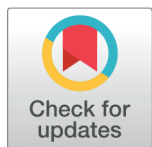


ORIGINAL ARTICLE

 OPEN ACCESS

Received: 16/07/2026

Accepted: 23/07/2026

Published: 17/08/2026

Citation: Rajagopal MR, Malathi S (2026) Hybrid Deep Reinforcement Learning with Kookaburra Optimization for QoS-Aware Bandwidth Allocation in GMPLS Optical Networks. Indian Journal of Science and Technology 19(30): 2140-2151. <https://doi.org/10.17485/IJST/v19i30.1055>

* **Corresponding author.**visitmalathi@gmail.com**Funding:** None**Competing Interests:** None

Copyright: © 2026 Rajagopal & Malathi. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](https://www.isee.in))

ISSN

Print: 0974-6846

Electronic: 0974-5645

Hybrid Deep Reinforcement Learning with Kookaburra Optimization for QoS-Aware Bandwidth Allocation in GMPLS Optical Networks

M R Rajagopal¹, S Malathi^{2*}

¹ A Veeriyar Vandayar Memorial Sri Pushpam College (Autonomous), Thanjavur-613504 (Affiliated to Bharathidasan University), Tamil Nadu, India

² Department of Computer Science, Swami Dayananda College of Arts & Science, Manjakkudi. Tiruvarur (dt), -612610 (Affiliated to Bharathidasan University), Tamil Nadu, India

Abstract

Objectives: To address dynamic bandwidth allocation with strict Quality of Service (QoS) requirements in Generalized Multi-Protocol Label Switching (GMPLS) optical networks under strain from internet services, real-time multimedia, and cloud infrastructure. **Method:** A Hybrid Deep Reinforcement Learning (Hyb-DRL) framework combined with the Kookaburra Optimization Algorithm (KkOA) for adaptive weight adjustment is proposed. Dynamically generated input data, including user request rates, queue lengths, server availability, and link stability metrics, were used to simulate real-world traffic. The Hyb-DRL agent learned optimal routing and bandwidth provisioning policies while KkOA optimized model weights for faster convergence and stability. **Findings:** The simulation results show that the suggested Hyb-DRL-KkOA algorithm performs better than the conventional bandwidth allocation algorithms. In contrast to conventional algorithms, it reduces the blocking probability, makespan, cost, and energy utilization while improving the throughput; therefore, providing enhanced quality of service (QoS). The proposed framework achieves a lower blocking probability by 78%, makespan by 64%, energy consumption by 51%, and operational cost by 47%. In addition to this, it provides better throughput performance by 69% than other conventional techniques. Moreover, it provided a delay of 0.0189 s, minimal energy consumption of 33 mJ, and maximal throughput of 950 Mbps. **Novelty:** A combination of reinforcement learning and meta-heuristic optimization leads to adaptive decision-making regarding routing and bandwidth allocation in the face of different traffic demands. The performance gain in terms of QoS is due to optimal utilization of network resources with low blocking probability, energy and operational cost. It provides a scalable and adaptive solution for high-speed, reliable data transmission in modern communication networks.

Keywords: GMPLS Optical Networks; Kookaburra Optimization Algorithm; Bandwidth Allocation; Quality of Service (QoS); Blocking Probability

1 Introduction

The rapid rise in cloud computing, Internet services, real-time multimedia applications, edge computing, and latency-critical communications has led to an increase in bandwidth requirements in GMPLS-based optical networks⁽¹⁾. Being the core of communication infrastructure systems, GMPLS networks should provide reliable, fast, and quality-of-service-aware transmission while efficiently using available optical resources. Nevertheless, due to the constantly changing characteristics of network traffic, as well as diverse needs of different services, bandwidth allocation becomes a complex optimization task. Therefore, intelligent management of optical resources that would be able to adapt to dynamically changing traffic becomes important^{(2), (3)}.

It is clear from the recent literature that a great deal of advancement has been made in the areas of intelligent bandwidth management and traffic engineering using machine learning, deep learning, reinforcement learning, graph neural networks, software-defined networking, and metaheuristic optimization^{(4), (5), (6), (7)}. The application of machine learning has helped in traffic prediction and routing by leveraging the learned behaviour of the network from previous traffic data, while the deep learning approach has provided better modelling of complex traffic properties. Furthermore, reinforcement learning has provided the ability to make decisions based on constant interaction with the network environment. Similarly, bio-inspired algorithms have been used to optimize routing, wavelengths, spectra, and resource allocation⁽⁸⁾.

However, despite these developments, current approaches suffer from certain critical disadvantages. The vast majority of optimization techniques are based on static heuristics, static optimization algorithms, or independently designed machine learning methods which fail to accommodate the high level of dynamism of the traffic environment^{(9), (10)}. Classical bandwidth allocation algorithms using heuristic methods tend to be prone to high levels of traffic blocking and poor resource usage due to a pre-defined routing policy. Deep learning techniques also typically involve complex off-line training sessions and fail to implement the mechanism of dynamic parameter optimization while the network is operating⁽¹¹⁾. On the other hand, while providing adaptive decision-making capabilities, reinforcement learning algorithms still have significant issues with convergence rates, policy stability, and scalability in the context of GMPLS networks⁽¹²⁾. In addition, there currently exist a considerable number of hybrid optimization algorithms which focus either on the performance of routing or on the process of resource allocation, failing to account for intelligent policy learning, adaptive parameter optimization, and multiple QoS parameters such as traffic blocking rate, throughput, latency, energy efficiency, and cost^{(13), (14)}.

Some of the illustrative studies are described below. Singh et al.⁽¹⁴⁾ used MET, MCT, and OLB algorithms for bandwidth allocation in GMPLS networks; however, the heuristics were non-adaptive under dynamic traffic conditions. Hadian et al.⁽¹⁵⁾ combined MPLS technology with routing and spectrum allocation approaches for efficient network resources management; however, the performance of the scheme degraded in case of restricted spectrum availability. Khan et al.⁽¹⁶⁾ suggested a machine learning-based QoS-aware SDN-enabled photonic switching architecture; however, the performance of this architecture was highly dependent on the quality and distribution of training data. Vasundhara et al.⁽¹⁷⁾ proposed a spectrum allocation algorithm based on deep learning; the algorithm proved to be effective in reducing spectrum fragmentation; however, it could not optimize latency and several other QoS parameters. Kim et al.⁽¹⁸⁾ proposed an OSNR-aware wavelength assignment algorithm for multicast routing; the efficiency of the algorithm was highly dependent on the topological properties of the network and OSNR evaluation accuracy. More recent works include the use of deep reinforcement learning for optical bandwidth reconfiguration⁽¹⁹⁾ and self-supervised learning for optical interconnections⁽²⁰⁾. However, both of the approaches require stable offline training and have limited flexibility under highly dynamic traffic conditions.

Currently, there is no single intelligent bandwidth allocation mechanism that incorporates adaptive reinforcement learning along with efficient global parameter optimization and multi-objective QoS optimization in large-scale dynamic GMPLS optical networks. The existing approaches either concentrate on learning the policies but do not consider adaptive optimization; or just optimize some of the QoS parameters. This study therefore proposes a new Hybrid Deep Reinforcement Learning with Kookaburra Optimization Algorithm (Hyb-DRL-KkOA) framework for QoS-aware bandwidth allocation in GMPLS optical networks, in an attempt to overcome the above-mentioned problems. In contrast to all other hybrid intelligent approaches, which focus on optimizing either the learning process or the routing problem separately, the proposed approach tries to achieve the best in both fields.

The significant contributions of this paper are highlighted as follows:

- A novel Hybrid Deep Reinforcement Learning and Kookaburra Optimization Algorithm (Hyb-DRL-KkOA) is introduced for adaptive QoS-aware bandwidth allocation in GMPLS optical networks.
- A new adaptive reinforcement learning method is designed to learn the routing and bandwidth allocation policy directly from varying states of the network.

- Kookaburra Optimization Algorithm is employed for the optimization of the DRL network parameters, thereby enhancing the convergence rate and learning ability of the allocation algorithm.
- Simultaneous optimization of multiple QoS parameters such as blocking probability, throughput, makespan, delay, energy consumption, and operational cost using the proposed framework is achieved.
- Simulation results reveal that the proposed framework outperforms the current state-of-the-art intelligent bandwidth allocation methods in terms of scalability, adaptability, and resource efficiency.

The novelty of this paper is highlighted as follows:

- The novelty of this research is to present a Hybrid Deep Reinforcement Learning-Kookaburra Optimization Algorithm (Hyb-DRL-KkOA) for optimizing both the routing strategies and the deep learning models for allocating bandwidth considering QoS parameters in a GMPLS optical network.
- Contrary to the current DRL methods, which concentrate mainly on routing strategy and bandwidth optimization, the current approach combines adaptive learning of the policy with the meta-heuristic approach for optimizing weights, leading to a rapid convergence rate, policy stability and optimal use of resources under dynamic traffic conditions.
- Moreover, the proposed framework optimizes multiple QoS parameters such as the blocking probability, delay, throughput, makespan, energy efficiency, and cost.

2 Methodology

This manuscript presents an efficient and smart bandwidth allocation mechanism in GMPLS optical networks based on a Hybrid Deep Reinforcement Learning (Hyb-DRL) architecture improved by the Kookaburra Optimization Algorithm (KkOA).

The suggested framework gathers all the necessary dynamic network state information, which consists of user demand, queues, availability of servers, and link states, to model the existing GMPLS network. Hyb-DRL agent uses continuous interaction with the network to learn the best policies regarding routing and bandwidth allocation, whereas the Kookaburra Optimization Algorithm (KkOA) helps in parameter optimization of the model. According to the obtained policy, effective allocation of the bandwidth and wavelength resources takes place considering the Quality of Service (QoS) constraints under different traffic scenarios. The result is better network performance characterized by reduced blocking probability, delay, power and cost, as well as increased throughput and resource usage. Figure 1 represents the workflow of the proposed framework for bandwidth allocation mechanism in GMPLS optical networks.

2.1 GMPLS Optical Network Initialization

The envisioned GMPLS optical network is installed across a two-dimensional geographical plane characterized by its network width and network length in kilometers. The simulation framework is designed to model an actual-world, scalable optical backbone facility. The network comprises two main entities: user nodes and server nodes. The user nodes are tasked with creating traffic demands, while the server nodes perform routing and bandwidth and link resource allocation. The nodes are located at fixed locations within the specified area. There are 21 user nodes (N1 to N21) dispersed uniformly and randomly over the network region, shown by blue circular markers throughout the figures. The server nodes (T1 to T7) are shown using red square markers and located along the right boundary of the network to model centralized control or edge cloud resource centers.

Figure 2(a) shows the Initial GMPLS Network Configuration, which is an example of the dynamic bandwidth request generated by the users' nodes that reflects the real-time traffic of the GMPLS optical network. Figure 2(b) depicts the process of path establishment, where the appropriate server is chosen based on available resources, link state and wavelengths, which allow adaptive routing and bandwidth allocation. This network setup allows several simultaneous requests and reflects the realistic scenario for the QoS-aware resource allocation, blocking probability, and network performance evaluation.

2.2 Hyb-DRL for Bandwidth Allocation in GMPLS Optical Network

After the GMPLS optical network initialization, where users and servers are positioned dynamically within a defined area, the Hybrid Deep Reinforcement Learning (Hyb-DRL) model is applied to perform intelligent bandwidth allocation. The model learns optimal policies for server and wavelength assignment based on network states, including traffic demand, queue sizes, and link conditions. Deep Reinforcement Learning (DRL) methods exhibited encouraging outcomes in decision-making under dynamic situations like network traffic management. The conventional DRL models, though, suffer from gradient explosion while learning based on high-dimensional and diverse network states. Therefore, we propose a Hybridized Deep Reinforcement Learning (Hyb-DRL) method to enhance decision stability and convergence rate during bandwidth allocation.

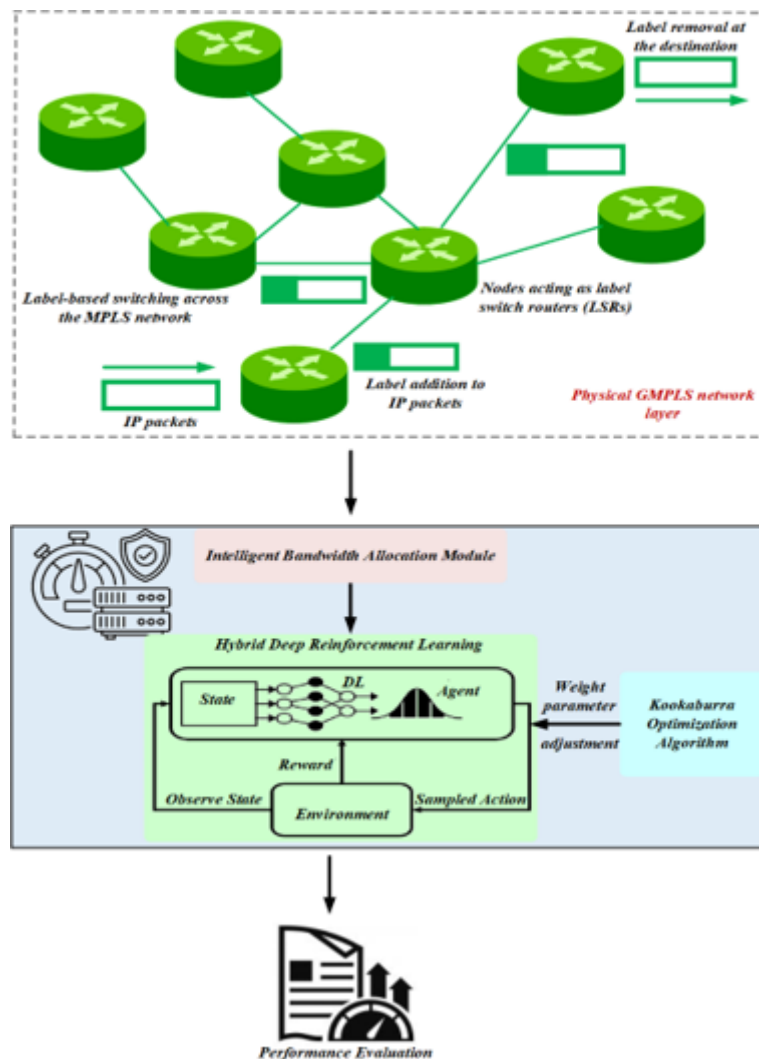


Fig 1. Workflow of the Proposed Framework for bandwidth allocation mechanism in GMPLS optical networks

The standard Q-learning (QL) framework establishes a policy by associating state-action pairs with corresponding Q-values $Q(s, a)$. These Q-values are recursively updated through environmental interactions,

$$Q(s, a) = Q(s, a) + \beta (z + \alpha \max_A Q(s, a) - Q(s, a)) \quad (1)$$

Where, β indicates the learning rate, α specifies the discount factor, z expresses the reward received, s, a denotes the current state and action, s, a is the next state and best action. The DQN components use an LSTM network to approximate the Q-function for continuous time-dependent states in the GMPLS network. The model captures sequential dependencies such as fluctuating queue sizes and user demands. Unlike static CNNs, the DQN incorporates experience replay memory $y_t = (s_t, a_t, z_t, s_{t+1})$, which stores transitions to improve learning stability and a target network z_t is periodically updated to stabilize the learning. Each training step samples mini-batches from y_t and the Q-values are updated by minimizing the temporal difference loss, which is computed using equation (2),

$$l_k(\phi_k) = E(s_t, a_t, z_t, s_{t+1}) [(m_k - Q(s_t, a_t; \phi_k))^2] \quad (2)$$

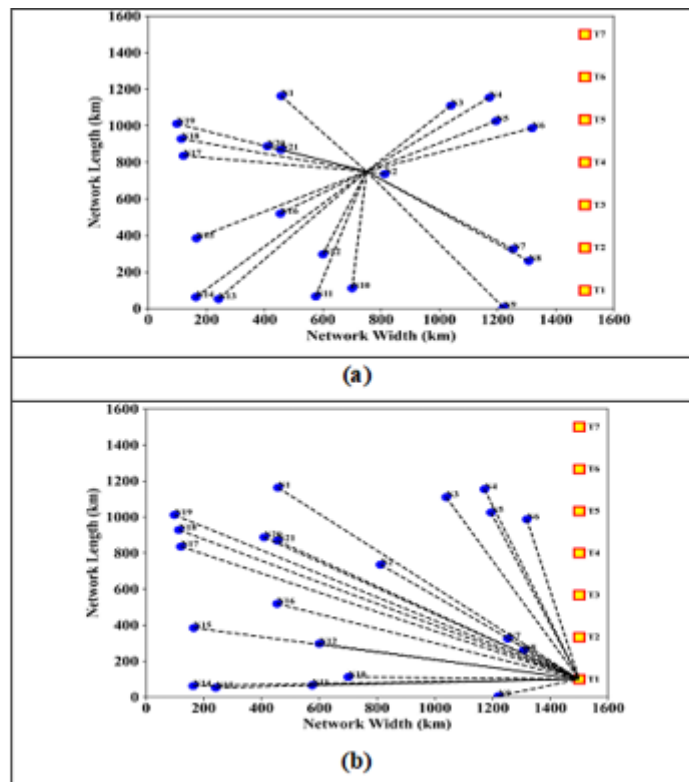


Fig 2. a) Initial GMPLS Network Configuration b) GMPLS Connection Establishment

Where, ϕ specifies the current network parameter. The above equation (2) helps to update $k^{th} \sqrt{x^2 + y^2}$ iteration value of ϕ and s_t indicates the non-lethal state, and the parameter m_k can be formulated as,

$$m_k = z + \alpha \max_a \hat{Q}(s_{t+1}, a; \tilde{\phi}) \quad (3)$$

The LSTM includes two fundamental parts: the Gating Unit (GU) and the Memory Unit (MU). The MU records and buffers temporal patterns in network conditions like traffic volume, queue length, and link metrics. The GU controls how this information travels across the network to impact bandwidth allocation decisions. The Input Gate (IG) regulates the entry of new traffic features into the MU for prospective decision-making. With sigmoid activation and previous memory, the gate outputs values between 0 and 1 in controlling the flow of information. The Output Gate screens corresponding memory content to produce the hidden states for the best server and wavelength selection.

$$s_k = \sigma(w_{ix}i_k + w_{ih}h_{k-1} + w_{id}d_{k-1} + B_s) \quad (4)$$

$$f_k = \sigma(w_{if}i_k + w_{fh}h_{k-1} + w_{fd}d_{k-1} + B_f) \quad (5)$$

$$d_k = f_k d_{k-1} + s_k \tanh(w_{id}i_k + w_{hd}h_{k-1} + B_d) \quad (6)$$

$$y_k = \sigma(w_{yx}i_k + w_{yh}h_{k-1} + w_{yd}d_k + B_y) \quad (7)$$

$$h_k = y_k \tanh(d_k) \quad (8)$$

Here, σ indicates the sigmoid activation function, and s , f , d , and y represents the input gate, forget gate, cell activation vector, and output gate, respectively. Moreover, h and w represents the hidden trajectory, and the weights between two units. The model faces difficulty in acquiring precise temporal behavior because of the irregularity of the previous state transitions. But the suggested LSTM-based Hy-DRL model successfully learns dynamic network behaviour under GMPLS conditions. To enhance convergence and prevent suboptimal learning, the Kookaburra Optimization Algorithm (KkOA) is utilized to fine-tune key hyperparameters. This allows the model to dynamically allocate bandwidth efficiently and adaptively in real time, dependent on learned and current states.

2.3 KkOA for Hyperparameter Tuning in Hyb-DRL-Based Bandwidth Allocation

In this proposed framework, a Kookaburra Optimization Algorithm (KkOA) is used to optimize the hyperparameters in the Hy-DRL model to allocate bandwidth efficiently in GMPLS optical networks. KkOA is a nature-inspired, general-purpose optimization algorithm that emulates the social behaviour and hunting tactics of kookaburras. The birds have intelligent prey detection behaviours whose analogies are used here to emulate hunting for optimal parameter settings in intricate learning systems such as Hy-DRL. In this, KOA is used to tune the important hyperparameters of the Hyb-DRL agent, such as the learning rate, reward discount factor, neural network weight initialization, and LSTM configuration. Optimization of these parameters greatly improves the convergence speed, learning stability, and overall network performance of the DRL model for bandwidth allocation. The stepwise procedure of KkOA is provided as follows,

Step 1: Initialization

The first set of kookaburra agents and a given set of Hy-DRL hyperparameter vectors are randomly initialized. Each kookaburra represents a potential configuration of hyperparameters.

Step 2: Random Generation

From the initialized parameter, each kookaburra randomly selects an initial configuration. This collection serves as the baseline solution on which performance for bandwidth allocation is evaluated.

Step 3: Fitness Function Calculation

Each configuration is assessed by a fitness function, computed according to performance metrics like blocking probability, throughput, makespan, cost, energy consumption, and delay,

$$FitnessFunction = f(\alpha, \eta, w) \quad (9)$$

Where, α specifies the learning rate, η indicates the activation function and w expresses the weight matrices of the Hyb-DRL model.

Step 4: Exploration (Hunting Strategy)

In KkOA, every kookaburra compares the hyperparameter settings (positions) of other kookaburras in the population, which have been successful with higher fitness values according to QoS performance in GMPLS networks. Such high-performing settings are contemplated as possible prey positions. Depending on this comparison, every kookaburra computes its candidate set of prey, as indicated in equation (10),

$$Pc_m = \{x_n : O_n < O_m \text{ and } n \neq m\} \\ m = 1, 2, 3, \dots, M \text{ and } n \in \{1, 2, 3, \dots, M\}$$

Where, the candidate prey set is denoted as Pc_m for m^{th} kookaburra, x_n is the position of kookaburra, O_n is the fitness value. Randomly, each kookaburra picks an object from its candidate list and approaches it by adjusting its position (i.e., parameter vector). The step is computed by employing equation (11),

$$x_{m,d}^{p1} = x_{m,d} + RW_L (PS_{m,d} - I.x_{m,d}), \\ m = 1, 2, 3, \dots, M \text{ and } d = 1, 2, 3, \dots, N$$

$$x_m = \begin{cases} x_m^{p1} & O_m^{p1} < O_m \\ x_m & \text{else} \end{cases} \quad (12)$$

Where, the new updated position is denoted as x_m^{p1} for the m^{th} kookaburra, its dimension is represented as $x_{m,d}^{p1}$, and its fitness value is denoted as O_m^{p1} . Also, the selected prey is $PS_{m,d}$ for m^{th} kookaburra with d^{th} dimension, W_L is the weight matrix, R, I is the random numbers and here R lies between $[0,1]$ and I is selected from the set $\{1,2\}$, N is the decision variable quantity, and M is the total number of kookaburras. This step supports adaptive exploration so that every agent can explore the parameter space effectively by taking advantage of improved solutions in the population, directly supporting enhanced convergence and bandwidth allocation effectiveness in the Hyb-DRL model for GMPLS optical networks.

Step 5: Exploitation Phase

In the KkOA, after the exploration process, every kookaburra goes into an exploitation process where it computes a new hyperparameter setting (i.e., a new location) in the vicinity of the best solution found so far. This refines the parameters in the vicinity of promising areas in the search space, for instance, near-optimal learning rates, weight initializations, or reward discount factors employed in the Hyb-DRL model for allocating bandwidth.

$$x_{m,d}^{p2} = x_{m,d} + \vartheta + \phi(1-2R) \frac{(Ub_d - Lb_d)}{t},$$

$$m = 1, 2, 3, \dots, M, d = 1, 2, 3, \dots, N, \text{ and } t = 1, 2, 3, \dots, T$$

$$x_m = \begin{cases} x_m^{p2} & O_m^{p2} < O_m \\ x_m & \text{else} \end{cases} \quad (14)$$

Where, the new updated position is denoted as x_m^{p2} for the m^{th} kookaburra, its dimension is represented as $x_{m,d}^{p2}$, and its fitness value is denoted as O_m^{p2} . Also, ϑ, ϕ is the activation function of Hyb-DRL, the counter of iterations is denoted as t and the total number of iterations are denoted as T .

Step 6: Return Best Optimal Solution

The one with the best fitness is chosen as the best set of Hyb-DRL parameters for the end training.

Step 7: Termination

Repeat Steps 3–6 until the stopping criterion is achieved. This KkOA-based tuning process significantly enhances the learning ability and efficiency of Hyb-DRL for bandwidth allocation, thus yielding lower blocking probability, lower energy consumption, and improved QoS performance in GMPLS optical networks.

2.4 Simulation Environment and Experimental Setup

The Hyb-DRL-KkOA framework is implemented and tested in a simulation environment coded in Python to analyze dynamic QoS-aware bandwidth allocation in GMPLS optical networks. As there is no publicly available benchmark dataset for this application, this work utilizes a simulation-based traffic generator that mimics actual GMPLS network behaviour. Dynamic connection requests are generated in the form of changes in the traffic arrival rate, bandwidth requirement, queue size, server availability, link usage, and link stability, thus providing different network load and QoS scenarios. Such an environment allows conducting an experiment that includes the evaluation of the proposed framework under various operational conditions. In particular, the experimental setup is represented by the GMPLS optical network where parameters such as network topology, link capacity, wavelength capacity, and traffic load are variable. The implementation of the proposed framework is done in Python, and the execution was carried out on a computer with an Intel(R) Core(TM) i5-4300M CPU, 4 GB RAM, and a 64-bit Windows operating system. Such a simulation, measured in terms of blocking probability, makespan, operational cost, energy consumption, throughput, and average delay under various conditions of traffic load, wavelengths, link utilization rate, and simulation duration.

The Deep Reinforcement Learning algorithm is trained using the hyperparameters listed in Table 1, which include learning rate, discount rate, batch size, replay memory size, frequency of updates for the target network, optimizer, activation function, number of hidden layers, and number of training episodes.

The parameters used in the Kookaburra Optimization Algorithm, such as population size, number of iterations, exploration and exploitation values, and stopping criterion, are given in Table 2. This list of parameters has been chosen based on prior experimentation to achieve stability and reliability of the learning process.

3 Results and Discussion

This section shows the simulation results comparing the proposed scheme with conventional schemes under the same network conditions. Of primary interest are performance metrics like blocking probability, cost, delay, throughput, energy utilization,

Table 1. Hyperparameters of Hyb-DRL

Hyperparameter	Value
Learning rate (α)	0.001
Mini-batch size	64
Exploration strategy	ε -greedy
Exploration decay	0.995
Minimum exploration rate	0.01
Target network update interval	100 episodes
Replay memory size	10,000
Initial exploration rate (ε)	1.0
Number of training episodes	500
Optimizer	Adam
Activation function	ReLU
Discount factor (γ)	0.99
Output activation	Linear
Hidden layers	2
Neurons per hidden layer	128
Maximum steps per episode	200

Table 2. Hyperparameters of KkOA

Hyperparameter	Value
Population size	30
Maximum iterations	100
Solution dimension	Equal to DRL trainable weights
Fitness function	Average cumulative reward
Selection strategy	Best fitness
Exploration coefficient	0.5
Exploitation coefficient	0.8
Weight update strategy	Adaptive
Stopping criterion	Maximum iterations
Best solution retention	Enabled
Random seed	42

and makespan. The results indicate the efficiency and efficacy of the proposed model as compared to MET-GMPLS⁽¹⁴⁾, RSA-EON⁽¹⁵⁾, and BDM-ML⁽¹⁶⁾ approaches.

The effect of wavelength availability and traffic load on the probability of connection blocking in a GMPLS optical network is depicted in Figure 3. From Figure 3(a) and Figure 3(b), it can be seen that there is a considerable decrease in blocking probability with increasing total available wavelengths across varying wavelength bandwidths (WB). This is because this wavelength increase facilitates better meeting of requirements. From graphs 3(c) and 3(d), it is evident that there exists a very strong correlation between network load and blocking probability in the case of high user density and very high user density. It is, therefore, very evident that wavelength allocation is an essential process in GMPLS networks.

The improvement in the blocking probability clearly shows how the combination of the adaptive policy learning technique with the Kookaburra-based parameter optimization technique results in efficient wavelength assignment across different traffic scenarios. It is worth noting that similar performance has also been observed for reinforcement learning-based optical networks. The main idea behind using such a technique was to demonstrate the advantages of adaptive decisions over heuristic-based techniques. The presented model outperforms these systems as it maintains lower blocking probability even at very high network loads.

Figure 4 displays the network cost against simulation time. From the graphed values, the Proposed Method always manages to have the lowest cost of operations relative to MET-GMPLS, RSA-EON, and BDM-ML. At the end of the simulation (50 ms), the proposed methodology has a cost value of approximately 2.0×10^8 while MET-GMPLS, RSA-EON, and BDM-ML have

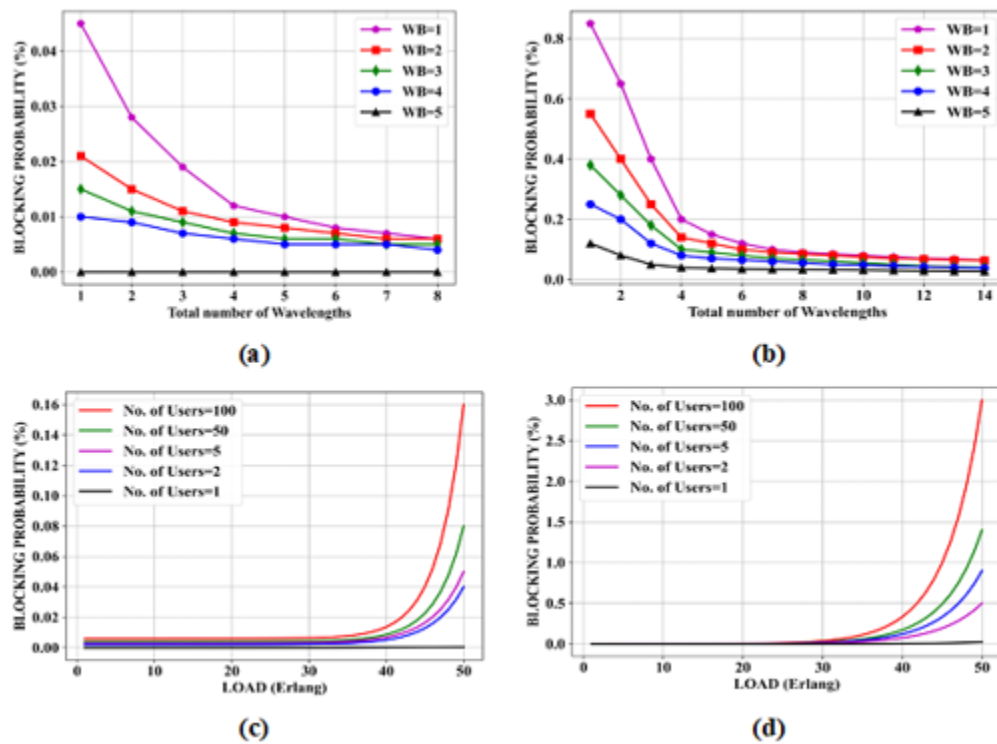


Fig 3. a) Blocking Probability vs. Wavelengths (1–8), b) Blocking Probability vs. Wavelengths (1–14), c) Blocking Probability vs. Load (High User Count) d) Blocking Probability vs. Load (Very High User Count)

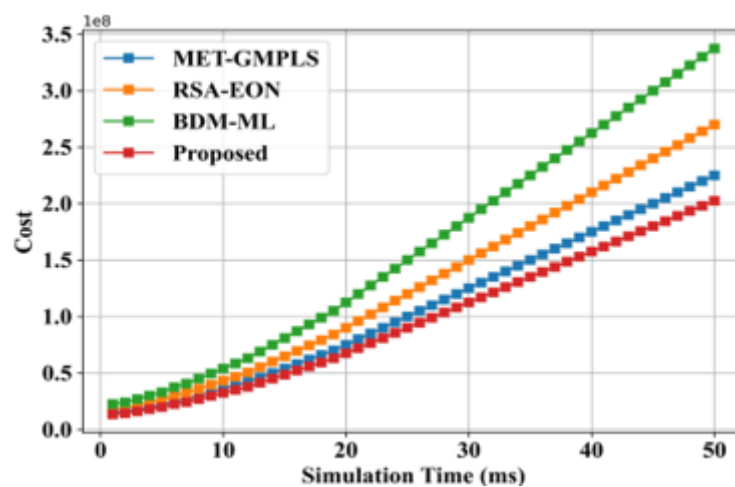


Fig 4. Analysis of Cost Vs Simulation Time

2.25×10^8 , 2.70×10^8 , and 3.35×10^8 , respectively. It corresponds to 11.1%, 25.9%, and 40.3% lower cost of operation relative to MET-GMPLS, RSA-EON, and BDM-ML, respectively. The decreasing cost indicates that the proposed Hyb-DRL-KkOA framework effectively utilizes bandwidth and optimizes routing in order to minimize resource consumption. These results indicate that the proposed method is more scalable and cost-efficient as compared to other methodologies, especially when the network demand increases. Lower operational cost is due to the intelligent allocation of bandwidth and path formation. This is consistent with prior resource optimization research conducted in GMPLS and elastic optical networking, where adaptive resource management minimized operational costs. The proposed framework lowers operational cost even more by combining reinforcement learning and adaptive weight optimization.

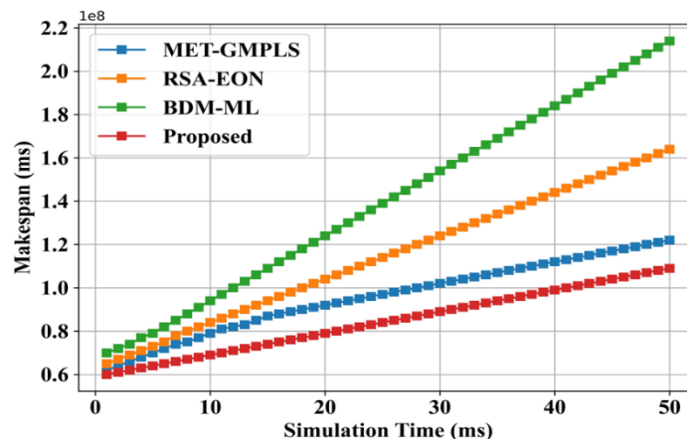


Fig 5. Analysis of Makespan Vs Simulation Time

Figure 5 plots makespan against simulation time. The Proposed Hyb-DRL-KkOA framework obtains the minimum makespan during the simulation period compared to MET-GMPLS, RSA-EON, and BDM-ML frameworks, indicating that the framework takes less time to complete its tasks as well as provisions bandwidth efficiently. With an increase in simulation time from 2 ms to 50 ms, the makespan of all methods increases; however, the proposed framework has the least rate of increment in makespan, making it scalable at an increased load of the network. At the end of the simulation period (50 ms), the proposed approach attains a makespan of about 1.08×10^8 while MET-GMPLS, RSA-EON, and BDM-ML obtain makespan values of approximately 1.21×10^8 , 1.63×10^8 , and 2.14×10^8 , respectively. This gives 10.7%, 33.7%, and 49.5% reductions in makespan compared to MET-GMPLS, RSA-EON, and BDM-ML frameworks, respectively. Reduced makespan indicates fast bandwidth allocation and routing, which are important for delay-sensitive optical communication systems. Such findings have been observed in intelligent traffic engineering research⁽²¹⁾. But the proposed hybrid deep reinforcement learning and Kookaburra optimization algorithm-based framework obtains better scheduling efficiency through adaptive policy learning and parameter optimization.

The performance of the suggested GMPLS optical network architecture is better when the problem of blocking probability minimization is taken into account in comparison with other resource allocation schemes. The findings of the research are consistent with the works that note the benefits of applying the dynamic routing and wavelength assignment methods. Unlike the traditional schemes that lack adaptability at the increased levels of traffic load, the developed model ensures the efficient use of resources and stability of operation in heavy traffic conditions. The achieved improvements prove the success of the proposed approach and its ability to solve the problem of existing GMPLS network optimization methods.

The comparative performance analysis of the proposed method versus the existing methods such as MET-GMPLS, RSA-EON, and BDM-ML about delay, energy consumption, and throughput is presented in Table 3. With the lowest value of delay equal to 0.0189 seconds, the proposed method shows its superior ability for making decisions and resource allocation as compared to other traditional methods. In addition, the proposed method consumes minimum energy, which is equal to 33 mJ. Besides, the proposed method achieves a maximum throughput of 950 Mbps. This is due to the application of adaptive deep reinforcement learning along with Kookaburra Optimization Algorithm in which both the decision making and parameter optimization take place simultaneously. It is unlike RSA-EON which mostly concentrates on routing and spectrum allocation. Also, it is unlike BDM-ML where parameter optimization does not exist. Due to this dynamic adaptation, the system performs much better when it comes to utilizing bandwidth, reducing congestion, and providing QoS.

Table 3. Comparative Performance Analysis

Parameters/Methods	MET-GMPLS	RSA-EON	BDM-ML	Proposed
Delay (seconds)	0.2051	0.097	0.0329	0.0189
Energy consumption (mJ)	46	36	40	33
Throughput (Mbps)	780	825	885	950
Blocking Probability	Higher	Moderate	Moderate	Lowest
Main Limitation	No adaptive learning	Optimizes only RSA	No parameter optimization	Joint routing and parameter optimization using Hyb-DRL and KkOA

From the experimental results, it is clear that the proposed Hyb-DRL-KkOA model improves bandwidth allocation performance in dynamic GMPLS networks. The model shows better results compared with MET-GMPLS, RSA-EON, and BDM-ML in terms of lower blocking probability, delay, operating cost, energy consumption, and increased throughput. In comparison with all previously developed approaches based on static optimization or offline learning, the proposed model includes adaptive Deep Reinforcement Learning combined with Kookaburra Optimization. As a result, the achieved results prove the applicability of the suggested model for QoS-aware bandwidth allocation.

4 Conclusion

This study developed an innovative method for adaptive routing and dynamic bandwidth allocation for GMPLS optical networks using the Hybrid Deep Reinforcement Learning (Hyb-DRL) technique along with the Kookaburra Optimization Algorithm (KkOA). As opposed to traditional optimization techniques where routing is done independently from other network parameters, the proposed model learned both optimal routing and dynamically modified network weights, thereby filling the gap in existing knowledge in adaptive network resource management in highly dynamic GMPLS networks. This approach differs from other approaches in that both the route decisions and the model parameters are optimized and give less delay of 0.0189 s, less energy consumption of 33 mJ, and high throughput of 950 Mbps. These results extend present knowledge by demonstrating the ability of deep reinforcement learning in combination with metaheuristic optimization to exhibit superior scalability and QoS capabilities under both low and heavy load network conditions. However, the model has been verified solely based on simulations with artificial traffic loads and thus, its scalability and capabilities for real-life implementation remain unknown. Thus, future studies may consider validating the model based on real traffic datasets, extending the scope of the study to multi-layer and multi-domain GMPLS networks, and considering SDN-enabled federated learning-based approaches.

5 Acknowledgements

Nil

Author details: M R Rajagopal is a Research Scholar; Dr. S. Malathi is a Research Supervisor- cum- Assistant Professor.

References

- 1) Malini P, Kavitha KR. An efficient deep learning mechanisms for IoT/Non-IoT devices classification and attack detection in SDN-enabled smart environment. *Computers & Security*. 2024;141:103818. <https://doi.org/10.1016/j.cose.2024.103818>.
- 2) Akilandeswari R, Malathi S. Correction to: Design and implementation of controlling with preventing DDOS attacks using bitcoin by Ethereum block chain technology. *Journal of Transportation Security*. 2023;16(1):7. <https://doi.org/10.1007/s12198-023-00265-1>.
- 3) Hema V. An energy efficient chaotic sparrow search algorithm based clustering protocol for wireless sensor network communications. *International Journal of Health Sciences*. 2022;6(S5):4174–4186. <https://doi.org/10.53730/ijhs.v6nS5.9538>.
- 4) Bermúdez JA, Morales P, Pempelfort H, Araya M, Jara N. Understanding deep reinforcement learning: Enhancing explainable decision-making in optical networks. *ICT Express*. 2025;11(5):969–973. <https://doi.org/10.1016/j.icte.2025.08.002>.
- 5) Chen J, Xiao W, Li X, Zheng Y, Huang X, Huang D, et al. A routing optimization method for software-defined optical transport networks based on ensembles and reinforcement learning. *Sensors*. 2022;22(21):8139. <https://doi.org/10.3390/s22218139>.
- 6) Radha B, Sarojini R. Advanced AI techniques for optimizing resource allocation in 6G networks: A hybrid diffusion-enhanced meta-deep-reinforcement-learning framework. *AIP Advances*. 2025;15:115306. <https://doi.org/10.1063/5.0296968>.
- 7) Srivastava R, Singh YN. Link spectrum status aware adaptive multi-channel routing and spectrum allocation algorithm in elastic optical network. *Computer Networks*. 2026;p. 112247. <https://doi.org/10.1016/j.comnet.2026.112247>.

- 8) Mangaiyarkarasi V, Malathi S. Intelligent Deep Learning-aided Malicious Node Detection and Encrypted Routing Optimization in Software-Defined Networking. *Indian Journal of Science and Technology*. 2026;19(18):1169–1181. <https://doi.org/10.17485/IJST/v19i18.367>.
- 9) Jiang J, Su Y, Cheng J, Shang T. Multi-link fragmentation-aware deep reinforcement learning RSA algorithm in elastic optical network. *Photonics*. 2025;12(7):634. <https://doi.org/10.3390/photonics12070634>.
- 10) Devi SS, Kirubanandasarathy N. An enhanced quality of transmission-aware optical networking based on deep reinforcement learning. *International Journal of Communication Systems*. 2025;38(13):e70204. <https://doi.org/10.1002/dac.70204>.
- 11) Rajagopal MR, Malathi S. An efficient resource provisioning and traffic load balancing in multiprotocol label switched network using optimized gated graph convolution neural network. *Wireless Networks*. 2025;31(5):3803–3823. <https://doi.org/10.1007/s11276-025-03967-7>.
- 12) Thummar R, Dhadhal D, Mishra V. Investigating the performance of 64-channel DWDM systems with high data rates up to 2.5 Tbps at varying power levels and bit rate. *Indian Journal of Science and Technology*. 2025;18(38):3103–3111. <https://doi.org/10.17485/IJST/v18i38.757>.
- 13) Panchal M, Prajapati P. FPA-DQN: A fairness- and pressure-aware dueling deep Q-network for adaptive traffic signal control using UAV-based trajectory data. *Indian Journal of Science and Technology*. 2025;18(28):2257–2272. <https://doi.org/10.17485/IJST/v18i28.735>.
- 14) Singh S, Wason A. Performance Evaluation of bandwidth allocating algorithms in generalized multi-protocol label switched optical networks to enhance Quality of Service. *IEEE Access*. 2022;10:40731–40742. <https://doi.org/10.1109/ACCESS.2022.3166912>.
- 15) Hadian A, Bagherian M, Ayat SS. A new mathematical model considering the multi-protocol label switching and the routing and spectrum allocation problems jointly in elastic optical networks. *Photonic Network Communications*. 2022;44(1):10–20. <https://doi.org/10.1007/s11107-022-00978-8>.
- 16) Khan I, et al. Optimized management of ultra-wideband photonics switching systems assisted by machine learning. *Optics Express*. 2022;30(3):3989–4004. <https://doi.org/10.1364/OE.442194>.
- 17) Vasundhara, Mandloi A. Deep learning for core allocation and fragmentation minimization in an elastic optical network with space division multiplexing. *Journal of Optics*. 2025;p. 1–24. <https://doi.org/10.1007/s12596-025-02566-1>.
- 18) Kim YW, Han TH. OWBM: OSNR-Aware Wavelength Allocation and Branching Methods for Multicast Routing in Custom Topology-Based Optical Network-on-Chips. *IEEE Access*. 2024. <https://doi.org/10.1109/ACCESS.2024.3384555>.
- 19) Geresu K, Gu H, Fadhel M, Wei W, Yu X. ODRAD: An optical wireless DCN dynamic-bandwidth reconfiguration with AWGR and deep reinforcement learning. *Optical Switching and Networking*. 2024;52:100771. <https://doi.org/10.1016/j.osn.2024.100771>.
- 20) Liu CY, Chen X, Li Z, Proietti R, Yoo SB. SL-Hyper-FleX: a cognitive and flexible-bandwidth optical datacom network by self-supervised learning. *Journal of Optical Communications and Networking*. 2022;14(2):A113–A121. <https://doi.org/10.1364/JOCN.439801>.
- 21) Monika, Singh S, Wason A. Performance investigations on data protection algorithms in generalized multi protocol label switched optical networks. *Scientific Reports*. 2023;13(1):425. <https://doi.org/10.1038/s41598-022-26942-0>.