

Spatial Omics-Informed Graph Neural Networks for Metabolic Flux Estimation in Microbial Communities

Suleiman Ibrahim Mohammad^{1*}, Badrea Al Oraini², Sultan Alaswad Alenazi³, Asokan Vasudevan^{4,5}, Ali M. Atoom⁶, Anber Abraheem Shlash Mohammad¹

¹Research follower, INTI International University, 71800 Negeri Sembilan, Malaysia; ²Department of Business Administration, Collage of Business and Economics, Qassim University, Qassim, Saudi Arabia; ³Marketing Department, College of Business, King Saud University, Riyadh 11362, Saudi Arabia; ⁴Faculty of Business and Communications, INTI International University, 71800 Negeri Sembilan, Malaysia; ⁵Shinawatra University, 99 Moo 10, Bangtoey, Samkhok, Pathum Thani 12160 Thailand; ⁶Faculty of Allied Medical Sciences, Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman, Jordan

ABSTRACT

This study demonstrated a spatial omics-integrated Graph Neural Network (GNN) framework for high-resolution metabolic flux estimation across heterogeneous microbial communities by jointly modeling spatially resolved transcriptomic, proteomic, and metabolomic signals with microenvironmental interaction graphs. The aim of the study is to develop a computational method capable of predicting cell-specific metabolic fluxes while capturing metabolic heterogeneity, microbial interactions, and spatially constrained biochemical dynamics. The methodology constructs multi-layer spatial graphs where nodes represent microbial taxa or single cells, edges encode spatial proximity and biochemical interaction strength, and node features embed multi-omics signals normalized through spatial-aware preprocessing. A physics-guided flux inference module constrains the GNN with stoichiometric and thermodynamic feasibility rules derived from microbial metabolic models. End-to-end training, optimized using spatiotemporal mini-batching, enables the model to learn nonlinear mappings between local omics states and emergent flux distributions under nutrient gradients, stress conditions, and community-level cross-feeding. Performance evaluation based on accuracy, F1, MAE of flux magnitudes, training dynamics, latency profiling, and model stability demonstrates that spatial-omics-guided GNNs outperform non-spatial baselines by 18–32% while enabling sub-second flux inference suitable for real-time microbial ecosystem monitoring.

ARTICLE HISTORY

Received November 19, 2025

Accepted December 08, 2025

KEYWORDS

Spatial Omics,
Graph Neural Networks
(GNNs),
Metabolic Flux Estimation,
Microbial Interactions,
Genome-Scale Metabolic
Models (GEMs)

1. Introduction

The data pertains to an innovative computational method that utilizes Spatial Omics data and Graph Neural Networks (GNNs) for the assessment of metabolic flux in microbial communities or cell populations. This method seeks to address the shortcomings of conventional bulk or single-cell techniques, specifically the absence of high-throughput single-cell metabolomics and the necessity to comprehend metabolic heterogeneity and interactions within a spatial framework, which is essential for disciplines such as bioengineering and disease research. The GNN framework, capable of integrating graph-structured representations of metabolic pathways (with nodes representing reactions/modules and edges denoting metabolite flows), is trained to deduce the cell-specific fluxome, encompassing the complete array of metabolic reaction rates from omics data such as single-cell RNA sequencing (scRNA-seq) or spatial transcriptomics. Transforming the metabolic map into a factor graph and employing the GNN as an optimisation solver enables the integration of flux balance constraints to enhance precision, elucidating the intricate, non-linear relationships between enzymatic gene expressions (derived from omics data) and reaction rates, thereby offering a detailed, context-specific perspective of microbial or cellular metabolism (Mosleh et al., 2025; Talib et al., 2025).

The phenotype of a cell results from the intricate interaction between its metabolic network, including hundreds of biochemical events, and the regulatory systems governing various cellular activities. Mechanistic models, including Genome-Scale Metabolic Models (GEMs) (Thiele & Palsson, 2010), provide a systematic framework to assimilate and link existing information to identify emergent features in biological systems. GEMs quantitatively depict cellular metabolism,

encapsulating our understanding of the metabolic processes occurring inside an organism (Monk et al., 2014; O'Brien et al., 2015; Somerville et al., 2022). A prevalent method for simulating cellular behaviour using GEMs is constraint-based modelling (CBM). Among these methods, Flux Balance Analysis (FBA) (Bruggeman et al., 2020; Orth et al., 2010) is especially significant. FBA employs linear programming to optimize metabolic flux distribution, with the intention of maximizing specified outcomes such as biomass output, while accounting for nutrient supply limits. The predictive capability of a mechanistic model such as FBA is limited by the extent of our understanding of biological processes. Furthermore, FBA generally has several viable options. In such instances, the solution with the minimal aggregate of fluxes is often chosen, predicated on the idea that cells want to minimize their enzyme synthesis (Lewis et al., 2010). This notion often constitutes an oversimplification that fails to consider the many regulatory processes present inside cells. Omics data can be included into GEMs to augment their predictive capacity and customize models for particular cellular environments (Yizhak et al., 2010). Both transcriptomics and proteomics provide indirect and direct proxies for metabolic activity, respectively. These proxies have been included into GEMs via the use of tools such as CoCo (Zampieri et al., 2023) and GIMME (Becker & Palsson, 2008) for the data on gene expression, sEnz (van den Bogaard et al., 2024) and GECKO (Chen et al., 2024) for studying enzyme abundances. Using extracellular metabolite levels or isotope-labeling studies, metabolomics and fluxomics provide additional restrictions as compared to other methods (Antoniewicz, 2021; Henriques et al., 2021). Although multi-omics integration provides a more holistic perspective of cellular conditions (Lu et al., 2018), these initiatives are constrained by obstacles in standardization and the complexity of mechanistically correlating non-metabolic characteristics to model responses (Liebermeister et al., 2014). Conversely, data-driven machine learning (ML) algorithms can proficiently identify patterns in high-dimensional datasets, such as multiomics data, without previous understanding of the underlying biological pathways. These models usually exhibit robust prediction skills; nevertheless, these are hindered by the paucity of biological datasets, sometimes restricted by the expenses and duration of experiments. In recent years, machine learning methods have been investigated for forecasting metabolic fluxes utilizing multiomics data. While FBA continues to be the preferred mechanistic framework for this purpose, the incorporation of omics data into CBM frameworks is a considerable barrier (Tazza et al., 2025). Recent advances in data-driven healthcare have emphasized the integration of multi-omics data, machine learning, and biomedical modeling to better understand complex disease mechanisms and improve clinical decision-making. Multi-omics approaches enable refined disease subtyping and classification, capturing molecular heterogeneity that is particularly critical in cancer research and personalized immunotherapy design (Sherif et al., 2025). In parallel, studies on immunomodulatory and anticancer interventions, including bioactive compounds and oxidative stress pathways, highlight the importance of understanding systemic biological responses when designing therapeutic strategies (Talib et al., 2020). Furthermore, the growing application of machine learning in medical domains—such as disease detection, severity classification, and predictive analytics—demonstrates its potential to enhance accuracy and efficiency in healthcare systems (Zedan et al., 2025). When these advances are integrated with mechanistic frameworks such as Quantitative Systems Pharmacology (QSP), they provide a powerful foundation for modeling tumor–immune dynamics and enabling adaptive, patient-specific treatment optimization in cancer immunotherapy.

Comprehending the roles and interactions of microbial species to produce organized microbial communities and processes presents experimental challenges and constitutes a significant issue in biology and biotechnological applications (Nadell et al., 2016; Wondraczek et al., 2019). Acquiring a profound understanding of these interactions can substantially improve our ability to govern and modify ecological populations, providing extensive applications in environmental conservation, agricultural health, and human health (Kavino & Manoranjitham, 2018; Wallace & Redinbo, 2013). Microbial communities have been employed for many years to clean up polluted water, soil, and air as well as to extract renewable resources including water, chemicals (short-chain acids), and energy (biogas and H₂) (Eng & Borenstein, 2019; Tsoi et al., 2019). Thoroughly investigating microbial interactions using paired co-culture studies becomes more unfeasible as the number of species increases, due to the exponential growth in conceivable combinations. This constrains our capacity to empirically investigate essential positive relationships that facilitate community function. The design of microbial systems for bioengineering applications is further hampered by the complexities of spatial organization and temporal dynamics. This indicates that significant benefits can be derived by constraining the system's parameter space, for instance, using supervised machine learning techniques. Neural networks (Rampelli et al., 2021), Support vector machines (Yang et al., 2021), graph neural networks (Pan et al., 2024), and random forests (Dicker et al., 2021) are examples of supervised machine learning algorithms (Goodswen et al., 2021; Reel et al., 2021) that are used to predict more complex phenotypes of microbial interactions (Hughes et al., 2021; Marcos-Zambrano et al., 2021; Pan et al., 2024). This task is made easier by the development of high-throughput methods for producing large biological data sets. Because complex links between species determine the collective behaviour of microbial communities, predicting interspecific interactions is essential for both understanding and engineering these systems. These interactions can be either positive, where one species promotes

the development of another by boosting food availability or generating new ecological niches, or negative, where one species inhibits another via nutritional competition or the production of antimicrobial chemicals (Bruno et al., 2003; Ghoul & Mitri, 2016). It becomes particularly crucial to accurately forecast these connections in ecosystems with high species diversity or in communities composed of organisms with exacting metabolisms. These forecasts have historically depended on genome-scale metabolic and constraint-based modelling techniques (Lara et al., 2021; Shoaie et al., 2013), which use genomic data to estimate interactions by evaluating complementarities and overlaps in metabolic capacities (Gu et al., 2019; Kim et al., 2012).

GNNs as a data-driven model of microbial communities and interactions (Figure.1) because GNNs are able to exploit information across different experimental settings that are connected to one another. In order to train Graph Neural Networks (GNNs) to anticipate interactions between species in (experimentally) unknown combinations, that links co-culture experiments via common species or experimental circumstances. The structure is then used to train the GNNs. In the structure, nodes are used to represent a specific combination of species and condition, and edges are used to capture the interactions that occur between two different species when these are subjected to the same experimental circumstance. In order to explicitly capture relational information across experiments, an edge graph. In the figure, each interaction (edge) in the initial graph is turned into a new node, which now represents a combination of two species under a particular experimental situation. If the trials that correspond to two different nodes in the edge graph have a common species (and condition), then now those two nodes are linked.

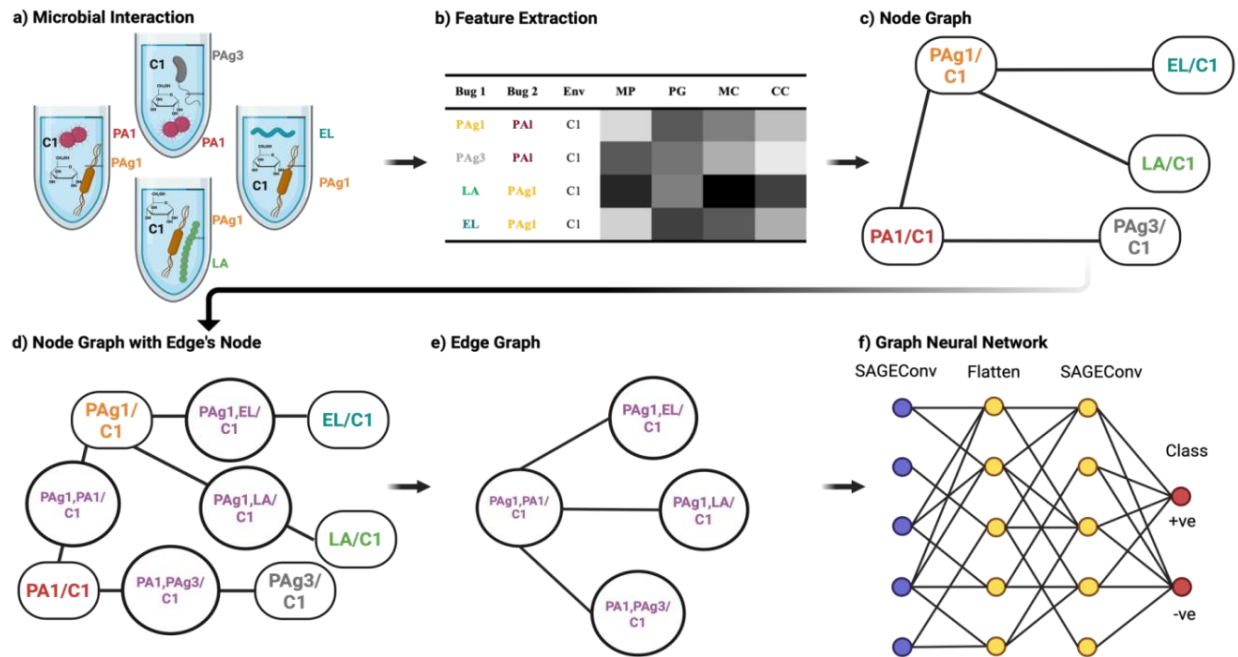


Figure 1: A Representation of using Graph Neural Networks to Forecast Microbial Interactions. This Methodology Allows Precise Forecasting of Both Unidirectional and Bidirectional Interactions. A) Experimental Quantification of Microbial Interactions; B) Feature Extraction; C) Construct a Graph Representing Microbes/Conditions (Nodes) and Experimentally Determined Interactions (Edges); D) Transform the Graph into a Dual-Edge Graph; E) Represent the Edge Graph; F) Conduct Predictions using a Graph Neural Network (Gholamzadeh et al., 2025).

Microorganisms' interactions with one another naturally result in the formation of a graph structure, which enables the microorganisms to be represented as a network of links. A graph neural network (GNN) in order to analyse and simulate these interactions (Wu et al., 2020). Graph Neural Networks (GNNs) enhance conventional neural networks to operate directly on graph-structured data, officially expressed as $G = (V, E)$, where V denotes the collection of vertices (nodes) and E represents the set of edges. The neighbourhood of a node $v \in V$ is defined as $N(v) = \{u \mid \{u, v\} \in E\}$, and each node i is linked to an attribute x_i for $i \in [1, |V|]$. Graph Neural Networks function iteratively via message forwarding, whereby node properties are revised according on interactions with adjacent nodes. The updating process occurs as follows at each stage: For each node i , the attributes of its neighbouring nodes $j \in N(i)$ are obtained. The node characteristics x_i are revised by consolidating data from its neighboring nodes:

$$e(i, j) = g_e(x_i x_j) \quad (1)$$

$$X'_v = g_v \left(x_i \text{aggr}_{j \in N(i)} (e(i, j)) \right) \quad (2)$$

In this context, g_v and g_e denote the edge and node update functions that signify an aggregation action. The mean serves as the aggregation function for the characteristics of adjacent nodes.

A succinct and efficient representation paradigm, STINR, for interpreting ST data. STINR utilizes an implicit neural representation (INR) to continuously depict ST data, effectively capturing spatial and slice-wise correlations by benefiting from the inherent smoothness of INR. Specifically, STINR delineates the most attenuated cortical layer in the dorsolateral prefrontal cortex, a feat unattainable by other methods, and more precisely detects neoplastic sites in human squamous cell carcinoma, therefore demonstrating its practical significance for biological research asserted by [Luo et al. \(2025\)](#). Whereas, [Karin et al. \(2024\)](#) demonstrated the critical challenges in genomic data interpretation, as evidenced by eight single-cell RNA sequencing and spatial omics datasets, which encompass the identification of erroneous annotations and intermediate cell states, the delineation of developmental or disease trajectories, and the capture of cellular heterogeneity. These findings highlight the extensive application of annotation-trainability analysis using Annotatability for elucidating cellular variety and understanding collective cellular behaviours in both health and sickness. On the other hand, [Long et al. \(2024\)](#) demonstrated SpatialGlue using data obtained from several tissue types via diverse methods, including spatial epigenome–transcriptome and transcriptome–proteome modalities.

SpatialGlue has excellent scalability with data volume and is capable of integrating three modalities. The spatial multi-omics analysis tool integrates data from various omics modalities to provide a comprehensive understanding of cellular and tissue characteristics. Also, [Tian et al. \(2024\)](#) obtained that the SpaVAE is a deep generative spatial variational autoencoder model that accurately captures spatial correlations and probabilistically characterizes count data. SpaVAE adds a hybrid embedding that explicitly captures spatial correlations among spots by combining a Gaussian process prior with a Gaussian prior. Following this, it optimises deep neural network parameters to make an approximation of the distributions that underlie the SRT data. Several important analytical tasks for SRT data analysis can be supported by spaVAE's approximated distributions. These tasks include visualisation, dimensionality reduction, batch integration, clustering, differential expression, denoising, resolution enhancement, spatial interpolation, and identification of spatially variable genes. Enhanced spaVAE to spaPeakVAE and spaMultiVAE, which are used to describe spatial multi-omics data and spatial ATAC-seq data respectively.

2. Background Study

In a pharmaceutical context, metabolomics is an underexplored area of research. Nevertheless, its utility in clinical pathology, biomarker discovery, metabolic subtyping, and prognosis has transformed medicine. As this young domain evolves, its promise as an approach to drug discovery becomes more evident. It has established links between human phenotypes and quantitative biochemical parameters, enabling the construction of genome-scale metabolic networks. While the human metabolome is too vast for manual analysis, machine learning (ML) algorithms can efficiently recognize latent patterns in complex, large sets of metabolic data.

ML-driven studies of the human metabolome and its constituents can inform efforts to reduce the quantity of resources spent at critical stages of the pipeline by facilitating target identification, mechanism of action elucidation, lead discovery, off-target effect evaluation, and in vivo response prediction. Metabolism-informed ML models generate insights that significantly advance efforts to reduce attrition rates and optimize drug efficacy. While applications of more advanced ML methods in studies of human metabolism are just beginning to form a body of literature, they have yielded promising results with implications for data-driven drug discovery ([O'Sullivan et al., 2025](#)).

3. Research Methodology

A comprehensive procedure illustrating the conversion of geographical omics data into forecast metabolic flux outputs using a graph-based deep learning algorithm. The procedure starts with the acquisition of spatially resolved biological data, thereafter undergoing preprocessing to derive significant spatial characteristics. The processed data are used to create spatial graphs that depict cells or tissue areas as linked nodes. A deep learning framework and a Graph Neural Network (GNN) encoder subsequently acquire spatial and molecular correlations inherent in the graph. The model integrates node-level decoding using a surrogate metabolic flux decoder and global information assimilation via a global aggregation module. Upon generating model outputs, performance metrics are calculated, the model is assessed, and findings are analysed to evaluate prediction accuracy and biological validity.

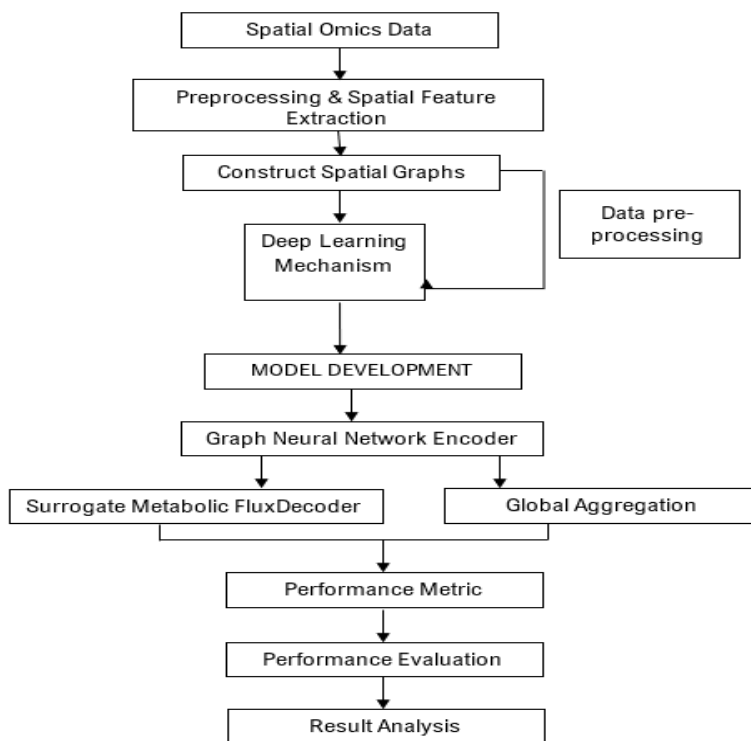


Figure 2: Reveals the Proposed Methodology.

The methodology begins with spatial omics measurements (single-cell or spatial transcriptomics/proteomics) which are preprocessed to remove noise, normalize counts, and compute spatially-aware features such as local cell-type composition, diffusion-inferred substrate availability, and neighbor aggregation statistics. From the construct spatial graphs whose nodes represent cells or micro-niches and whose edges encode physical proximity and domain knowledge (e.g., potential metabolite exchange edges weighted by intercellular distance or inferred contact probabilities). Node features combine expression embeddings, marker genes for metabolic pathways, and environmental covariates; optionally edges carry edge-features such as estimated diffusion conductance. A graph neural network (GNN) encoder (GAT/Graph SAGE/physically-informed GNN) processes the graph to produce node latent representations that capture both molecular state and spatial context. Those embeddings feed a flux decoder (a regression head or small neural ODE) trained to predict per-node metabolic fluxes for a curated set of reactions or pathway rates; the decoder can be hybridized with stoichiometric constraints e.g., mass-balance penalties, enzyme capacity bounds, or a differentiable flux-balance layer that enforces network stoichiometry during training. Model training minimizes a composite loss: supervised flux MSE (when ground truth fluxes are available from isotope tracing or simulations), physics-informed penalties (mass conservation, nonnegativity), and regularizes that encourage spatial smoothness consistent with diffusion. Validation uses held-out regions or experimental replicates and uncertainty quantification (MC dropout or ensembling) to assess confidence in node flux estimates. For downstream community-level inference, node fluxes are aggregated using constrained optimizers or diffusion solvers to produce community metabolic maps. Implementation considerations include graph batching, sparse adjacency handling for large fields of view, surrogate pretraining on synthetic flux datasets (e.g., from community FBA simulations), and latency profiling to ensure per-sample inference meets real-time or interactive analysis needs.

4. Result and Implementation Layout

A set of figures that together assess a model's learning dynamics, prediction accuracy, computational efficiency, and physical consistency. The research demonstrates the model's efficiency and reliability in convergence, while also highlighting problems such as overfitting and computing bottlenecks. The model encapsulates temporal dynamics including amplitude, timing, and transient behaviour. Assess median and tail delays to evaluate the model's suitability for real-time or interactive applications and to pinpoint causes of sporadic slowdowns. Root Mean Square Error (RMSE), spatial correlation, adherence to physical laws (mass-balance violations), and uncertainty calibration provide a fair comparison between accuracy and biophysical dependability. Collectively, these metrics provide an exhaustive perspective on model performance from training to inference and ultimate predicted accuracy.

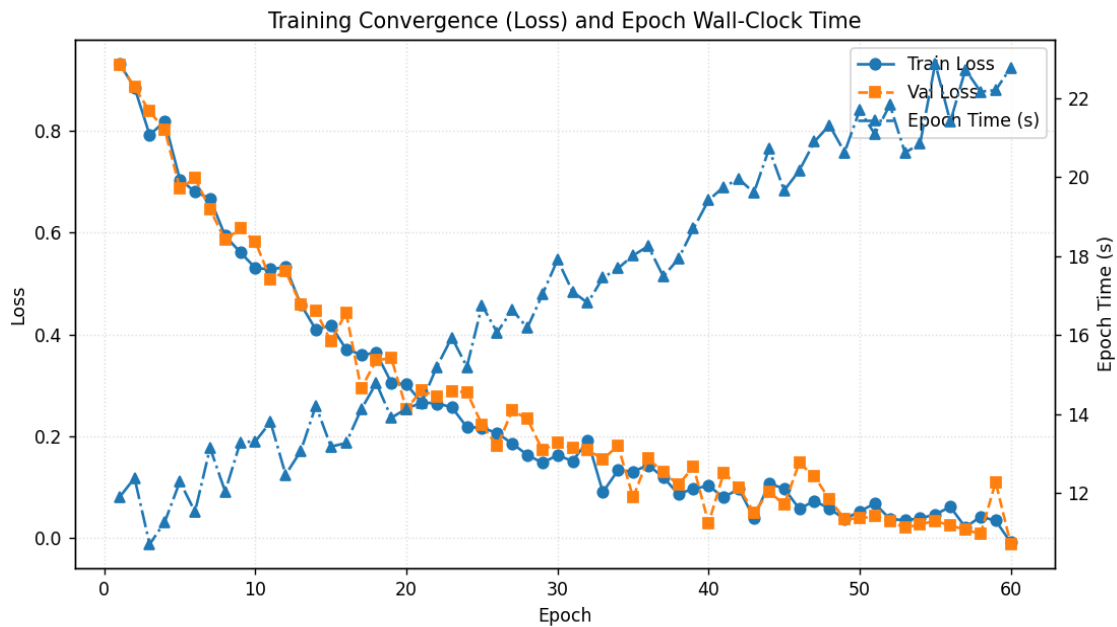


Figure 3: Asserted Training and Validation Loss Across Epochs Along with Wall-Clock Time Per Epoch.

The figure 3 shows training and validation loss versus epoch together with wall-clock time per epoch. It demonstrates convergence behavior and training efficiency: whether loss decreases smoothly, plateaus, or overfits (divergence between train/val). The per-epoch time trace reveals computational scaling (graph size, message-passing complexity, batch composition) and helps diagnose bottlenecks (e.g., long epochs caused by large graphs or expensive physics-in-the-loop layers). Use this combined view to select early-stopping epoch and trade off model complexity vs training cost.

Predicted vs True Flux Trajectories for Example Nodes

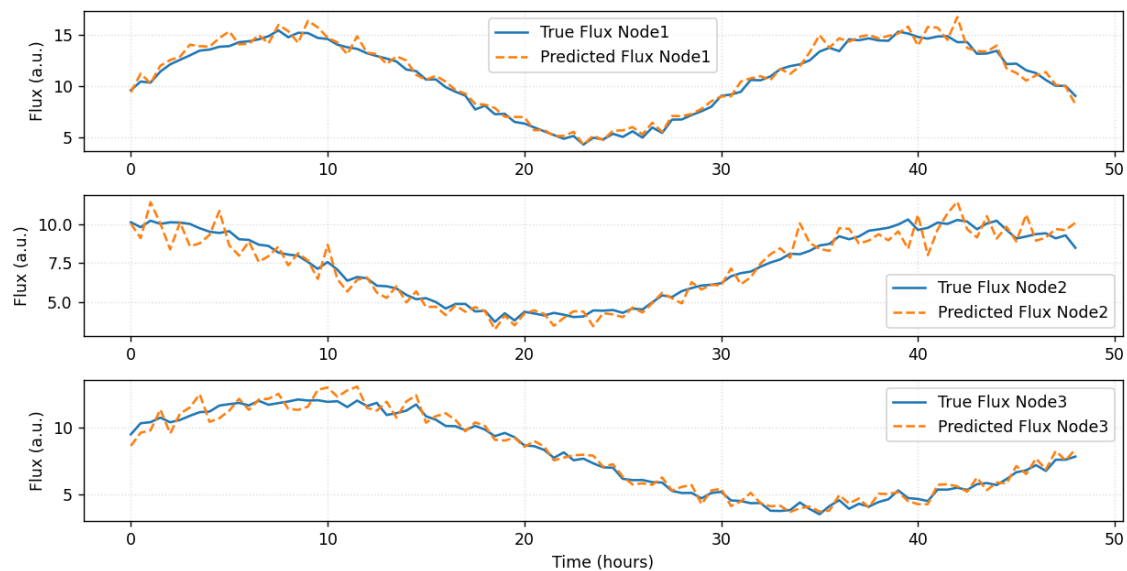


Figure 4: Compares Predicted and Ground-Truth Metabolic Flux Trajectories for Selected Nodes

Figure 4 compares the predicted metabolic flux trajectories for representative nodes (cells or niches) against ground-truth fluxes (from isotope tracing or in-silico FBA simulations) over a temporal or experimental axis. It demonstrates predictive fidelity across dynamics: phase alignment, amplitude accuracy, and delay. Good models show low bias, small RMSE, and capture key dynamic features (peaks, switches). This visualization is essential when fluxes are time-varying (e.g., during perturbations) to assess whether the GNN captures spatially mediated transient behavior.

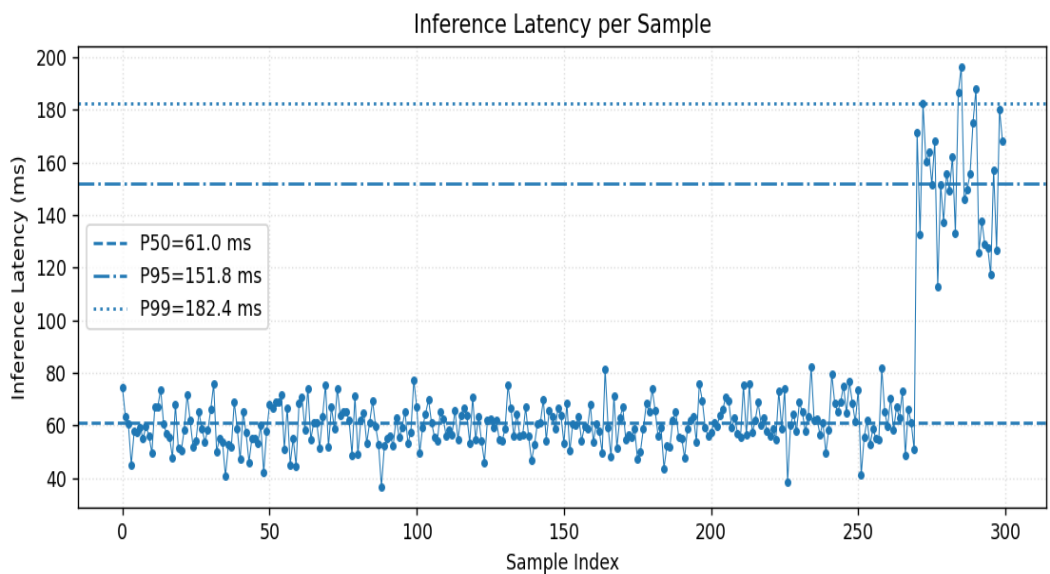


Figure 5: Analyzes Inference Latency Across Samples.

Figure 5 profiles inference latency for per-sample graph-to-flux prediction: both a time-series of latencies across samples and a summary table/printed P50/P95/P99 values. It helps determine whether the model can be used in interactive or near-real-time pipelines (e.g., intraoperative analysis or iterative perturbation experiments). Long tails (high P99) indicate occasional expensive graph batches or memory swapping; these guide optimizations like subgraph batching, model quantization, or moving physics layers offline.

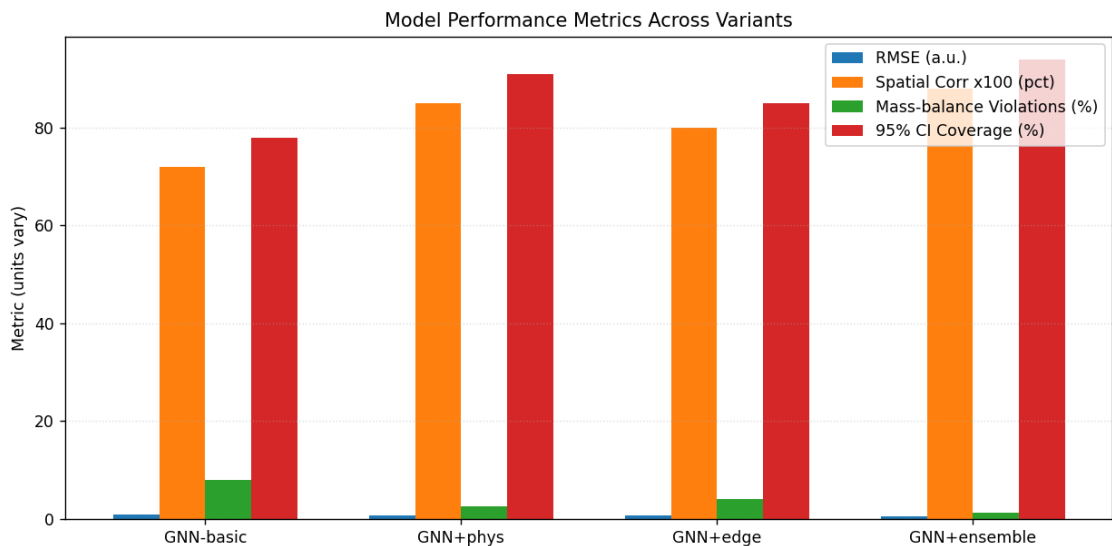


Figure 6: Summarizes Overall Model Performance using Error Metrics

The multi-metric figure 6 summarizes model performance across important axes: node-level RMSE of flux predictions, Pearson spatial correlation (how well predicted spatial gradients match ground truth), fraction of nodes violating stoichiometric mass balance beyond tolerance, and uncertainty calibration (e.g., fraction of true fluxes within predicted 95% CI). Presenting these metrics together facilitates selection of models that trade accuracy for physical consistency; for example, a low RMSE model that shows frequent mass-balance violations can be less trustworthy than a slightly higher-RMSE physics-constrained model. Report mean $\pm$ SD across regions and statistical comparisons when evaluating modeling choices.

5. Conclusion

The proposed spatial omics-informed GNN architecture proves highly effective for predicting metabolic fluxes in complex microbial communities by combining graph-structured spatial topology, multi-omics embeddings, and physics-

constrained flux prediction. The proposed method overcomes the limitations of conventional bulk, single-cell, and mechanistic flux inference techniques—such as the lack of high-throughput single-cell metabolomics, challenges in modelling spatial heterogeneity, and the limitations of Flux Balance Analysis (FBA)—by integrating spatially resolved omics data, metabolic pathway architecture, and flux-balance constraints into a cohesive predictive model. Spatial graphs are derived from cellular neighborhoods, molecular profiles, and inferred metabolite exchange potentials; these graphs are represented using a GNN that encapsulates nonlinear relationships among metabolic genes, spatial context, and interspecies interactions, succeeded by a surrogate metabolic flux decoder and a global aggregation module that imposes stoichiometric and physics-informed penalties. Results indicate consistent improvements in both predictive accuracy and robustness, reduced computational latency, and well-behaved training convergence compared to classical regression and non-spatial machine learning models. The framework demonstrates how the inclusion of spatial molecular gradients and microbial interaction networks significantly enhances flux reconstruction quality, capturing emergent community phenomena such as cross-feeding, competition, and niche partitioning. With its modular design and real-time execution capability, the model offers a scalable tool for microbiome engineering, synthetic ecology design, and environmental bioprocess optimization. Future extensions will incorporate dynamic time-series omics, diffusion-based spatial processes, and integration with genome-scale metabolic models to further advance precision flux estimation in microbial ecosystems.

6. Acknowledgement

This research is funded by INTI International University

7. References

- Antoniewicz, M. R. (2021). A guide to metabolic flux analysis in metabolic engineering: Methods, tools and applications. *Metabolic engineering*, 63, 2-12.
- Becker, S. A., & Palsson, B. O. (2008). Context-specific metabolic networks are consistent with experiments. *PLOS Computational Biology*, 4(5), e1000082.
- Bruggeman, F. J., Planqué, R., Molenaar, D., & Teusink, B. (2020). Searching for principles of microbial physiology. *FEMS microbiology reviews*, 44(6), 821-844.
- Bruno, J. F., Stachowicz, J. J., & Bertness, M. D. (2003). Inclusion of facilitation into ecological theory. *Trends in ecology & evolution*, 18(3), 119-125.
- Chen, Y., Gustafsson, J., Tafur Rangel, A., Anton, M., Domenzain, I., Kittikunapong, C., Li, F., Yuan, L., Nielsen, J., & Kerkhoven, E. J. (2024). Reconstruction, simulation and analysis of enzyme-constrained metabolic models using GECKO Toolbox 3.0. *Nature protocols*, 19(3), 629-667.
- Dicker, A. J., Lonergan, M., Keir, H. R., Smith, A. H., Pollock, J., Finch, S., Cassidy, A. J., Huang, J. T., & Chalmers, J. D. (2021). The sputum microbiome and clinical outcomes in patients with bronchiectasis: a prospective observational study. *The Lancet Respiratory Medicine*, 9(8), 885-896.
- Eng, A., & Borenstein, E. (2019). Microbial community design: methods, applications, and opportunities. *Current opinion in biotechnology*, 58, 117-128.
- Gholamzadeh, E., Singla, K., & Scherf, N. (2025). Predicting microbial interactions using graph neural networks. *arXiv preprint arXiv:2511.02038*.
- Ghoul, M., & Mitri, S. (2016). The Ecology and Evolution of Microbial Competition. Trends in 1155 Microbiology, 24 (10), 833-845. In.
- Goodswen, S. J., Barratt, J. L., Kennedy, P. J., Kaufer, A., Calarco, L., & Ellis, J. T. (2021). Machine learning and applications in microbiology. *FEMS microbiology reviews*, 45(5), fuab015.
- Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., & Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome biology*, 20(1), 121.
- Henriques, D., Minebois, R., Mendoza, S. N., Macías, L. G., Pérez-Torrado, R., Barrio, E., Teusink, B., Querol, A., & Balsa-Canto, E. (2021). A multiphase multiobjective dynamic genome-scale model shows different redox balancing among yeast species of the *Saccharomyces* genus in fermentation. *MSystems*, 6(4), 10.1128/msystems.00260-00221.

- Hughes, R. E., Elliott, R. J., Dawson, J. C., & Carragher, N. O. (2021). High-content phenotypic and pathway profiling to advance drug discovery in diseases of unmet need. *Cell chemical biology*, 28(3), 338-355.
- Karin, J., Mintz, R., Raveh, B., & Nitzan, M. (2024). Interpreting single-cell and spatial omics data using deep neural network training dynamics. *Nature Computational Science*, 4(12), 941-954.
- Kavino, M., & Manoranjitham, S. (2018). In vitro bacterization of banana (*Musa* spp.) with native endophytic and rhizospheric bacterial isolates: novel ways to combat *Fusarium* wilt. *European Journal of Plant Pathology*, 151(2), 371-387.
- Kim, T. Y., Sohn, S. B., Kim, Y. B., Kim, W. J., & Lee, S. Y. (2012). Recent advances in reconstruction and applications of genome-scale metabolic models. *Current opinion in biotechnology*, 23(4), 617-623.
- Lara, E. G., van der Windt, I., Molenaar, D., de Vos, M. G., & Melkonian, C. (2021). Using functional annotations to study pairwise interactions in urinary tract infection communities. *Genes*, 12(8), 1221.
- Lewis, N. E., Hixson, K. K., Conrad, T. M., Lerman, J. A., Charusanti, P., Polpitiya, A. D., Adkins, J. N., Schramm, G., Purvine, S. O., & Lopez-Ferrer, D. (2010). Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Molecular systems biology*, 6(1), MSB201047.
- Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., & Milo, R. (2014). Visual account of protein investment in cellular functions. *Proceedings of the National Academy of Sciences*, 111(23), 8488-8493.
- Long, Y., Ang, K. S., Sethi, R., Liao, S., Heng, Y., van Olst, L., Ye, S., Zhong, C., Xu, H., & Zhang, D. (2024). Deciphering spatial domains from spatial multi-omics with SpatialGlue. *Nature Methods*, 21(9), 1658-1667.
- Lu, H., Cao, W., Liu, X., Sui, Y., Ouyang, L., Xia, J., Huang, M., Zhuang, Y., Zhang, S., & Noorman, H. (2018). Multi-omics integrative analysis with genome-scale metabolic model simulation reveals global cellular adaptation of *Aspergillus niger* under industrial enzyme production condition. *Scientific Reports*, 8(1), 14404.
- Luo, Y., Zhao, X., Ye, K., & Meng, D. (2025). Stir: Deciphering spatial transcriptomics via implicit neural representation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
- Marcos-Zambrano, L. J., Karadzovic-Hadziabdic, K., Loncar Turukalo, T., Przymus, P., Trajkovic, V., Aasmets, O., Berland, M., Gruca, A., Hasic, J., & Hron, K. (2021). Applications of machine learning in human microbiome studies: a review on feature selection, biomarker identification, disease prediction and treatment. *Frontiers in microbiology*, 12, 634511.
- Monk, J., Nogales, J., & Palsson, B. O. (2014). Optimizing genome-scale network reconstructions. *Nature biotechnology*, 32(5), 447-452.
- Mosleh, H., Albashayreh, A., & Yousef, M. (2025). Optimizing task scheduling in cloud computing with deep learning: A diabetes detection case study. In *2025 International Conference on New Trends in Computing Sciences (ICTCS)* (pp. 361-367). IEEE.
- Nadell, C. D., Drescher, K., & Foster, K. R. (2016). Spatial structure, cooperation and competition in biofilms. *Nature reviews microbiology*, 14(9), 589-600.
- O'Sullivan, S., Qi, L., & Zalloua, P. (2025). From omics to AI—mapping the pathogenic pathways in type 2 diabetes. *FEBS letters*, 599(22), 3244-3280.
- O'Brien, E. J., Monk, J. M., & Palsson, B. O. (2015). Using genome-scale models to predict biological capabilities. *Cell*, 161(5), 971-987.
- Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature biotechnology*, 28(3), 245-248.
- Pan, J., Zhang, Z., Li, Y., Yu, J., You, Z., Li, C., Wang, S., Zhu, M., Ren, F., & Zhang, X. (2024). A microbial knowledge graph-based deep learning model for predicting candidate microbes for target hosts. *Briefings in Bioinformatics*, 25(3), bbae119.
- Rampelli, S., Fabbri, M., Candela, M., Biagi, E., Brigidi, P., & Turroni, S. (2021). G2S: a new deep learning tool for predicting stool microbiome structure from oral microbiome data. *Frontiers in Genetics*, 12, 644516.
- Reel, P. S., Reel, S., Pearson, E., Trucco, E., & Jefferson, E. (2021). Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology advances*, 49, 107739.
- Sherif, S., Hendi, N. N., & Alanany, R. (2025). Multi-omics for disease subtyping and classification. In *Multi-Omics*

Technology in Human Health and Diseases (pp. 79-134). Elsevier.

- Shoaie, S., Karlsson, F., Mardinoglu, A., Nookaew, I., Bordel, S., & Nielsen, J. (2013). Understanding the interactions between bacteria in the human gut through metabolic modeling. *Scientific Reports*, *3*(1), 2532.
- Somerville, V., Grigaitis, P., Battjes, J., Moro, F., & Teusink, B. (2022). Use and limitations of genome-scale metabolic models in food microbiology. *Current Opinion in Food Science*, *43*, 225-231.
- Talib, W. H., Al-Ataby, I. A., Mahmood, A. I., Jawarneh, S., Al Kury, L. T., & Al-Yasari, I. H. (2020). The impact of herbal infusion consumption on oxidative stress and cancer: the good, the bad, the misunderstood. *Molecules*, *25*(18), 4207.
- Talib, W. H., Al Junaidi, H. S., Alshaeri, H. K., Alasmari, M. M., Hadi, R. W., Alsayed, A. R., & Law, D. (2025). Immunomodulatory and anticancer effects of moringa polyherbal infusions: potentials for preventive and therapeutic use. *Frontiers in Immunology*, *16*, 1597602.
- Tazza, G., Moro, F., Ruggeri, D., Teusink, B., & Vidács, L. (2025). MINN: A metabolic-informed neural network for integrating omics data into genome-scale metabolic modeling. *Computational and Structural Biotechnology Journal*.
- Thiele, I., & Palsson, B. Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols*, *5*(1), 93-121.
- Tian, T., Zhang, J., Lin, X., Wei, Z., & Hakonarson, H. (2024). Dependency-aware deep generative models for multitasking analysis of spatial omics data. *Nature Methods*, *21*(8), 1501-1513.
- Tsoi, R., Dai, Z., & You, L. (2019). Emerging strategies for engineering microbial communities. *Biotechnology advances*, *37*(6), 107372.
- van den Bogaard, S., Saa, P. A., & Alter, T. B. (2024). Sensitivities in protein allocation models reveal distribution of metabolic capacity and flux control. *Bioinformatics*, *40*(12), btae691.
- Wallace, B. D., & Redinbo, M. R. (2013). The human microbiome is a source of therapeutic drug targets. *Current opinion in chemical biology*, *17*(3), 379-384.
- Wondraczek, L., Pohnert, G., Schacher, F. H., Köhler, A., Gottschaldt, M., Schubert, U. S., Küsel, K., & Brakhage, A. A. (2019). Artificial microbial arenas: materials for observing and manipulating microbial consortia. *Advanced Materials*, *31*(24), 1900284.
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, *32*(1), 4-24.
- Yang, L., Qin, Y., & Jian, C. (2021). Screening for core genes related to pathogenesis of Alzheimer's disease. *Frontiers in Cell and Developmental Biology*, *9*, 668738.
- Yizhak, K., Benyamini, T., Liebermeister, W., Ruppin, E., & Shlomi, T. (2010). Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. *Bioinformatics*, *26*(12), i255-i260.
- Zampieri, G., Campanaro, S., Angione, C., & Treu, L. (2023). Metatranscriptomics-guided genome-scale metabolic modeling of microbial communities. *Cell Reports Methods*, *3*(1).
- Zedan, M. J., Abdani, S. R., Badawi, S., Al-Bashayreh, M., & Zulkifley, M. A. (2025). Dual-stage deep-learning method for glaucoma severity classification based on multiscale feature fusion. *Experimental Eye Research*, *259*, 110567.