Comparison of Spectral and Energy Efficiency Metrics using Measurements in a LTE-A Network

Sandrine Boumard, Ilkka Harjula, Teemu Kanstrén, and Seppo J. Rantala
VTT Technical Research Centre of Finland, Email: name.surname@vtt.fi

***Abstract*—Energy and spectral efficiencies are key metrics to assess the performance of networks and compare different configurations or techniques. There are many ways to define those metrics, and the performance indicators used in their calculation can also be measured in different ways. Using an LTE-A network, we measure different performance indicators and the metrics' outputs are compared. Modifying the transmitted output power, the bandwidth, and the number of base stations, different network configurations are also compared. As expected, the measurements show that increasing the bandwidth increases the throughput more than it increases the energy consumption. Results clearly show that using inappropriate indicators can be misleading. The power indicator should include all energy consumed and the throughput should be dependent on the traffic, taking into account the idle time of the network, if any. There is a need to include more performance indicators into the metrics, especially those related to quality of service.**

I. Introduction

With the increasing attention to energy savings and green radio techniques, energy efficiency (EE), how well the energy resource is used, is one of the key metrics used to assess the performance of future networks. On the other hand, spectral efficiency (SE), how well the spectrum resource is used, is also one of the criteria utilised to define communication systems' performance. Both EE and SE are listed as objectives for future development of IMT 2020 and beyond [1], which can be seen as a regulatory framework for developing Fifth Generation (5G) systems. Therefore, it is important to understand the effect of the performance indicators used to calculate SE and EE to select the most appropriate ones.

The performance of Long Term Evolution (LTE) networks has been studied extensively from different points of view. In addition to numerous analytical and simulation-based studies, field tests have been carried out varying from nation-wide mobile network measurements [2] and real-life urban area drive tests [3] to smaller scale small cell indoor and outdoor measurements [4] and to indoor measurements focusing, for example, in the performance of specially devised setup of coordinated small cells [5]. Some studies exist that focus on the monitoring and modelling of the energy consumption of different networks using real-life measurements. For example, an energy consumption monitoring network for Third (3G) and Fourth Generation (4G) networks was presented in [6]. In [7], an energy consumption model for macrocell and small cell base stations was developed, and it was validated with power measurements of actual base transceiver stations (BTS) and applied for LTE, Worldwide Interoperability for Microwave Access (WiMAX) and High Speed Packet Access (HSPA) networks. A measurement-based model of energy consumption in 3G femtocells was proposed in [8]. The energy consumption of Global System for Mobile Communications (GSM) and Universal Mobile Telecommunications System (UMTS) networks were studied in [9], for example.

Both EE and SE are well known metrics and they have been widely discussed in the scientific literature. EE can be defined as the number of bits that can be transmitted per unit of consumed energy. SE is usually defined by the throughput of the system per unit of bandwidth. A more detailed look at the definition of these metrics in the literature, the performance indicators used to define these metrics, and how they are used in this study, is given in Section II. Although the EE and SE metrics themselves are well defined, there is still a lack of agreement in the scientific community in the definition of the performance indicators used in the metrics. In this paper, we demonstrate how the selection of the throughput and power performance indicators affects the EE and SE results by measuring these indicators for different network configurations in a real network of LTE-A small cells.

This paper is organised as follows. The energy and spectral efficiency metrics are defined in Section II, by first introducing the definitions from the literature and then the metrics used in the measurements. Then the network deployment and test cases are described in Section III. The measurements results and their analysis are presented in Section IV. Conclusions are drawn in Section V.

II. Energy and Spectral Efficiency Metrics

A. *Metrics Definition*

To compare different wireless networks configurations, key performance indicators (KPIs) and metrics need to be carefully defined. Many performance indicators can be measured: throughput, delays, energy consumption, etc. They can be measured at the network level or at the user level. Combining the measurements into a metric that encapsulates the network performance is not an easy task. There are numerous metrics defined in the literature. In the remaining text, the term power consumption is used even though power cannot be consumed, only energy can. This is to keep in line with the literature and the fact that power is the measured value.

The SE is defined as bits per second per Hz [15]

$$SE = \frac{R}{B}, \tag{1}$$

where R is the throughput in bit/s and B is the occupied bandwidth.

The most commonly used EE metric is defined as [10]

$$EE = \frac{R}{P}, \quad (2)$$

where P is the transmitted power. Alternatively, the power consumption can be used.

An overview of the standardized energy metrics as of 2012 is presented in [11]. The energy consumption rating (ECR), although not standardized, is often used in the literature and is defined as the ratio of the measured peak power in watt by the maximum throughput in bit/s, under full buffer conditions. For wireless networks, the ECR can be formulated as the ratio of the energy consumed by the amount of transmitted data. It is also termed normalized energy. The energy efficiency rate (EER) is the inverse of the ECR. Similarly, the telecommunication energy efficiency ratio (TEER) is defined as the ratio of the useful work by the power, for example Mbps/watt. The ITU has defined an energy metric for the wireless access networks as the ratio of power by the subscriber traffic area, in watt/bps/m^2. This is the metric used in [12], in watt/(bit/s)/km^2 or watt/erlang/km^2, where the power is the radiated power. The authors conclude that while these metrics seem to be adequate for comparison, they are in fact misleading for wireless systems, since they do not take into account performance and other energy consuming elements in the network. Another metric, more suitable for rural areas, is defined as the ratio of the base station coverage by the average site power consumption, in km^2/watt. In urban areas, the ratio of the number of subscribers on the average busy hour by the average site power consumption can be used, in subscribers/watt. In [13] another metric is defined for EE in watt/km^2. In [14] the area spectral efficiency (ASE) is defined as the cell average spectral efficiency, in bit/s/Hz/km^2. The area energy efficiency (AEE) is similar to the ASE, in bit/joule/km^2. Some of the EE and SE metrics found in the literature are summarized in Table I.

In a network composed of several nodes, the issue of estimating and averaging the metric for the whole network arises. Indeed, the mean of the ratios of two variables is not equal to the ratio of their means. The choice is based on what information needs to be extracted from this calculation. The overall network efficiency of a network is usually calculated as the ratio of the sums of the averages at each network node. For example, the overall AEE is the ratio of the sum of the averages at each network node or cell [10].

As SE and EE improvements are conflicting goals in a wireless system, a trade off must be made. An attempt is made in [15] to combine the two metrics into one. An energy spectral efficiency (ESE) and SE trade-off (EST) metric is defined as

$$EST(P) = [SE^{\text{norm}}(P)]^w \times [ESE^{\text{norm}}(P)]^{1-w}, \quad (3)$$

where $w \in [0,1]$ is the preference for SE, and the SE and ESE metrics are normalized by the maximal ESE and SE over the range of possible transmitted power P. The ESE metric is defined as the ratio of SE to the total power consumption.

One drawback of the SE and EE metrics is that these metrics do not show the other KPIs and the tradeoff in terms of quality of service (QoS) to all customers of the network [11] [12]. The choice of the metric can also affect the solution for network dimensioning, since "maximizing the energy efficiency is not equivalent to minimizing the energy consumption unless capacity and coverage requirements of the system are carefully considered" [13].

The QoS requirements could be included in a cost function similar to (3). The throughput of $x\%\%$ of users could be included, or the delay of $x\%\%$ of users, depending on the traffic. Percentile $x\%\%$ indicates the value below which a percentage x of observations in a group of observations fall. The weight or preference for each element of the cost function is obviously difficult to assess. Besides, there could also be a minimum throughput or maximum delay that would also eliminate any solution to the EE-SE trade-off that does not provide these minimum requirements.

TABLE I
SOME OF THE EE AND SE METRICS FOUND IN THE LITERATURE.

Metrics	Definition	Notations and comments
SE [15]	$\frac{R}{B}$	R aggregate throughput (bit/s) B occupied bandwidth (Hz)
EE [10]	$\frac{R}{P}$	power P can be power consumed P_C or transmitted P_T (W)
ESE [15]	$\frac{R}{B \times P}$	
ASE [14]	$\frac{R}{B \times A}$	A area (unit of area)
EE_A [13]	$\frac{P_C}{A}$	
AEE [10]	$\frac{R}{P \times A}$	
ECR [11] 1/EER	$\frac{P_{\text{peak}}}{R_{\text{max}}}$	P_{peak} measured peak power (W) R_{max} maximum throughput (bit/s)
ECR_cell [11]	$\frac{P_C \times T}{D}$	T time (s) D number of bits transmitted during T
EE_ITU [12]	$\frac{P_C}{R \times A}$	
EE_rural [11]	$\frac{A_c}{P_S}$	A_c base station coverage P_S average site power consumption
EE_urban [11]	$\frac{N_s}{P_S}$	N_s number of subscribers on the average busy hour
EST [15]	$[SE^n(P)]^w$ $\times [ESE^n(P)]^{1-w}$	$w \in [0,1]$ normalized SE and ESE

B. Measurement of SE and EE Metrics

The most commonly used metrics, defined in (1) for SE and in (2) for EE are used. Even if the metric is defined, there is a lack of agreement in the definition of the performance indicator used in this metric. We will use here the following indicators:

Throughput R is the sum of the throughput in each cell. There are two ways to define the throughput

1) $R1$ is the Packet Data Convergence Protocol (PDCP) Service Data Unit (SDU) data volume on eUu interface in the downlink, divided by the period during which it is measured.
2) $R2$ is the PDCP SDU data volume on eUu interface in the downlink, divided by the period during which it was sent, which is the number of Transmit Time Intervals (TTI) in downlink with at least one User Equipment (UE) scheduled to transmit user plane data.

Bandwidth B is the network bandwidth, the total occupied spectrum in the network. If two cells transmit at different carrier frequencies, the bandwidth is the sum of their respective bandwidths.

Power P is the sum of the powers. There are two ways to defined the power

1) $P1$ is the transmitted power.
2) $P2$ is the power consumed.

III. Network Deployment and Test Cases

In this section, the measurement setup is introduced by describing the deployment of the network and the equipment used. We also describe the test cases, including the network parameter settings and traffic patterns used to load the network.

A. Network Deployment

The measurements were carried out with commercial off-the-shelf (COTS) equipment. Three small cell BTSs were deployed in an office building complex connected via corridors, and eight mobile phones were used as UEs, spread over the coverage area of the BTSs and kept in static locations. The BTSs were connected to 5GTN [16], a research network providing the core network functionalities. The power consumption of the BTSs were measured with COTS equipment.

The map of the BTSs' locations and the UEs' locations is shown in Fig. 1. The locations of the BTSs are marked with red stars, and the locations of the UEs are marked with yellow stars. The BTS A and C are the furthest away from each other, and BTS B is in the middle. BTS A was located on the 2nd floor, while the others were on the ground floor. The BTS B was switched off in some test cases. The UEs were placed in a manner that while BTS B was shut down, the UEs closest to it were still in the coverage area of either BTS A or C.

The power consumption $P2$ is measured at the BTSs only. The commercial tools used are Smart-me Plugs [17]. The device connects directly to the socket. It measures the power consumed by the device attached to it in either one minute or 15 minutes intervals and saves the data to a cloud server. Since the transmitted power $P1$ cannot be measured, it is calculated using the transmitted power parameter set at the BTS and the period of time the BTS is transmitting, the percentage of TTIs in downlink with at least one UE scheduled to transmit user plane data during the period of observation.

Fig. 1. Map of the pico cells and user phone placement for the final demo on site.

B. Test Cases

A block diagram of the high-level concept of the measurements is shown in Fig. 2. During the measurements, the network parameters were altered, such as carrier frequencies, bandwidths, and transmission powers of the base stations, or even the number of BTS in the network. Various parameters, such as throughput and power consumption of the BTS, were collected from the network under different traffic loads. The focus was on downlink performance. The BTSs operate at 2.6 GHz, band 7, in frequency division duplexing mode, and use the same bandwidths both in uplink and downlink directions. The configuration for transmission is dynamic open loop MIMO, with two transmitting antennae.

Five test cases were defined:

1) All BTSs share the same carrier frequency and use a 5 MHz bandwidth.
2) All BTSs use a 5 MHz bandwidth but BTS B uses a different carrier frequency than BTSs A and C.
3) All BTSs share the same carrier frequency and use a 10 MHz bandwidth.
4) BTS B is off and all other BTSs share the same carrier frequency and use a 5 MHz bandwidth.
5) BTS B is off and all other BTSs share the same carrier frequency and use a 10 MHz bandwidth.

Furthermore, for each test case, two subcases were defined, one with high transmission power (24 dBm maximum transmission power at the BTS), and one with low transmission power (17 dBm maximum transmission power at the BTS). The transmission power indicated is per antenna, hence the actual transmitted power is double this value.

To create different types of traffic going through the network, File Transfer Protocol (FTP) and User Datagram Protocol (UDP) transmissions were used. The FTP transmission mimics a FTP1 traffic pattern [18], with a 2 MB file downloaded every 5 seconds to the UE. The UDP traffic was created by using iperf3 [19], and it corresponds to a full buffer type of traffic. In order to make sure that we are not measuring the edge effects but a steady state situation, each transmission instance was set so that it lasted for one hour. Furthermore, each transmission instance was repeated twice.

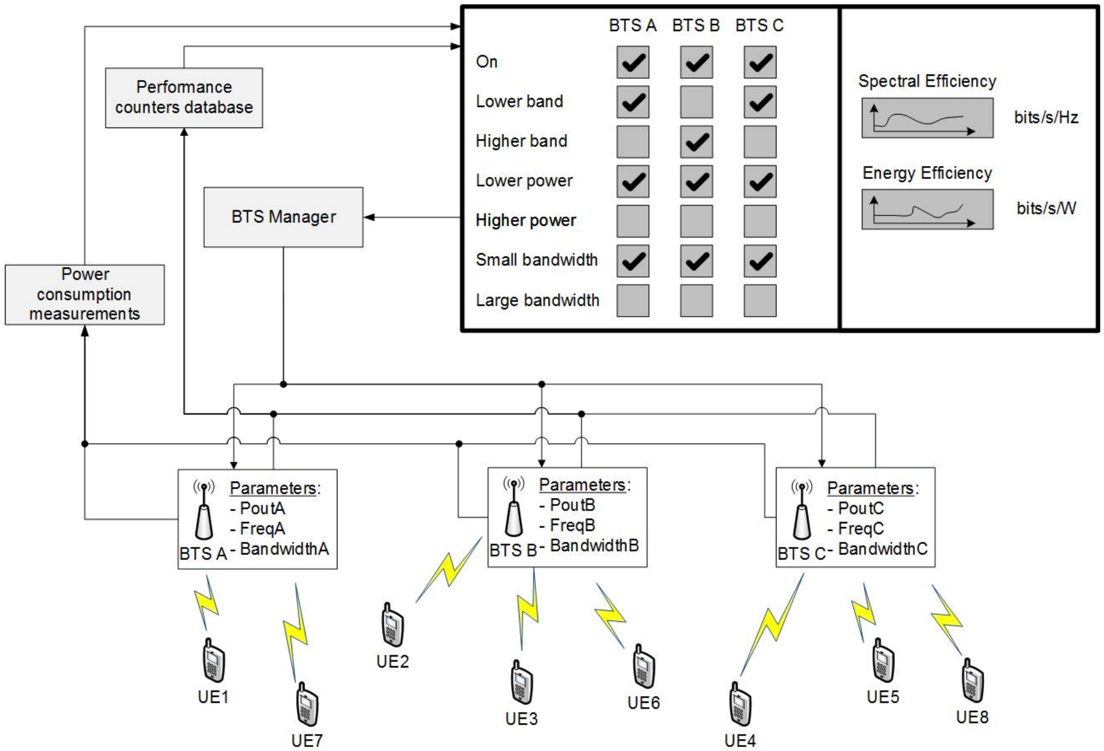

Fig. 2. High level specification of the measurements with all the building blocks involved.

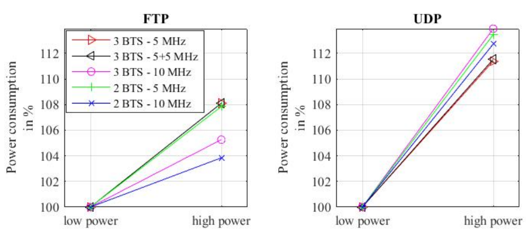

Fig. 3. The variation is in percentage for the power consumption, when the 100 % reference is the low power value for the same test case.

IV. Measurements and Analysis

The measurements and their analysis are presented in this section. Measurements from the BTS performance counter database and the Smart-me plugs are collected every minute. Those measurements are averaged over a 45 minutes sliding window. The values are taken at the plateau region of the average curves. The averaging helps also with the lack of timing reference between the BTS performance counter database and the Smart-me plugs measurements.

A. Metrics Comparison

Fig. 3 shows the variation of the consumed power between low and high transmitted power mode for all test cases and for the two different traffic patterns. The increase in power consumption between low and high transmitted power modes for the FTP case is twice as small as in the UDP traffic case. This is expected since the FTP traffic uses only part of the resources. Indeed, the traffic volume is less than the air interface capacity. The figure shows the part of the transmitted power in the total consumed power and the effect of the traffic on the consumed power.

Fig. 5 and 6 show the throughput and SE for all test cases and traffic patterns, for low transmitted power and high transmitted power, respectively. The test cases using 5 and 10 MHz bandwidth have been separated to clarify the results. There is clearly not a big difference between low and high transmitted power cases. Both throughput calculation methods are also shown on the figures. Using $R2$, the traffic pattern does not affect the results. There is only a slight increase of throughput in the 10 MHz in some cases between FTP and UDP. This may be due to the resource allocation method that may favour network SE over user QoS in the full buffer case, whereas in FTP, there is time to send data to all users and thus the network SE may suffer from the effect of the low throughput users. However, this also may be due to the measurements inaccuracy since the transmitted power does not seem to affect the throughput. When using $R1$, there is a clear difference between the low FTP traffic and the high UDP traffic in terms of both throughput and SE. In terms of SE, the 5 MHz, all BTSs on the same carrier frequency, is the best choice in the 3 BTSs case. When only 2 BTSs are used, 5 MHz and 10 MHz lead to similar SE. The 10 MHz network configuration does not seem to have taken advantage of the lower overhead when using higher bandwidth, in LTE-A. Reducing the interference, in test case 2, only leads to a small throughput increase for UDP, compared to test case 1. When considering the throughput, the best choice is of course more bandwidth and more BTSs. When looking at SE, only adding BTSs helps.

Fig. 7 and 8 show the EEs for all test cases and traffic patterns, for low transmitted power and high transmitted power, respectively. The bottom figures refer to the EE calculated using $P1$ and the top figures using $P2$. Using $P2$, the results do not show too much of a difference between low and high power, the transmitted power is a small portion of the power consumed at the BTSs. The difference between low and high transmitted powers being 5 folds, the results for UDP traffic using $P1$ clearly show this ratio, since the throughputs between low and high transmitted powers are roughly the same. However, regardless of the power used in the EE metric, putting aside the actual values, the relation between the results of different test cases stays roughly the same at high and lower transmitted powers. Focusing, for example, on the low transmitted power results, in Fig. 7, there is no advantage in using more BTSs for UDP traffic, especially with a 10 MHz bandwidth, whether $P1$ or $P2$ are used. For the FTP traffic, it is not so clear-cut and the 3 BTSs cases seem to have almost the same EE than their equivalent 2 BTSs cases for $P2$. With $P1$, for FTP, using 3 BTSs seems to be more energy efficient, setting test case 2 aside. Overall, 10 MHz bandwidth remains the best choice using both $P1$ and $P2$.

Making measurements using COTS equipment comes with issues that cannot be ignored. The following list highlights the issues that may affect the results:

- There is no control over which BTS the users connect to. In some measurements, users kept switching from one BTS to another during some transmissions. The switching also happened between test cases, as it is shown in Fig. 4, which shows the number of users connected to each BTS for the test cases where all BTSs are on, for low and high transmitted power. This could partially explain why test case 2, in which there is less interference, did not lead to a higher throughput than test case 1, for example.
- There is no knowledge of the resource allocation algorithms used, since the BTSs are COTS. This may affect the results, especially in the UDP traffic case. The QoS requirements for the traffic pattern need monitoring.
- UDP traffic is not full buffer, when there is only one user connected to a BTS, not all TTIs were used. This can be seen in the different results for $R1$ and $R2$ in UDP for test case 2, high transmitted power, Fig. 6. There seems to be a bottleneck somewhere in the system tampering with the intended UDP traffic.

B. Choice of Performance Indicators

Regarding the power indicator, $P1$ leads to misleading conclusions since the transmitted power is only a small part of the consumed power. There is a non-negligible power consumed at the BTS apart from the transmitted power. Thus $P1$ is not the right choice when comparing different techniques or settings for EE improvement.

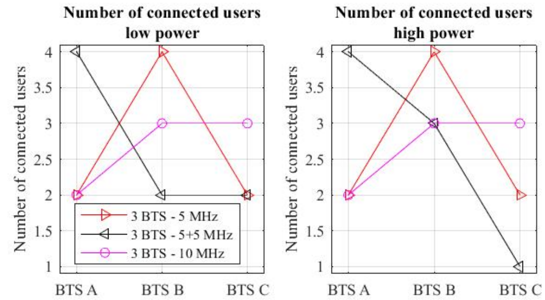

Fig. 4. Number of users connected to each BTS in cases where all 3 BTSs are used.

Regarding the throughput indicator, $R2$ is not so dependent on the traffic unlike $R1$, and hence can indicate the performance of the BTS, regardless of the traffic. Since the consumed power is only slightly dependent on the traffic pattern, there is no big difference between the EEs of the FTP and UDP traffic patterns when using $R2$ and $P2$. The same can be said of the EE using $R1$ and $P1$, since both indicators are dependent on traffic. Focusing on $P2$, the most appropriate power indicator, $R1/P2$ leads to a traffic dependent EE and $R2/P2$ to a traffic independent EE, in our case. However, looking closely at the test cases' results, the conclusions drawn using $R2$ are misleading. It looks like for FTP, using 3 BTSs is more energy efficient than using 2, whereas this is clearly not the case, since all users have received their data in both cases and 2 BTSs consume less energy than 3. This is shown in the EE results obtained with $R1$.

When looking at a metric inclusive of QoS, the power and throughput indicator should be dependent on the traffic used for the measurements in order to take into account all possible aspects of the transmission. For example, a low traffic load will show better QoS results than a high traffic load at a cost of lower SE and EE. This shows the EE and SE metrics should be traffic dependent to enable a fair comparison.

V. Conclusions

The measurements have shown that using a larger bandwidth leads to better EE, for a similar SE. Using a larger bandwidth increases the throughput much more than it increases the energy consumption.

The comparison between metrics have shown that an inappropriate choice of indicators can lead to the wrong conclusions. The energy consumption is not proportional to the traffic and this should be reflected in the EE, hence the total power consumed should be used instead of the transmitted power. The throughput should not be traffic independent and should take into account the idle time of the system, since the system will consume power during these idle moments. However, this independence is important when solely evaluating the BTS's performance.

As it has been stated many times in the literature, the SE and EE are limited by the fact that they do not take into account KPIs related to QoS. The network deployment cannot be solely based on SE and EE trade-off, but must also take into account the user experience. Cost functions are a practical way to include all important parameters. The chosen QoS parameter could be the throughput of $x\%\%$ of users, or the delay of $x\%\%$ of users, depending on the traffic. Additionally, the QoS parameter can have a cut-off value under which no solution to the EE-SE trade-off can be accepted. In the future, the measurements could be further developed to include the monitoring of the QoS at the UEs. The SE-EE trade-off metric could then be tested and weights' values tested.

Acknowledgment

This work was done for the ESEC project that is partly funded by Business Finland, the Finnish Funding Agency for Innovation (decision number 1938/31/2015). The authors would like to thank their colleagues Adrian Kotelba and Aarne Mämmelä for the fruitful discussions on metrics and their comments. They would also like to thank the reviewers for their relevant comments.

References

[1] "IMT Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond", Recommendation ITU-R M.2083-0.

[2] J. Hyun *et al.*, "High-end LTE service evolution in Korea: 4 years of nationwide mobile network measurements," in *2017 13th Int. Conf. on Network and Service Manage. (CNSM)*, Tokyo, pp. 1–7, 2017.

[3] S. Avallone *et al.*, "Smartphone-based measurements of LTE network performance," in *2017 IEEE Int. Instrumentation and Measurement Technol. Conf. (I2MTC)*, Turin, pp. 1–6, 2017.

[4] J. Beyer et al., "Performance Measurement Results Obtained in a Heterogeneous LTE Field Trial Network," in *2013 IEEE 77th Veh. Technology Conf. (VTC Spring)*, Dresden, pp. 1–5, 2013.

[5] J. Weitzen, R. Wakim and E. Webster, "Comparing RSRP, CQI, and SINR measurements with predictions for coordinated and uncoordinated LTE small cell networks," in *2015 IEEE Int. Conf. on Microwaves, Commun., Antennas and Electronic Syst. (COMCAS)*, Tel Aviv, pp. 1–5, 2015.

[6] A. Capone, S. D'Elia, I. Filippini, A. E. C. Redondi and M. Zangani, "Modeling Energy Consumption of Mobile Radio Networks: An Operator Perspective," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 120–126, Aug. 2017.

[7] M. Deruyck, W. Joseph, and L. Martens, "Power Consumption Model for Macrocell and Microcell Base Stations," *Wiley Trans. Emerging Telecommun. Technol.*, vol. 25, no. 3, pp. 320–33, Aug. 2012.

[8] R. Riggio and J. L. Douglas, "A Measurement-Based Model of Energy Consumption in Femtocells," in *IFIP Wireless Days 2012*, Dublin, Ireland, Nov. 2012.

[9] J. Lorincz, T. Garma, and G Petrovic, "Measurements and Modelling of Base Station Power Consumption under Real Trafc Loads," *MDPI Sensors*, vol. 12, no. 4, pp. 4281–4310, Mar. 2012.

[10] W. Wang and G. Shen, "Energy efficiency of heterogeneous cellular network," in *Proc. IEEE VTC'10 Fall*.

[11] H. Hamdoun *et al.*, "Survey and applications of standardized energy metrics to mobile networks," *Ann. Telecommun.*, vol. 67, no. 3–4, pp. 113–123, 2012.

[12] A. D. Gandhi and M. E. Newbury, "Evaluation of the energy efficiency metrics for wireless networks," *Bell Labs Tech. J.*, vol. 16, no. 1, pp. 207–215, Jun. 2011.

[13] S. Tombaz, K. W. Sung, and J. Zander, "On metrics and models for energy-efficient design of wireless access networks," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 649–652, Dec. 2014.

[14] Y. Xin *et al.*, "Area spectral efficiency and area energy efficiency of massive MIMO cellular systems," *IEEE Trans. on Veh. Technol.*, vol. 65, no. 5, pp. 3243–3254, May 2016.

[15] L. Deng *et al.*, "A unified energy efficiency and spectral efficiency tradeoff metric in wireless networks," *IEEE Commun. Lett.*, vol. 17, pp. 55–58, Jan. 2013.

[16] E. Piri et al., "5GTN: A test network for 5G application development and testing," *2016 European Conf. on Networks and Commun. (EuCNC)*, Athens, pp. 313–318, 2016.

[17] Smart-me - Smart energy solutions. Available: https://www.smart-me.com/

[18] 3GPP2, "CDMA2000 Evaluation Methodology," C.R1002-0 revision 0, December 2004.

[19] iPerf3. Available: https://iperf.fr/

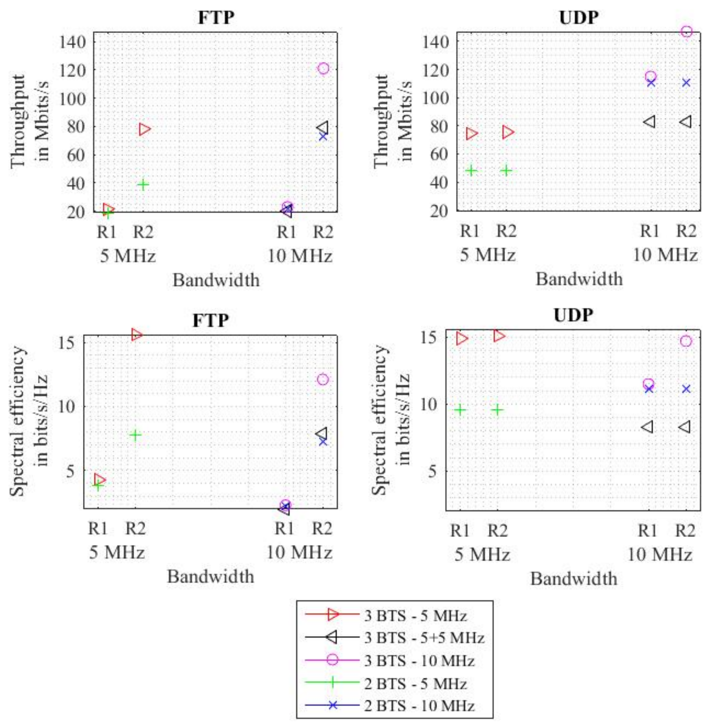

Fig. 5. Variation in the network throughput and spectral efficiency performance between small and large bandwidth usage in the low transmitted power case.

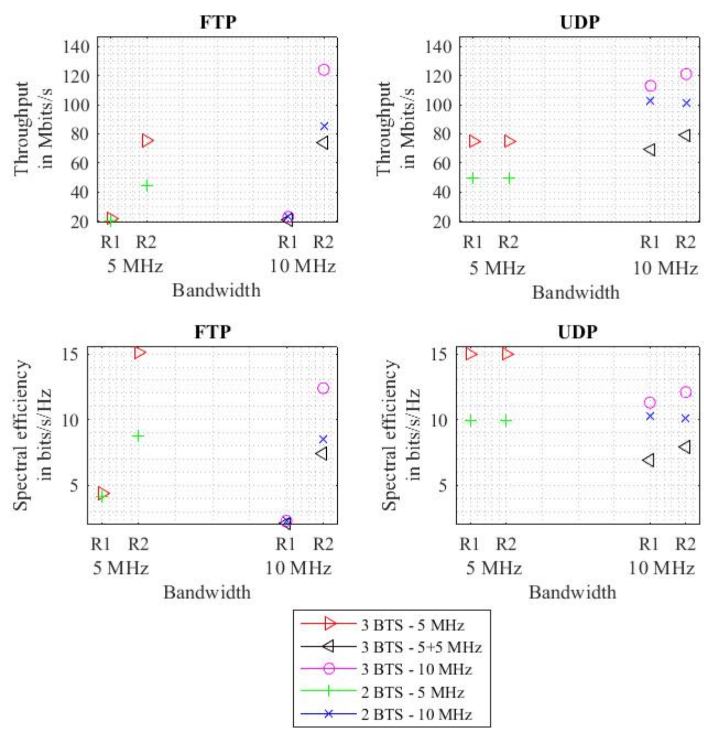

Fig. 6. Variation in the network throughput and spectral efficiency performance between small and large bandwidth usage in the high transmitted power case.

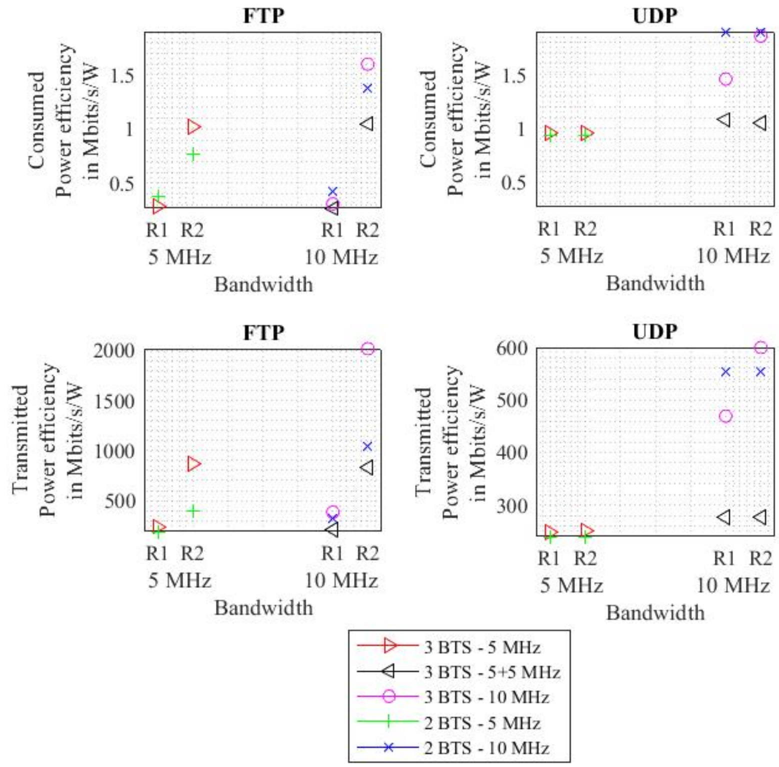

Fig. 7. Variation in the network power efficiency performance between small and large bandwidth usage in the low transmitted power case.

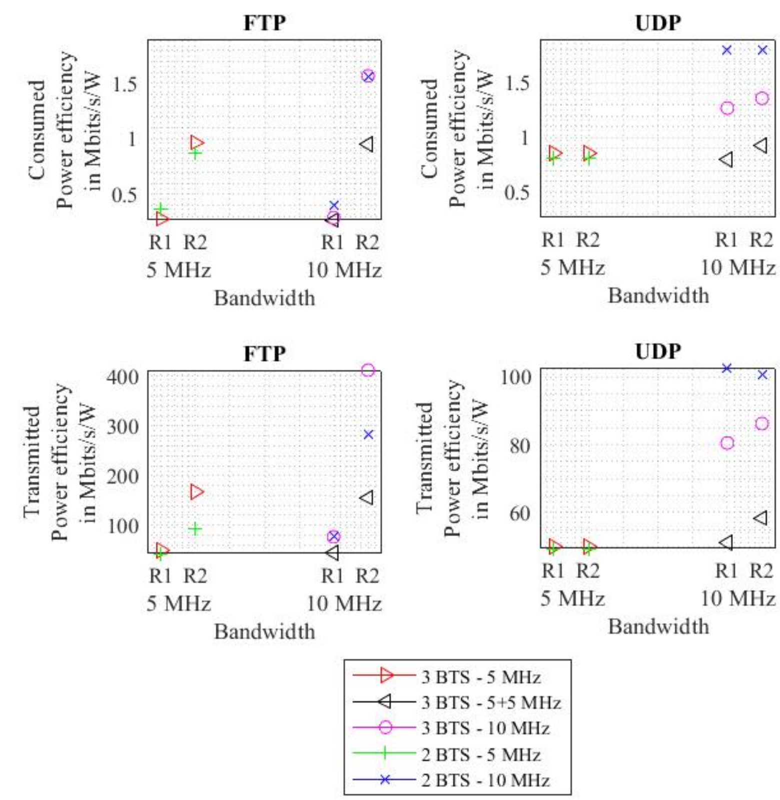

Fig. 8. Variation in the network power efficiency performance between small and large bandwidth usage in the high transmitted power case.