

Article

An EMF-Aware Intelligent Framework for Adaptive SSB Periodicity Control in 5G Networks

Keze Li , Michael S. Mollel , Olaoluwa Popoola , Muhammad Ali Imran  and Yusuf Sambo 

James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, UK;
michael.mollel@sartify.com (M.S.M.); olaoluwa.popoola@glasgow.ac.uk (O.P.);
muhammad.imran@glasgow.ac.uk (M.A.I.); yusuf.sambo@glasgow.ac.uk (Y.S.)

* Correspondence: 23577611@student.gla.ac.uk

Abstract

Synchronization Signal Blocks (SSBs) support initial access, synchronization, and beam management in fifth-generation networks. However, the periodic transmission of SSBs contributes to background electromagnetic field (EMF) exposure, while increasing the SSB periodicity may increase the waiting time experienced by newly arriving User Equipment (UE). This paper proposes an EMF-aware adaptive SSB periodicity control framework that integrates Long Short-Term Memory (LSTM)-based demand prediction with a Proximal Policy Optimization (PPO) controller. Aggregated Internet usage measurements collected by Telecom Italia in Milan in 2013 at 10 min resolution are transformed into a traffic-derived UE arrival proxy. The LSTM model forecasts the proxy for the subsequent control interval, and the PPO policy selects a SSB periodicity from 5, 10, 20, 40, 80, and 160 ms. The optimization reward combines normalized SSB EMF power density and normalized aggregate UE waiting time, thereby avoiding dimensional and numerical-scale inconsistencies between the two objectives. The LSTM and PPO models are developed using six chronological weeks of data and evaluated without parameter updates over a seven-day held-out period. The prediction-driven controller achieves a mean SSB EMF power density of $3.67 \times 10^{-5} \text{ W/m}^2$ and a weekly average waiting time of 19.22 ms per person. Relative to the fixed 20 and 40 ms configurations, the proposed controller reduces mean EMF power density by approximately 65.4% and 30.8%, respectively. Its waiting time is close to that of the fixed 40 ms configuration and substantially lower than that of the fixed 80 ms configuration. Although the fixed 80 ms configuration provides lower EMF power density, it incurs considerably greater waiting time. The results show that the proposed framework provides an adaptive intermediate operating point between the lower waiting time of short fixed periodicities and the lower EMF exposure of long fixed periodicities.



Academic Editors: George Vardoulas,
Ioannis Bartsiokas and Panagiotis K.
Gkonis

Received: 18 June 2026

Revised: 21 July 2026

Accepted: 28 July 2026

Published: 6 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: electromagnetic field (EMF) exposure; synchronization signal block (SSB); 5G new radio; reinforcement learning

1. Introduction

The deployment of fifth-generation (5G) mobile networks is significantly altering wireless communication technologies, facilitating unprecedented data speeds and enhanced connectivity. This technological evolution promises progress in multiple sectors, including healthcare, automotive, entertainment, and industrial automation, enabling advances such as remote surgery, driverless automobiles, immersive virtual reality, and smart manufacturing. Among the numerous enabling components of 5G networks, Synchronization Signal

Blocks (SSBs) play an important role by providing essential reference signals that support initial access, synchronization, and beam management [1].

SSBs are specifically designed to deliver critical reference signals that assist User Equipment (UE) in executing fundamental activities such as cell detection, synchronization, and effective beam management [2]. In 5G New Radio (NR), SSBs are fundamental to initial access procedures, influencing the efficacy of cell selection, handover operations, and beam alignment precision [3,4]. Effective beam management enabled by SSBs guarantees dependable connections, minimizes interference, and enhances network resource allocation, hence directly impacting overall network performance and user Quality of Experience (QoE). Because user activity may be distributed unevenly across beam directions and vary over time, adaptive allocation of SSB transmission resources has been investigated to improve initial-access efficiency compared with uniform SSB allocation [5].

Nonetheless, with the increasing deployment of 5G infrastructure, apprehensions over Electromagnetic Field (EMF) exposure from wireless networks have escalated. Public concerns of the possible health dangers of EMF exposure have increased significantly, leading regulatory agencies globally to establish rigorous guidelines and recommendations to mitigate health hazards linked to EMF radiation. International regulations, such as those published by the International Commission on Non-Ionizing Radiation Protection (ICNIRP), have imposed specific thresholds for maximum permissible EMF levels to protect public health [6]. Compliance with these rigorous limits is crucial for regulatory adherence and for improving the public acceptance of 5G technology, promoting socially responsible network rollout.

SSB transmissions, marked by their periodic broadcasting, directly affect the electromagnetic field exposure levels encountered by the public. Our earlier measurement-based study [7] quantified the EMF power density caused by SSB transmissions and highlighted the potential to reduce unnecessary SSB broadcasts to minimize EMF exposure. Shorter SSB periods, characterized by frequent signal transmission, typically enhance communication reliability and decrease UE initial access latency, consequently ensuring seamless connectivity and improved user satisfaction. Nonetheless, these recurrent transmissions increase EMF power density levels, potentially intensifying public exposure. Conversely, extending the delay between SSB transmissions significantly lowers EMF exposure but may adversely affect network performance by prolonging initial access waiting times, leading to unsatisfactory user experiences. Consequently, a complex and significant trade-off occurs between reducing EMF emissions and maintaining high-quality user connectivity, which requires advanced solutions.

To effectively tackle this complex trade-off, an innovative reinforcement learning (RL)-based approach is proposed that can dynamically select the appropriate SSB periodicity. Our approach predicts future user-demand levels using traffic-derived estimates by applying long short-term memory (LSTM) neural networks [8]. These networks are trained using the Milan city dataset, which provides urban internet-usage measurements at a 10 min resolution. In this work, a chronological seven-week period from November to December 2013 is used. The first six weeks are reserved for LSTM and PPO model development, while the subsequent seventh week is held out exclusively for the final evaluation of the integrated framework [9]. No observations from the test week are used during model training or parameter selection.

The established RL framework uses these predictions to adjust SSB periodicities to reduce EMF exposure while concurrently keeping user waiting times within acceptable limits. In contrast to static SSB configurations, the proposed framework provides an adaptive mechanism that responds to temporal variations in the traffic-derived demand proxy.

Our principal contributions in this paper include the following: (1) the development and validation of a robust predictive model for accurately forecasting estimated UE arrival proxy utilizing data from Milan Internet Traffic Dataset; (2) the design of a RL framework that dynamically and autonomously adjusts SSB periodicity to minimize public exposure to EMF while keeping a low average UE access delay; and (3) extensive, realistic simulations to assess the effectiveness of the proposed approach. The results demonstrate that the proposed prediction-driven controller reduces SSB EMF power density relative to the fixed 20 ms and 40 ms configurations while maintaining substantially lower waiting time than the fixed 80 ms configuration. The adaptive controller therefore provides an intermediate EMF–delay operating point rather than outperforming every fixed periodicity in both objectives.

The subsequent sections of this work are structured as follows: Section 2 offers a comprehensive analysis of the current research concerning SSB periodicity optimization, electromagnetic field control tactics in 5G implementations, and the utilization of reinforcement learning methodologies in wireless networks. Section 3 presents our system model, the fundamental assumptions, and the exact problem formulation. Section 4 details our proposed RL algorithm, predictive modeling approach, and implementation strategy. Section 5 contains the performance evaluation, encompassing simulation configurations, outcomes, and analysis. Finally, Section 6 concludes the work by presenting our findings.

2. Related Works

Recent studies have highlighted the important role of SSBs in supporting critical tasks such as cell discovery, synchronization, and beam management. More broadly, synchronization itself is essential in 5G systems, where accurate time and frequency alignment is necessary to achieve high data rates, low latency, and reliable communication. Prior works have examined the role of SSBs in enabling reliable connectivity, particularly under high-mobility conditions. For example, the authors in [10] proposed a joint optimization of SSB beam planning and Unmanned Aerial Vehicle (UAV)-based cell selection using Deep Reinforcement Learning (DRL) to reduce handover failures along aerial highways. In a related study, the same authors developed a DRL framework for SSB beam selection in massive MIMO highway scenarios, improving UAV cell association and link reliability under dynamic mobility. These studies demonstrate the applicability of deep reinforcement learning to SSB beam planning, beam selection, and mobility-aware cell association. However, their decision variables concern the spatial configuration of SSB beams and the serving-cell selection process rather than the temporal periodicity of SSB transmission. They therefore provide the relevant background on learning-based SSB control but are not treated as prior solutions to the EMF-aware SSB periodicity-selection problem considered in this work [11].

Notable, 3GPP NR defines fixed periodic configurations for SSB transmissions, typically ranging from 5 ms to 160 ms [12]. While such static settings simplify implementation, they fail to adapt to temporal variations in user activity. As a result, short fixed periods (e.g., 20 ms) reduce access latency but also increase background EMF exposure, whereas long periods (e.g., 80 ms or 160 ms) reduce EMF but significantly increase user wait times. Our earlier study [7] experimentally quantified this trade-off, demonstrating that shorter SSB periodicities lead to higher power density and throughput, while longer ones reduce EMF at the cost of degraded access performance. These findings demonstrate the EMF and access delay trade-off associated with fixed SSB configurations and motivate the investigation of traffic-aware periodicity selection.

The growing deployment of 5G infrastructure has been accompanied by increased public scrutiny regarding EMF exposure, with many studies focusing on assessing health risks associated with continuous exposure to radiofrequency EMF. International bodies such

as the ICNIRP have provided guidelines emphasizing the critical need to maintain public EMF exposure within safe limits [6]. Accordingly, numerous researchers have addressed EMF management specifically in the context of emerging 5G technologies. For example, Chiaraviglio et al. reviewed the EMF exposure implications of massive MIMO deployments, underscoring the heightened potential for exposure near densely deployed base stations and reviews the challenges of deploying 5G under strict EMF regulations, demonstrating through real-world case studies how EMF saturation and legal constraints significantly impact network planning, service quality, and infrastructure costs, highlighting that most current approaches address EMF exposure through offline planning or static constraints rather than adaptive, traffic-aware control mechanisms [13].

In parallel, the distinctive transmission characteristics of 5G, including massive MIMO, directional beamforming, and operation over wider frequency ranges, have motivated research on both EMF-aware network design and improved exposure-assessment methods. Elzanaty et al. [14] examined misconceptions and unresolved questions concerning 5G EMF exposure and proposed EMF-aware cellular-network design approaches intended to reduce human exposure while maintaining communication performance. In contrast, the study in [15] focused on the limitations of conventional exposure-assessment methods when applied to 5G systems and investigated modeling approaches suited to beamformed and spatially varying transmissions. These studies establish the broader need for EMF-aware network operation and reliable exposure assessment, but they do not consider adaptive SSB-periodicity selection.

In response to these limitations, machine learning techniques, particularly RL, have increasingly been applied to manage complex trade-offs in wireless communication systems. RL is well suited for such problems because it optimizes decisions dynamically through repeated interaction with a time-varying environment. For example, paper [16] investigated deep learning frameworks to enhance physical-layer functions in 5G, including Non-Orthogonal Multiple Access, massive MIMO, and mmWave, showing that neural networks can significantly improve channel estimation and resource allocation under dynamic conditions. Paper [17] formulated handover control in a 5G cellular network as a reinforcement-learning problem, in which a centralized agent selected handover actions using UE radio measurement reports to improve long-term mobility performance. Collectively, these works highlight the potential of RL for adaptive and context-aware resource management. However, most prior applications have focused on beamforming, spectrum allocation, or mobility management, leaving open questions about how RL can be leveraged for control signals such as SSBs.

Adaptive modification of SSB periodicity has previously been discussed in the context of network energy saving and traffic dependent operation. For example, time-domain energy saving techniques can increase the SSB periodicity during low activity conditions to create longer opportunities for gNodeB (gNB) sleep operation [18]. These existing approaches primarily target network energy consumption rather than public EMF exposure. In contrast, the present study investigates traffic-aware SSB periodicity selection using a LSTM predictor and a PPO controller, with normalized SSB EMF power density and aggregate user waiting time included explicitly in the optimization objective.

3. System Model

We consider a 5G macro-cell deployment where a base station periodically broadcasts SSB to support cell discovery, synchronization, handover procedures, and beam management. The 3GPP specification defines discrete transmission periodicities:

$$T_{\text{SSB}} \in \{5, 10, 20, 40, 80, 160\} \text{ ms} \quad (1)$$

where shorter periods improve access responsiveness but increase signaling overhead and EMF exposure, while longer periods reduce signaling overhead and EMF at the cost of higher user access delay.

During each SSB burst, the gNB transmits one or more SSBs over different beams using beam sweeping. The average user access delay is therefore directly proportional to the chosen T_{SSB} , as defined in the previous subsection. The choice of SSB periodicity thus constitutes the key decision variable in our system model.

3.1. SSB EMF Power Density Calculation

While SSBs are critical for network operation, their continuous transmission results in persistent background EMF exposure. As illustrated in Figure 1, each SSB is transmitted within a burst window of one radio frame and can be beamformed toward multiple directions in Figure 2, resulting in periodic directional exposure patterns.

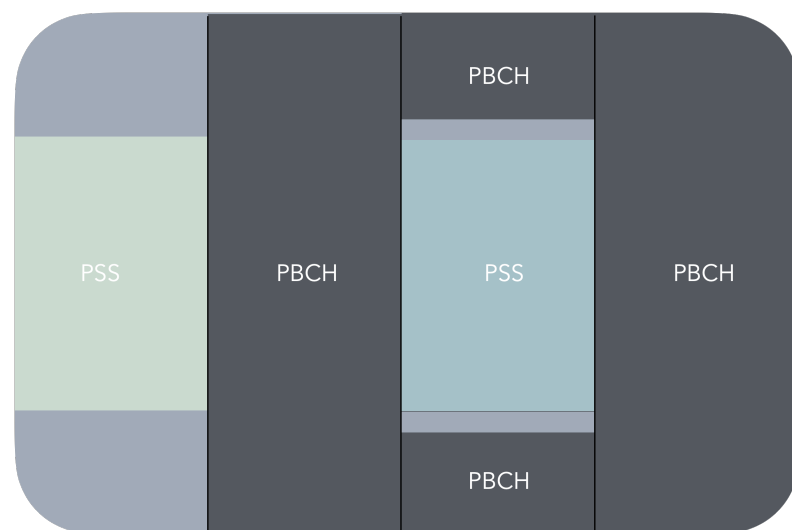


Figure 1. Structure of a Synchronization Signal Block (SSB) in 5G NR [7].

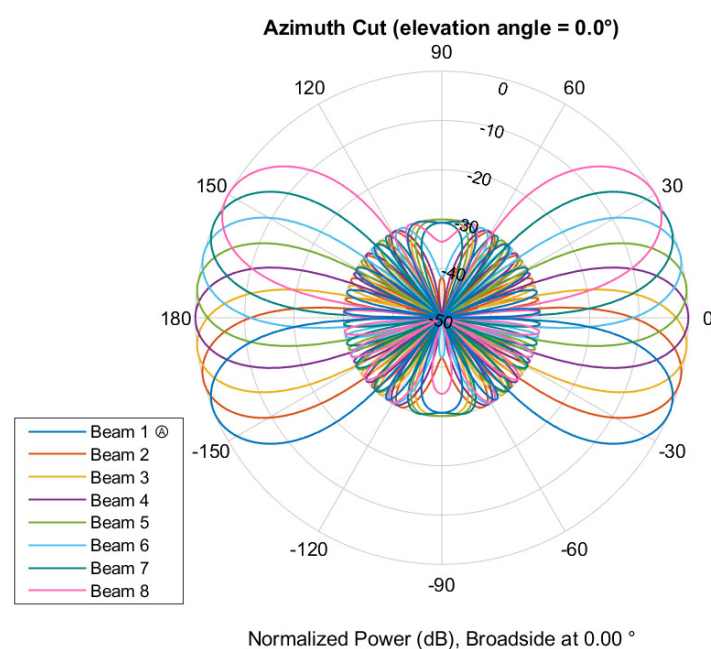


Figure 2. Beam pattern. Normalized azimuth radiation patterns of the eight SSB transmit beams at an elevation angle of 0° . Each coloured curve represents one beam steering direction.

Each SSB comprises the Primary Synchronization Signal (PSS), Secondary Synchronization Signal (SSS), Physical Broadcast Channel (PBCH), and the associated PBCH demodulation reference signals. The SS/PBCH block occupies 240 contiguous subcarriers in the frequency domain and four OFDM symbols in the time domain [19]. The total energy transmitted by a SSB is proportional to the received Synchronization Signal Reference Signal Received Power (SS-RSRP), scaled over time. In this work, we quantify the time-averaged EMF power density resulting from SSB transmissions based on physical-layer parameters, as derived from our prior measurement-based study [7]. The time averaged received power contribution associated with one SSB beam is modeled as follows:

$$P_{\text{SSB}} = \text{SS-RSRP} \cdot \frac{N_{\text{sym}} \cdot t_{\text{sym}}}{T_{\text{SSB}}} \quad (2)$$

where we have the following:

- SS-RSRP is the received power per SSB resource;
- N_{sym} is the number of OFDM symbols per SSB (typically 4);
- t_{sym} is the duration of one symbol (e.g., 33.3 μs at 30 kHz sub carrier spacing);
- T_{SSB} is the SSB transmission periodicity.

To capture multi-beam operation, the total average received power from all L beams is computed as follows:

$$\langle P_{\text{SSB}} \rangle = \sum_{i=1}^L \text{SS-RSRP}_i \cdot \frac{N_{\text{RE}} \cdot t_{\text{sym}}}{T_{\text{SSB}}} \quad (3)$$

where N_{RE} denotes the total number of resource elements occupied by the SSB block (e.g., 127 + 240 + 223 + 240 in different symbols), yielding a duration scaling factor of approximately 1.38.

In the physical layer simulation used in this work, this distance dependent propagation effect is represented through the SS-RSRP evaluated at the receiver location. The received SSB power is subsequently converted into power density using the effective aperture of the receiving antenna.

$$P_d = \frac{4\pi}{\lambda^2 G_r} \langle P_{\text{SSB}} \rangle \quad (4)$$

where P_d is the received power density, λ is the wavelength, and G_r is the receiver antenna gain. In our implementation, the EMF contribution of SSB transmissions is calculated via this model at a reference location (e.g., $R = 50$ m), using simulation-calibrated SS-RSRP values and antenna patterns. The factor $\lambda^2 G_r / (4\pi)$ represents the effective aperture of the receiving antenna. Therefore, the obtained simulated SS-RSRP values can be converted into incident power density at the receiver location.

3.2. Traffic-Derived UE Arrival Proxy

To enable anticipatory and EMF-aware control of SSB periodicity, the proposed framework requires an indicator of future user demand. Since the Milan Internet Traffic Dataset does not directly record UE cell-entry events, aggregate traffic measurements are transformed into a traffic-derived UE arrival proxy, which serves as the input demand signal for the prediction and control framework. This subsection outlines the data transformation pipeline.

The demand data are obtained from the “Telecommunications–SMS, Call, Internet–MI” dataset released by Telecom Italia through Harvard Dataverse [9,20]. The dataset contains spatially aggregated telecommunications activity recorded over Milan during November and December 2013. In this study, only the Internet-activity measure, aggregated over geographical grid squares and 10 min intervals, is used, which is resampled into 5 ms. The

dataset does not contain UE-level trajectories, physical cell-entry events, or random-access measurements. As direct UE entry information is unavailable, we estimate the number of users entering the coverage area based on the assumption that each user contributes a fixed average amount of traffic. Although this assumption simplifies the relationship between aggregate traffic volume and user arrivals, it provides a practical approximation for constructing a temporally varying demand signal from publicly available data.

Let $I(t)$ denote the aggregate Internet-usage value recorded during the 10 min interval indexed by t . The corresponding traffic-derived estimate of the number of active UEs is computed as follows:

$$\hat{N}_{\text{UE}}(t) = \left\lfloor \frac{I(t)}{\bar{I}_{\text{user}}} \right\rfloor \quad (5)$$

where $\bar{I}_{\text{user}}$ denotes the assumed average traffic contribution per user during one 10 min interval. In all experiments, the assumed average traffic contribution was set to $\bar{I}_{\text{user}} = 0.0325$ traffic units per user per one 5 ms resampled interval. This value is a modeling assumption used to scale the aggregated Internet-activity sequence and is not inferred from direct UE-level measurements.

To construct a proxy for new UE arrivals, the positive first-order difference between consecutive active-UE estimates is calculated as follows:

$$\Delta N_{\text{UE}}(t) = \max(N_{\text{UE}}(t) - N_{\text{UE}}(t - \Delta t), 0) \quad (6)$$

This operation captures the net increase in UE population, discarding departures or idle periods. To align with the control interval of the RL agent (10 min), we further apply a rolling sum over a 10 min window:

$$N_{\text{enter}}(t) = \sum_{k=t-10\text{min}}^t \Delta N_{\text{UE}}(k) \quad (7)$$

This transformation yields a smooth and high-resolution time series representing estimated UE ingress behavior in dense urban conditions. Since traffic demand varies across users, applications, and time periods, the derived arrival sequence is interpreted as an approximation. Nevertheless, the proposed framework is independent of the specific estimation method and can readily incorporate direct network measurements, such as active UE counters or random-access statistics, when available.

3.3. Access Delay

When a UE enters during epoch t , it must wait until the next SSB transmission configured by the period T_{SSB} . Let $u_{t,i} \in [0, \Delta)$ denote the i -th UE's intra-epoch entry offset (ms). The per-UE wait $w_{t,i}$ equals the time to the next SSB burst and is computed as

$$w_{t,i} = (T_{\text{SSB}} - (u_{t,i} \bmod T_{\text{SSB}})) \bmod T_{\text{SSB}} \quad (8)$$

The per-epoch total and average waits are

$$W_t = \sum_{i=1}^{N_{\text{enter}}(t)} w_{t,i} \quad (9)$$

$$\hat{w}_t = \frac{W_t}{\max(1, N_{\text{enter}}(t))} \quad (10)$$

where W_t denotes the total waiting time accumulated by all UEs that enter the coverage area during epoch t , and $\hat{w}_t$ represents the average per-UE waiting time in epoch t .

3.4. Optimization Problem Formulation

The objective of this work is to determine the optimal SSB periodicity T_{SSB} that balances user access performance with EMF exposure. On one hand, shorter T_{SSB} reduces average waiting time but increases background EMF radiation; on the other hand, longer T_{SSB} lowers EMF exposure but leads to larger access delays.

Accordingly, the problem can be expressed as the minimization of a weighted cost function.

To eliminate the dimensional and numerical-scale differences between the delay and EMF objectives, both quantities are normalized before being combined. The normalized EMF power density is defined as

$$\tilde{P}_d(t) = \frac{P_d(t)}{P_{d,5ms}} \quad (11)$$

where $P_{d,5ms}$ is the power density associated with the 5 ms SSB periodicity, which is the maximum power density among the considered SSB configurations.

The normalized total waiting time is defined as

$$\tilde{W}_t = \frac{W_t}{(160 \text{ ms})N_{\max}} \quad (12)$$

where $N_{\max}$ is a fixed normalization constant equal to the maximum traffic-derived UE arrival count observed in one 10 min interval of the dataset, i.e., $N_{\max} = \max_t N_{\text{enter}}(t)$. Since each UE wait is bounded by the maximum considered SSB periodicity of 160 ms, $(160 \text{ ms})N_{\max}$ represents the maximum aggregate waiting time used for normalization. Consequently, $\tilde{P}_d(t)$ and $\tilde{W}_t$ are dimensionless and bounded within $[0, 1]$.

$$\begin{aligned} \min_{\pi} \quad & \mathbb{E} \left[w_a(1 - \alpha)\tilde{P}_d(t) + w_b\alpha\tilde{W}_t \right] \\ \text{s.t.} \quad & T_{SSB} \in \{5, 10, 20, 40, 80, 160\} \text{ ms} \\ & \tilde{W}_t \geq 0, \quad \tilde{P}_d(t) \geq 0, \end{aligned} \quad (13)$$

where $\tilde{P}_d(t)$ and $\tilde{W}_t$ denote the normalized EMF power density and normalized total waiting time, respectively. The parameter $\alpha \in [0, 1]$ controls the principal trade-off between the two objectives: increasing α places greater emphasis on reducing user waiting time, whereas decreasing α places greater emphasis on reducing EMF exposure. The coefficients w_a and w_b are fixed objective weights, and π denotes the SSB periodicity selection policy.

This formulation captures the inherent trade-off in adaptive SSB configuration: minimizing EMF exposure requires infrequent SSB transmissions, while minimizing user access delay requires frequent transmissions. As a result, static fixed periodicity configurations cannot adapt their operating point to temporal changes in the demand proxy.

4. EMF-Aware Adaptive SSB Periodicity Control Framework

In this section, we present the proposed EMF-aware adaptive SSB periodicity control framework that integrates traffic prediction and RL into a unified decision-making pipeline. Building upon the system model and problem formulation described in Section 3, our approach is designed to proactively adjust the SSB transmission periodicity according to predicted user arrival patterns while minimizing both EMF exposure and access delay. The framework consists of three tightly coupled components: (i) a LSTM-based predictive module that forecasts estimated UE arrival proxy from real-world urban traffic data; (ii) a custom reinforcement learning environment that models the interaction between SSB periodicity, EMF power density, and user access delay; and (iii) a PPO-based decision agent that learns an optimal traffic-aware SSB selection strategy. The workflow, shown in

Figure 3, provides an end-to-end process that enables the gNB to dynamically adjust its SSB configuration with foresight rather than reactively.

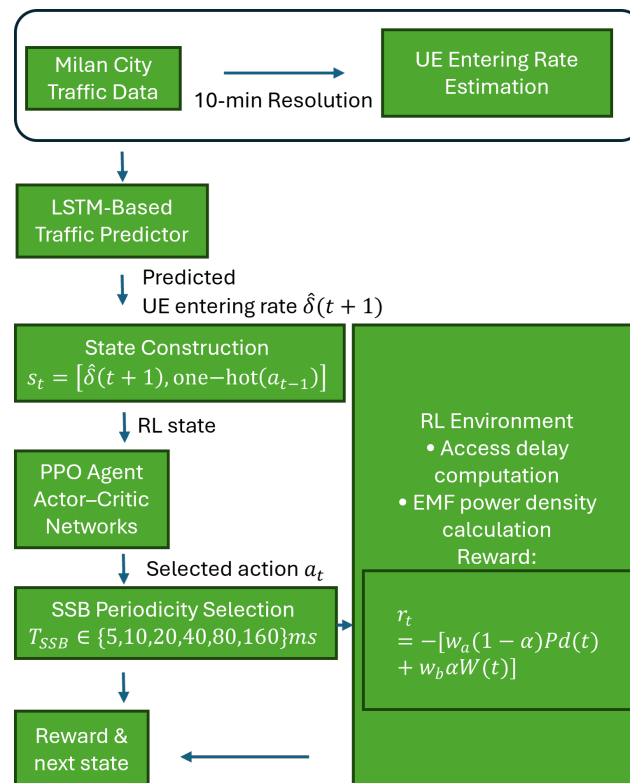


Figure 3. Overall workflow of the proposed adaptive SSB periodicity control framework. Aggregated Internet usage data are transformed into a traffic-derived UE arrival proxy and forecast using a LSTM predictor. The predicted demand proxy is used by a PPO agent operating within a custom RL environment to select a SSB periodicity that balances EMF exposure and user access latency.

4.1. System Workflow Overview

The proposed framework operates by tightly coupling traffic prediction, EMF calculation, and RL to achieve proactive and adaptive T_{SSB} control. Figure 3 illustrates the end-to-end workflow. The process begins with aggregated urban Internet-usage traces, which are transformed into a traffic-derived UE arrival proxy and forecast using a LSTM-based predictor. These predicted values provide a forward-looking estimate of the traffic-derived demand level for the next control interval.

The predicted entering rate, together with the previous SSB configuration, forms the input state to a custom RL environment. This environment computes two key performance indicators for each candidate T_{SSB} : (i) the EMF power density generated by SSB transmissions and (ii) the UE access delay induced by the periodicity choice. Based on this state and the evaluated reward, a PPO agent selects the SSB periodicity to be applied for the next interval.

During PPO training, the environment computes the resulting reward and records the state transition for subsequent policy updates. During final evaluation, the trained PPO parameters remain fixed, and the reward is recorded only for performance assessment. The environment computes the resulting reward, closes the feedback loop, and updates the PPO policy parameters. By iterating through this predictive and adaptive loop, the system learns a traffic-aware SSB control policy that minimizes EMF exposure while maintaining acceptable user access performance.

4.2. LSTM-Based User-Entry Prediction Model

Forecasting future variations in the traffic-derived UE arrival proxy enables proactive SSB periodicity control. Rather than relying solely on instantaneous traffic observations, the proposed system employs a LSTM recurrent neural network, originally introduced by Hochreiter and Schmidhuber [8], to capture temporal dependencies and nonlinear fluctuations in the traffic-derived demand proxy. The predictor operates on aggregated Internet-usage traces from the Milan dataset, which are transformed into traffic-derived UE arrival-proxy sequences as described in Section 3.2.

To enhance temporal feature richness, the predictor augments each traffic observation with engineered calendar and cyclical features: hour of day, day of week, time of day sine and cosine encodings, weekend indicators, and business hour flags. These features enable the model to capture both short-term variations and longer-term temporal patterns, particularly the daily cycles present in urban telecommunications activity.

The supervised learning dataset is constructed using a sliding-window mechanism. Given a lookback window of $L = 144$ time steps (one day of 10 min intervals), the input tensor for each sample is as follows:

$$\mathbf{X}_t = [\mathbf{f}_{t-L}, \mathbf{f}_{t-L+1}, \dots, \mathbf{f}_{t-1}] \quad (14)$$

where $\mathbf{f}_k$ contains both data traffic values and deterministic time features at index k . The target variable is the UE arrival proxy at the next interval:

$$y_t = \delta(t) \quad (15)$$

The LSTM predictor comprises two recurrent layers containing 64 and 32 hidden units, respectively. The first LSTM layer returns the complete sequence to the second recurrent layer. A dropout rate of 0.2 is applied after each LSTM layer, followed by a fully connected layer with 16 rectified linear unit neurons and a single linear regression output:

$$\hat{\delta}(t+1) = f_{\text{LSTM}}(\mathbf{X}_t; \Theta) \quad (16)$$

where Θ denotes the set of trainable network parameters. The model is optimized using the Adam optimizer with mean-squared error (MSE) loss. Early stopping is used to prevent overfitting. The training and validation sequences are both drawn chronologically from the six-week development partition, with the final portion of this partition reserved for validation. The held-out seventh week is not used for model fitting, early stopping, or hyperparameter selection.

For final evaluation, a walk-forward inference strategy is adopted to preserve temporal causality. At each interval of the held-out test week, the predictor uses only the most recent observations available up to that interval to generate the next-step forecast. Future observations from the test week are never included in the input window. The LSTM parameters remain fixed throughout this evaluation period.

The final output of this module is the predicted UE arrival proxy $\hat{\delta}(t+1)$, which is passed directly into the RL environment as one of the primary components of the state representation. By providing forecasts of the traffic-derived demand proxy, the LSTM predictor enables the PPO agent to anticipate changes in the input demand signal, reduce unnecessary SSB transmissions during low-proxy periods, and adapt the SSB periodicity according to the predicted operating condition.

4.3. Integration of LSTM Prediction Within the RL Environment

The predicted traffic-derived UE arrival proxy generated by the LSTM module serves as a key input to the reinforcement learning environment, enabling forecast-driven rather than purely reactive T_{SSB} control. Unlike traditional SSB configuration schemes that rely solely on instantaneous load measurements, the proposed framework incorporates future traffic estimates to anticipate both high-load and low-load conditions. This predictive capability allows the controller to reduce unnecessary SSB transmissions during low proxy intervals while selecting shorter periodicities when the predicted demand proxy increases.

In each control interval t , the LSTM predictor outputs the one-step-ahead estimate $\hat{\delta}(t+1)$ of the UE arrival rate. This predicted value is combined with the previous SSB periodicity to form the RL state vector:

$$s_t = [\hat{\delta}(t+1), \text{one-hot}(a_{t-1})] \quad (17)$$

where a_{t-1} denotes the T_{SSB} selected during the previous interval. The RL environment then uses this state to simulate the resulting EMF exposure, compute the per-UE and total access delay, and evaluate the reward corresponding to each candidate action.

By incorporating the predicted UE arrival proxy directly into the state, the RL agent is able to learn traffic-aware behaviors. During low-load intervals, the agent learns to increase the T_{SSB} to reduce background EMF exposure, while in high-load intervals it adapts by selecting shorter periodicities to maintain acceptable access delay. This predictive RL formulation offers a significant advantage over purely reactive schemes, which can only respond to traffic changes after they occur.

The LSTM and PPO components are trained and evaluated using strictly separated chronological partitions. During PPO training, the environment uses only traffic-derived arrival sequences and corresponding forecasts from the six-week development period. During final evaluation, the trained LSTM and PPO parameters are frozen, and the framework processes the held-out seventh week sequentially. At each test interval, the PPO agent receives forecasts generated from information available up to that time. This procedure prevents observations from later test intervals from influencing earlier predictions or control decisions.

Through this tight coupling between prediction and control, the proposed system achieves forward-looking SSB adaptation that reduces EMF exposure, improves traffic responsiveness, and maintains stable communication performance across a wide range of urban traffic conditions.

4.4. Reinforcement Learning Environment Design

To enable the PPO agent to learn an adaptive SSB periodicity policy, we design a custom RL environment that accurately emulates the interaction between T_{SSB} , P_d , and W_t . The environment operates on 10 min control intervals and is implemented as a Gym-compatible environment, allowing seamless integration with the PPO training pipeline. Its design follows the access-delay model and EMF formulation introduced in Section 3.

4.4.1. State Representation

At the beginning of each interval t , the environment constructs the RL state vector:

$$s_t = [\hat{\delta}(t+1), \text{one-hot}(a_{t-1})] \quad (18)$$

where $\hat{\delta}(t+1)$ is the predicted UE arrival proxy obtained from the LSTM predictor, and a_{t-1} encodes the previously applied SSB periodicity. This representation captures both future

traffic intensity and recent control actions, enabling the PPO agent to learn temporal consistency and switching behavior.

4.4.2. Action Space

The action space is discrete and corresponds exactly to the SSB transmission periodicities defined in 3GPP NR:

$$\mathcal{A} = \{5, 10, 20, 40, 80, 160\} \text{ ms} \quad (19)$$

Each action represents the SSB periodicity selected by the controller for the subsequent 10 min control interval. The environment assumes that the selected configuration becomes effective at the beginning of that interval. This abstraction isolates the effect of T_{SSB} on EMF exposure and access delay and does not model the signaling procedure required to apply the configuration in a deployed network.

The action applied in the environment represents a control decision rather than an instantaneous physical layer switch in an operational network. In practice, changes to the serving-cell SSB periodicity would require coordination with the applicable RRC configuration and may interact with system information scheduling, measurement configuration, and random-access procedures. Such changes would therefore be applied on a slower, semi-static timescale and would be subject to reconfiguration latency, UE compatibility, and vendor-specific implementation constraints. The 10 min control interval used in this study is intended to represent a supervisory network-management timescale over which such reconfiguration could be completed; the signaling exchange itself is not modeled [12].

4.4.3. Environment Dynamics

Upon receiving action a_t , the environment retrieves all UE entry timestamps within the current window and computes the per-UE wait times based on the SSB-access model:

$$w_{t,i} = (T_{\text{SSB}} - (u_{t,i} \bmod T_{\text{SSB}})) \bmod T_{\text{SSB}} \quad (20)$$

where $u_{t,i}$ is the intra-interval entry offset. The total waiting time and the average per-UE waiting time are subsequently obtained from (9) and (10), respectively. Simultaneously, the P_d associated with the selected periodicity is derived from the SSB power model introduced earlier. In this manner, the environment produces two performance metrics for each interval: W_t and P_d .

4.4.4. Reward Function

The environment returns a scalar reward that encodes the EMF-latency trade-off. Since the total waiting time and EMF power density have different physical units and numerical scales, the normalized quantities defined in (11) and (12) are used in the reward calculation.

$$r_t = - \left[w_a(1 - \alpha)\tilde{P}_d(t) + w_b\alpha\tilde{W}_t \right] \quad (21)$$

where $\alpha \in [0, 1]$ controls the principal trade-off between user waiting time and EMF exposure. Increasing α assigns greater importance to reducing aggregate waiting time, whereas decreasing α places greater emphasis on reducing SSB-induced EMF power density. The coefficients w_a and w_b are fixed scaling factors used in the implemented reward function to adjust the relative contributions of the two normalized objectives. Because $\tilde{P}_d(t)$ and $\tilde{W}_t$ are dimensionless and normalized to comparable ranges, these parameters express the intended optimization preference rather than compensating for differences in physical units.

4.5. Deep Reinforcement Learning Model

In this section, we present the deep reinforcement learning (DRL) framework used to learn an adaptive SSB periodicity policy that jointly minimizes EMF exposure and access delay. Unlike value-based methods such as Q-learning or Double Deep Q-Network (DDQN), the proposed framework employs PPO, an on-policy policy-gradient algorithm that constrains successive policy updates through a clipped surrogate objective [21]. This makes PPO well suited for traffic-dependent SSB control, where the agent must learn a stochastic, load-aware policy based on fast-varying user arrival patterns.

4.5.1. RL Framework

The adaptive T_{SSB} selection problem is formulated as a sequential decision-making task in which an agent interacts with the 5G cell environment and learns a policy that balances EMF exposure and user access latency. At every 10 min interval t , the environment provides the agent with a state vector s_t that encapsulates the predicted user entering rate and the most recent SSB configuration. Based on this observation, the agent selects a SSB periodicity $a_t \in \mathcal{A}$, where $\mathcal{A} = \{5, 10, 20, 40, 80, 160\}$ ms. Within the simulation environment, the selected periodicity is applied to the downlink SSB transmission structure for the entire 10 min control interval.

After the action is executed, the environment transitions to a new state s_{t+1} and computes a scalar reward r_{t+1} , which combines the normalized EMF power density and normalized aggregate waiting time produced during the current control interval. This interaction defines a Markov decision process (MDP) characterized by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$, where $\mathcal{S}$ is the continuous state space, $\mathcal{A}$ is the discrete action set of candidate SSB periodicities, $\mathcal{P}$ governs the transition dynamics imposed by real user-arrival patterns, and $\mathcal{R}$ encodes the EMF–latency trade-off.

The goal of the RL agent is to learn a stochastic policy $\pi_\theta(a|s)$, parameterized by θ , that maximizes the expected discounted return

$$J(\theta) = \mathbb{E}_{\pi_\theta} \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \right] \quad (22)$$

where $0 < \gamma \leq 1$ is the discount factor. A stochastic policy is essential for SSB control, as optimal periodicity depends on time-varying traffic conditions in which multiple actions may yield comparable reward. Deterministic or greedy decision rules often fail in such non-stationary environments, particularly when the entering rate follows strong daily cycles.

PPO restricts harmful policy deviations while still enabling sufficiently large improvements during training. This robustness is critical for the SSB control problem, where the agent must learn consistent long-term behavior across thousands of traffic cycles while adapting smoothly to abrupt fluctuations in user arrival rates.

4.5.2. State, Action, and Reward Design

The performance of the PPO agent largely depends on the quality of the state representation, the action granularity, and the reward shaping strategy. In the proposed framework, these components are designed to enable the agent to balance EMF exposure and user access delay while adapting to the highly non-stationary entering patterns observed throughout the day.

(1) State Space: Each state s_t is represented by a 7-dimensional feature vector:

$$s_t = [\hat{\delta}(t+1), \text{one-hot}(a_{t-1})] \quad (23)$$

The first component represents the LSTM-predicted UE arrival proxy for the next control interval, while the one-hot encoding of the previous action a_{t-1} provides the policy with context regarding the recent SSB periodicity configuration. This compact state representation captures both anticipated short-term traffic load and the temporal dependence induced by previous configuration choices.

(2) Action Space: The agent selects one of six discrete SSB periodicities:

$$\mathcal{A} = \{5, 10, 20, 40, 80, 160\} \text{ ms} \quad (24)$$

which correspond to the valid periodicities defined in 3GPP NR specifications. Short periodicities (e.g., 5–20 ms) improve initial-access responsiveness but increase EMF power density due to more frequent SSB transmissions. Conversely, longer periodicities (e.g., 80–160 ms) reduce EMF exposure but increase connection waiting time. The agent must therefore learn a policy that selects actions appropriate for the current traffic level while minimizing unnecessary oscillations.

4.5.3. PPO Actor–Critic Architecture

PPO is well suited for this control problem as it ensures stable policy updates through a clipped surrogate objective while retaining sufficient flexibility to explore alternative SSB configurations during fluctuating load conditions.

Figure 4 illustrates the two-network design used in the proposed system. Both the actor and critic networks adopt a multilayer perceptron (MLP) architecture consisting of two fully connected hidden layers, each with 64 units and rectified linear unit (ReLU) activation. This structure strikes an effective balance between model expressiveness and computational efficiency, enabling the policy to operate in real time.

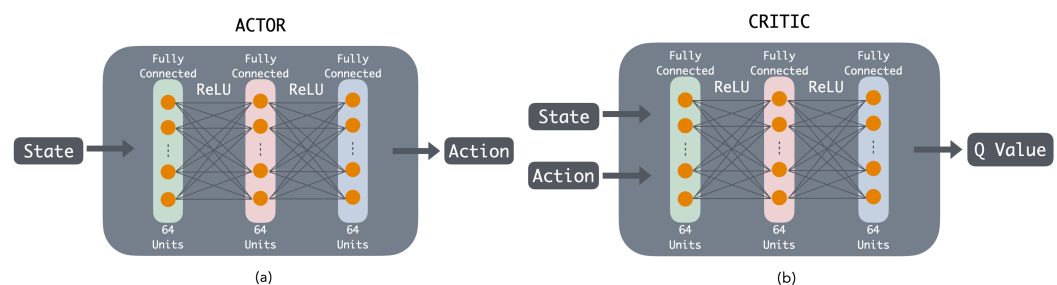


Figure 4. Illustration of the PPO actor–critic architecture used for adaptive SSB periodicity control. (a) The actor network receives the system state and outputs a categorical distribution over the six SSB actions. (b) The critic network estimates the state value $V(s)$ that guides advantage estimation during training. Both networks use two fully connected hidden layers with 64 ReLU units each.

(1) Actor Network: As shown in Figure 4a, the actor maps the input state s_t to a categorical distribution over the six candidate SSB periodicities. Formally, the actor computes the following:

$$\pi_{\theta}(a|s) = \text{softmax}(W_{\pi}h + b_{\pi}) \quad (25)$$

where W_{π} and b_{π} denote the weight matrix and bias vector of the actor output layer, respectively, and h denotes the shared hidden representation learned by the network. The stochastic nature of the policy π_{θ} enables smooth adaptation to changing traffic conditions, allowing the agent to explore longer or shorter SSB periodicities depending on environmental uncertainty.

Serving as a low-variance baseline, the actor parameters are updated by maximizing the clipped surrogate objective

$$L^{\text{CLIP}}(\theta) = \mathbb{E}[\min(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t)] \quad (27)$$

where $r_t(\theta)$ denotes the probability ratio between the new and old policies [21]:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} \quad (28)$$

The clipping parameter ϵ prevents excessively large policy updates, thereby stabilizing training and avoiding performance collapse.

In parallel, the critic network is updated by minimizing the squared error between its predicted value $V_\mu(s_t)$ and the Monte Carlo return R_t :

$$L^V(\mu) = \mathbb{E}[(V_\mu(s_t) - R_t)^2] \quad (29)$$

where the return is defined as

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad (30)$$

The value-loss gradient stabilizes advantage estimation and ensures reliable learning signals for the actor update.

During training, the actor and critic parameters are optimized using stochastic gradient descent over several epochs per rollout. Entropy regularization is added to encourage exploration, particularly in low-load regions where multiple SSB periodicities may yield comparable reward. This training regimen enables PPO to discover a robust SSB periodicity policy that remains effective under strongly time-varying user arrival patterns.

The characteristics of the SSB periodicity control problem make PPO particularly well suited compared to value-based methods such as Q-learning or DDQN. First, the traffic load and user arrival patterns exhibit strong non-stationarity due to daily and weekly cycles. This requires a learning algorithm capable of adapting smoothly to evolving data distributions. The stochastic policy $\pi_\theta(a|s)$ maintained by PPO enables the agent to adjust its SSB selection strategy continuously without relying on brittle greedy decisions.

Second, SSB periodicity selection involves a discrete but highly nonlinear action space, where different SSB periods produce complex trade-offs between EMF power density and user access delay. PPO's clipped surrogate objective allows the policy to improve steadily even when reward gradients are noisy or when multiple actions yield similar outcomes. This stability is critical for avoiding oscillatory behavior, such as rapid switching between short and long periodicities, which could degrade system performance.

Third, the actor–critic structure of PPO enables efficient variance reduction through advantage estimation, allowing the agent to learn from sparse or noisy reward signals. This is especially beneficial in EMF-aware optimization, where the immediate EMF level may not fully capture the long-term effect of sustained SSB configurations. The critic provides a learned baseline that guides the actor toward policies that reduce long-term EMF exposure while maintaining acceptable access latency.

Finally, unlike off-policy value-based algorithms that rely on replay buffers, PPO operates on-policy and is therefore naturally synchronized with the LSTM-based traffic predictor. The predictor and PPO agent evolve together, allowing the system to incorporate real-time forecasting and adapt transmission periodicity proactively. This synergy results in a robust and responsive control mechanism capable of delivering lower EMF power density without sacrificing user experience.

4.6. End-to-End Forecast-in-the-Loop SSB Control

The components described in the preceding subsections operate jointly to provide a real-time adaptive SSB periodicity control mechanism. In this final step, we integrate the LSTM-based traffic predictor, the RL environment, and the PPO decision agent into a unified forecast-in-the-loop control process. At each 10 min interval, the system first forecasts the upcoming UE arrival proxy, constructs the RL state, and selects T_{SSB} based on the learned policy. The selected periodicity determines the access delay experienced by arriving users and the resulting EMF power density. These measurements are combined into a reward that guides policy improvement during training. Algorithm 1 summarizes the complete operational pipeline, showing how prediction and control are tightly coupled to enable proactive, traffic-aware, and EMF-efficient SSB configuration.

Algorithm 1 Forecast-in-the-loop adaptive SSB periodicity control

```

1: Input: Trained LSTM predictor  $f_{LSTM}$ , trained PPO policy  $\pi_\theta$ , maximum traffic-derived
   arrival count  $N_{max}$ , and reference power density  $P_{d,5ms}$ 
2: Output: Adaptive SSB periodicity decisions  $T_{SSB}(t)$ 
3: for each control interval  $t$  do
4:   Observe historical traffic features  $\mathbf{X}_t$ 
5:   Predict the next-step traffic-derived UE arrival estimate:
      $\hat{\delta}(t+1) = f_{LSTM}(\mathbf{X}_t)$ 
6:   Construct the RL state:
      $s_t = [\hat{\delta}(t+1), \text{one-hot}(a_{t-1})]$ 
7:   Select an action:
      $a_t \sim \pi_\theta(\cdot | s_t)$ 
8:   Map the selected action to a SSB periodicity:
      $T_{SSB}(t) \in \{5, 10, 20, 40, 80, 160\} \text{ ms}$ 
9:   Apply  $T_{SSB}(t)$  during control interval  $t$ 
10:  Compute the waiting time of each arriving UE:
      $w_{t,i} = (T_{SSB}(t) - (u_{t,i} \bmod T_{SSB}(t))) \bmod T_{SSB}(t)$ 
11:  Compute the total waiting time:
      $W_t = \sum_{i=1}^{N_{enter}(t)} w_{t,i}$ 
12:  Compute the EMF power density  $P_d(t)$  associated with  $T_{SSB}(t)$ 
13:  Normalize the total waiting time:
      $\tilde{W}_t = \frac{W_t}{(160 \text{ ms})N_{max}}$ 
14:  Normalize the EMF power density:
      $\tilde{P}_d(t) = \frac{P_d(t)}{P_{d,5ms}}$ 
15:  Compute the reward:
      $r_t = -[w_a(1 - \alpha)\tilde{P}_d(t) + w_b\alpha\tilde{W}_t]$ 
16:  if the framework is operating in training mode then
17:    Observe the next state  $s_{t+1}$ 
18:    Store the transition  $(s_t, a_t, r_t, s_{t+1})$ 
19:  end if
20: end for
21: return  $\{T_{SSB}(t)\}$ 

```

5. Results and Evaluation

This section presents a comprehensive performance evaluation of the proposed forecast-in-the-loop adaptive SSB periodicity control framework. The objective is to examine how effectively the system balances EMF exposure and user access delay when operating under realistic, time-varying traffic conditions derived from the Milan dataset.

To this end, we assess both the individual components of the architecture and the end-to-end behavior of the integrated LSTM–RL pipeline. The evaluation comprises five parts: (i) characterization of the EMF power density associated with different SSB transmission periodicities; (ii) performance analysis of the LSTM-based user-entry prediction model; (iii) training dynamics and policy behavior of the PPO agent; (iv) an investigation of the EMF–latency trade-off emerging from adaptive SSB control; and (v) benchmarking of the proposed method against static SSB configurations, a myopic rate-aware controller, and an oracle upper bound with perfect future information. Together, these results demonstrate the benefits of incorporating traffic forecasting and reinforcement learning into EMF-aware network control.

5.1. Experimental Setup

The development data include all observations from 4 November to 15 December 2013, while the held-out test period extends from 16 December to 22 December 2013. Each observation represents one 10 min interval. The LSTM input window contains 144 consecutive observations, corresponding to the preceding 24 h. In addition to the Internet-activity value, the predictor uses sine and cosine encodings of hour of day, minute of hour, and day of week.

The predictor comprises two LSTM layers with 64 and 32 hidden units, respectively. A dropout rate of 0.2 is applied after each recurrent layer, followed by a fully connected layer containing 16 ReLU neurons and a single linear output. The model is trained using the Adam optimizer with mean-squared error loss, a batch size of 32, and a maximum of 50 epochs. Early stopping is applied with a patience of 10 epochs while restoring the weights associated with the lowest validation loss. The NumPy and TensorFlow random seeds are both set to 42.

All input variables are transformed to the interval $[0, 1]$ using a Min–Max scaler fitted exclusively to the development data. Prediction over the test week is recursive: the first forecast uses the final 144 development observations, and each subsequent forecast uses the previously predicted Internet activity value together with the calendar features of the next timestamp. The LSTM therefore predicts aggregated Internet activity, after which the predicted sequence is transformed into the traffic-derived arrival proxy used by the PPO controller.

The reinforcement learning environment runs with the same 10 min control granularity as the dataset. At each control interval, the environment receives (i) the traffic-derived UE Arrival Proxy expected to enter the cell in the next window and (ii) a one-hot encoded representation of the SSB periodicity applied in the current interval. These two quantities jointly form the observation that is fed to the PPO agent. Based on this observation, the environment computes the resulting UE access delay using the analytical model and evaluates the induced EMF power density using the physical-layer formulation described in Section 3. The agent then selects a SSB periodicity from the discrete action set $\{5, 10, 20, 40, 80, 160\}$ ms to be applied for the subsequent interval.

The PPO agent is implemented using the Stable-Baselines3 reinforcement-learning library [22]. The policy adopts a multilayer perceptron actor–critic architecture with two fully connected hidden layers, each containing 64 ReLU neurons. The discount factor is set to $\gamma = 0.99$, and generalized advantage estimation is applied using $\lambda_{\text{GAE}} = 0.95$. Policy updates employ a clipping threshold $\epsilon = 0.20$ and are optimized using the Adam optimizer with a learning rate of 5×10^{-5} . Each training run spans 6480 interaction steps, corresponding to several full traversals of the diurnal traffic pattern and ensuring exposure to a wide range of load conditions. For all experiments, the reward parameters were fixed at $w_a = 5$, $w_b = 1$, and $\alpha = 0.4$. Here, w_a scales the normalized SSB-induced power-density

term, w_b scales the normalized aggregate waiting-time term, and α controls the relative emphasis placed on reducing waiting time.

SSB-induced EMF power density is computed according to Equations (2)–(4), using calibrated SS-RSRP values from our prior measurement study [7] and a reference observation point located 50 m from the transmitting antenna. This configuration is representative of typical public-exposure distances in dense urban deployments and provides a fair and consistent basis for comparing all baseline and RL-driven SSB periodicity strategies.

All simulations, including LSTM training, walk-forward prediction, environment rollout, and PPO optimization, are executed entirely on an NVIDIA RTX-3080 GPU, which accelerates both the neural-network prediction tasks and the RL training loop.

The evaluation compares five controller configurations: fixed SSB periodicities of 20, 40, and 80 ms; the proposed prediction-driven controller; and an observed-proxy controller. The prediction-driven controller supplies the LSTM-predicted traffic-derived arrival proxy to the trained PPO policy, whereas the observed-proxy controller supplies the corresponding observed proxy value to the same policy. The observed-proxy controller is included as a reference for examining the effect of prediction.

5.2. Characterization of SSB-Induced EMF Power Density

We begin by examining the EMF exposure generated by SSB transmissions under different periodicity configurations. Understanding this relationship is essential, as SSB periodicity determines the duty cycle of broadcast signals and therefore directly influences background exposure levels experienced by nearby users.

The time-averaged power density associated with a given SSB period is computed using the physical-layer model, which incorporates calibrated SS-RSRP measurements and the beam-sweeping structure outlined in Section 3. For each valid SSB periodicity $T_{SSB} \in \{5, 10, 20, 40, 80, 160\}$ ms, we calculate the corresponding time-averaged power density from the simulated SS-RSRP at a reference receiver location 50 m from the gNB antenna. The same receiver location, antenna configuration, and propagation conditions are used for all SSB periodicities.

Figure 6 illustrates the sharp inverse relationship between T_{SSB} and the resulting EMF power density. Shorter transmission intervals (5–10 ms) lead to higher time-averaged power density because the SSB bursts are repeated more frequently. Increasing the periodicity from 5 ms to 160 ms reduces the calculated SSB-induced power density from 4.23×10^{-4} W/m² to 1.3×10^{-5} W/m², corresponding to an approximately 32-fold reduction under the same simulated receiver conditions. These numerical values are summarized in Table 1, which confirms that SSB periodicity is a dominant factor governing baseline downlink EMF levels.

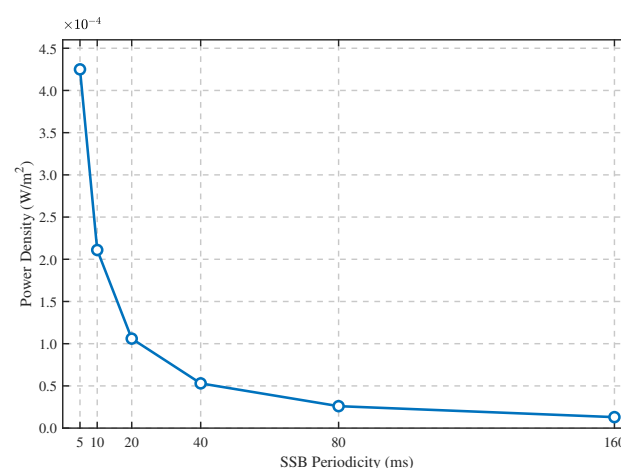


Figure 6. Relationship between SSB periodicity and power density.

Table 1. Time-averaged EMF power density at 50 m for various SSB periodicities.

SSB Periodicity (ms)	Power Density (W/m ²)
5	0.000423
10	0.000211
20	0.000106
40	0.000053
80	0.000026
160	0.000013

5.3. LSTM Internet Traffic Prediction

Accurate Internet traffic forecasting provides the demand information required by the proposed controller. Figure 7 compares the one-step-ahead LSTM predictions with the observed Internet traffic during the first day of the held-out evaluation week. At this stage, the LSTM predicts the aggregated Internet traffic sequence rather than the number of arriving UEs. The predicted traffic is subsequently converted into a traffic-derived UE arrival proxy, which is supplied to the controller for adaptive SSB periodicity selection. Therefore, the resulting UE arrival proxy represents an estimate derived from network traffic and should not be interpreted as a direct measurement of physical UE entries.

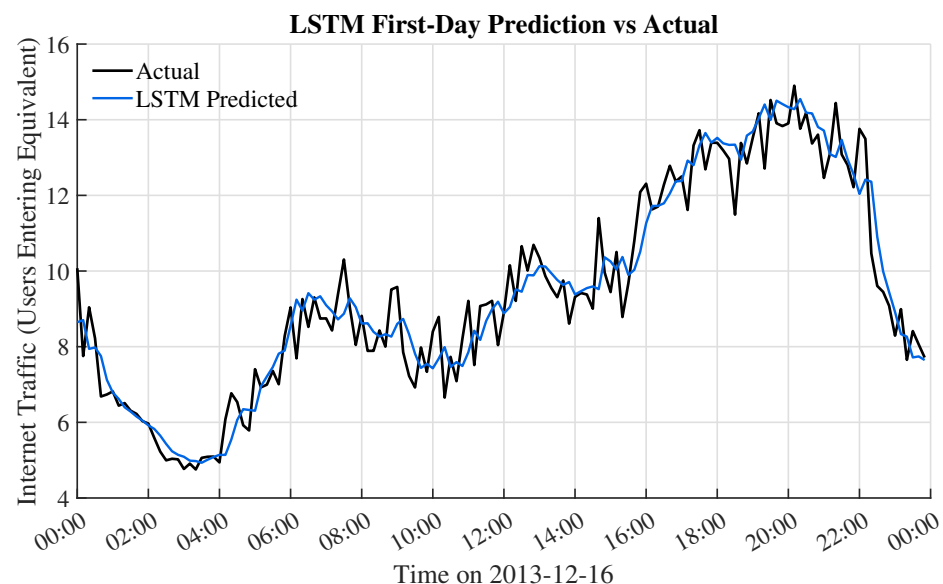
**Figure 7.** LSTM prediction of internet traffic during the first day of the evaluation week.

Table 2 reports the mean absolute error (MAE), mean squared error (MSE), and root-mean-squared error (RMSE) of the Internet traffic predictions for each day of the seven-day evaluation period. The overall MAE, MSE, and RMSE are 0.7168, 0.8748, and 0.9353, respectively. These results quantify the one-step-ahead forecasting accuracy of the LSTM model before the predicted traffic is converted into the UE arrival proxy used by the proposed controller.

Table 2. Daily internet traffic prediction errors of the LSTM model over the evaluation week.

Metric	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Overall
MAE	0.5840	0.7839	0.6856	0.6196	0.7667	0.7481	0.8297	0.7168
MSE	0.5685	1.0083	0.8199	0.6487	0.9758	0.9213	1.1809	0.8748
RMSE	0.7540	1.0041	0.9055	0.8054	0.9878	0.9599	1.0867	0.9353

From a system perspective, the most important observation is that the LSTM forecasts are sufficiently accurate for downstream reinforcement learning. Since the RL policy relies only on short-term load trends rather than exact point predictions, small forecasting deviations do not degrade performance. This robustness is confirmed in later sections, where the RL agent driven by predicted arrivals achieves nearly identical behavior to the agent driven by traffic-derived arrivals.

5.4. RL Training Behavior and Learned Policy Characteristics

We now examine the learning dynamics of the PPO agent and the structure of the resulting adaptive SSB periodicity policy. Throughout training, the agent interacts with the custom environment described in Section 4, receiving predicted user-entry rates, selecting a SSB periodicity, and observing the resulting wait time and EMF penalties. The goal is to identify a policy that delivers low EMF exposure during light-load periods while ensuring minimal access delay under high traffic.

Figure 8 presents the evolution of the average episodic reward over 6480 training steps. The reward increases rapidly during the first few episodes, indicating fast discovery of high-performing periodicity patterns. Beyond approximately six episodes, the learning curve stabilizes with small fluctuations, suggesting convergence to a consistent policy.

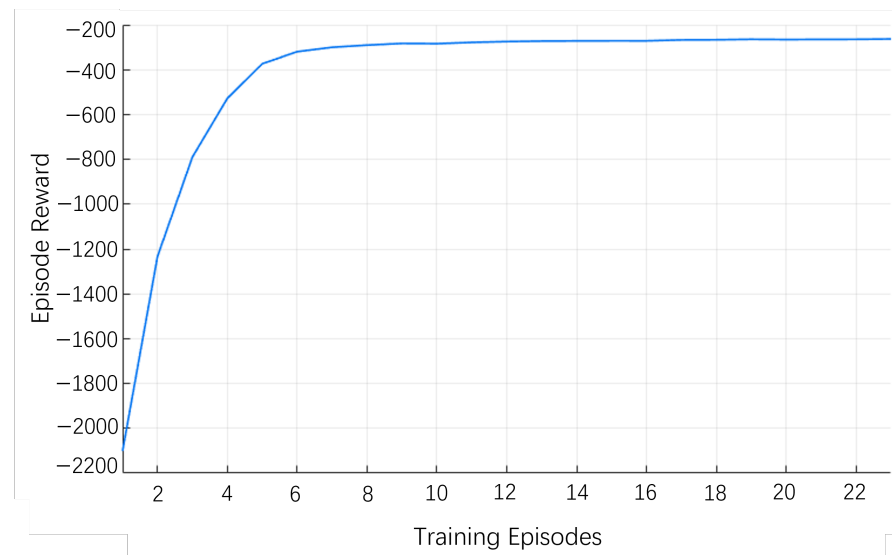


Figure 8. Average reward over training episodes.

To gain insight into the structure of the policy itself, we analyze the frequency distribution of SSB periodicities selected during evaluation on the test week. As shown in Figure 9, the learned policy exhibits a bimodal behavior: short periodicities (e.g., 20 ms) are selected during high-traffic intervals to minimize UE access delay, while long periodicities (e.g., 160 ms) dominate during low-load intervals to reduce unnecessary EMF exposure. Intermediate periodicities appear less frequently, indicating that the PPO agent has learned to exploit the extremes of the design space rather than oscillating among similar configurations.

To further quantify the relationship between predicted load and chosen actions, Figure 10 shows a two-dimensional heat map mapping the predicted user-entry rate to the probability of selecting each SSB periodicity. Short intervals are preferred when the predicted user-entry count exceeds the 95th percentile of historical values, while long intervals dominate in the lowest quartile. This monotonic load–action alignment demonstrates that the learned policy is context-aware and responds adaptively to traffic variations rather than resorting to a static pattern.

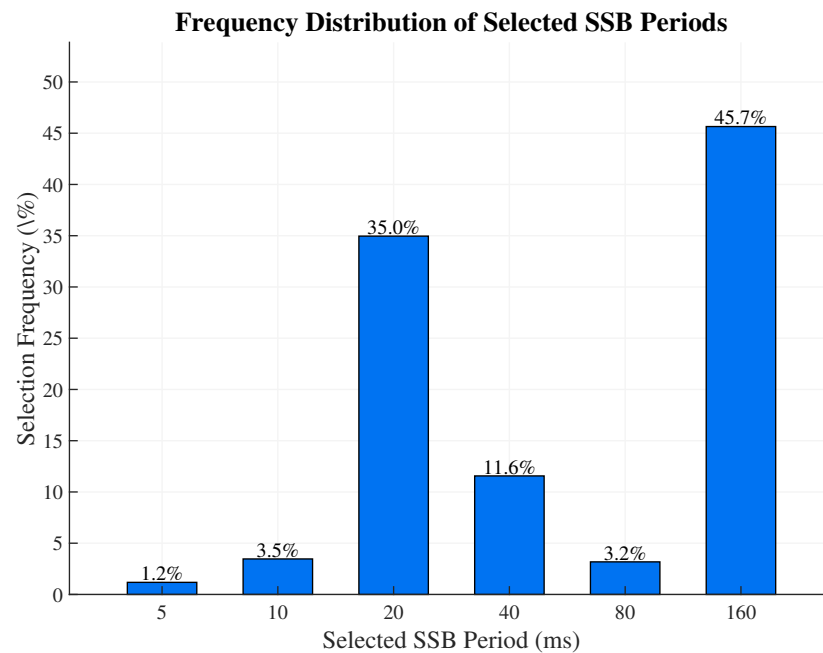


Figure 9. Frequency distribution of the selected SSB periods during deployment.

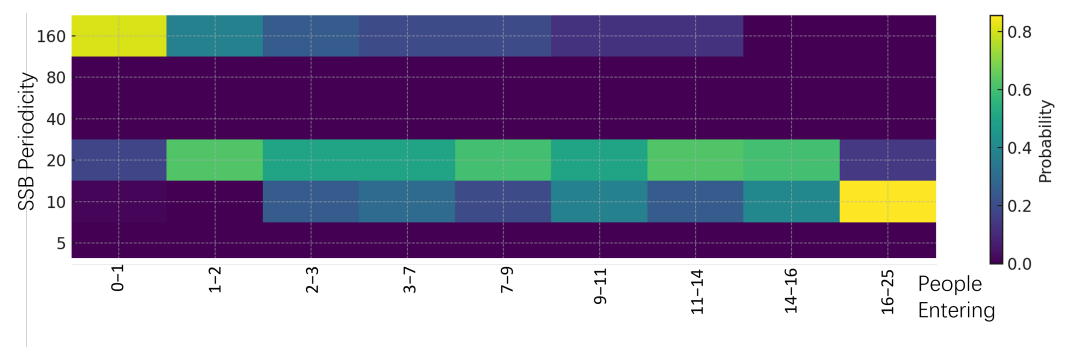


Figure 10. Heat map of predicted UE arrival proxy versus selected SSB periodicity.

We additionally evaluate the temporal stability of the policy by measuring the number of action switches per hour. Excessive switching may indicate policy instability or sensitivity to prediction noise. Figure 11 presents the switching frequency of the PPO agent over the seven-day evaluation period. The number of switches varies between zero and six per hour, with most intervals exhibiting approximately three to five switches. Although occasional intervals contain no switching, the policy generally updates the selected SSB periodicity several times per hour. These results indicate that the agent responds dynamically to variations in the predicted traffic demand.

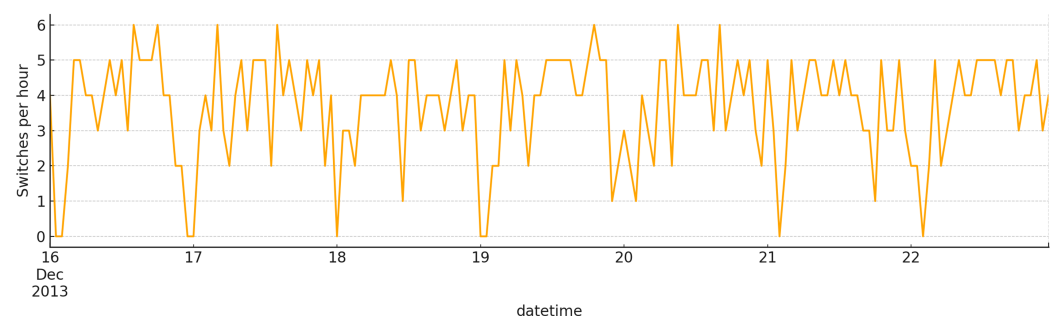


Figure 11. Action switch rate.

Overall, these results indicate that the PPO agent successfully learns an interpretable and traffic-aware control strategy. The policy converges consistently across training runs, selects actions consistent with the underlying EMF–latency trade-off, and remains stable over time. This provides strong evidence that reinforcement learning is an effective mechanism for adaptive SSB periodicity control in dynamic urban environments.

5.5. Seven-Day EMF Latency Trade-Off Comparison

To evaluate whether the performance of adaptive SSB periodicity control remains consistent under different daily traffic conditions, we calculate the average SSB EMF power density and average UE waiting time per person for each day of the complete seven-day test period. Figures 12 and 13 present the daily results for the proposed prediction-driven controller, the same trained policy supplied with the actual traffic-derived arrival sequence, and the fixed 20, 40, and 80 ms SSB configurations.

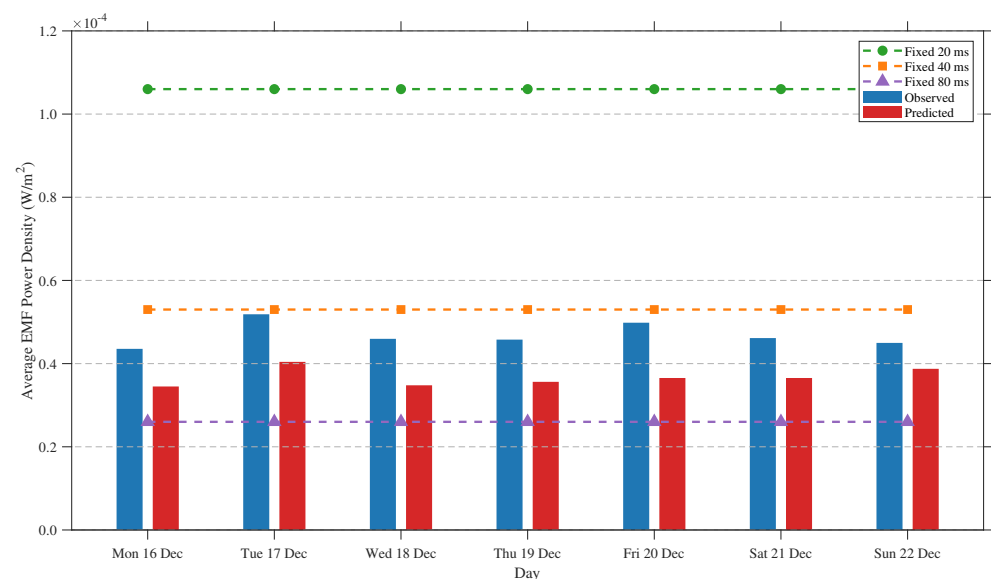


Figure 12. Daily average SSB-induced EMF power density over the complete seven-day test period for the fixed-periodicity and adaptive SSB control strategies.

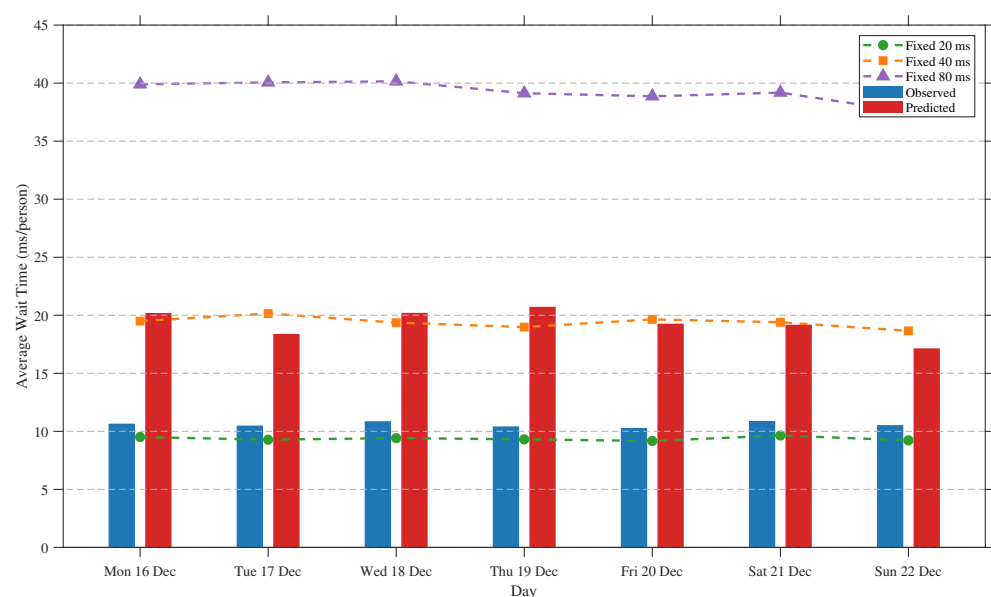


Figure 13. Daily average UE waiting time per person over the complete seven-day test period for the fixed-periodicity and adaptive SSB control strategies.

The prediction-driven controller represents the intended operating mode of the proposed framework, in which the PPO policy receives the demand proxy forecast generated by the LSTM model. The observed proxy controller uses the same trained PPO policy but receives the corresponding observed traffic-derived arrival sequence. This reference configuration removes prediction error and allows its effect on the control decisions to be examined. The fixed 20, 40, and 80 ms configurations are included to represent static SSB operation.

Figure 12 shows that the prediction-driven controller maintains a lower daily average SSB EMF power density than the fixed 20 and 40 ms configurations throughout the complete test week. Its values remain between those of the fixed 40 and 80 ms configurations. The fixed 20 ms configuration produces the highest power density, while the fixed 80 ms configuration produces the lowest value among the static baselines. The observed proxy controller produces higher power density than the prediction driven controller on each test day.

Figure 13 shows the corresponding daily average waiting time per person. The fixed 20 ms configuration produces the lowest waiting time, with daily values close to 10 ms, whereas the fixed 80 ms configuration produces values of approximately 37 to 40 ms. The prediction-driven controller maintains daily average waiting times of approximately 18 to 21 ms, which are substantially lower than the fixed 80 ms results and generally close to the fixed 40 ms results. The observed proxy controller produces lower waiting time than the prediction-driven controller.

The complete numerical results are reported in Tables 3 and 4. Over the seven-day test period, the prediction-driven controller achieves a mean SSB EMF power density of $3.67 \times 10^{-5} \text{ W/m}^2$ and a weekly average waiting time of 19.22 ms per person. The observed proxy controller achieves a mean power density of $4.69 \times 10^{-5} \text{ W/m}^2$ and a weekly average waiting time of 10.59 ms per person.

Table 3. Daily average SSB-induced EMF power density over the seven-day test period ($\times 10^{-5} \text{ W/m}^2$).

Method	16 DEC	17 DEC	18 DEC	19 DEC	20 DEC	21 DEC	22 DEC	Mean
Fixed 20 ms	10.60	10.60	10.60	10.60	10.60	10.60	10.60	10.60
Fixed 40 ms	5.30	5.30	5.30	5.30	5.30	5.30	5.30	5.30
Fixed 80 ms	2.60	2.60	2.60	2.60	2.60	2.60	2.60	2.60
Observed-proxy controller	4.40	5.20	4.60	4.60	5.00	4.60	4.50	4.69
Prediction-driven controller	3.40	4.00	3.50	3.60	3.70	3.70	3.90	3.67

Table 4. Daily average UE waiting time per person over the seven-day test period.

Method	16 DEC	17 DEC	18 DEC	19 DEC	20 DEC	21 DEC	22 DEC	Weekly Mean
Fixed 20 ms	9.51	9.28	9.42	9.31	9.17	9.64	9.22	9.36
Fixed 40 ms	19.50	20.15	19.36	18.98	19.65	19.40	18.66	19.39
Fixed 80 ms	39.90	40.06	40.16	39.13	38.87	39.18	37.53	39.22
Observed-proxy controller	10.66	10.50	10.87	10.42	10.29	10.89	10.54	10.59
Prediction-driven controller	20.19	18.39	20.21	20.73	19.28	19.18	17.15	19.22

These results show that prediction error changes the SSB periodicities selected by the trained PPO policy. The prediction-driven controller selects an operating point with lower EMF power density and higher waiting time than the observed proxy controller. Since the demand of a future interval is unavailable when the control decision is made, the prediction-driven results represent the practical operating condition of the proposed framework.

6. Conclusions

This paper presented an EMF-aware framework that integrates LSTM-based demand prediction with PPO-based SSB periodicity control in a 5G network simulation. Aggregated Internet usage measurements from the Milan dataset were transformed into a traffic-derived UE arrival proxy and used to evaluate the controller over a seven-day held-out period. The proposed controller selected among SSB periodicities of 5, 10, 20, 40, 80, and 160 ms according to the predicted demand proxy, while the reward jointly considered normalized SSB power density and aggregate user waiting time.

Over the complete test week, the prediction-driven controller achieved a mean SSB power density of $3.67 \times 10^{-5} \text{ W/m}^2$ and a weekly average waiting time of 19.22 ms per person. Compared with the fixed 20 ms configuration, this corresponds to an EMF power density reduction of approximately 65.4%. Compared with the fixed 40 ms configuration, the reduction is approximately 30.8%, while the average waiting time remains close to that of the fixed 40 ms baseline. The fixed 80 ms configuration achieves a lower EMF power density than the proposed controller, but it results in a substantially higher average waiting time. Therefore, the proposed method provides an intermediate operating point between the low delay of shorter fixed periodicities and the lower EMF exposure of longer fixed periodicities, rather than outperforming every static configuration in both objectives.

The observed proxy controller achieved a lower waiting time of 10.59 ms per person but a higher mean power density of $4.69 \times 10^{-5} \text{ W/m}^2$. This difference shows that prediction errors influence the SSB periodicities selected by the trained PPO policy. Nevertheless, the prediction-driven configuration represents the intended operating condition because the demand of a future control interval is not available when the SSB periodicity decision is made.

The proposed method is evaluated as a network controller. The Milan dataset does not directly measure physical UE entries, and the derived arrival sequence should therefore be interpreted as a traffic-based demand proxy. Practical implementation would additionally require consideration of SSB configuration signaling, reconfiguration latency, system information scheduling, and vendor specific constraints. Future work will evaluate the framework using direct network measurements, extend the model to multiple cells, and investigate practical configuration procedures under changing traffic and propagation conditions.

Author Contributions: Conceptualization, K.L. and Y.S.; methodology, K.L.; software, K.L.; validation, K.L., M.S.M. and O.P.; formal analysis, K.L.; investigation, K.L.; data curation, K.L.; writing—original draft preparation, K.L.; writing—review and editing, K.L., M.S.M., O.P., M.A.I. and Y.S.; visualization, K.L.; supervision, O.P., M.A.I. and Y.S.; project administration, Y.S.; funding acquisition, M.A.I. and Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the UK Engineering and Physical Sciences Research Council (EPSRC) through the CHEDDAR project under Grant EP/X040518/1 and the CHEDDAR TMF Uplift project under Grant EP/Y037421/1.

Data Availability Statement: The Milan telecommunications dataset used in this study is publicly available through Harvard Dataverse at <https://doi.org/10.7910/DVN/EGZHFV>. The processed data generated during this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors gratefully acknowledge the support provided through the CHEDDAR and CHEDDAR Uplift projects.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Omri, A.; Shaqfeh, M.; Ali, A.; Alnuweiri, H. Synchronization procedure in 5G NR systems. *IEEE Access* **2019**, *7*, 41286–41295.
- Giordani, M.; Polese, M.; Roy, A.; Castor, D.; Zorzi, M. Standalone and Non-Standalone Beam Management for 3GPP NR at mmWaves. *IEEE Commun. Mag.* **2019**, *57*, 123–129. <https://doi.org/10.1109/MCOM.2019.1800384>.
- Parkvall, S.; Dahlman, E.; Furuskar, A.; Frenne, M. NR: The New 5G Radio Access Technology. *IEEE Commun. Stand. Mag.* **2017**, *1*, 24–30. <https://doi.org/10.1109/MCOMSTD.2017.1700042>.
- Document TS 38.213, 3GPP. NR; Physical Layer Procedures for Control. Available online: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3215> (accessed on 30 May 2022).
- Perera, A.I.; Manosha, K.B.S.; Rajatheva, N.; Latva-aho, M. An Initial Access Optimization Algorithm for Millimeter-Wave 5G NR Networks. In *Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6. <https://doi.org/10.1109/VTC2020-Spring48590.2020.9128625>.
- ICNIRP. Guidelines for limiting exposure to electromagnetic fields (100 kHz to 300 GHz). *Health Phys.* **2020**, *118*, 483–524.
- Li, K.; Mollel, M.S.; Popoola, O.; Sambo, Y.; Imran, M.A. Analysis of Electromagnetic Field Radiation From Synchronisation Signal Block (SSB) in 5G Systems. In *Proceedings of the 2023 5th International Conference on Computer Communication and the Internet (ICCCI)*; IEEE: Piscataway, NJ, USA, 2023; pp. 185–189. <https://doi.org/10.1109/ICCCI59363.2023.10210151>.
- Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Telecom Italia. Telecommunications—SMS, Call, Internet—MI. Harvard Dataverse, V1. 2015. Available online: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/EGZHFV> (accessed on 27 July 2026).
- Bernabè, M.; López-Pérez, D.; Piovesan, N.; Geraci, G.; Gesbert, D. Optimal SSB Beam Planning and UAV Cell Selection for 5G Connectivity on Aerial Highways. In *Proceedings of the GLOBECOM 2024—2024 IEEE Global Communications Conference*; IEEE: Piscataway, NJ, USA, 2024; pp. 2040–2045. <https://doi.org/10.1109/GLOBECOM52923.2024.10901503>.
- Bernabè, M.; López-Pérez, D.; Piovesan, N.; Geraci, G.; Gesbert, D. Massive MIMO for Aerial Highways: Enhancing Cell Selection via SSB Beams Optimization. *IEEE Open J. Commun. Soc.* **2024**, *5*, 3975–3996. <https://doi.org/10.1109/OJCOMS.2024.3418339>.
- 3rd Generation Partnership Project. NR; Radio Resource Control (RRC); Protocol Specification; Technical Specification TS 38.331, Version 17.2.0, Release 17; 3rd Generation Partnership Project (3GPP): Sophia Antipolis, France, 2022.
- Chiaraviglio, L.; Cacciapuotì, A.S.; Martino, G.D.; Fiore, M.; Montesano, M.; Trucchi, D.; Melazzi, N.B. Planning 5G Networks Under EMF Constraints: State of the Art and Vision. *IEEE Access* **2018**, *6*, 51021–51037. <https://doi.org/10.1109/ACCESS.2018.2868347>.
- Elzanaty, A.; Chiaraviglio, L.; Alouini, M.S. 5G and EMF exposure: Misinformation, open questions, and potential solutions. *Front. Commun. Netw.* **2021**, *2*, 635716.
- Pawlak, R.; Krawiec, P.; Żurek, J. On Measuring Electromagnetic Fields in 5G Technology. *IEEE Access* **2019**, *7*, 29826–29835. <https://doi.org/10.1109/ACCESS.2019.2902481>.
- Huang, H.; Guo, S.; Gui, G.; Yang, Z.; Zhang, J.; Sari, H.; Adachi, F. Deep Learning for Physical-Layer 5G Wireless Techniques: Opportunities, Challenges and Solutions. *IEEE Wirel. Commun.* **2020**, *27*, 214–222. <https://doi.org/10.1109/MWC.2019.1900027>.
- Yajnanarayana, V.; Rydén, H.; Hévízi, L. 5G handover using reinforcement learning. In *Proceedings of the 2020 IEEE 3rd 5G World Forum (5GWF)*; IEEE: Piscataway, NJ, USA, 2020; pp. 349–354.
- Ericsson. Find the Gaps: Boost Your Network's Energy Performance with Time Domain Techniques. 2024. Available online: <https://www.ericsson.com/en/blog/2024/11/find-the-gaps-boost-your-networks-energy-performance-with-time-domain-techniques> (accessed on 27 July 2026).
- Document TS 38.211, 3GPP. Physical Channels and Modulation. Available online: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213> (accessed on 30 May 2022).
- Barlacchi, G.; De Nadai, M.; Larcher, R.; Casella, A.; Chitic, C.; Torrisi, G.; Antonelli, F.; Vespignani, A.; Pentland, A.; Lepri, B. A Multi-Source Dataset of Urban Life in the City of Milan and the Province of Trentino. *Sci. Data* **2015**, *2*, 150055. <https://doi.org/10.1038/sdata.2015.55>.
- Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal Policy Optimization Algorithms. *arXiv* **2017**. arXiv:1707.06347. <https://doi.org/10.48550/arXiv.1707.06347>.
- Raffin, A.; Hill, A.; Gleave, A.; Kanervisto, A.; Ernestus, M.; Dormann, N. Stable-Baselines3: Reliable Reinforcement Learning Implementations. *J. Mach. Learn. Res.* **2021**, *22*, 1–8.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.