

Dynamic Power Allocation for Energy-Efficient 5G Networks Using Load Prediction with RF and SARIMA

Fadi Jawdat Ghosna

Tartous University

Mostfa ALBDAIR

`mostfa.albdair@unisq.edu.au`

University of Southern Queensland

Laila Mohammad Esmaeil

Tartous University

Maher Ali Ibrahim

Tartous University

Aqeel Sahi

University of Southern Queensland

Naji Ibrahim Mohamad

Tartous University

Research Article

Keywords:

Posted Date: May 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9332122/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Dynamic Power Allocation for Energy-Efficient 5G Networks Using Load Prediction with RF and SARIMA

Fadi Jawdat Ghosna, Email: fadighosna@tartous-univ.edu.sy¹, Mostfa Albdaire, Email: mostfa.albdair@unisq.edu.au^{2,3}, Laila Mohammad Esmaeil, Email: LailaIsmael@tartous-univ.edu.sy¹, Maher Ali Ibrahim, Email: MaherIbrahim@tartous-univ.edu.sy⁴, Aqeel Sahi, Email: aqeel.sahi@unisq.edu.au^{3,5}, and Naji Ibrahim Mohamad, Email: NajiMohammad@tartous-univ.edu.sy¹

¹Communication Technology Engineering Department, College of Information and Communication Technology Engineering, Tartous University, Tartous 96318, Syria

²Engineering Faculty, Misan University, Misan, Iraq

³Department of Cybersecurity, Engineering Technical College, Al-Ayen Iraqi University, Iraq

⁴Information Technology Engineering Department, College of Information and Communication Technology Engineering, Tartous University, Tartous 96318, Syria

⁵School of Science, Engineering and Digital Technologies, University of Southern Queensland, QLD, Australia

Abstract—For 5G networks, designed to handle a range of applications, such as enhanced mobile broadband (eMBB), ultra reliable low-latency communications (URLLC), and massive machine-to-machine communications (mMTC), energy efficiency is a key performance criterion. It is essential to achieve an optimal balance between energy efficiency and other important elements, such as latency, throughput, connection density, and reliability. This applies to both base stations and user equipment (UEs). This work presents a new energy saving model. A technique is developed to save energy consumption in 5G networks based on predicted load values using two load prediction algorithms. This is achieved by allocating power to access points (APs) based on a dynamic threshold selected based on mathematical data analysis and varying with network load changes. Furthermore, optimal threshold values and threshold change steps are identified that achieve the highest energy savings in the studied network. MATLAB and GraphPad are the primary tools for simulation, analysis, and visualization. A significant improvement in energy efficiency was achieved, with acceptable energy savings of approximately 20% at certain threshold values and threshold changes. Overall, the study highlights the effectiveness of machine learning algorithms in load prediction, making them promising strategies for reducing energy consumption in 5G systems, improving performance and efficiency while maintaining core performance metrics. It also provides valuable insights into future developments in energy-efficient 5G technologies.

I. INTRODUCTION

The rapid deployment of 5G networks has revolutionized mobile communications, delivering unparalleled speed, connectivity, and reliability. As 5G infrastructure grows, its power requirements increase, raising serious energy efficiency concerns for customers and service providers alike. The growing number of Internet of Things (IoT) devices, smartphones, and autonomous systems means that optimizing 5G network power is becoming even more critical. The dynamic requirements and use cases of 5G—which include (eMBB), (mMTC), and ultra reliable, and (URLLC)—often place a significant burden on traditional power management techniques. Fortunately, new advances in artificial intelligence (AI), particularly in the areas of deep learning and machine learning, are opening up new ways

to address many of the challenges facing various systems, and have been adopted in 5G networks to address several issues, including energy-related ones. Real-time energy savings for both user devices (UEs) and base stations can be achieved through deep learning methods, which use data-driven models to predict and respond to network conditions [1]

The next sections of this study are organized as follows: Related Work, methods and materials, the model and steps, simulation results and evaluation, conclusion, recommendations, and references.

II. RELATED WORK

Energy efficiency in wireless communication systems has been a major research focus for many years, leading to various strategies aimed at decreasing power consumption in cellular networks. In the realm of 4G and earlier technologies, traditional methods for conserving energy included resource allocation, base station sleep modes, and power control for user equipment. However, the intricate and varied nature of 5G networks necessitates more advanced solutions [2]. Numerous studies have explored energy-saving methods for 5G networks, some based on machine learning and others not. For example, researchers [2] introduced the concept of Green 5G, which aims to reduce operational energy consumption by encouraging the use of energy-efficient devices, network densification, and the integration of renewable energy sources. Strategies such as coordinated sleep patterns, small cell deployment, and massive MIMO have been explored to reduce energy consumption at base stations while ensuring quality of service. However, these technologies often react to changes rather than immediately adapt to 5G traffic dynamics.

The work presented by [3] demonstrated that the use of machine learning (ML) techniques for wireless network power management has gained increasing popularity recently. By an-

analyzing traffic patterns, supervised learning algorithms have enabled networks to predict times of low demand and adjust resource allocation appropriately. Similarly, reinforcement learning (RL) has been used to control power in multi-user environments, enabling the system to develop optimal power-saving strategies based on feedback. The application of deep learning in the context of 5G energy management is still a nascent field. Research has shown that deep learning can improve energy consumption in base stations. Some studies, such as [4], have relied on convolutional neural networks (CNNs) to efficiently schedule resources, while others have investigated recurrent neural networks (RNNs) to predict traffic patterns over time to enhance energy efficiency. Despite these strides, there is still room to improve the use of deep learning techniques in comprehensive energy saving strategies in 5G networks, particularly those that can be adapted in real-time across both user equipment and infrastructure. The study in [5] proposed the Radio Resource Control (RRC) INACTIVE STATE technology, which included the addition of a new state for the user device to the two states of connected (RRC) and idle (RRC). This state was named INACTIVE STATE, and the device switches to this state when there is no data traffic, returning to the connected state when data arrives within a short period. The advantage of this technology is that the transition from the inactive to connected state takes a shorter time than the transition from idle to connected state. However, its disadvantage is that small data cannot be transmitted in the inactive state. In [6], this study presents an improvement on the traditional cell switch on/off technique previously used to conserve energy in 5G systems. Cell states were switched between on and off based on the traffic exchanged, with some cells being off and others operating, depending on the load associated with each cell. However, switching off cells can degrade system performance, as some cells that remain active require additional traffic to serve their own traffic, which means greater energy consumption. The researchers in this study proposed the switch on/off technique based on machine learning techniques and developed a model of the general architecture of this technique. This model consists of a public module whose task is to collect data, followed by a decision module responsible for conserving energy by measuring performance indicators and setting the threshold accordingly. The output is then sent to the final module, called the execution module, which is responsible for updating the network management database. This technique has proven effective in reducing energy consumption. The study [7] proposed the idea of using Long Short-Term Memory (LSTM) technology to predict data center load. This technology is a smart neural network architecture used in deep learning, primarily for short-term load prediction. LSTM can process not only single data points such as images, but also data sequences such as video or speech, as LSTM consists of multiple memory cells. The results of this study demonstrated high accuracy in load prediction, making it one of the most important technologies used in this field. Researchers [8] have introduced a new technique for predicting users' hourly and daily load. This method overcomes some of the limitations of traditional prediction methods, such as the difficulty of handling nonlinear data, and the limitations of machine learning techniques in feature extraction, such as the difficulty of interpreting training results and

limited learning capacity. The proposed technique, called Deep Cycle Neural Network (DCNN), aims to apply deep learning to analyze user energy data to accurately predict short-term load. High-dimensional features were used to predict the total daily load, then predict the hourly load by selecting low dimensional features. The results demonstrated that the proposed technique simplifies the prediction process and improves accuracy. It combines a deep learning Convolutional Neural Network (CNN) algorithm with the LSTM load prediction technique. Given the importance of energy efficiency for 5G and 6G, and with the expected presence of a trillion IoT devices equipped with non-rechargeable batteries, researcher [9] provided a comprehensive assessment of the cost of achieving the desired latency and reliability of radio communications when applying energy-saving techniques. They also provided a comprehensive review of the most important energy-saving features of the 3GPP standard, such as energy-efficient RRC idle state, early migration procedures, small data transfer without RRC state transition, and other techniques. Among the most important recommendations resulting from this study are that PDCCH skipping is inappropriate for critical services that require very low latency, such as URLLC services. Furthermore, it is important to ensure that collisions do not occur in grant-based uplink transmissions to achieve the required latency gains. The work presented by [10] provides an overview of the evolution of energy-saving technologies supported by the 5G NR standards, taking into account other key performance indicators such as throughput and latency, in addition to highlighting the main trends in green communications development and future improvements beyond 5G. Many scientific studies that addressed the topic of energy saving in 5G and post-5G networks were discussed and reviewed in the work presented by researchers [11], where the advantages and disadvantages of each were discussed, as well as the basic limitations of the available solutions such as computational cost, scalability, and deployment complexities. The research aims to design a system that combines two algorithms: one for load prediction and the other for energy saving, with the aim of reducing energy consumption, improving performance, and increasing efficiency in 5G network communication systems. The performance of two load prediction algorithms was tested, the first one is not based on machine learning (Seasonal Autoregressive Integrated Moving Average (SARIMA)), and the second one is based on machine learning (Random Forest (RF)). An energy-saving technique was developed that relies on a dynamic threshold that varies with network and load conditions, relying on data analysis and determining the most appropriate statistical distribution.

III. METHODS AND MATERIALS

This research was conducted based on recent scientific studies and research related to 5G Cellular systems and technologies, as well as scientific articles on load prediction techniques, their types, and characteristics. The results of these studies were taken into consideration, and the MATLAB programming environment and GraphPad software were used to program and analyse the system under study, display, and compare the results. The implementation steps consisted of the following stages:

1. Obtaining 5G network data for the study.

2. Selecting two load prediction algorithms and applying them to the studied data.
3. Analysing the predicted data using GraphPad.
4. Optimizing the energy-saving technique and applying it to the studied network.
5. Calculating the achieved energy savings.
6. Presenting and analysing the results.

A. Network and traffic model in 5G network

Mathematical modelling of traffic and networks in 5G systems focuses on understanding data traffic patterns, user mobility, and overall network performance with the goal of optimizing resources such as bandwidth, latency, and energy efficiency. Given that 5G networks manage a variety of traffic types, such as enhanced mobile broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable low-latency communications (URLLC), the mathematical models need to reflect the unique requirements and characteristics of each of these services.

A.1 Network Model in 5G

According to [12], the capacity of a communication link can be modelled using Shannon's capacity theorem, which gives the theoretical maximum data rate over a noisy channel:

$$C = B \log_2 \left(1 + \frac{S}{N} \right) \quad (1)$$

B. Used algorithms:

B.1 SARIMA:

At certain points in time, the time series may be non-stationary due to variations in the average values over time within the seasonal range under study. For this reason, a seasonal ARIMA model, called SARIMA, was developed to combine seasonal and non-seasonal factors. It is abbreviated as SARIMA (p, d, q) × (P, D, Q)_S [13]. Where:

- The symbols (b, d, q) are defined in the ARIMA model.
- The symbols (P, D, Q) are the seasonal values of the symbols (b, d, q).
- The symbol S indicates the time period over which the seasonal pattern repeats.

The SARIMA model $(p, d, q) \times (P, D, Q)_S$ is represented by the following equation:

$$\phi(B) \Phi(B^S) \Delta^d \Delta_S^D y_t = c + \theta(B) \Theta(B^S) W_t [13] \quad (2)$$

Where:

- D is the seasonal difference of order D , defined as $\Delta_S^D = (1 - B^S)^D$,
- $\phi(B)$ and $\theta(B)$ are the non-seasonal autoregressive (AR) and moving average (MA) polynomials of order p and q , respectively,
- $\Phi(B^S)$ and $\Theta(B^S)$ are the seasonal AR and MA polynomials of order P and Q , respectively.

The SARIMA model is one of the most important models used to predict future values of time series, due to its accuracy and optimality in most cases. It relies heavily on previously observed values in the time series under study, in order to provide optimal predictions with minimal errors. Accordingly, this model was chosen to study the proposed system [13].

B.2 RF:

Leo Breiman developed the term "random forest" (RF) to describe an ensemble of unpruned classification or regression trees, which are generated by randomly selecting specific samples from training data. The prediction process is the result of a majority vote, or in other words, it is a set of predictions within the ensemble. The authors of [14] claim that random forests are characterized by the following:

- Fast training and prediction.
- Can be used for high-dimensional problems.
- Ease of parallel implementation.
- Naturally handles both regression and multi-class classification.
- Accounts for missing values.
- Unsupervised learning.

C. The model and steps:

C.1 Phase 1:

There are many types and patterns of data traffic in 5G networks, which may be exchanged simultaneously by multiple devices in varying quantities. Therefore, the idea of predicting data traffic and network load emerged with the aim of precisely controlling and managing radio resources, controlling energy consumption in the system, and minimizing losses as much as possible. In the field of load prediction, many algorithms have been developed, and extensive research and studies have been conducted to compare and analyze their performance. Some of these algorithms rely on machine learning, while others do not, and they differ in terms of specifications, testing methods, and accuracy. This phase included the implementation of two types of load prediction algorithms: the first is a developed technique that does not rely on machine learning, namely SARIMA, and the second is a machine learning-based technique, namely RF. The load in the studied network was predicted using these two algorithms and the prediction accuracy of each was measured. Then, the values of the predicted cell load data were analyzed using GraphPad. The aim of the analysis was to find the most appropriate probability distribution for the data and determine the most appropriate initial threshold value based on the resulting distribution.

Call Detail Records (CDRs) are created by the Telecom Italia cellular network in Milan, with the aim of recording and tracking user activity for billing and administration purposes. The data is obtained through accounts applied to these records, which include call, SMS and internet services. The simulation included measuring some parameters for both algorithms, which are:

- MAE: Mean absolute error
- MAPE: Mean Absolute Percentage Error
- RMSE: Root Mean Square Error

The simulation results were as shown in Figure 1

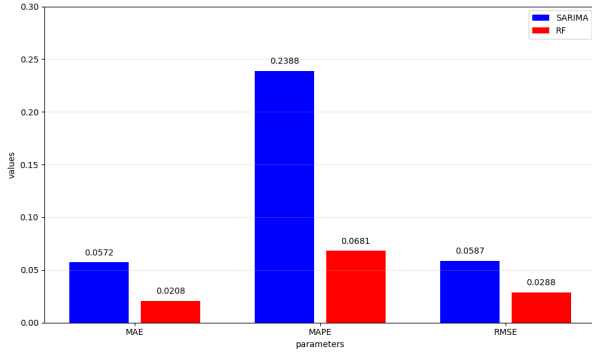


Fig. 1: MAE, MAPE, and RMSE parameters for RF and SARIMA

We note from the values that in the case of load prediction in the studied 5G network, the RF algorithm performs better than SARIMA. It is 3% better in terms of MAE and RMSE, and 17% better in terms of MAPE. This makes it more suitable and effective when building a model or system to reduce energy consumption in a 5G network based on predicted load values. It is also worth noting that SARIMA's performance is largely acceptable. Ultimately, RF and SARIMA are two different techniques used in various fields of machine learning and data analysis. The choice between RF and SARIMA depends on the nature of the data and the specific problem being solved. The RF algorithm is preferred when the dataset is large and diverse, and when we need to handle multiple features. On the other hand, SARIMA is preferred when working with time series data with a seasonal or cyclical pattern and where temporal relationships are evident.

C.2 Phase 2:

After the cell load prediction phase, the GraphPad data analysis software was used to find the closest distribution to the normalized predicted data. Based on the analysis results, an initial value was determined for the dynamic threshold used in the third phase, which is the energy consumption reduction phase. The results of the data analysis process resulted in two types of possible distributions: (Sum of two Gaussians) distribution and the (Sum of two Lorentzian) distribution, for both SARIMA and RF.

- **Sum of two Lorentzian:** It is a continuous probability distribution that describes many phenomena, such as spectral lines and resonance. It is characterized by being symmetrical, resembling a bell, and having heavy tails. It extends farther out to the edges, making it prone to having values far from the center [15].
- **Sum of two Gaussians:** The Gaussian distribution is a continuous probability distribution in statistics and science. It is also characterized by its symmetrical "bell" shape and is used to describe many natural and random phenomena (such as measurement errors or exam scores). This model is well known for its ability to model data that contain more than one central cluster or pattern, and unlike the Lorentzian, the tails of the sum of Gaussians decline more quickly, just like the single Gaussian [16].

The resulting curves are shown in the diagrams below.

Figure 2 shows the results of data analysis when the RF algorithm is used to predict load. Figure 2a shows the blue line, which represents a mathematical function resulting from the (sum of the two Lorentzian) functions. This function is used to model the expected load of the cell. The red dots represent experimental or simulated data for the expected load of the cell. Figure 2b shows the blue line, which represents a mathematical function resulting from the (sum of the two Gaussians) functions. This function is used to model the expected load of the cell. The red dots represent experimental or simulated data for the expected load of the cell.

Figure 3 shows the results of data analysis when the SARIMA algorithm is used to predict load. Figure ?? and Figure3b shows the same results as the previous ones, but in the case of the SARIMA algorithm.

We notice that the curve resulting from the (sum of the two Lorentzian) functions gives a shape resembling a "saddle" or "double dome", and it appears that the predicted data generally follow the shape of the theoretical curve, with some slight variations. Also, we notice that the curve resulting from the (sum of the two Gaussian) functions gives a shape resembling a "hill" or "single peak," and the predicted data generally seem to follow the shape of the theoretical curve as well. Both functions (Lorentz and Gauss) show a good fit to experimental data, but there may be slight differences in how accurately each represents the data. The mathematical function is chosen based on how well it matches the experimental data and the objectives of the study.

Based on a visual analysis of the curves, it can be seen that the most appropriate distribution for the studied data is the sum of two Gaussians. This is because the shape of the curve resulting from the sum of two Gaussian functions best matches the shape of the experimental data, showing a single peak resembling a "hill" shape, and the data generally follows this pattern [16]. Because the data under study does not contain many extreme values, the Lorentz distribution is not the best choice. Furthermore, the mean and variance of this distribution cannot be defined, which is why we rely on the Gauss distribution as the most appropriate distribution for the data under study [15]. From [16], the basic equations for the sum of two Gaussian distributions are given as follows:

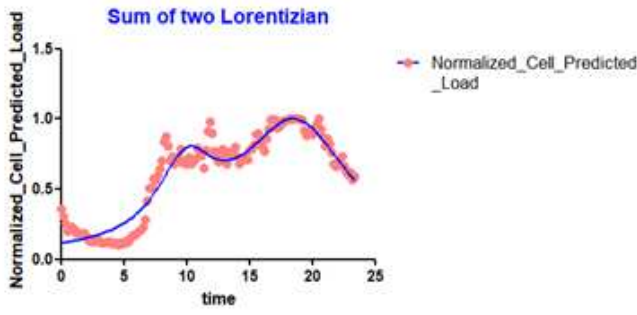
$$\text{one} = \text{Amplitude}_1 \cdot \exp\left(-0.5 \left(\frac{X - \text{Mean}_1}{SD_1}\right)^2\right) \quad (3)$$

$$\text{two} = \text{Amplitude}_2 \cdot \exp\left(-0.5 \left(\frac{X - \text{Mean}_2}{SD_2}\right)^2\right) \quad (4)$$

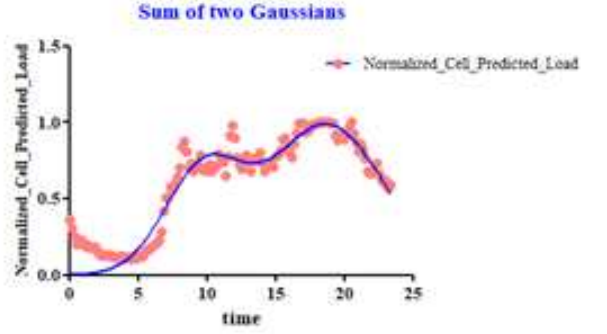
$$Y = \text{one} + \text{two} \quad (5)$$

where:

- *one* and *two* represent the first and the second Gaussian function.
 - SD_1 and SD_2 are the standard deviations for each function.
- The goal of determining the most appropriate distribution for the data was to use it to determine the initial threshold value for the next stage. The mean values of each of the two Gaussian functions were taken, their average was calculated, and adopted as the initial value for the dynamic threshold, for both the RF and SARIMA algorithms.

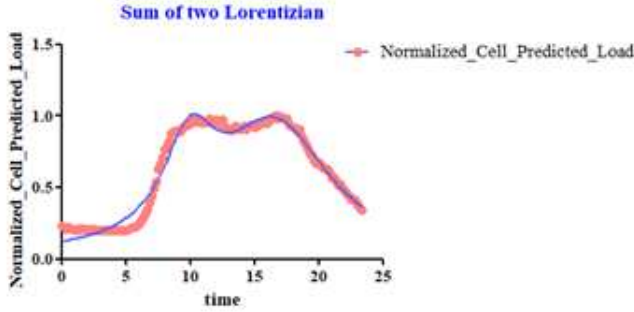


(a) Comparing with the “sum of the two Lorentzian” distribution

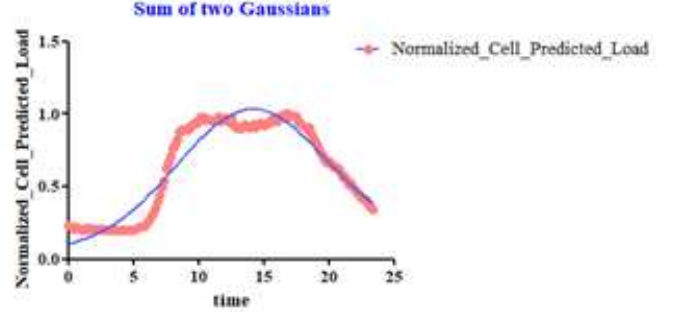


(b) Comparing with the “sum of the two Gaussians” distribution

Fig. 2: Analysis of the normalized cell predicted load by RF



(a) Comparing with the “sum of the two Lorentzian” distribution



(b) Comparing with the “sum of the two Gaussians” distribution

Fig. 3: Analysis of the normalized cell predicted load by SARIMA

In the case of using RF to predict load, the results of the data analysis were as shown in Table 1:

TABLE I: The results of the data analysis predicted using RF

Sum of two Gaussians	Value
Mean1	9.668
Mean2	18.68
Sum	28.34
Average	14.17

In the case of using SARIMA to predict load, the results of the data analysis were as shown in Table 2:

TABLE II: The results of the data analysis predicted using SARIMA

Sum of two Gaussians	Value
Mean1	14.02
Mean2	15.05
Sum	29.07
Average	14.53

C.3 Phase 3:

This phase involved developing an energy optimization model using MATLAB, targeting base stations (BSs) by dynamically adjusting network parameters based on real-time traffic conditions. This model aims to improve energy efficiency while maintaining overall network performance. The proposed strategy involves dynamic power allocation to the BS. In other words, the dynamic threshold actively adapts to changes in load. The network uses the dynamic threshold as a mechanism for adaptive resource control, becoming more conservative when the load is high and more lenient when the load is low. This approach aims to balance resource efficiency with ensuring quality of service for users at different times of the day.

The power consumed by the BS in a densely populated network can be expressed as:

$$P_i = \sum_i S_i \left(P_{\text{idle}} + \frac{1}{\eta} P_{\text{trans},i} \right) \quad (6)$$

where:

- $P_{\text{trans},i}$: the transmission power of the i -th station,
- P_{idle} : the base station power consumption in non-transmitting state,

- η : the efficiency coefficient,
- S_i : the signal.

The following are the constraints that must be met to ensure the effectiveness of the proposed model: The user's received power must meet the minimum data transfer rate:

$$\sum_{i=1}^N x_i \cdot \text{SINR}_{ij} \geq R_j \quad (7)$$

The transmitted power must be within the permissible range:

$$P_{\min} \leq P_i \leq P_{\max} \quad (8)$$

The primary goal is to ensure that the power is as low as possible:

$$\text{Min } P_{\text{total}} = \sum_{i=1}^N S_i \left(P_{\text{idle}} + \frac{1}{\eta} P_i \right) \quad (9)$$

C.4 Power Allocation Algorithm Steps:

After initial setup, parameter adjustment, and cell load prediction, the predicted load is normalized, and the load trend (load tendency) is then calculated. This is followed by dynamically adjusting the threshold based on the load and load tendency, and then dynamically allocating power in proportion to the threshold. Finally, the amount of energy saved is calculated using the dynamic power allocation mechanism, in addition to determining the optimal parameter values that provide the greatest energy savings.

To calculate the dynamic threshold, follow these steps:

- The dynamic threshold is initialized and parameters are defined, such as the threshold change step and the initial startup threshold value, which is equal to the mean of the first and second centers of the Gaussian distribution.
- For each time step:
- The previously Normalized Cell Predicted Load (NCPL) is compared with the previously

Normalized Full Network Predicted Load (NFNPL), along with their slopes.

If the conditions are as follows:

- If $NCPL < NFNPL$ and $CPLT < FNPLT$, decrease the threshold by $2 \times \text{threshold_change_step}$.
- If $NCPL \geq NFNPL$ and $CPLT < FNPLT$, decrease the threshold by $\text{threshold_change_step}$.
- If $NCPL < NFNPL$ and $CPLT \geq FNPLT$, increase the threshold by $\text{threshold_change_step}$.
- Otherwise, increase the threshold by $2 \times \text{threshold_change_step}$.

The threshold is fixed in the range between 0 and 1, the dynamic threshold value is stored for each time step, and finally the percentage of power allocation is plotted over time: the power allocation is displayed as a percentage with the optimal values of the parameters used. The flowchart in Figure 4 illustrates the steps of the previous algorithm.

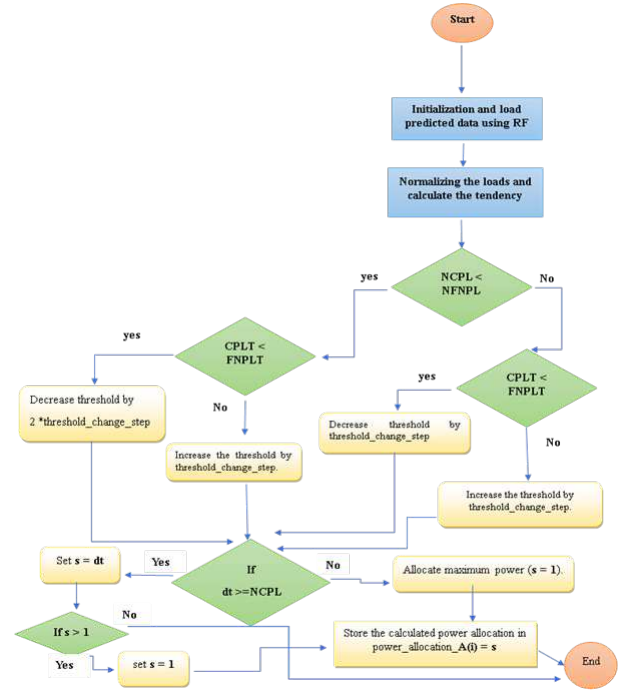


Fig. 4: Power Allocation Algorithm flowchart

IV. SIMULATION RESULTS AND EVALUATION:

In this section, we present the results of the proposed algorithm and the resulting energy savings improvement for the studied network. The dataset used in this study was obtained from Harvard University and contains information on telecommunications activities in the city of Milan for billing and network management purposes. Telecom Italia's mobile network generates call detail records (CDRs), from which this data was obtained.

A. The first case: RF algorithm

Figure 5 illustrates how the Normalized Cell Predicted Load (NCPL) values change and fluctuate as the dynamic threshold changes over time. Blue Curve (NCPL): This curve shows the predicted and measured load on a 5G network cell over time (in hours). Higher values indicate increased network usage, such as increased active users or higher data volume. Lower values indicate decreased network usage.

Red Curve (DT - Dynamic Threshold): This curve represents a dynamic threshold that changes over time. This threshold represents the point at which the network may begin implementing resource management or usage regulation mechanisms to avoid congestion and ensure quality of service. This threshold may depend on various factors such as cell capacity, the required quality of service, or network policies.

Based on the comparison between the load and the dynamic threshold, energy is allocated. When the load is low, high-energy values can be allocated and all resources allocated. Conversely, when the load is high and the network is under pressure, energy and resources must be managed effectively to avoid problems.

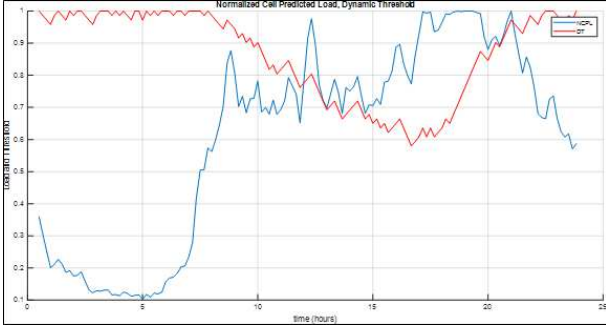


Fig. 5: NCPL changes with DT in the case of RF

In general, Figure 6 shows how the power allocation percentage (PA) changes over time (in hours). Interpreting the different phases of the curve:

- The period from hour 0 to approximately hour 8: During this period, we observe that the cell load (NCPL) is very low, well below the relatively high dynamic threshold (DT). In this case, the system is able to allocate almost all of the power (as seen in the previous curve at 100)
- The period from approximately hour 8 to approximately hour 13: Here, we witness a sharp increase in the cell load (NCPL). The load begins to approach the dynamic threshold and then significantly exceeds it. This increasing congestion in the network, which exceeds a specific capacity (represented dynamically by the threshold), is likely the main reason for the gradual decrease in the power allocation ratio we observed earlier. The system temporarily reduces the percentage of power allocated to each user or service to manage congestion and prevent network collapse when demand exceeds the available capacity as determined by the dynamic threshold.
- Period from approximately 13 hours to approximately 20 hours: During this period, the cell load (NCPL) drops significantly and rapidly, becoming well below the dynamic threshold (DT). This reduction in network demand allows the system to lift the constraints imposed on power allocation, which explains the rapid return of the power allocation ratio to nearly 100
- Period from approximately 20 hours to the end of the period: Here, we see that the cell load (NCPL) begins to rise again, approaching the dynamic threshold (DT), and sometimes slightly exceeding it. This is consistent with the slight fluctuations in the power allocation ratio seen in the previous curve during this period. When the load increases and approaches the permissible limit (the dynamic threshold), the system may slightly adjust the power allocation to maintain stable performance.

V. RESULTS DISCUSSION

Comparing the two curves, it becomes clear that the power allocation ratio in a 5G system is likely a dynamic response mechanism to changes in cell load. When cell load increases and approaches or exceeds the threshold, the power allocation ratio is reduced to manage congestion and ensure service continuity. When the load decreases and falls below the threshold, the system switches back to allocating more power.

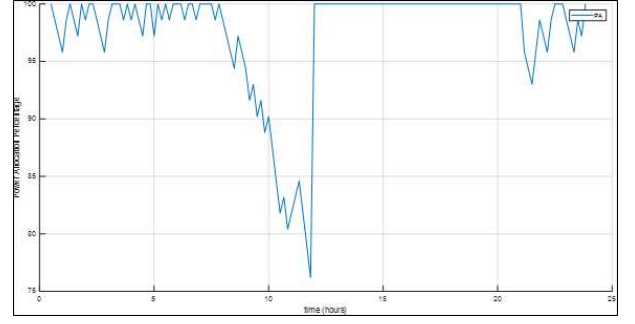


Fig. 6: The power allocation percentage (PA) over time

The output of the energy consumption reduction algorithm is shown in Figure 7 and Table 3 where the best values for the parameters that give the greatest energy savings were found depending on the startup threshold and the step value. The three-dimensional curve represents the simulation results of the proposed algorithm and illustrates the relationship between the starting threshold values, the threshold step size, and the resulting energy savings. We note the following:

X-axis: Represents the optimal starting threshold value for the algorithm. This threshold determines when the algorithm begins activating energy-saving mechanisms based on a specific level of load or activity in the network. Values range from 0 to 1. Values close to 0 mean that the algorithm starts early at low load levels, while values close to 1 mean that it only starts at high load levels.

Y-axis: Represents the optimal step size for changing the threshold. It shows the magnitude of these increases or decreases in the threshold. Values range from 0 to 1. Smaller values represent subtle and gradual changes in the threshold, while larger values represent more drastic changes.

Z-axis: Represents the optimal energy savings achieved.

A. Curve Notes:

- High-saving areas: We notice a clear peak (in yellow and light green) in a specific part of the curve. This indicates an optimal or near-optimal combination of starting threshold values and step size that achieves the greatest energy savings. At the point indicated by the index (X: 0.001, Y: 0.253, Z: 20.3546), the algorithm appears to achieve high energy savings.
- Presence of flat or minimally variable regions: Some regions of the curve appear relatively flat, meaning that changing the values of one of the axes (start threshold or step size) within this range may not significantly impact the energy savings.

TABLE III: Simulation results in the case of RF and sum of two Gaussians

startup_threshold_best	threshold_change_step_best	power_saving_best
0.253	0.001	20.35%

TABLE IV: Simulation Results in the case of SARIMA and Sum of two Gaussians

startup threshold best	threshold change step best	power_saving_best
0.248	0	19.73%

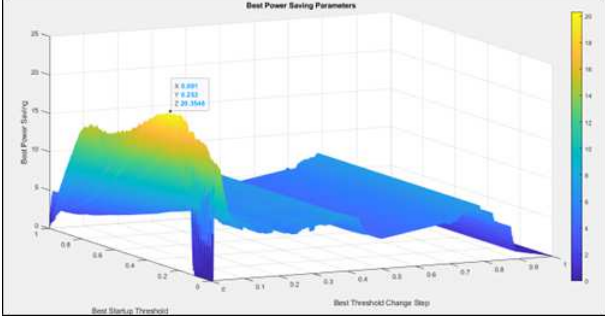


Fig. 7: The greatest energy savings with the best values for the parameters

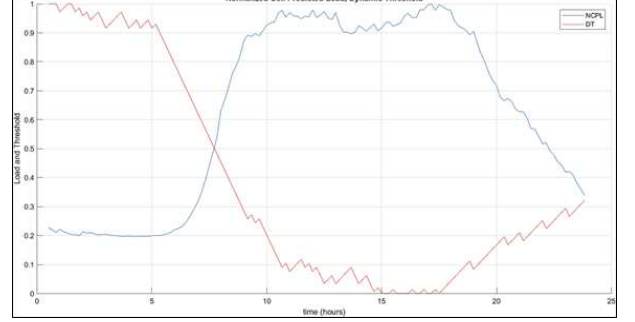


Fig. 8: NCPL changes with DT in the case of SARIMA

B. The second case: SARIMA algorithm

As in the previous case, the proposed energy-saving algorithm was applied to the SARIMA load prediction algorithm for comparison and performance analysis. The results were as follows: Figure 8 shows the change in the NCPL values for the dynamic threshold over time, where we see: The period from 0 to approximately 6: During this period, the cell load is relatively low and stable at around 0.2. In contrast, the dynamic threshold (DT) is high, approaching 1. This indicates that the network is underutilized and that there is ample capacity available before the network needs to activate resource management mechanisms.

The period from approximately 6 to approximately 9: The cell load begins to rise sharply and rapidly, reaching its peak at approximately 1:00 PM. During this period, the dynamic threshold decreases significantly, reaching its lowest levels, almost exactly when the load is at its peak. This behavior of the dynamic threshold indicates that the network is adjusting the threshold based on expected or actual load and becomes more stringent in managing resources during peak times.

The period from approximately 9:00 PM to approximately 18: The cell load continues to remain relatively high with some fluctuations. The dynamic threshold remains low during most of this period, indicating that the network remains in a more conservative resource management mode due to high demand.

Period from approximately 18 hours to approximately 24 hours: Cell load begins to gradually decrease. Conversely, the dynamic threshold begins to rise again as the load decreases. By the end of the period, the load returns to low levels, and the dynamic threshold rises again to high values.

In general, the Figure 9 shows how the power allocation percentage (PA) changes over time, and the different phases of the curve are: Before approximately 7 hours (high and stable power allocation): During this period, the cell load (blue curve) was relatively low, and the dynamic threshold (red curve) was high. When the load is low and the threshold is high, there is no significant pressure on the system, and thus a high percentage of power can be allocated stably.

Between 7 hours and around 8.5 hours (sharp drop in power allocation): This is the most interesting part. We notice that this sharp drop in power allocation roughly coincides with the period when the cell load starts to increase significantly, while the dynamic threshold continues to decrease. This indicates that the system, as the load began to approach the low dynamic threshold, temporarily reduced power allocation to this cell to manage resources more efficiently or to avoid congestion as the load continued to increase. This could have redirected some of the power to other cells that needed it more at that moment.

After approximately 8.5 hours (sharp rise in power allocation): After the sudden drop, power allocation quickly returns to a high level. This coincides with the cell load continuing to increase and exceeding the low dynamic threshold. The reason for this sudden spike may be that the system recognized the increased need for power as the load increased significantly, and thus reallocated more power to meet this increased demand.

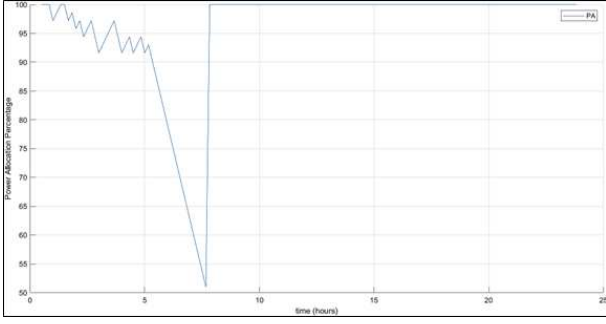


Fig. 9: The power allocation percentage (PA) over time

The output of the power consumption minimization algorithm is shown in Figure 10 and Table 4 where the best values of the parameters that give the greatest power savings are found depending on the startup threshold and step value, in the case of using SARIMA for load prediction. Looking at the curve, we can observe a clear peak in energy savings. There is a clear region in the curve (located approximately at X values close to 0 and Y values around 0.25) where the highest energy savings of 19.73 (approaching 20) are achieved. This indicates that there is an optimal combination of starting threshold values and threshold step changes that results in maximum energy efficiency according to this algorithm and the load data predicted using SARIMA.

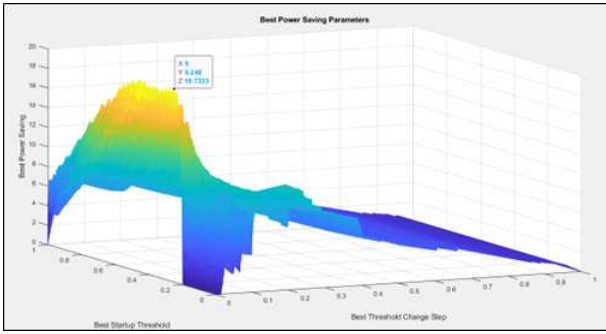


Fig. 10: The greatest energy savings with the best values for the parameters

VI. CONCLUSION

Based on the results obtained in this study, several important conclusions can be drawn.

First, the threshold step size has a significant impact on energy savings. There exists an optimal range of step sizes that achieves the best performance. Very small step sizes may result in a slow response to variations in network load, while excessively large step sizes can introduce unnecessary fluctuations in energy savings and system behaviour.

Second, the start-up threshold plays a critical role in determining energy efficiency. Varying the start threshold reveals that there is a specific range of values that leads to improved energy savings. Extremely low start thresholds may trigger energy-saving mechanisms prematurely, even when the network load is low, potentially degrading performance without yielding substantial savings. Conversely, excessively high start thresholds

may delay the activation of energy-saving mechanisms, resulting in missed opportunities to reduce power consumption during moderate load conditions. Consequently, analysing the traffic characteristics and identifying an appropriate probability distribution are essential steps in determining an effective start threshold value.

The study further demonstrates that, in 5G systems, the adoption of intelligent resource allocation mechanisms is beneficial for balancing energy efficiency with the dynamic load demands of the network. This balance can be achieved through the use of a dynamic threshold-based power management strategy. When the cell load is relatively low and the threshold is high, a higher percentage of power is allocated. As the load increases and approaches a lower dynamic threshold, power allocation is temporarily reduced as a resource management measure. Once the load exceeds the threshold, the system returns to allocating a higher level of power to satisfy the increased demand.

Finally, it should be noted that the presented results are specific to the considered scenario and modelling assumptions. Energy savings may vary across different 5G network environments, and the optimal configuration may also depend on additional factors not addressed in this study, such as Quality of Service (QoS) requirements and user experience. As a result, a trade-off between energy efficiency and network performance may exist and should be carefully considered in practical deployments.

A. Recommendations

Based on the findings of this work, the following recommendations can be drawn:

- When developing energy reduction techniques for 5G networks, it is essential to consider both the individual cell load and the overall network load, as well as their respective temporal variations and slopes.
- Random Forest (RF) and Seasonal Autoregressive Integrated Moving Average (SARIMA) are distinct techniques commonly used in machine learning and data analysis. The results of this study indicate that their performance is comparable for the investigated scenario, as both achieved similar energy savings. Consequently, the choice between RF and SARIMA should depend on the characteristics of the data and the specific problem. RF is more suitable for large and heterogeneous datasets involving multiple features, whereas SARIMA is preferable for time-series data with clear seasonal or cyclical patterns.
- The proposed dynamic thresholding approach demonstrates acceptable energy savings in 5G networks. In particular, appropriate combinations of the start-up threshold and switching step can be selected to maximise energy efficiency and may serve as a reference for similar network scenarios.
- Implementing power-reduction algorithms during peak and congestion periods may be beneficial, as network load is significantly higher during these times. This approach can help mitigate the additional overhead introduced by algorithmic complexity, computational processes, and resource consumption.

REFERENCES

- [1] A. K. Bhat, V. Gojanur, and R. Hegde, "5g evolution and need: A study," in *International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO)*, 2015, pp. 1–4.
- [2] O. Shurdi, L. Ruci, A. Biberaj, and G. Mesi, "5g energy efficiency overview," *European Scientific Journal*, vol. 17, no. 3, p. 315, 2021.
- [3] D. Ailyn and A. Samy, "Machine learning applications in wireless network optimization and management," 2024.
- [4] A. Mughees, M. Tahir, A. Sheikh, and A. Ahad, "Towards energy efficient 5g networks using machine learning: Taxonomy, research challenges, and future research directions," *IEEE Access*, vol. 8, pp. 187 498–187 522, 2020.
- [5] "Nr and nr-ran overall description; stage 2 (release 15)," 3GPP, Tech. Rep. TS 38, December 2018.
- [6] Y. Gao, J. Chen, Z. Liu, B. Zhang, Y. Ke, and R. Liu, "Machine learning based energy saving scheme in wireless access networks," in *International Wireless Communications and Mobile Computing (IWCMC)*, 2020, pp. 1573–1578.
- [7] Q. Guo *et al.*, "Research on lstm-based load prediction for edge data centers," in *IEEE International Conference on Computer and Communications (ICCC)*, 2018, pp. 1825–1829.
- [8] W. Zhu, Y. Zeng, Z. Kang, and J. Fu, "Deep learning based short term load prediction in smart grids," in *IEEE International Conference on Electronic Information and Communication Technology (ICEICT)*, 2020, pp. 674–678.
- [9] A. A. Esswie, "Power saving techniques in 3gpp 5g new radio: A comprehensive latency and reliability analysis," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 66–71.
- [10] Y.-N. R. Li, M. Chen, J. Xu, L. Tian, and K. Huang, "Power saving techniques for 5g and beyond," *IEEE Access*, vol. 8, pp. 108 675–108 690, 2020.
- [11] A. Ichimescu, N. Popescu, E. Popovici, and A. Toma, "Energy efficiency for 5g and beyond 5g: Potential, limitations, and future directions," *Sensors*, vol. 24, p. 7402, 2024.
- [12] O. Rioul and J. Magossi, "Shannon's formula and hartley's rule: A mathematical coincidence?" *American Journal of Physics*, 2014.
- [13] J. Arlt, P. Trcka, and M. Arltová, "The problem of the sarima model selection for the forecasting purpose," *Statistika*, vol. 97, pp. 25–32, 2017.
- [14] I. Alawe, A. Ksentini, Y. Hadjadj-Aoul, and P. Bertin, "Improving traffic forecasting for 5g core network scalability: A machine learning approach," *IEEE Network*, 2018.
- [15] G. Major, V. Fernandez, N. Fairley, and M. Linford, "A detailed view of the gaussian–lorentzian sum and product functions and their comparison with the voigt function," *Surface and Interface Analysis*, vol. 54, 2021.
- [16] T. Betlehem and A. J. Coulson, "Distribution of the sum of a complex gaussian and the product of two complex gaussians," in *Australian Communications Theory Workshop (AusCTW)*, 2012, pp. 126–129.