where the number of subscribers is saturated, operator revenue has been flat. Users expect faster services with higher data rates, but are rarely willing to pay additional money. Operators must hence provide the growing data rates at constant cost. It is often claimed

that energy efficiency was not considered in earlier mobile telecom generations, while in fact, it is the contrary: increasing energy efficiency has been one enabler for the rapid data traffic growth.

Energy cost is not the only driver to increase energy efficiency. In addition, global warming is a direct result of the greenhouse gas emissions caused by power consumption in general. The European Commission has correspondingly set three key objectives, known as the "20-20-20" targets: a 20 % reduction in European Union (EU) greenhouse gas emissions from 1990 levels, an increase of the share of EU energy consumption produced from renewable resources to 20 %, and a 20 % improvement in the EU's energy efficiency [32]. Measuring the progress towards these goals is based on accurate data on energy consumption and emissions supported by several telecom power consumption and efficiency test standards developed by ITU-T, European Telecommunications Standards Institute (ETSI), the Alliance for Telecommunications Industry Solutions (ATIS), etc.

5G will introduce several new network features having a large impact on network energy consumption and efficiency. For instance, massive MIMO and antenna beam steering will be essential to increase link budget and compensate fading in particular for millimetre-wave carrier frequencies, and in general to increase spectral efficiency per area. However, the introduction of such new solutions can potentially increase power consumption. Moreover, IoT services involving a huge number of connected devices and increasing coverage requirements, as well as services related to ultra-high reliability and very high data rates and hence requiring simultaneous connection via multiple frequency bands, pose a serious challenge towards an improved energy efficiency.

15.3.2 Energy efficiency metrics and models

To quantify network energy efficiency, the following metrics are often used [33][34][35][37]:
- **Energy per bit**, especially used in urban environments, where the planning of the network is usually capacity-constrained. Here, E stands for consumed energy in a given observation period measured at Medium Access Control (MAC) layer:

$$\lambda_I = \frac{E}{I} \ [J / bit]$$

Under certain circumstances also throughput vs. power consumption [bps/W] can be applied as efficiency indicator.
- **Power per area unit**

$$\lambda_A = \frac{P}{A} \ [W / m^2]$$

typically applicable in suburban or rural environments, where the planning of the network is mainly constrained by the achieved coverage (A is the area coverage).

Both metrics are often applied by academic organizations, while for reporting and product evaluation purposes the inverse measures are usually used:
- **Number of delivered bits per energy** [bit/J] is another common metric to assess equipment and operational network energy efficiency in some environmental standards from ITU-T, ETSI, etc.
- **Coverage area per daily energy consumption** [m^2/J] or [m^2/Wh] is applied as energy efficiency parameter for operational mobile networks, while the previous metrics are often

used also in simulated scenarios to estimate the energy efficiency without real measurements of the involved parameters.

In order to evaluate the energy efficiency of a network in a simulated scenario, power consumptions models of the elements in the network are needed. One of the first widespread power consumption models was developed in the OPERA-net project [36], which was initiated in 2008 as one of the first international projects dealing specifically with mobile network energy efficiency. One of its key objectives was to develop metrics and KPIs for mobile network efficiency. Three operational sites (rural, sub-urban and urban) were selected and equipped with several power meters to allow detailed measurements of the power consumption of the different elements of the BSs sites. Simultaneous temperature and load measurements (i.e., the amount of DL data per cell) allowed analyzing variations during the day and the correlation between power consumption, data rate and temperature. Based on these measurements, a BS model was created, which described the power consumption of different configurations (i.e., number of sectors, number of transmitters per sector, maximum installed radio frequency (RF) power and actual load) but was at the same time simple enough to be directly applied in a network planning tool. This enabled simultaneous simulations of network capacity, coverage and power consumption for different network configurations. The power consumption of the BS was calculated as:

$$P_{BS} = \tau P_{P} + n P_{TRX} + (k_1 P_{RF1} + \cdots + k_n P_{RFn})/c$$

where, $k_1 \ldots k_n$ are load factors describing the fraction of available RF power transmitted in sectors $1 \ldots n$, with n being the number of installed sectors, τ denotes the increment of installed baseband processing capacity, P_P is the power consumption of the processing unit, while P_{TRX} describes the power consumption of the radio module, P_{RFn} represents the maximum RF output power of sector n, and c is a direct current (DC) to RF conversion slope parameter. This model is suited particularly to analyze the efficiency of different network configurations within a network planning tool, e.g., the effect of an increasing number of cells per macro site or the addition of small cells in a macro layer. The model requires the knowledge of practical BS power consumption parameters for different load levels, which can be derived from the BS manufacturer's data sheets. It should be noted that ETSI has created a standard to measure power consumption of BSs at different load levels [35].

Although a significant development effort was spent to decrease the power consumption as a function of the load in active mode, today's BSs show relative large and fixed power consumption level already in idle mode. BSs sleep modes, where a single cell or even a complete BS is put in hibernation, play an important role to minimize network power consumption during low load levels and periods. However, the broadcast and pilot channel requirements of current radio systems limit the time were sleep modes can be activated.

To overcome this, the 5G system will introduce novel signalling channel approaches, and 5G BSs will be specifically designed to enter different sleep modes that are characterized by a different extent of functionality deactivation and consequently different extents of power savings and recovery times. For instance, it will be possible to switch fast to a sleep mode with medium power saving, allowing to switch back again fast, as opposed to hibernating the BS, involving a longer recovery time. Consequently, a new power model, which allows to model different levels of component (de-)activation and time-variant power consumption, is therefore needed.

This topic was in the past first investigated by the EARTH project [38], which addressed the global environmental challenge by studying and introducing effective solutions to lower energy consumption and increase energy efficiency of mobile broadband communication systems, without affecting users' perceived QoS and system capacity. Some interesting results from the project are collected in [39]. More specifically, some of the achievements from the EARTH project were architectures, network solutions and deployment strategies to achieve the goals of energy efficiency increase in future networks, as well as an in-deep analysis of the energy consumption sources in BSs. The project is for instance well known for the introduction of the so-called *power model* of a BS, representing a detailed investigation of the energy consumed in the different parts constituting a radio BS hardware [40].

The power model concept was carried on and refined by the GreenTouch consortium [41] which investigated energy efficiency aspects from 2010 to 2015. The ambitious goal of GreenTouch was to evolve the cellular network to ensure an increase of energy efficiency by a factor of 1000 comparing to the level in 2010. To achieve this goal, GreenTouch studied improvements in energy efficiency of mobile, fixed and core networks.

The most recent power model proposed for 5G in METIS-II [13] considers power consumption behaviour for various deployments, parameterization capability and flexibility. It allows describing the actual power consumption behaviour of the whole BS under different deployment solutions and network statuses.

As it can be seen in Figure 15-10, a BS's instantaneous power consumption is basically proportional to the bandwidth load level λ with a constant power spectrum density α_{PSD}. As the load level grows, the overall power consumption of the BS increases accordingly. P_{max} is the maximum output power, while P_0 is the power consumption at the minimum non-zero output power due to load-independent operation. When the load is low, the BS can switch to micro-discontinuous transmission, which means that instead of continuous operation the BS is rapidly switched into sleep mode for a very short interval, as anticipated in the initial part of this section. During this time, the power consumption will further decrease to P_{sleep}, which denotes BS power consumption in a sleep mode.
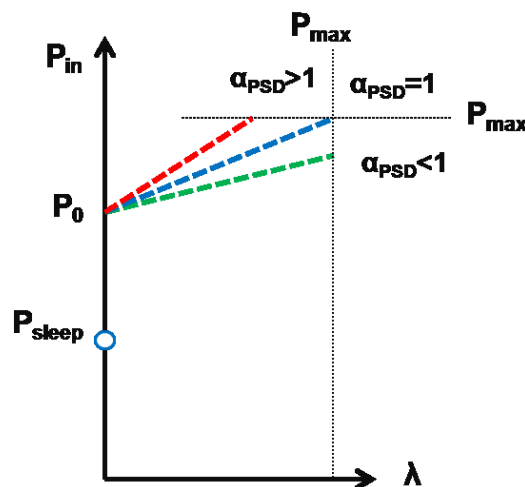

Figure 15-10. Illustration of power consumption behaviour of a BS with a constant power spectrum density.

Note that the actual power consumption of the BS is tightly connected with the BS transmit power, or equivalently, to the power spectrum density ratio α_{PSD}, which is defined as the ratio of the actual power spectrum density to the one with maximum transmit power multiplied with the total bandwidth. Based on Figure 15-10, the overall power consumption behaviour of a BS is calculated as:

$$P_{BS} = \begin{cases} n\left(P_0 + \Delta_p P_{max} \lambda \alpha_{PSD} + P_1 \lambda\right), & 0 < \lambda < 1 \\ nP_{sleep}, & \lambda = 0 \end{cases}$$

where n is the number of sectors in the BS, Δ_p is the slope of the load dependent power consumption largely determined by the radio unit efficiency, and P_1 is baseband related power consumption.

The models described so far require relatively detailed knowledge of the BS, but at the same time omit some or all the site-specific factors. Practical network deployments include different variations of BSs: some are optimized for area coverage, others for high capacity; some are designed for extreme climate conditions and others for indoor use only. Specific models are developed for academic purposes to allow a simulation with additional environmental parameters but without increasing the model complexity. All these models have in common that they focus on BSs only, without considering that what is really important for the operators is the power consumed in the whole site, including also all the other equipment necessary for the operation of the network, such as backhaul and fronthaul, see Chapter 7. The countless variants of sites with their specific needs and equipment are very difficult to be taken into account. The simplest way to take the site elements into account is to define a *site efficiency* as the average BS power consumption divided by the average total site power consumption (or alternatively the inverse, which is called the power usage effectiveness).

In this sense, the following sections describe two different set of measurements of the energy efficiency. First, we consider the energy efficiency measured in laboratory environments, typically for the BSs and the most widely used equipment in the network, as detailed later. Secondly, we consider the measurement of the energy efficiency of a whole network in a live environment.

15.3.3 *Energy efficiency metrics and product assessment in the laboratory*

As a first step to improve and measure the efficiency of the separate telecom equipment of a network, specific laboratory test standards have been developed for BSs, routers, etc. The standards allow assessing the efficiency for equipment in standalone mode under defined laboratory conditions.

In 2009, the ETSI Environmental Engineering (EE) Technical Committee published the specification on the energy efficiency of mobile radio BSs for Global System for Mobile Communications (GSM) and Wideband Code Division Multiple Access (WCDMA). This standard is regularly updated to cover the latest development in BS technology [34]. In 2013, the specification on the energy efficiency of routers and switch equipment for core, edge and access routers followed [42]. The defined metric is based on the so-called energy efficiency ratio of equipment and is defined as the throughput obtained with 1 W of power. A set of weights is given to consider different load levels of the equipment as well.

In 2014, a specification for the energy efficiency of mobile core network and radio access control equipment was released [43], covering metrics and measurement methods.

Network performance including its power consumption can be modelled and simulated. However, simulations always depend on assumptions that might be different from real operating conditions. These simulations usually cover the telecom equipment, but it is rarely possible to include all the support equipment (like air-conditioning, security systems, lightning, networking equipment related to fronthaul or backhaul etc.) installed at the different sites. Laboratory and field measurements are therefore unavoidable to measure the actual network performance.

Ultimately, because of the complex nature of telecom networks, product-based methods cannot describe the actual mobile network efficiency. Many features that are influencing network efficiency are not visible in a stand-alone test of the BS in a laboratory, but appear when considering the whole network. Such effects are particularly visible between different network generations and, of course, will have to be considered also when 5G equipment will be measured and evaluated with respect to previous generations.

15.3.4 Numeric network energy efficiency evaluation

The standardization activity for the evaluation of RAN energy efficiency started in 2012 with the publication of a technical report [44] that paved the way for the activities that followed and were summarized by the publication of the standard [35]. This standard, which has been also adopted as a recommendation by the ITU-T Study Group 5, presents metrics and methods to measure in a live environment the energy efficiency of mobile networks, including coexisting 2G, 3G and 4G systems.

The metrics proposed in this standard are twofold: there is the metric based on the ratio between throughput and energy consumed to deliver that throughput, and a metric based on coverage (i.e. the area covered by the network) and energy consumed by hardware providing this coverage, similar to what was described in Section 15.3.2, but applied to real networks. This duality is intended to cater for cases where the network is deployed for capacity purposes and those for coverage reasons mainly.

The method is based on a set of measurements (i.e., energy, data volume, coverage) made directly on field in the network under test, where the network is split into so-called "partial" networks, which are manageable in terms of number of requested measurements.

To extend the application of the method to wider networks, an extrapolation method is proposed. The specification is not yet applicable to 5G networks, but the activity to evaluate the extension of the specification to 5G networks is ongoing in order to suggest a set of additions and modifications to the existing specifications to cover also the new system.

Following the overall setup specified in the mentioned standards, in order to estimate the network efficiency of the 5G networks, a network energy efficiency process is proposed in [13]. This process captures the assessment of both rural and dense urban deployments, as well as a 24 hours timeframe, to account for the spatial and temporal fluctuations of traffic. This approach allows to compare 5G solutions aiming at coverage-limited (macro and rural BSs) and capacity-limited deployments (micro BSs and small cells). The main idea of the process is to prove that 5G can provide similar power consumption w.r.t. to the traffic witnessed by early 4G deployments, even considering the massive traffic uptake expected in 2020 and beyond. The defined procedure consists of the following steps:

1. Calculate the expected traffic volume density for a 5G dense urban deployment and estimate corresponding packet inter-arrival time (IAT);
2. Scale the obtained IAT to account for different load levels of three periods calculated for traffic profiles proposed in [40];
3. Repeat steps 1 and 2 for rural 5G network deployments, considering different experienced user data rates;
4. Derive the total 5G radio network power consumption at a given load via simulations based on calculated IATs and load points;
5. Repeat steps 1-4 for baseline 4G deployments, considering 1000x lower traffic volume and different deployments of 4G versus 5G systems;
6. Integrate obtained results with network-specific weights (which can be different, e.g., from country to country) and compare 5G power consumption to 4G, to derive the overall energy efficiency improvement.

In [13], the network energy efficiency performance of the dense urban information society use case was evaluated, and the outcome is captured in Figure 15-11 and 15-12.
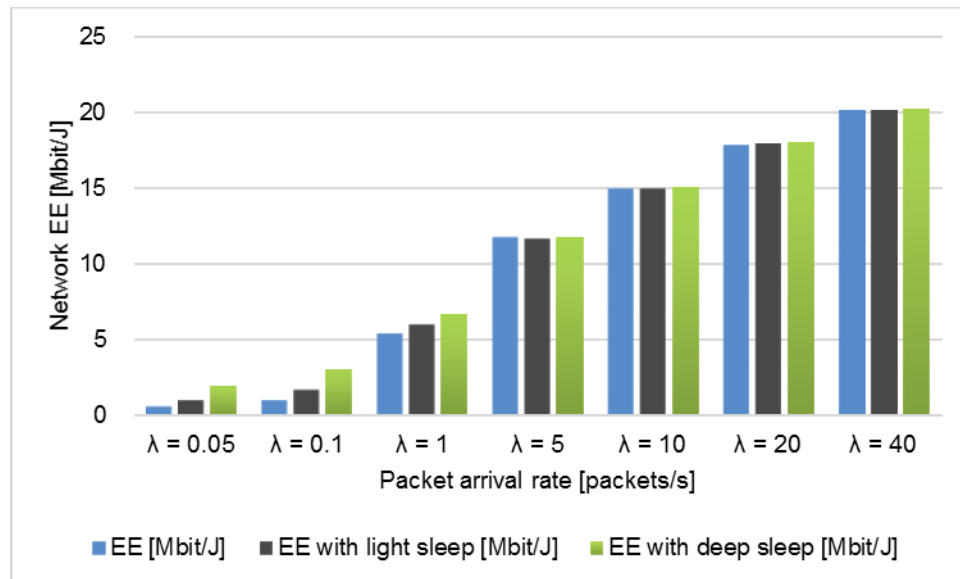

Figure 15-11. RAN energy efficiency for the dense urban information society use case.
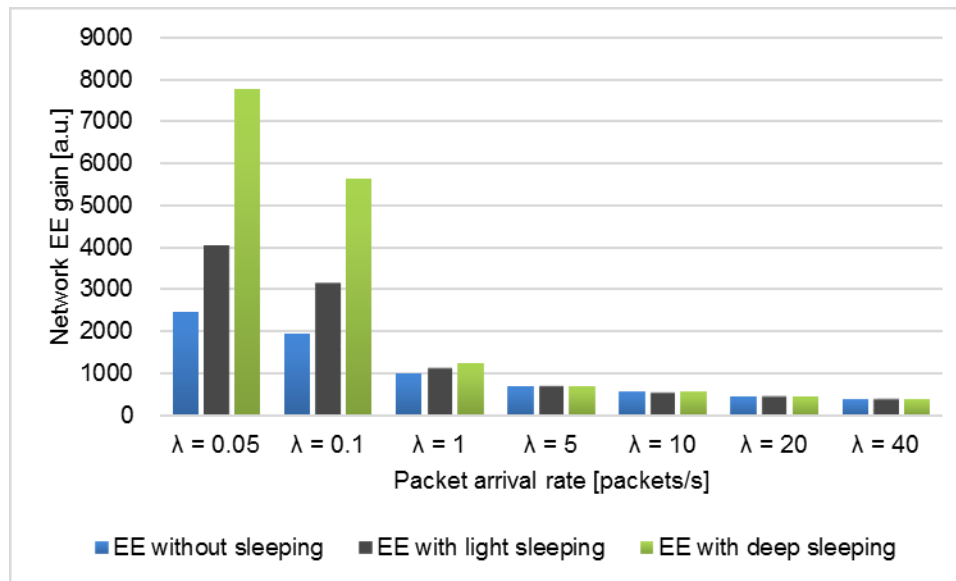
Figure 15-12. RAN energy efficiency gain for the dense urban information society use case over a baseline 4G deployment.

From Figure 15-11, it can be observed that higher 5G RAN energy efficiency performance is expected for higher traffic load levels, as more traffic can be delivered while the ratio of load-independent static power consumption could be reduced accordingly. In addition, advanced sleeping strategies can achieve further performance gains, especially in low load scenarios. Figure 15-12 proves that when the load level is low, a high performance gain of 5G over 4G can be achieved, and similarly this gain increases as more advanced sleeping strategies are implemented. Even when the system load level is very high, the improvement is noticeable, which is mainly due to the introduction of small cells and more energy efficient hardware for 5G.

15.4   Techno-economic evaluation and analysis of 5G deployment

This section covers the economic assessment of the deployment of a new technology such as 5G both in terms of OPEX and CAPEX, by describing a methodology for this assessment and presenting the overall techno-economic evaluation and analysis.

15.4.1 Economic assessment of new technology deployment in mobile networks

Before introducing a new technology such as 5G, a financial assessment to support the decision about its launch is needed. Economic studies must analyze rational economic criteria in order to make the deployment decision and to choose the best scenario for making the technology, being not only profitable in a long term, but also profitable in a relatively short period of time.

The economic analysis to judge the profitability of a new technology deployment is generally based on a free cash flow analysis: the difference between revenues and savings on one side and expenditures on the other side. The new technology is economically interesting for a company when the cash flow generated by income and/or savings overpasses the expenditures.

In practice, the net present value (NPV) is the most often used indicator to analyze the profitability of a new technology deployment. The NPV is the cumulative discounted free cash

flow over period of time under economic analysis. The technology is profitable when the NPV is positive.

The payback period is another frequently used indicator which measures the necessary time for the new technology to become profitable. It is the date when the cumulative discounted free cash flow becomes positive.

Generally speaking, an MNO can expect additional revenues generated by a new technology like 5G from new customer subscriptions, higher subscription fee and induced savings in the network operations, e.g. energy saving of new generation of equipment.

Expenditures are typically divided into capital expenditures (CAPEX) and operating expenditures (OPEX), where CAPEX include all expenditures that an operator invests for the initial setup of the network, while OPEX are recurrent expenditures to operate, run and maintain the network.

15.4.1.1   CAPEX

The main items constituting the CAPEX of a mobile network are cost related to
- Radio spectrum licensing;
- Site building or site acquisition;
- Purchase of antennas and feeders;
- Purchase of radio access network equipment (e.g., BSs);
- Purchase of core network equipment;
- Transport network building.

In general, the site building in the CAPEX model covers site design, site engineering, site research, site acquisition, civil work, mast purchase and installation, housing, non-telecommunication equipment installation and commissioning, and the purchase of power supply equipment. The civil work includes concrete plinth and steel beam as foundation, support of mast and radio cabinet, fence and access path to the site, etc. The housing refers to purchase, delivery and installation of shelter, air conditioning, fire protection system and site adaptation. The electricity power equipment corresponds to electricity feeding and connection, purchase, delivery and installation of generator, fuel tank, DC power cabinet etc.

CAPEX differs for a site built on rooftop and a site built on a green field. A green field site is not situated in the existing architecture but built on the ground. A green field site's infrastructure comprises more elements, and it is often more expensive than a rooftop site. Similarly, two values of site building cost needed to be considered depending on whether it is an existing site shared with other previous generations of mobile technologies, or a totally new site dedicated for the new technology.

Finally, it should be highlighted that a major element of the CAPEX is the cost of transport network building, i.e. the costs of fronthaul and backhaul networks from the BS to the first aggregation point, for which either wireless or wireline technologies can be utilized, as detailed in Chapter 7.

15.4.1.2   OPEX

The main items constituting the OPEX are site OPEX, which may be shared among different cellular generations, energy cost (cf. Section 15.3), transport network OPEX, billing and sales cost.

The site OPEX consist of site rent, labor cost and diverse maintenance costs, e.g. related to mast maintenance for macro sites, supplier maintenance, site security, site caretaking, etc. Labor cost is defined as the cost for the operator's internal staff working directly on the operation, maintenance and supervision of the network. The transport network OPEX reflects the rent paid to potential third parties for the rental of transport network infrastructure, e.g., optical fiber fronthaul or backhaul connectivity, see Chapter 7.

15.4.2 Methodology of 5G deployment assessment

The previous generation of cellular networks, i.e. 4G, was predominantly designed for mobile broadband users. In this context, provisioning the targeted mobile broadband experience in terms of coverage and capacity has been the main objective in the evolution of mobile networks. Usually, only the trade-off between the targeted mobile service experience on one side and the network CAPEX and OPEX on the other side was investigated. However, the increasing adoption of mMTC services brings new challenges to traditional cellular network signalling mechanisms and control plane system capacity. Therefore, even if 5G eMBB services can be assessed by making usage of classic methods, new techno-economic assessment methodology should be developed for mMTC services.

A methodology for 5G deployment assessment proposed in [45] consists of following steps:
- 5G traffic forecast;
- Estimation of 5G revenue;
- Dimensioning of 5G networks;
- Assessment of deployment scenarios;
- Techno-economic analysis.

For the **5G traffic forecast** it is assumed that eMBB traffic demand per area unit is equal to the average mobile data usage per user, times the number of users per area unit. Mobile data usage is the amount of data sent and received per user during one month. For the mMTC services, traffic forecast is based on the activity predicted per day and area, the number of devices per service, and the time the service needs to access the network and to transmit its payload.

Three classes of mMTC services have been defined in [45]. The indoor mMTC/IoT service class is foreseen for the stationary sensors deployed indoors. The outdoor services class represents sensors and actuators deployed outside, possibly involving mobility. The third class represents services that in all cases require device mobility.

Traffic profiles are defined for three cases: low load, baseline and high load, where the low and high load cases represent a load that is ten times lower or higher than the baseline, respectively. The values considered for eMBB and for mMTC are listed in Table 15-2.

Table 15-2. Parameters for eMBB and mMTC traffic profiles.

Service	Parameters	High Load	Baseline load	Low Load
eMBB	Monthly traffic in GB/month/ subscriber (heavy usage)	100	50	20

	Monthly traffic in GB/month/ subscriber (medium usage)	30	15	8
	Monthly traffic in GB/month/ subscriber (low usage)	4	2	1
	Share of total eMBB users [%] (heavy usage)	70	20	10
	Share of total eMBB users [%] (medium usage)	20	50	20
	Share of total eMBB users [%] (low usage)	10	30	70
mMTC	MTC services penetration rate [%]	68	48	28
	CAGR of MTC penetration [%]	39	29	19
	Device arrival window [number of frames][1]	5	100	200
	Event cogeneration factor[2]	0.0028	0.00028	0.000072
	NB-IoT control plane usage rate per frame[3]	0.25	0.5	0.8

To assess **5G revenues**, the average revenue per user or per unit (ARPU) for each device type is estimated. The estimation takes into account the pricing "tactics" of the operators, the evaluation of customers' willingness to pay, the traffic per customer, the 5G deployment schedules and the vertical market analysis.

IoT revenues assumptions take into account the development of mMTC and URLLC applications. In terms of technology, machine-to-machine (M2M) applications currently rely mainly on 2G/2.5G. By 2020, 4G will have a significant position, with estimated 65.2 % compound annual growth rate (CAGR) at global level, between 2015 and 2020. Commercial launches of connected and autonomous cars will stimulate the 5G take-off from 2021 onwards, as they require more accuracy and reactivity in data treatment, as detailed in Chapter 14.

mMTC services' ARPU is expected to be relatively low, as so-called low power wide area operators are currently setting the tariffs for the lower end of the market. Estimations of the ARPU from mMTC also take into account the trends observed from main mobile operators in Europe, South Korea and in the USA, and are depicted for the next years in Table 15-3. It should be kept in mind that mMTC is attractive for mobile operators due to the reduced churn of these subscriptions and the fact that customers pay for connectivity and not for the traffic volume.

Table 15-3. ARPU estimates for EU 28 countries for 2020-2025 period.

Year	2020	2021	2022	2023	2024	2025
mMTC	5.0	5.0	4.0	3.5	3.0	2.5

1 defined as the number of devices that can arrive within that window.
2 defining how many events are active and reserve some resource at given point of time.
3 As NB-IoT is using the same resources for control and data transmission, this factor indicates how many resources can be occupied by the CP at a given point of time

URLLC	300	297	294	291	288	285

For eMBB services, the legacy **dimensioning methods** can be used for 5G. For mMTC, a particular attention should be paid to event-driven services like smart grid, disaster management, earthquake or flood detection, etc., because they can cause an access crunch by triggering huge number of devices in a limited geographical area to send incident reports at the same time. In such conditions, the traffic and network access patterns will be different from those experienced in current human-centric service networks.

From mMTC perspective, the number of Physical Random Access Channel (PRACH) slots required to achieve an appropriate access rate should be estimated first. PRACH dimensioning impacts only the uplink, while the maximum number of devices that attempt to connect within the retransmission period will also impact the downlink. In fact, more spectrum should be allocated to physical downlink control channels to meet required performance expectations.

Finally, **5G deployment scenarios** have to be assessed. Based on experience from previous cellular generations, the first step of 5G deployment is assumed to be a 5G roll-out in existing macro cell sites. The small cell deployment will follow the macro cell deployment when all macro cell frequency bands are overloaded. A maximum ratio of outdoor small cell number to a macro site is assumed, while the percentage of area covered by indoor small cells, the average number of floors and the small cell coverage surface have to be set as parameters in order to define the limit of indoor small cell density. If all macro and small cell radio resources are utilized, then the macro site densification is considered as the last technical option due to its high cost. Considered frequency bands are:

- Macro: 700 MHz FDD, 3500 MHz TDD and 2600 MHz TDD
- Small cells: 2600 MHz TDD, 3500 MHz TDD and 30 GHz TDD

The upper limit of an outdoor small cell capacity is estimated according to its spectrum efficiency, and its lower limit by the traffic in its coverage. The capacity of an indoor small cell is assumed to be equal to the traffic of an eMBB user for a home/office indoor small cell, and equal to its spectrum efficiency times its frequency bandwidth for a hotspot indoor small cell.

15.4.3 Techno-economic evaluation and analysis of 5G deployment

Following the assumptions and the methodology described in the previous subsection, a typical European dense urban area of 4 km^2 is now considered for a 5G deployment techno-economic analysis, as detailed in [45]. Note that only eMBB and mMTC services are taken into account in this analysis.

Based on the methodology and traffic forecast described in Section 15.4.2, the analysis shows that macro cells will provide enough capacity during the first years of 5G roll-out. Nevertheless, after this period of time, a large number of small cells, especially for home/office indoor usage will be necessary to deliver the forecasted traffic. Comparing to other mobile technologies, small cells are of particular interest in 5G. The very big traffic volume per eMBB user in a dense urban information society will necessitate a high small cells deployment cost, which will take a much larger part of total CAPEX and OPEX than in the precedent mobile technology generations. Under the assumption that 5G eMBB ARPU remains the same as in precedent generations, or even decreases with time, the large amount of additional small cells will make it more challenging for

MNOs to make 5G deployments profitable. Figure 15-13 represents the cumulative discounted cash flow of an MNO sharing a market with a different numbers of MNOs. The calculation has been made assuming each MNO has the equal share of a market. For the considered dense urban area, the cumulative discounted cash flow of an MNO will become positive several years after the beginning of the 5G deployment.
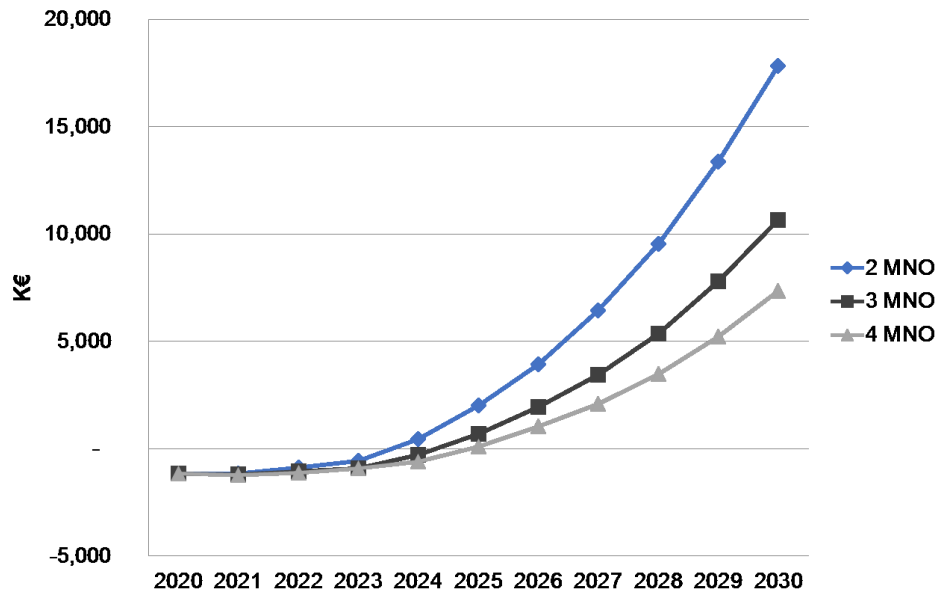

Figure 15-13. Cumulative discounted cash flow of an MNO with different numbers of MNOs in the area [45].

Only in isolated business districts where the population density during daytime is much higher than in an average European dense urban area, the macro cell densification will be required, mainly due to indoor traffic. Since it is practically impossible to densify the existing macro network in such areas, the indoor traffic offload by alternative radio solutions such as Wi-Fi should be envisioned. Regarding mMTC, it is noted that mMTC consumes little radio resources. Its incremental expenditure is in consequence very low. As a result, the mMTC contribution to MNO cash flow will be positive.

15.5   Summary

A fair evaluation of 5G performance, energy efficiency and techno-economic aspects is one of the major steps that have to be taken to answer key design questions for the new generation of cellular technology. This chapter has presented an evaluation framework for the qualitative and quantitative assessment of 5G. Comparing to 4G, such framework has to address new use cases, such as mMTC and URLLC, and also reflect the trend of going towards more diversified and heterogeneous deployments and operations in higher frequency regimes. First evaluation results indicate that 5G will bring significant gains in all generic use cases, e.g., handling an order of magnitude higher traffic densities in wide area eMBB scenarios, 10+ year operations for mMTC devices with a single battery, and 99.999 % reliability for URLLC services.

5G is expected to bring a major improvement in energy efficiency of the network infrastructure, therefore this chapter has introduced some key projects and standards that aim at this goal. In order to assess energy efficiency correctly, proposed models cater for temporal and spatial variations of the traffic as well as for static and dynamic power consumption. An exemplary assessment that was done for dense urban deployments indicates that 5G is able to bring energy efficiency improvements that at least follow the traffic growth, i.e. allowing for a flat overall energy consumption over time.

Finally, this chapter has provided a techno-economic analysis of 5G. Based on traffic forecasts and ARPU, a methodology has been developed and applied to a European dense urban area to analyze the 5G cash flow of a mobile network operator and identify the factors impacting it. It was concluded that for the considered dense urban area, the cumulative discounted cash flow of an MNO will become positive about 4-6 years after the beginning of 5G deployment.

15.6   References

[1]   ITU-R M.2135, "Guidelines for evaluation of radio interface technologies for IMT-Advanced", 2008

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

[3]   3GPP TR 36.814, "Further advancements for E-UTRA physical layer aspects (Release 9)", Mar. 2010

[4]   3GPP TR 36.839, "Mobility enhancements in heterogeneous networks (Release 11)", Dec. 2012

[5]   Y. Srinivasan, J. Zhuang, L. Jalloul, R. Novak and J. Park, "IEEE 802.16m Evaluation Methodology Document (EMD)", IEEE 802.16m Broadband Wireless Access Working Group, IEEE 802.16m-08/004r5, Jan. 2009

[6]   ITU-R WP 5D, Draft new report ITU-R M.[IMT-2020.TECH PERF REQ], "Minimum requirements related to technical performance for IMT-2020 radio interfaces", Feb. 2017

[7]   ITU-R WP 5D, Draft new Report ITU-R M.[IMT-2020.EVAL], "Guidelines for the evaluation of radio interface technologies for IMT-2020", June 2017

[8]   ITU-R WP 5D, Document IMT-2020/2 Rev. 1, "Submission, evaluation process and consensus building for IMT-2020", Feb. 2017

[9]   3GPP TR 38.913, "Study on Scenarios and Requirements for Next Generation Access Technologies (Release 14)", March 2016

[10]   Y. Wang, J. Xu and L. Jiang, "Challenges of System-Level Simulations and Performance Evaluation for 5G Wireless Networks", IEEE Access, vol. 2, pp. 1553-1561, Dec. 2014

[11]   5G PPP FANTASTIC-5G project, Internal Report IR2.1, "Use cases, KPIs and requirements", Oct. 2015

[12]   5G-PPP White Paper, "Living document on use cases and performance evaluation", Version 1, April 2016

[13]   5G PPP METIS-II project, Deliverable D2.3, "Performance evaluation framework", Feb. 2017

[14]   5G PPP mmMAGIC project, Deliverable D1.1, "Use case characterization, KPIs and preferred suitable frequency ranges for future 5G systems between 6 GHz and 100 GHz", Nov. 2015

[15] 5G PPP SPEED-5G project, Deliverable D3.2, "SPEED-5G enhanced functional and system architecture, scenarios and performance evaluation metrics", June 2016

[16] Aalto University et al., White Paper, "5G Channel Model for bands up to 100 GHz", Dec. 2015, see http://www.5gworkshops.com/5GCM.html

[17] 3GPP TR 36.873, "Study on 3D channel model for LTE (Release 12)", June 2015

[18] 3GPP TR 36.942, "Radio Frequency (RF) system scenarios (Release 13)", Jan. 2016

[19] 3GPP TR 36.931, "Radio Frequency (RF) requirements for LTE Pico Node B (Release 13)", Jan. 2016

[20] 3GPP TR 36.885, "Study on LTE-based V2X Services (Release 14)", June 2016

[21] F. Schaich, B. Sayrac, S. Elayoubi, I.-P. Belikaidis, M. Caretti, A. Georgakopoulos, X. Gong, E. Kosmatos, H. Lin, P. Demestichas et al., "Fantastic-5g: flexible air interface for scalable service delivery within wireless communication networks of the 5th generation", Transactions on Emerging Telecommunications Technologies, Wiley, vol. 27, no. 9, pp. 1216–1224, Sept. 2016

[22] NGMN, White Paper, "NGMN 5G Initiative White Paper", Feb. 2015

[23] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime", IEEE Transactions on Information Theory, vol. 59, no. 10, pp. 6441–6463, Oct. 2013

[24] M. Kurras, L. Thiele, and G. Caire, "Multi-Stage Beamforming for Interference Coordination in Massive MIMO Networks", Asilomar Conference on Signals, Systems and Computers, Nov. 2015

[25] 5G PPP FANTASTIC-5G project, Deliverable D2.1, "Air interface framework and specification of system level simulations", May 2016

[26] B. Mouhouche and M. Al-Imari, "Optimization of Delivery Time in Broadcast with Acknowledgement and Partial Retransmission", IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB 2016), June 2016

[27] 5G PPP METIS-II project, Deliverable D2.1, "Preliminary performance evaluation framework", Jan. 2016

[28] D. Martín-Sacristán, C. Herranz and J. F. Monserrat, "Traffic Safety in the METIS-II 5G Connected Cars Use Case: Technology Enablers and Baseline Evaluation", European Conference on Networks and Communications (EuCNC 2017), June 2017

[29] I.-P. Belikaidis, A. Georgakopoulos, P. Demestichas, B. Miscopein, M. Filo, S. Vahid, B. Okyere and M. Fitch, "Multi-RAT Dynamic Spectrum Access for 5G Heterogeneous Networks: The SPEED-5G Approach", IEEE Wireless Communications Magazine, Special Issue on "Dynamic Spectrum Management for 5G", July 2017

[30] Nokia, "Flatten network energy consumption technology vision 2020", White Paper, Aug. 2015

[31] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper", Feb. 2017

[32] Commission of the European communities, "Limiting Global Climate Change to 2 degrees Celsius. The way ahead for 2020 and beyond", Jan. 2007, see https://ec.europa.eu/clima/policies/strategies/2020_en

[33] 5G PPP METIS-II project, Deliverable D1.1, "Refined scenarios and requirements, consolidated use cases, and qualitative techno-economic feasibility assessment", Jan. 2016

[34] ITU-T Recommendation L.1310, "Energy efficiency metrics and measurement methods for telecommunication equipment", Version 2.0, Aug. 2014

[35] ETSI ES 202 706, "Measurement method for power consumption and energy efficiency of wireless access network equipment", Version 1.4.0, Oct. 2014

[36] CELTIC Opera-NET project, see https://www.celticplus.eu/project-operanet/

[37] ETSI ES 203 228, "Environmental Engineering (EE); Assessment of mobile network energy efficiency", Version 1.2.1, Apr. 2017

[38] FP7 EARTH project, see https://www.ict-earth.eu/

[39] A. Galis and A. Gavrasm, "The Future Internet - Future Internet Assembly 2013: Validated Results and New Horizons", Springer, 2013

[40] FP7 EARTH project, Deliverable D2.3, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown", Version 2, Jan. 2012

[41] GreenTouch, White Paper, "Improving the nationwide energy efficiency in 2020 by more than a factor of 10000 in relation to the 2010 reference scenario", Aug. 2015

[42] ETSI ES 203 136, "Environmental Engineering (EE); Measurement methods for energy efficiency of router and switch equipment", Version 1.1.1, May 2013

[43] ETSI ES 201 554, "Environmental Engineering (EE); Measurement method for Energy efficiency of Mobile Core network and Radio Access Control equipment", Version 1.2.0, May 2014

[44] ETSI TR 103 117, "Environmental Engineering (EE); Principles for Mobile Network level energy efficiency", Version 1.1.1, Nov. 2012

[45] 5G PPP METIS-II project, Deliverable D1.2, "Quantitative techno-economic feasibility assessment", June 2017