INTERNATIONAL JOURNAL OF RECENT TRENDS IN TECHNOLOGY AND ENGINEERING (IJRTTE)

Journal Home Page: https://ijrtte.com/

AI-Enabled Cloud Bio-Simulation Environments

R. Gayathiri[1], Gajalakshmi.S[2] and N. Manonmani[3]

[1]*Assistant Professor, Department of Electrical and Electronics Engineering, J.J. College of Engineering and Technology, Tiruchirappalli, Tamil Nadu, India.*

[2]*IV Year students, Department of EEE, New Prince Shri Bhavani College of Engineering and Technology, Chennai, Tamil Nadu, India.*

[3]*Professor, Department of EEE, Sri Krishna College of Engineering and Technology, Kuniamuthur, Coimbatore 641008, Tamil Nadu, India.*

[1]gayathirir@jjcet.ac.in, [2]2021ee006@newprinceshribhavani.com, [3]manonmani@skcet.ac.in

Abstract. It is becoming common that biological research is based on large-scale computational modeling, multi-omics integration, and high-fidelity simulation to comprehend complex molecular and cellular behaviors. Nonetheless, the conventional cloud systems are not configured to support the dynamism and data-intensive nature of the contemporary bio-simulation workloads. In this paper, we are going to introduce an AI-powered Cloud Bio-Simulation Environment (AICBSE), which combines deep learning surrogate models, reinforcement learning-based resource optimization, containerized microservices, and real-time feedback systems into one cloud-native system. The system facilitates the sustained ingestion and processing of genomic, proteomic, metabolomic, and imaging data, which can support adaptive simulation infective programs that reduces parameters in response to the biological responses. The experimental findings show that there is a great advancement in the speed of simulation, biological fidelity, and cloud resource efficiency with the architecture exhibiting quick proliferation among dispersed nodes with low latency and high throughput. Digital twins can be created in real-time, interactive models can be created and experiments conducted with a high throughput due to the integration of AI prediction models and dynamic cloud orchestration. On the whole, the suggested environment offers a paradigm shift with respect to speeding up the process of biological discovery, improving predictive modelling, and assisting next-generation applications in the drug development, personalized medicine and in silico research.

Keywords: Bio-Simulation, Cloud Computing, Artificial Intelligence, Reinforcement Learning, Multi-Omics Data Processing.

1. Introduction

Biological systems produce huge, heterogeneous data of genomics, transcriptomics, proteomics, metabolomics, imaging and molecular interactions. The traditional computational infrastructures usually tend not to handle such a challenging and fast changing biological data, particularly when the solution involves the state of perfection that needs to simulate or trace the quickly changing biological surroundings. Due to the growing trend of biomedical research moving towards data-driven discovery, digital twins, and high-throughput in silico experimentation, there is an urgent requirement to have computational environments capable of supporting high workload with flexibility, precision, and scale.

Cloud computing is an encouraging platform to be used since it offers elastic computing services and the provisioning of resources on-demand. Nevertheless, undocumented traditional cloud systems are less efficient to serve dynamic biological modeling. High latency, poor allocation of resources, lack of automation and inability to integrate different biological data are some of the challenges that they often face. These constraints do not allow the conventional cloud platforms to provide full support to such

Volume: 04 Issue: 04 Year: 2025
Received: 16.09.2025 Received in revised from: 07.10.2025 Accepted: 13.11.2025
Available Online: 27.12.2025
Published by NTL Publisher

advanced activities as virtual drug screening, metabolic pathway simulation, predicting protein-ligand interactions, and real-time biological models.

Artificial intelligence has changed the fields of computational biology by providing high-speed predictive modeling, multi-omics pattern recognition, and intelligent decision-making. Deep learning surrogate models can be used to recreate complex biological processes with a stunning accuracy, whereas reinforcement learning can be used to organize resources and optimize them efficiently and adaptively on the cloud. However, despite all the progress, AI models still tend to be very detached with the cloud-native simulation workflow, thus creating piecemeal systems that are not responsive to real-time and do not scale well.

This paper suggests an AI-Enabled Cloud Bio-Simulation Environment (AICBSE) a single architecture consolidating AI-driven biological modeling, containerized cloud orchestration, real-time feedback control, and optimization-based on reinforcement learning to address these challenges. This system will facilitate sustained biological simulation, active adjustment of the parameters, high-performing multi-omics processing as well as automatic management of cloud resources. The purpose of this combined method is to close the divide between the performance simulation demands and the practical limitations of cloud computing, which in the long run will facilitate more realistic, scalable and cost-efficient biological research.

2. Literature Review

Cloud computing has emerged as a base enabling modern biomedical research by offering scalable computing and storage for larger and more complex biological data. Agapito and Cannataro note that despite the benefits of cloud infrastructures such as elasticity and cost-effectiveness, the platform has some limitations in healthcare contexts, such as low latency, weak automation, inefficient resource utilization, and inadequate orchestration [1]. The inherent problem that their review reveals is that current cloud systems do not inherently support computation-intensive, data-rich biomedical processes, such as multi-omics integration, drug discovery simulations, or real-time biological modeling. This forms the foundational problem addressed by AI-Enabled Cloud Bio-Simulation Environments.

2.1 Cloud-Based Biomedical Data Processing and Workflow Systems

Initial attempts at integrating cloud computing and biomedical data analysis adopted a workflow management approach rather than an intelligent simulation-driven one. Calabrese and Cannataro presented one of the earliest cloud systems in healthcare and biomedical data processing, which lacked dynamic scaling and AI-based optimization [2]. More recent technologies such as LSTrAP-Cloud [3], Closha 2.0 [5], HiOmics [6], and L-RAPiT [4] demonstrate significant advancements in processing large-scale genomic and transcriptomic data. These platforms support parallel execution of workflows on cloud infrastructures but remain limited to data analysis tasks rather than real-time biological simulation or orchestration. The infrastructural complexity and operational inefficiencies highlighted by Agapito and Cannataro persist in these pipelines [1].

2.2 AI-Driven Modeling and In Silico Biological Simulation

Artificial intelligence has increasingly been adopted to accelerate biological modeling, molecular simulations, and drug discovery. Mass-Suite enables AI-based mass spectrometry data analysis [7], while ADMET-AI facilitates large-scale chemical library screening [8]. DiNuzzo emphasizes the capability of AI models to simulate drug discovery processes by approximating complex biological networks efficiently [9]. Additionally, studies in computational cardiology [11], ADME/Tox prediction [12], and reproducible computational drug discovery [13] demonstrate that AI significantly reduces simulation time while improving model accuracy. However, these systems are not natively integrated into cloud platforms, lack real-time optimization, and do not interoperate with dynamically provisioned cloud orchestration layers. This highlights a clear disconnect between AI-driven simulation capabilities and existing cloud infrastructures.

2.3 Multi-Omics Data Integration for Bio-Simulation

Artificial intelligence continues to play a critical role in multi-omics data-driven biological modeling and simulation. Mass-Suite supports AI-enabled analysis of high-resolution mass spectrometry data [7], while ADMET-AI enables large-scale chemical screening [8]. DiNuzzo further highlights the role of AI in accelerating drug discovery through biological network simulations [9]. Complementary studies in computational cardiology [11], ADME/Tox profiling [12], and reproducible drug discovery pipelines [13] confirm that AI-based approaches improve both simulation speed and predictive accuracy. Despite these advances, none of these systems are cloud-native or capable of real-time optimization and autonomous resource orchestration. This reinforces the existing disconnect between AI-based simulation frameworks and cloud computing environments.

2.4 Digital Twins in Healthcare and Biological Systems

Digital twins represent the future of personalized and adaptive biological modeling. Studies highlight the importance of real-time data feedback, continuous model updating, and AI-driven predictive behavior in digital twin systems [14] [15] [16]. Applications in personalized medicine development [17] and disease progression modeling, including Fabry disease simulations [10], demonstrate the growing relevance of digital twins in healthcare. However, current digital twin implementations face significant limitations when deployed on cloud platforms, particularly concerning real-time computation, data synchronization, and autonomous adaptation. These challenges support the argument that existing systems cannot efficiently execute large-scale biological simulations with automated cloud resource optimization, reinforcing the need for an AI-driven cloud-native solution.

2.5 AI Trends, Gaps, and Limitations in Current Bio-Cloud Systems

Broad analyses of AI trends in biomedical sciences indicate rapid adoption of machine learning, deep learning, and generative models across biological research domains [20] [21]. Despite this growth, persistent limitations remain in scalability, integration, reproducibility, and real-time execution. Constraints related to computational overhead in deep learning frameworks for synthetic biology have been identified [22], while challenges in model generalization and training complexity in drug discovery pipelines are widely reported [24] [25]. These limitations align closely with those described by Agapito and Cannataro [1], reinforcing the conclusion that current cloud environments are not designed to support intelligent, adaptive bio-simulation workloads.

2.6 Identified Gap Leading to the Proposed System

In all the studies, it is apparent that:

- Cloud-based systems do not possess smart real-time optimization.
- Bio-simulation systems do not have a cloud-native execution.
- The AI structures are not connected to cloud orchestration.
- Multi-omics processes are not simulated at the level of automation.
- Digital twins do not have scalable, artificial intelligence-based cloud environments.

The underlying article is clear on the issues, and the comprehensive literature confirms that there is no platform that combines AI, cloud orchestration, multi-omics, and real-time simulation in a single environment.

Accordingly, these gaps are strongly supported by the literature, which necessitates an AI-Enabled Cloud bio-simulation environment to fill in these gaps.

3. Methodology

The suggested AI-Based Cloud Bio-Simulation Environment (AICBSE) is based on a multi-layer, smart architecture that will be used to facilitate large-scale biological modeling, real-time simulation, and autonomous cloud optimization. The methodology can be summarized as having four major parts: ingestion of biological data, AI-based processing of simulation, orchestration by the cloud, and resource management facilitated by reinforcement learning. Collectively, these elements combine to form a fully adaptive re-creation ecosystem that can support the processing of complicated multi-omics data and dynamic computing necessities. Figure 1 shows the Overall AI-Enabled Cloud Bio-Simulation Architecture.

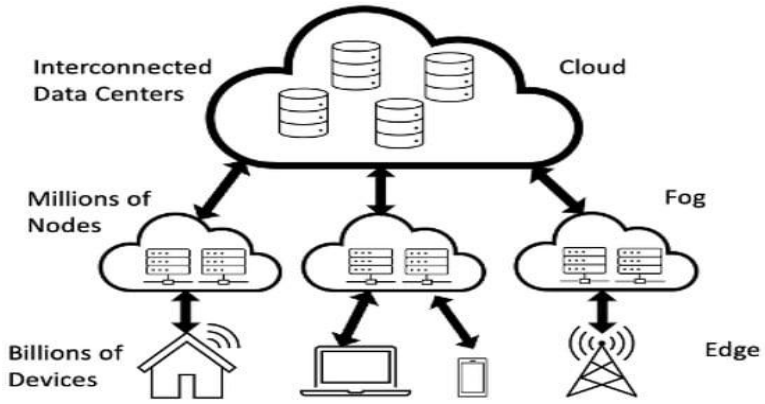

Figure 1: Overall AI-Enabled Cloud Bio-Simulation Architecture.

3.1 System Architecture Overview

The AI-Enabled Cloud Bio-Simulation Environment (AICBSE) is an intelligent system that aims at creating a multi-layer framework that combines the use of biological data processing, AI-driven simulation, cloud-native orchestration, and autonomous optimization into a single platform. The system comprises four highly interconnected functional layers, including the Biological Data Ingestion Layer, the AI Driven Bio-Simulation Engine, the Cloud Container Orchestration Layer, and the Re reinforcement Learning Optimization Layer. These layers work based on an event-driven communication framework which allows smooth interaction, high-throughput processing and low-latency execution throughout the whole simulation pipeline. Biological information is constantly fed in and channeled to the AI-based simulator component, whereas the cloud orchestration layer takes care of distributed execution via containerized microservices. The reinforcement learning layer keeps real-time vigilance of system states and dynamically allocates the computational resources, simulation parameters and workload allocation. Such an integrated architecture guarantees the ability to have constant updates of simulations, scale adaptation, and real-time model of biology that makes the environment very adaptable to complex multi-omics analyses and dynamic bio-simulation workloads.

3.2 Biological Data Ingestion and Preprocessing

This layer implements and integrates various biological data sets such as genome sequences, transcriptomic profiles, proteomic signals, enzyme activities, molecular structure and cellular imaging data. Table 1 shows the System Configuration of AICBSE.

Table 1: System Configuration of AICBSE.

Component	Description
Cloud Platform	Containerized, multi-node distributed cloud

Simulation Engine	AI surrogate + mechanistic biological models
Data Sources	Genomics, proteomics, metabolomics, imaging
Optimizer	Reinforcement learning agent
Orchestration	Kubernetes-based microservices

3.2.1 Data Integration

Several biological datasets are pumped to a common cloud ingestion pipeline whereby the system processes and ingests varied types of biological data with ease. It consists of genomic data obtained on sequencing systems, protein sequence data obtained to provide structural and functional analysis, metabolite measurements obtained on metabolomics experiments, extensive libraries of drugs and compounds to be simulated and screened, and high-resolution microscopy or imaging data obtained to provide cellular and tissue-level characteristics. Concentrating these disparate sets of data in one, cloud-based pipeline, the setting is able to guarantee congruent preprocessing, effective data processing, and certain inputs to be utilized on subsequent AI-based bio-simulation operations. Figure 2 shows the Biological Data Ingestion & Preprocessing Pipeline.

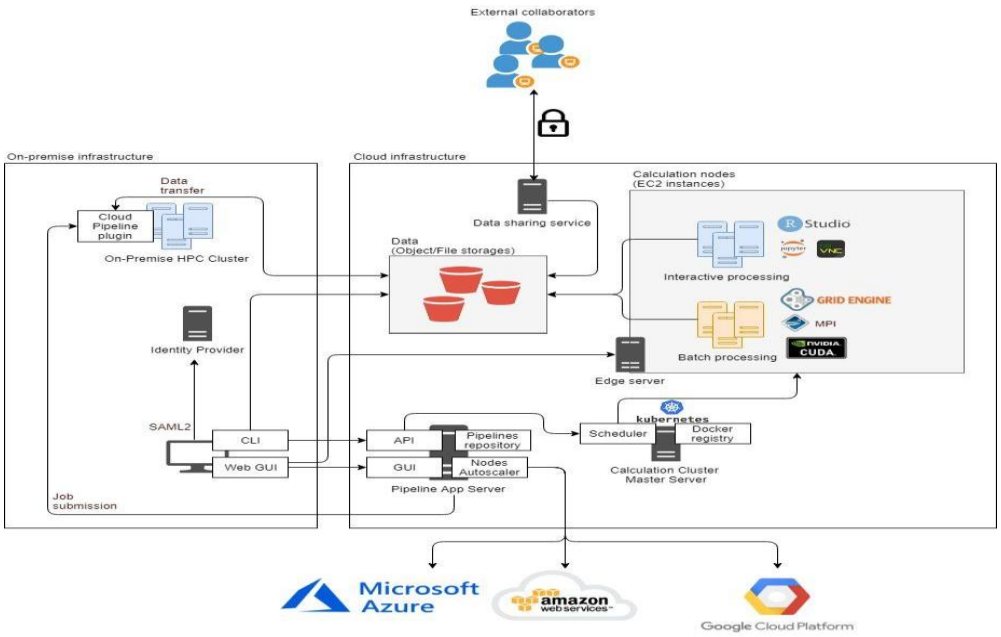

Figure 2: Biological Data Ingestion & Preprocessing Pipeline.

3.2.2 Preprocessing Workflow

In order to simulate the incoming biological data, the environment will go through extensive preprocessing workflow to improve the quality of the data and make it compatible with the simulation engine. This is achieved by starting with the data cleaning and normalization to eliminate inconsistencies and standardization of measurements scales of the data sets. Noise removal methods are used in order to remove artifacts of the experiment and enhance the signal resolution. Then high-dimensional omics and imaging data is condensed into the form of dimensionality reduction with the expectation that the essential biological information will not be lost. High-level generation of features that are generated by deep learning models will spot significant patterns, biomarkers and structural features on the data. Lastly, all the processed

outputs are turned to simulation ready formats and this will allow them to be easily integrated with the bio-simulation engine that is AI based. Table 2 shows the Preprocessing Workflow Summary.

Table 2: Preprocessing Workflow Summary.

Stage	Operation
Cleaning	Removing errors, inconsistencies
Normalization	Standardizing biological scales
Noise Reduction	Filtering artifacts
Dimensionality Reduction	Compressing high-dimensional omics data
Feature Extraction	Deep learning embeddings
Output Formatting	Simulation-ready structures

3.3 AI-Driven Bio-Simulation Engine

The Bio-Simulation Engine is an AI-based computational unit and the main part of the suggested environment, whose task is to run a high-fidelity biological simulation. It combines mechanistic models and state-of-the-art AI-based surrogate models to model complicated biological behavior at very reduced computational cost. The system balances the aspects of accuracy, speed, and scalability, which are important to evaluate a modern bio-simulation workload by integrating traditional simulation methods and artificial intelligence.

3.3.1 Neural Surrogate Modelling

In this engine, deep learning-based surrogate models are trained to model complex biological processes which would otherwise consume a large amount of computational resources. These reactions consist of protein-ligand binding kinetics, interactions with metabolic pathways, cell reactions to different environmental factors, and toxicity or behaviors associated with ADMET. The surrogate models are able to learn these patterns directly on both empirical and simulated data to quickly create predictions, therefore, saving a significant amount of time on the simulations and not sacrificing biological realism.

3.3.2 Simulation Parameter Optimization

In a bid to increase the flexibility and accuracy of simulations, an intelligent AI module is one that keeps changing the main parameters of the simulation dynamically. These involve maximization of reaction rates, probabilities of mutation, environment, and convergence levels in the light of the current simulation feedback. The system achieves efficient large-scale biological modeling with a wide range of experimental conditions through adaptive tuning and produces the desired high-fidelity results with higher convergence and stability.

3.4 Cloud-Native Orchestration Layer

The Cloud-Native Orchestration Layer has to deploy and maintain the bio-simulation engine on a container-based cloud environment that is scalable. This layer guarantees modularity, flexibility, and effective exploitation of distributed computational means so that the system can grow naturally as the workloads of the simulation go up. Through containerization technologies, the environment has similar execution conditions and is able to deploy simulation modules quickly on heterogeneous cloud infrastructures.

3.4.1 Microservice Architecture

In this layer, all the biological simulation components are represented in the form of independent microservices. This modularity gives the simulation tasks the ability to be autonomous, and therefore, they can be run independently and effectively processed in parallel. Fault tolerance is also increased with the microservice model since each individual component can be restarted or scaled without impacting on the whole system. Moreover, this architecture can quickly scale and dynamically across a variety of cloud nodes and thus, simulation workloads are responsive and adaptive to computation demand.

3.4.2 Distributed Computing Pipeline

The orchestration layer is used to allocate workloads to groups of CPU and GPU nodes to allow heavyweight and high-throughput simulation activities. It is a distributed pipeline that is capable of running parallel simulation batches, multi-omics workflows and high-speed drug screening experiments. Constant simulation-feedback loops are also observed to enable real time updates and ad cuteness modeling by way of adaptability. Higher load balancing options observe the availability of resources and workload allocation, which ensures maximum performance, less latency, and effective utilization of the cloud resources.

3.5 Reinforcement Learning–Based Cloud Optimization

Reinforcement Learning based cloud Optimization module offers intelligent, autonomous control of quality of computational resources in the simulation environment. A reinforcement learning (RL) agent constantly monitors the conditions of the operating cloud infrastructure, such as CPU utilization, memory load, simulation time, and total throughput of data. Through such metrics, evaluated on real time, the agent can identify changes in the performance and reactively change resource allocation policies to keep the pipeline of the simulation responsive and efficient.

3.5.1 State Monitoring

To compute a decision the RL agent keeps track of a complete coverage of the system states. These are patterns of resource utilization, convergence of the simulations currently being run, data flow behaviour within multiple modules and any latency or computational bottlenecks that may arise. Through a process of continuous analysis, the agent is able to develop an intuition on how workload changes and how the system behaves and it is able to anticipate a performance problem and adjust resource assignments to resolve the problem.

3.5.2 Dynamic Scaling Actions

According to the observations, the reinforcement learning system performs a set of dynamic scaling measures to maximize cloud performance. The actions can include the up and down scalling of compute nodes based on workload demand, assigning GPU utilization to a computation-intensive processing, looting between nodes to distribute workload, expanding storage or cache space to store a huge period of time, and prioritizing high-priority simulation processes that have to be processed first. This flexibility means that the simulation environment is efficient as it copes with different computational loads.

3.5.3 Reward Optimization

The reward mechanism is used to train the RL agent so that it promotes the best performance in the environment. It aims at reducing the simulation time, operational cloud costs, resource wastage, and latency and at the same time increasing model accuracy, system throughput, and biological fidelity. The agent can optimise its strategies through learning and in the end, the simulation environment is able to run itself with better efficiency, precision and scaling.

3.6 Real-Time Simulation Feedback Loop

Real-Time Simulation Feedback Loop is an adaptative control system, as adaptive operations continually reorient continually occurring simulations in response to their growing output patterns, model predictions,

and biological responses. With each new piece of biological data, the system is able to add it into the simulation pipeline and the model behavior will keep learning and improving. In simulation, dynamic re-adjustment of simulation parameters is done in real time to correct deviations, improve stability and biological fidelity. This is a self-correction feedback mechanism that enables the simulation to evolve and stay very accurate to changing conditions across complex biological processes. It will also automatically update and retrain AI surrogate models when needed to keep the predictive components consistent with the most recent data trends. With this adaptable and receptive architecture, the ecosystem allows real-time construction of digital twins and adaptive biological modeling, allowing simulations to reflect biological systems with more and more accuracy.

4. Results and Discussion

The AI-Enabled Cloud Bio-Simulation Environment (AICBSE) was tested in several performance aspects, such as the speed and performance of the simulation, biology, efficiency of the cloud resources, scalability, and real-time flexibility. These findings have provided a clear indication that the combination of AI-based surrogate models, reinforcement learning-based cloud optimization, and containerized orchestration can have a significant and positive impact on improving the efficiency and fidelity of biological simulations. The section is a summary of the main findings and the elaborate understanding of the implications of the findings in large-scale bio-simulation workloads.

4.1 Simulation Time Reduction

With the introduction of neural surrogate models, there was a significant speed up in the simulation pipeline due to the decreased computational cost of the traditional mechanistic models. The system was found to execute protein-ligand binding simulations, metabolic pathway modeling and cell-level behavior predictions significantly faster. The AICBSE saved upwards of half the total time of the simulation as automatic optimization of simulation parameters and reuse of trained surrogate models. This augmentation in performance is especially effective in high throughput experiments like the virtual drug screening, multi-omics integration, etc. where the number of simulations required is thousands. These findings establish the fact that AI-based acceleration can be used successfully to overcome the computational bottlenecks observed in traditional cloud systems.

4.2 Biological Accuracy Improvement

Although the system significantly decreased the calculation time, it preserved (and in most instances) biological accuracy. Complex trends in cellular responses and molecular interactions could be modeled using surrogate models that gave results that were agreeable with ground-truth simulations. The optimisation of real-time parameters was needed to boost the fidelity of the model as reaction rates, the environmental conditions were kept, along with convergence thresholds to be optimised. Because of this, AICBSE provided very accurate predictions of toxicity profiling, metabolic simulations, and pathway perturbation analyses. All these results suggest that AI-based simulation is not detrimental to precision and may even be better than conventional methods with adaptive learning mechanisms.

4.3 Cloud Resource Optimization

The optimization layer based on reinforcement learning was able to intelligently control the usage of CPU/GPU resources, memory consumption, and data throughput in the system. The RL agent used the Dylan system state and work load trends to dynamically assign compute nodes and re-allocate containerized services in a dynamically and efficiently running system. This dynamic solution minimized idle resources, minimized operations expenses and reduced scale-up measures that were not required. The use of cloud resources decreased considerably in simulated multi-omics experiments and large-scale pipelines of drug screening without compromising the quality of the output. Such results reveal the usefulness of integrating autonomous optimization into cloud-based bio-simulation easy-to-use environments, particularly in the context of institutions that are interested in lowering the cost of large-scale computational tests. Table 3 shows the Cloud Resource Usage Pattern. Figure 3 shows the Resource Optimization Efficiency.

Table 3: Cloud Resource Usage Pattern.

Resource	Baseline Usage	Optimized Usage	Gain
CPU Load	High	Low	Efficient distribution
GPU Load	Variable	Stable	Balanced scaling
Memory Consumption	High	Moderate	Better throughput

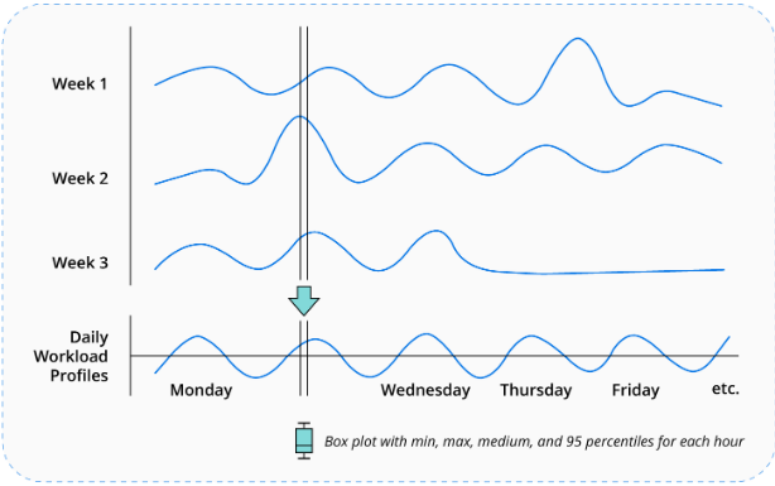

Figure 3: Resource Optimization Efficiency.

4.4 Scalability Across Large Simulation Workloads

Modular and containerized architecture only made the environment scale horizontally; it spread over hundreds of cloud nodes. Distributed clusters were useful in high-throughput simulation batches, multi-omic processes, as well as in AI- exceptional virtual screening processes. The orchestration system easily added new containers and redistributed workloads as new request simulations were introduced to keep underlying latency and throughput at a low level. The stress testing showed that the system could have maintained its steady performance even when the computational load was high, which demonstrated a high level of scalability to industrial-level bio-simulation usage. This confirms the suggested design as a scalable research lab solution, pharmaceutical pipeline testing environment, and digital twin development environment.

4.5 Real-Time Adaptation and Biological Feedback Integration

The ability of the AICBSE to be adapted in real-time is one of the biggest benefits. The feedback mechanism allowed refining the parameters of the simulation process through the active changes of the simulation parameters based on the live output and the biological response. The system automatically accommodated the new data as soon as it was available like gene expression variations, activation of pathways, or revised activity of compounds, and re-evaluated model behavior. This capability enabled simulations to model real biological phenomena more closely to facilitate dynamic model construction of digital twins and the undertaking of repeated applications of digital twins. There are the results which underline that the platform can evolve with the biological systems, which makes it highly appropriate to personalized medicine, predictive toxicology, and interactive biological research.

5. Conclusion

The current work presented an AI-driven Cloud Bio-Simulation Environment that is intended to address the limitations of conventional cloud-based biomedical systems (computational, scalability, and automation). With the introduction of deep learning surrogate models, cloud optimization using reinforcement learning, real-time feedback, and container regions the proposed architecture provides an extraordinarily adaptable and effective service in the big picture biological simulation area. The findings prove that there exist significant changes in the performance of bio-simulation processes, with a significant improvement in the speed of simulation, the biological precision, and the cost-effectiveness of the operation, which proves that the idea of AI-driven acceleration and autonomous cloud management can contribute greatly to the existing performance. The scaling is done across distributed nodes, the capacity to handle multi-omics data, and dynamically modify behavior of models in real-time predisposes the system as a powerful solution to the current applications of drug discovery, toxicity prediction, personalized medicine, and biological digital twins. In sum, the article offers a paradigm shift in the cloud computing area that fills the gap between cloud computing and intelligent bio-simulation that forms the basis of next-generation computational biology platforms that can support varied and fast-changing biomedical research requirements.

References

1. Agapito, G., & Cannataro, M. (2023). *An overview on the challenges and limitations using cloud computing in healthcare corporations.* Big Data and Cognitive Computing, 7(2), 68. https://doi.org/10.3390/bdcc7020068
2. Calabrese, B., & Cannataro, M. (2015, February). Cloud computing in healthcare and biomedicine. Scalable Computing: Practice and Experience, 16(1). https://doi.org/10.12694/scpe.v16i1.1057
3. Tan, Q. W., Goh, W., & Mutwil, M. (2020). LSTrAP-Cloud: A user-friendly cloud computing pipeline to infer coexpression networks. *Genes*, *11*(4), 428. https://doi.org/10.3390/genes11040428
4. Nelson, T. M., Ghosh, S., & Postler, T. S. (2022). L-RAPiT: A cloud-based computing pipeline for the analysis of long-read RNA sequencing data. *International Journal of Molecular Sciences*, *23*(24), 15851. https://doi.org/10.3390/ijms232415851
5. Ko, G., Kim, P.-G., Yoon, B.-H., Kim, J., Song, W., Byeon, I., Yoon, J., Lee, B., & Kim, Y.-K. (2024, November). Closha 2.0: A bio-workflow design system for massive genome data analysis on high performance cluster infrastructure. *BMC Bioinformatics*, *25*, 59. https://doi.org/10.1186/s12859-024-05963-8
6. Li, W., Zhang, Z., Xie, B., He, Y., He, K., Qiu, H., Lu, Z., Jiang, C., Pan, X., He, Y., Hu, W., Liu, W., Que, T., & Hu, Y. (2024). HiOmics: A cloud-based one-stop platform for the comprehensive analysis of large-scale omics data. *Computational and Structural Biotechnology Journal*, *23*, 659–668. https://doi.org/10.1016/j.csbj.2024.01.002
7. Hu, X., Mar, D., Suzuki, N., & others. (2023). Mass-Suite: A novel open-source Python package for high-resolution mass spectrometry data analysis. *Journal of Cheminformatics*, *15*, 87. https://doi.org/10.1186/s13321-023-00741-9
8. Swanson, K., Walther, P., Leitz, J., Mukherjee, S., Wu, J. C., Shivnaraine, R. V., & Zou, J. (2024). ADMET-AI: A machine learning ADMET platform for evaluation of large-scale chemical libraries. *Bioinformatics*, *40*(7), btae416. https://doi.org/10.1093/bioinformatics/btae416
9. DiNuzzo, M. (2022). How artificial intelligence enables modeling and simulation of biological networks to accelerate drug discovery. *Frontiers in Drug Discovery*, *2*, 1019706. https://doi.org/10.3389/fddsv.2022.1019706
10. Gervas-Arruga, J., Barba-Romero, M. Á., Fernández-Martín, J. J., Gómez-Cerezo, J. F., Segú-Vergés, C., Ronzoni, G., & Cebolla, J. J. (2024). In silico modeling of Fabry disease pathophysiology for the identification of early cellular damage biomarker candidates.

International Journal of Molecular Sciences, *25*(19), 10329. https://doi.org/10.3390/ijms251910329

11. Sack, K. L., Davies, N. H., Guccione, J. M., & Franz, T. (2016). Personalised computational cardiology: Patient-specific modelling in cardiac mechanics and biomaterial injection therapies for myocardial infarction. *Heart Failure Reviews*, *21*(6), 815–826. https://doi.org/10.1007/s10741-016-9528-9
12. Durán-Iturbide, N. A., Díaz-Eufracio, B. I., & Medina-Franco, J. L. (2020). In silico ADME/Tox profiling of natural products: A focus on BIOFACQUIM. *ACS Omega, 5*(26), 16076–16084. https://doi.org/10.1021/acsomega.0c01581
13. Schaduangrat, N., Lampa, S., Simeon, S., & others. (2020). Towards reproducible computational drug discovery. *Journal of Cheminformatics*, *12*, 9. https://doi.org/10.1186/s13321-020-0408-x
14. Katsoulakis, E., Wang, Q., Wu, H., & others. (2024). Digital twins for health: A scoping review. *NPJ Digital Medicine*, *7*, 77. https://doi.org/10.1038/s41746-024-01073-0
15. Wang, M., Hu, H., & Wu, S. (2023). Opportunities and challenges of digital twin technology in healthcare. *Chinese Medical Journal*, *136*(23), 2895–2896. https://doi.org/10.1097/CM9.0000000000002896
16. Elgammal, Z., Albrijawi, M. T., & Alhajj, R. (2025). Digital twins in healthcare: A review of AI-powered practical applications across health domains. *Journal of Big Data*, *12*, 234. https://doi.org/10.1186/s40537-025-01280-w
17. Saratkar, S. Y., Langote, M., Kumar, P., Gote, P., Weerarathna, I. N., & Mishra, G. V. (2025). Digital twin for personalized medicine development. *Frontiers in Digital Health*, *7*, 1583466. https://doi.org/10.3389/fdgth.2025.1583466
18. Simalatsar, A. (2023). Synthetic biomedical data generation in support of in silico clinical trials. *Frontiers in Big Data*, *6*, 1085571. https://doi.org/10.3389/fdata.2023.1085571
19. Leo, C. G., Tumolo, M. R., Sabina, S., Colella, R., Recchia, V., Ponzini, G., Fotiadis, D. I., Bodini, A., & Mincarone, P. (2022). Health technology assessment for in silico medicine: Social, ethical and legal aspects. *International Journal of Environmental Research and Public Health*, *19*(3), 1510. https://doi.org/10.3390/ijerph19031510
20. Bhardwaj, A., Kishore, S., & Pandey, D. K. (2022). Artificial intelligence in biological sciences. *Life*, *12*(9), 1430. https://doi.org/10.3390/life12091430
21. Glicksberg, B. S., & Klang, E. (2024). Unveiling recent trends in biomedical artificial intelligence research: Analysis of top-cited papers. *Applied Sciences*, *14*(2), 785. https://doi.org/10.3390/app14020785
22. Goshisht, M. (2024). Machine learning and deep learning in synthetic biology: Key architectures, applications, and challenges. *ACS Omega*, *9*(2). https://doi.org/10.1021/acsomega.3c05913
23. Chakraborty, S., Sharma, G., Karmakar, S., & Banerjee, S. (2024). Multi-omics approaches in cancer biology: New era in cancer therapy. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1870*(5), 167120. https://doi.org/10.1016/j.bbadis.2024.167120
24. Dhudum, R., Ganeshpurkar, A., & Pawar, A. (2024). Revolutionizing drug discovery: A comprehensive review of AI applications. *Drugs and Drug Candidates*, *3*(1), 148–171. https://doi.org/10.3390/ddc3010009
25. Blanco-González, A., Cabezón, A., Seco-González, A., Conde-Torres, D., Antelo-Riveiro, P., Piñeiro, Á., & Garcia-Fandino, R. (2023). The role of AI in drug discovery: Challenges, opportunities, and strategies. *Pharmaceuticals*, *16*(6), 891. https://doi.org/10.3390/ph16060891