nature portfolio

Peer Review File

Adaptable Graph Region for Optimizing Performance in Dynamic System Long-term Forecasting via Time-Aware Expert

Corresponding Author: Professor Ping Xiang

Parts of this Peer Review File have been redacted as indicated to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Review Comment
The manuscript uses the GAttention model, designed to address the challenges of low computational efficiency and slow processing speeds in multibody dynamics (MBD) systems. The model employs graph-based representations and a graph attention mechanism to capture stiffness relationships between components, facilitating scalable predictions. Additionally, it integrates multi-scale time series analysis, auxiliary learning strategies, and a hybrid training approach to improve temporal understanding and prediction accuracy, particularly for long-term structural responses in train-bridge coupled systems (TBCs).

1. The abstract mentions that "the GAttention model demonstrates superior accuracy and robustness." Please quantify the results to provide a clearer comparison with existing models. Additionally, the term "scalable predictions" should be better defined—does it refer to the model's ability to handle increasingly complex systems, or its efficiency across different scales of time or structure?

2. The review of neural network models for train-bridge coupling analysis in the introduction appears incomplete. A more thorough literature review is recommended to ensure all relevant approaches are considered and properly discussed.

3. In Figure 2, there are some errors in the labeling and color scheme of the Graph Representation structure. Please verify and correct these issues to improve clarity and consistency in the figure.

4. It is recommended to change the title of Table 2 from "predictive performance" to "Mean Absolute Error (MAE)" for clarity. In engineering, there is often more focus on the maximum response of the system; therefore, evaluating the absolute maximum error would also be beneficial. Additionally, please clarify whether the MAE values in Table 2 refer to the bridge's dynamic response or the vehicle's dynamic response. Furthermore, an explanation is needed for why the errors in various models increase as the number of spans grows. Finally, why is the MAE in Table 3 higher than the results in Table 2?

5. In line 179, it is stated that “Fig. 3 illustrates the update method for the graph representation structure of TBCs [20].” However, reference [20] does not provide relevant information regarding train-bridge coupling systems. Please verify whether the citation is correct.

6. Figure 3 presents the graph structure for TBCs, but this is not reflected in Figure 4. Please ensure that Figure 4 corresponds to the updated representation and is consistent with Figure 3.

7. In line 226, the manuscript states, "Since GAttention outputs all predictions at once, its efficiency is much higher than that …," and in line 232, it mentions, "a distill block [17] is added to the encoder, …, reducing the model’s computational load and enhancing its operational efficiency." Additionally, multiple sections highlight the high efficiency of the GAttention model. It is recommended that the authors add a section in the manuscript comparing the computational efficiency of the GAttention model with other models, as well as with the efficiency of traditional train-bridge motion equation computations.

8. It would be beneficial to label the four main steps of the GAttention model in Figure 7 for clarity. Line 276 mentions, “Second, both the original and downsampled time series are input into their respective encoders,” but the original time series is not labeled in Figure 7. Please ensure this is clarified.

9. In Figure 12, the three vertical plots in each column are not differentiated by labels. Additionally, the figure only provides the vehicle's dynamic response. Can the GAttention model also predict the bridge's dynamic response?

10. The train model parameters, bridge model parameters, and track irregularity spectra are not introduced in the manuscript. It would be helpful to include these details for a clearer understanding of the experimental setup.

11. Formatting issues:
In the abstract, TBCs should be spelled out in full on first mention.
Table 1 contains two identical column headers; it is recommended to differentiate them. Similar issues are present in Tables 2 and 3. Additionally, please clarify the meanings of labels 3, 4, …, 11 in Table 2.
Ensure consistency in the spelling of "GAttention" and "GATtention" throughout the manuscript.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
This study proposes a hybrid deep learning approach for scalable predictions in Multibody Dynamics Systems. Specifically, it suggests a graph-based representation of dynamical systems and performs a multi-scale time series analysis, auxiliary learning strategies, and a hybrid training approach to improve long-term prediction accuracy and scalability. This research falls within the scope of the Nature Communications journal. However, after reading the manuscript carefully, I cannot agree to its publication. The reasons are stated in the following:

Major comments:
1. The manuscript’s content, scope, and structure are not clear. The title and abstract suggest that a new deep learning approach for accurate and scalable predictions in Multibody Dynamics (MBD) Systems will be presented. Yet, there is no explanation as to what the term “Multibody Dynamics Systems” refers to. In fact, this term is only used in the abstract and conclusions of the manuscript. Then, the introduction of the manuscript stresses the current challenges of simulating MBD systems, such as high computational cost, but no analysis or comparison with existing algorithms is offered in this respect. Besides, the introduction presents primarily the work of the authors and not many other relevant works in the field.
The manuscript also focuses on the case of train-bridge coupled systems, which, however, is not mentioned either in the abstract or in the introduction, resulting in an abrupt transition to the topic in Section 2.2. That should have been mentioned earlier so that it is easier for the reader to follow. However, an application to solely one so-called MBD system, such as the train-bridge coupled system used herein, is not appropriate to prove the superiority of the proposed approach against existing algorithms and, more importantly, the applicability to different problems, which is in the scope of the Nature Communications journal.
2. The case study chosen for the demonstration of the proposed algorithm is not presented sufficiently. First of all, there is no description of the system used, i.e., the properties of the train and the bridge and what is the type of the bridge, or at least of the main properties of the system, for example, the frequencies of the train and the bridge. It is not clear why one node is chosen for each wagon of the train and three nodes for each span of the bridge. It is also unclear what is the feature vector of each node, what are the datasets of Table 1, what is the degree of required accuracy, and why only the number of spans of the bridge is used as a changing parameter for such a complicated coupled system. Changes in the speed of the train, the dynamic properties of the wagons, the physical properties of the bridge, or the contact characteristics between the train and bridge are more crucial for the dynamics of such a coupled system, especially in the case of long-term predictions, which the present study stresses as critical.
Finally, no detail is given as to how the other approaches that are used for comparison (LSTM, GNN, GATs, GATs+LSTM, GNBlock) are implemented, which is crucial to prove the superiority of the proposed method against existing approaches.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
Major concerns:
1. Although the analysis of train-bridge systems may encounter computational problems (if a fined finite element model is applied), numerous previous studies have demonstrated that appropriate simplification and advanced modeling techniques are effective in lowering the computational cost while maintaining a satisfied prediction accuracy. For example,
- Homaei, H., Stoura, C. D., & Dimitrakopoulos, E. G. (2024). Extended Modified Bridge System (EMBS) method for decoupling seismic vehicle-bridge interaction. Earthquake Engineering & Structural Dynamics, 53(13), 4054-4075.

- Zeng, Q., Yang, Y. B., & Dimitrakopoulos, E. G. (2016). Dynamic response of high speed vehicles and sustaining curved bridges under conditions of resonance. Engineering Structures, 114, 61-74.
- Du, X. T., Xu, Y. L., & Xia, H. (2012). Dynamic interaction of bridge–train system under non-uniform seismic ground motion. Earthquake Engineering & Structural Dynamics, 41(1), 139-157.

2. The utility of graph-based machine learning models to represent nonlinear structural systems to prompt response prediction is not new. For instance,
- Ref [8] in this manuscript.
- Chou, Y. T., Kuo, P. C., Li, K. Y., Chang, W. T., Huang, Y. N., & Chen, C. S. (2024). Inductive graph-based long short-term memory network for the prediction of nonlinear floor responses and member forces of steel buildings subjected to orthogonal horizontal ground motions. Earthquake Engineering & Structural Dynamics.

3. Given the stated dynamic systems on page 3, the last paragraph of the Introduction, the authors should clarify the definition of 'complex dynamic systems' and existing challenges more rationally.
- 'Although the current GATtention model is primarily applied to structural response analysis in civil engineering, its efficient computational capabilities and strong nonlinear adaptability make it potentially expandable to other fields, such as weather forecasting and financial market analysis, which also require handling complex dynamic systems and large volumes of time series data.'

4. On page 3, the last paragraph of Introduction, the perspective in SHM is interesting, however, this paper did not present any outcomes.
- 'The GATtention model can further develop its real-time data processing capabilities, achieving faster responses and higher real-time prediction accuracy, especially in applications such as structural health monitoring and immediate disaster response.'

5. On page 5, lines 117-127, regarding the modeling object, it appears that the system and numerical model used in this study are simple cases of actual train-bridge systems and simplified models of this system, respectively. As a result, the computational cost seems not to be a critical problem, as mentioned in the Introduction.
- 'Each train carriage is abstracted as a train node, and each span of the bridge is abstracted into three bridge nodes'

6. Do the generated earthquake ground motions based on the Cloug-Penzien model only include the y and z components, while the x component is excluded? What is the reason for neglecting this component?

7. If the proposed machine learning model (or approach) holds the primary innovation, this paper can be considered to be submitted to a journal/conference with ML subjects.

8. In Figure 12, the maximum acceleration response is 6×10^(-1) m/s^2. Do I understand correctly that this represents a small earthquake event? In this context, the implication or the challenge of predicting structural/train responses is questionable.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)
No code available for review for the paper.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The author has provided detailed explanations and revisions, I have no other comments.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
The authors have made a tremendous effort to improve the manuscript based on all reviewers' comments. To this end, the quality and context of the manuscript can now be considered for publication in the Nature Communications journal. However, I still have some comments and suggestions described in the following:

Comments:
1. In lines 305-317, the features used for the train, bridge, and pier nodes in the GNN are presented. However, the rationale behind the chosen features is not clearly discussed. It is mentioned that real-time responses from the bridge are unavailable, while such responses from the train are available, which is reasonable depending on the application. Nonetheless, the assumption that displacements of the train, on top of accelerations, are available in real time may not always be practical. Thus, further justification or clarification is needed.
2. Section 4.1 is somewhat unclear and would benefit from a more structured explanation. The authors state that the dataset is "derived from MATLAB computations" but do not specify what this entails. Details about the models used for the vehicle, bridge, and piers—including their physical parameters and assumptions—are missing. Additionally, the statement that "experimental verification shows that the discrepancies between these results and those obtained from structural simulations are negligible" lacks supporting evidence and should be substantiated with references or quantitative comparisons.
There is further confusion regarding the description of seismic waveform data preprocessing. While the manuscript refers to "collected" seismic data and discusses preprocessing steps such as denoising and baseline correction, it appears that no actual experimental signals are involved in this study. The authors should clarify this part and avoid including parts that are out of context in the manuscript. Finally, the last paragraph of this section seems unnecessarily long and somewhat out of place, with a disproportionate focus on dataset downsampling that could be more concisely stated.
3. The manuscript tends to repeat claims about the model's effectiveness, especially in Section 4.2, where long-term prediction capabilities are mentioned multiple times without accompanying detail. Statements such as "the significant performance improvement of the GAttention model is primarily attributed to its precise simulation of real-world data structures and its high adaptability" are not well supported by quantitative analysis or specific evidence. Such claims should be either substantiated or more conservatively stated.
4. The ablation study of Section 4.3 aims to assess the contribution of individual modules to the model's performance. However, the methodology is not described in sufficient detail. For instance, it is unclear whether only the graph structure was altered or if other components, such as the fusion graph convolution or time-aware experts, were also varied. Likewise, the formation of regions is not clearly defined. How is the central node selected, what determines region size, and do regions overlap?
Moreover, in Figure 13, it is unclear whether the comparison models (e.g., LSTM, GNN) were applied to the full graph or to the regionalized version. A more comprehensive analysis, including a combined ablation, e.g., multiple modules removed simultaneously, would offer deeper insight into which components are most critical to performance. The study lacks a clear, quantitative summary of which modules drive improvement.
5. Inconsistencies in terminology and repetitive phrasing make the manuscript harder to follow. For example, in line 727, Table 3 is said to demonstrate the Fusion Graph Convolution (FGC), but this table does not directly relate to that topic. Abbreviations such as FGC are also not consistently defined upon first mention. Additionally, lines 704-718 repeat ideas previously stated, reducing the overall clarity and conciseness. These issues should be revised for improved readability.

(Remarks on code availability)
The authors have shared their code with an accompanying README file.

Reviewer #3

(Remarks to the Author)
Comments on NCOMMS-24-76796A - A Deep Learning Model for Long-Term Time Series Forecasting of Dynamic Systems

Although the authors have attempted revisions addressing previous concerns, the reviewer maintains that the manuscript provides only minor contributions to civil/mechanical engineering. The manuscript appears more suitable for an industry-oriented journal rather than Nature Communications.

Therefore, the reviewer does not recommend publication based on the following major concerns:

1. The knowledge contribution and novelty are unclear. From the perspective of AI, hybrid training, multi-scale time series analysis and graph representation are well-established concepts. Their integration in this work lacks substantial novelty. From the civil/mechanical engineering standpoint, the train-bridge coupled behaviors and related analyses have also been extensively studied. The manuscript fails to clearly identify existing knowledge gaps or explicitly clarify its unique contributions to either field.

2. The study cases are of limited impact. Despite the broad range of dynamic system problems presented in the introduction, the actual scenarios studied—train-bridge systems under earthquakes and traffic flow systems—are common and well-researched examples. The train-bridge system studied here is relatively simple, failing to adequately demonstrate the method's potential or general applicability.

3. The improvement of the proposed approach appears largely incremental. According to Tables 3–5, the proposed method improves accuracy by approximately 10% compared to existing methods. Such gains suggest that traditional AI/ML techniques remain competitive. The authors should further justify the practical value of these incremental improvements.

Some minor concerns:

1. Validation based solely on simulation data may not convincingly demonstrate practical effectiveness. Authors should strengthen validation with laboratory or field test data and explicitly evaluate performance under various measurement noise levels.

2. Mean Absolute Error (MAE) alone is insufficient as a metric for performance evaluation. It is strongly recommended to benchmark the proposed method against advanced AI/ML techniques used for general time-series prediction beyond the limited scope of bridge engineering.

3. It is recommended to include at least one more complex scenario to demonstrate broader applicability. Specifically, address multi-target prediction scenarios (e.g., simultaneous multiple train crossings), generalization to different structural configurations, and explicitly discuss computational efficiency. Although such capabilities are claimed, validation and discussion are currently inadequate.

4. To fully demonstrate prediction accuracy, comparison of time-series predictions in the frequency domain is highly recommended. Additionally, authors should clearly state the method's performance limitations and underlying assumptions.

5. Page 8, Line 223: Verify whether "Figure 4" should be correctly cited as "Figure 2".

6. In Figure 3, clarify "GATtension." Similarly, on Line 400, verify if "GAttentionas" should be corrected to "GAttention."

7. Equations (15) and (16) lack explicit definitions of variables "y" and "z".

8. Figure 12 appears in the manuscript without explicit citation or discussion in the main text.

9. Figures 9–11 lack detailed explanations for each subplot, making it difficult to interpret the response and behavior of the train-bridge systems under earthquake scenarios. Additional clarification is required.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The comments to the authors have been adequately addressed. I do not have any more comments.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
The manuscript has been revised to a satisfactory level for publication.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career

Researchers who co-review manuscripts.

(Remarks on code availability)

Responses to Reviewers' Comments

Dear Editor and Reviewers,

First of and foremost, we would like to express our sincere thanks and appreciation to the editors and reviewers for your professional handling and review of the manuscript. Your comments and constructive suggestions were very helpful in improving the quality of our manuscript and helped us identify several shortcomings in earlier versions of the manuscript. We have used these valuable suggestions as a guide to address these shortcomings where appropriate. The paper has now been revised following the reviewers' comments, changes in the manuscript have been marked in "blue". In the reply to comments, all the changes are marked in "yellow". Please check the following responses.

Thank you very much!

Best regards,
PingXiang

Reply to the reviewer's comment (Reviewer#1)

General comments:

Authors' Reply:

Your valuable comments have greatly helped to improve the quality of our article. However, there are some comments that we cannot fully understand, whether you have pasted some comments wrongly or we have misunderstood your comments. We have done our best to reply to your comments and hope that we have not misunderstood your meaning. If there are any unclear points raised, we hope you can let us know during the next revision and we will carefully revise the article to your satisfaction. Thank you very much!

Comment 1:

The abstract mentions that "the GAttention model demonstrates superior accuracy and robustness." Please quantify the results to provide a clearer comparison with existing models. Additionally, the term "scalable predictions" should be better defined—does it refer to the model's ability to handle increasingly complex systems, or its efficiency across different scales of time or structure?

Authors' Reply:

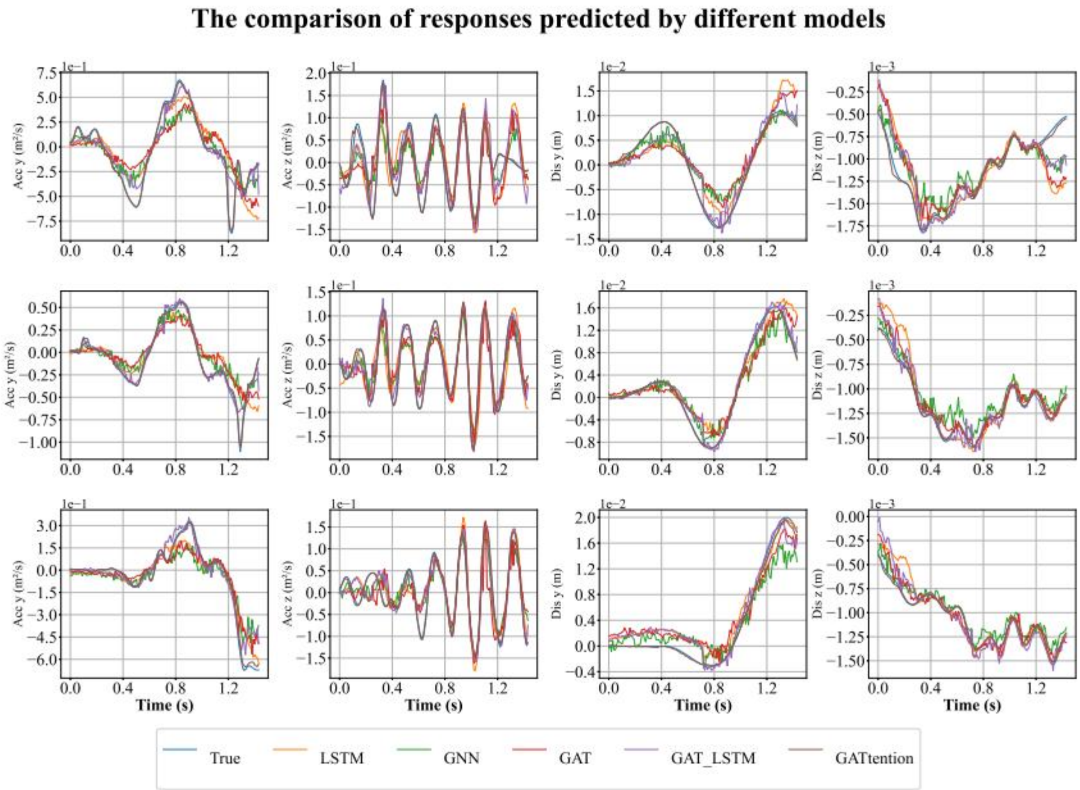
Thank you very much for your valuable feedback. In response to your suggestions, we have made several revisions to the paper. While we included quantitative comparisons in the paper, we realized that the results were not clearly presented in a visually intuitive way. To address this, we have added comparative charts of different models in the article, which will help to present the performance and advantages of each model more clearly under various conditions. This not only enhances the readability of the paper but also allows readers to better understand the differences and strengths of the models.

Additionally, regarding the term "scalable," we have modified it as per your recommendation. The original term "scalable" might have been somewhat misleading when describing the model, so we replaced it with "Structure-extensible." This new term more accurately reflects the model's capability, as it can make accurate long-term sequence predictions across different structures. Specifically, the model in this paper can effectively predict for bridges with different spans, even though the TBC (Bridge Vibration Control) system's graph structure changes. By using a unified model, we can still ensure prediction accuracy. This revision better highlights the model's flexibility and adaptability when faced with various structures, ensuring consistency and stability in predictions across different types of structures.

Once again, thank you for your detailed feedback. It has helped us refine the content and expression of the paper. We believe these improvements will enhance the scientific quality and readability of the paper, and more clearly convey the core ideas of our research.

which we refer to as $\textbf{Structure-extensible prediction}$. This means that the model can be extended to accommodate different structures.

At the same time, we compared the prediction results of other models with GATtention model under the hybrid training approach, as shown in the following figures:



Comment 2:

The review of neural network models for train-bridge coupling analysis in the introduction appears incomplete. A more thorough literature review is recommended to ensure all relevant approaches are considered and properly discussed.

Authors' Reply:

Thank you very much for your valuable suggestions on the introduction. Based on your feedback, we have supplemented the relevant content and included a more comprehensive literature review. The new discussion covers the application of existing neural network models in train-bridge coupling analysis and provides a detailed comparison of the advantages and disadvantages of different approaches.

For train-bridge dynamic systems, numerous advanced computational models have emerged for train-bridge dynamic systems. Zhai's foundational theory of train-bridge coupled \cite{zhai1, zhai2, zhai3} has provided a solid base for the calculation of train-bridge dynamics. Xiang introduced a high-speed train model under wind load

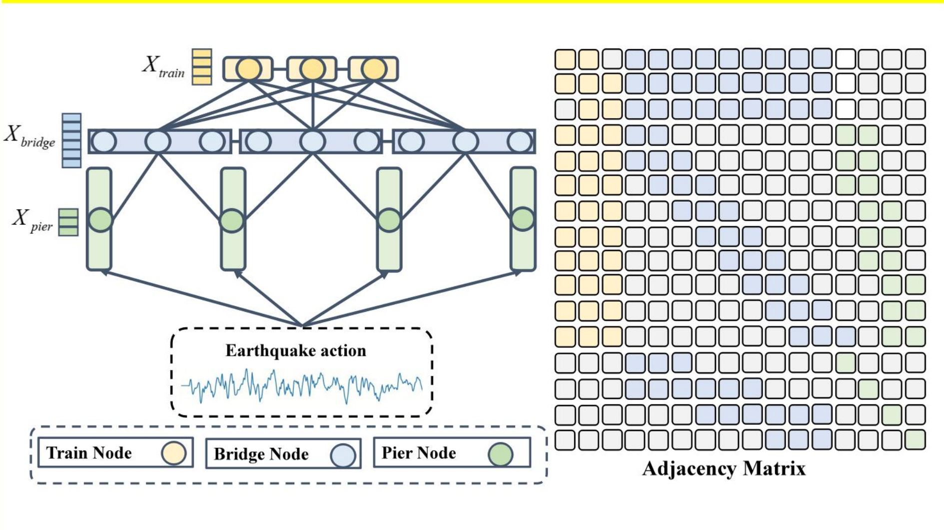
\cite{Xianghuoyue1, Xianghuoyue2, Xianghuoyue3}, addressing a gap in the field. Elias G. Dimitrakopoulos proposed an improved method for decoupling the vehicle-bridge interaction under seismic effects \cite{Elias1, Elias2, Elias3}, thereby improving computational efficiency.. These optimized models and methods have significantly advanced the study of train-bridge coupling. However, the issue of computational efficiency in real-time applications remains a challenge. To address this, several deep learning models have been proposed for more efficient computation.The neural network models discussed above exhibit limitations in handling dynamic train-bridge systems, particularly when dealing with the continuously evolving bridge and train structures. These changes cite{Sun2024AUG6} make it challenging to accurately capture the impact of such variations on system responses. Consequently, graph-based neural networks have emerged as a promising research area, as they can more effectively represent the dynamic relationships between various components of the system.X. Peng et al. proposed the GN Block model \cite{xiang2024adaptive}, a highly adaptive graph neural network designed to address complex structural response prediction tasks, especially in dynamic train-bridge coupled systems. This model is capable of adjusting its network structure and learning strategy in real-time, thereby enhancing both adaptability and prediction accuracy.Similarly, P. Zhang et al. developed an adaptive graph neural network model \cite{zhang2024novel} that not only adapts to known structures but also autonomously predicts the responses of unknown structures. This adaptability allows the model to handle various types of train-bridge systems and deliver accurate predictions.These graph-based neural networks \cite{Elias4, 2020arXiv200511650W, 2017arXiv170904875Y, 2019arXiv190600121W} address the limitations of traditional neural networks, which struggle to fully leverage topological information. By aggregating information from nodes and edges, they capture the intricate relationships between different system components, thereby improving prediction accuracy and computational efficiency, especially in complex and dynamically changing structures. The application of these models has significantly advanced research in train-bridge dynamics, driving it towards greater efficiency and precision.

Comment 3:

In Figure 2, there are some errors in the labeling and color scheme of the Graph Representation structure. Please verify and correct these issues to improve clarity and consistency in the figure.

Authors' Reply:

Thank you very much for your feedback. After verification, we confirmed that there were indeed some issues with Figure 2, and we have corrected these errors in the text. Once again, we greatly appreciate your valuable suggestions.



Comment 4:

It is recommended to change the title of Table 2 from "predictive performance" to "Mean Absolute Error (MAE)" for clarity. In engineering, there is often more focus on the maximum response of the system; therefore, evaluating the absolute maximum error would also be beneficial. Additionally, please clarify whether the MAE values in Table 2 refer to the bridge's dynamic response or the vehicle's dynamic response. Furthermore, an explanation is needed for why the errors in various models increase as the number of spans grows. Finally, why is the MAE in Table 3 higher than the results in Table 2?

Authors' Reply:

Thank you for your valuable suggestions; your feedback is very constructive. Based on your suggestion, we have revised the table title to "The Mean Squared Error values of different models." This change makes the title more intuitive and clear, better showcasing the performance of the various models and helping readers quickly understand the core content discussed in the article.Additionally, in response to the concerns you raised, we have added relevant content to the article, providing detailed answers to these questions. We conducted an in-depth analysis of the phenomenon where the model's loss increases as the number of iterations grows. Once again, thank you for your valuable suggestions. Your feedback has played a crucial role in optimizing the structure of the article and improving the accuracy and readability of the content. We hope these revisions make the article clearer and more accessible, providing more valuable insights for the research.

Table 1: TBC datasets for each span

Dataset	Bridge Span	Total Length	Time Interval	Total Length after downsampling	Time Interval after downsampling
TBC [26]	3	42654	5×10^5	2140	0.001
	4	50477	5×10^5	2530	0.001
	5	58300	5×10^5	2920	0.001
	6	66122	5×10^5	3310	0.001
	7	73945	5×10^5	3700	0.001
	8	79546	5×10^5	3970	0.001
	9	89591	5×10^5	4480	0.001
	10	97414	5×10^5	4880	0.001
	11	105236	5×10^5	5270	0.001

Model	Span							
	3	4	5	6	7	9	10	11
LSTM[27]	0.194	0.214	0.237	0.232	0.193	0.214	0.232	0.239
GNN[28]	0.302	0.326	0.380	0.376	0.302	0.336	0.367	0.381
GATs[22]	0.186	0.195	0.224	0.221	0.178	0.199	0.204	0.211
GATs + LSTM	0.182	0.193	0.222	0.220	0.180	0.194	0.221	0.198
GNBlock[14]	0.215	0.230	0.318	0.284	0.230	0.252	0.235	0.281
GAttention(Ours)	**0.087**	**0.092**	**0.103**	**0.114**	**0.111**	**0.117**	**0.116**	**0.112**

Table 5: The Mean Squared Error values of different models under the hybrid training approach

Model	Span							
	3	4	5	6	7	9	10	11
LSTM[27]	0.235	0.247	0.282	0.279	0.264	0.290	0.296	0.308
GNN[28]	0.341	0.370	0.396	0.398	0.352	0.386	0.393	0.388
GATs[22]	0.209	0.241	0.246	0.249	0.242	0.276	0.263	0.266
GATs + LSTM	0.184	0.207	0.212	0.207	0.226	0.256	0.240	0.248
GNBlock[14]	0.284	0.257	0.262	0.297	0.326	0.246	0.280	0.258
GAttention(Ours)	**0.110**	**0.129**	**0.133**	**0.134**	**0.125**	**0.116**	**0.124**	**0.121**

At the same time, we also observed a phenomenon: as the number of spans in the bridge increases, the model's loss value continues to rise. This phenomenon can be explained by the following reasons: as the number of spans increases, the number of bridge nodes in the graph structure also increases. However, since these bridge nodes remain unchanged in the time dimension, each node, after going through the graph convolution operation, is influenced by more bridge nodes. In other words, as the number of spans increases, the connections between nodes in the graph become more complex, and the range of information propagation expands, causing the features of each node to be more affected by other bridge nodes.This phenomenon can be seen as a form of "regularization" applied to the model. As the number of nodes increases, the model faces greater complexity and more extensive information interactions during the training process. As a result, the model's learning process becomes more challenging, leading to an increase in the loss value.

The model trained using hybrid training approach can adapt to different structures, but its performance is still somewhat inferior to that of a model trained on a single structure. This is primarily due to the need for the model to adjust its parameters to accommodate the features of various structures during mixed training. To achieve an adequate fit across diverse structures, the model must strike a balance between different data sets, which results in lower performance on a specific structure compared to a model trained exclusively on that structure.

During the training process, the model adjusts its parameters to account for the differences in bridge structures. While this adjustment enhances the model's adaptability, it may also limit the model's ability to fully capture and optimize certain structural features, thus affecting its accuracy. Therefore, although mixed training improves the model's general applicability, it sacrifices the ability to fine-tune and optimize for specific structures, leading to slightly lower accuracy compared to models trained solely on a single structure.

Comment 5:

In line 179, it is stated that “Fig. 3 illustrates the update method for the graph representation structure of TBCs [20].” However, reference [20] does not provide relevant information regarding train-bridge coupling systems. Please verify whether the citation is correct.

Authors’ Reply:

Thank you very much for pointing out the error. This mistake was not promptly removed during the writing process, but it has now been corrected.

Comment 6:

Figure 3 presents the graph structure for TBCs, but this is not reflected in Figure 4. Please ensure that Figure 4 corresponds to the updated representation and is consistent with Figure 3.

Authors' Reply:

Thank you very much for pointing out the issue. We greatly value your feedback, but it seems there may be a slight misunderstanding here. The core of the message-passing process depicted in Figure 4, within the graph structure, actually emphasizes that when a node wants to share information with other nodes, the information must pass through multiple layers of message aggregation. The focus of this process is the layer-by-layer transmission and integration of information. Through multiple graph convolution operations, the information gradually expands and merges, thereby enabling effective communication and collaboration between nodes.

We sincerely appreciate your attention, and we hope that this clarification helps you better understand the message-passing mechanism within the graph structure. If you have any further questions or suggestions, please feel free to raise them. We would be more than happy to engage in further discussion and improvement.

Comment 7:

In line 226, the manuscript states, "Since GAttention outputs all predictions at once, its efficiency is much higher than that …," and in line 232, it mentions, "a distill block [17] is added to the encoder, …, reducing the model's computational load and enhancing its operational efficiency." Additionally, multiple sections highlight the high efficiency of the GAttention model. It is recommended that the authors add a section in the manuscript comparing the computational efficiency of the GAttention model with other models, as well as with the efficiency of traditional train-bridge motion equation computations.

Authors' Reply:

Thank you very much for your valuable suggestions. This paper mainly emphasizes the efficiency improvements of GAttention in three aspects:

1. **Adoption of Encoder-Only Architecture**: Compared to traditional step-by-step network architectures, GAttention adopts an Encoder-Only architecture that inputs all results at once. This architecture has been proven to be more efficient in the deep learning field because it allows for global processing of the data in a single step, avoiding the redundancy of layer-by-layer computation and significantly reducing the computational

overhead, thus improving overall processing speed.

2. **Introduction of Distill Block to Reduce Temporal Length**: In the attention mechanism, reducing the number of queries (i.e., temporal length) can significantly enhance efficiency. By introducing the Distill Block, GAttention effectively reduces the temporal length of the input data, which in turn reduces the computational load, improving overall efficiency. From a theoretical perspective, shortening the temporal length directly alleviates the computational burden during the attention mechanism execution, a point that has been validated by multiple studies as highly effective.
3. **Regional Graph Representation**: To further improve efficiency, we followed your suggestion and specifically included experiments to validate the effectiveness of the localized graph representation. Experimental results show that using the localized graph representation increases processing speed by 4 to 20 times compared to not using it. This result verifies the significant advantages of the localized graph representation in enhancing computational efficiency and accelerating information propagation.

Thank you once again for your valuable suggestions. Your feedback has been instrumental in improving our paper. If you have any further questions or suggestions, we would be more than happy to discuss and refine the work further.

Table 6: Original Graph Structure

Dataset	Node Count	Memory	Speed
PEMS04	307	39.79Mb	1.004s
PEMS08	170	23.159Mb	8.526s

Table 7: Regional Graph Structure

Dataset	Node Count	Memory	Speed
PEMS04	129	16.72Mb	0.269s
PEMS08	70	9.53Mb	0.441s

Table. \ref{tab:tab6} and Table. \ref{tab:tab7} illustrate the original graph structure and the structure after regionalization, respectively. It is evident that regionalization significantly reduces the size of the graph structure. Specifically, memory consumption decreases by nearly 60\%, and computational efficiency improves by a factor of 4 to 20. This change is primarily due to the regional graph representation's ability to effectively compress the graph's scale, thereby reducing computational resource usage. These reductions in memory and improvements in efficiency are particularly critical when handling large-scale datasets, as they allow the model to run more efficiently, particularly in real-time applications and large-scale systems, where they significantly enhance data processing speed and system responsiveness.

Comment 8:

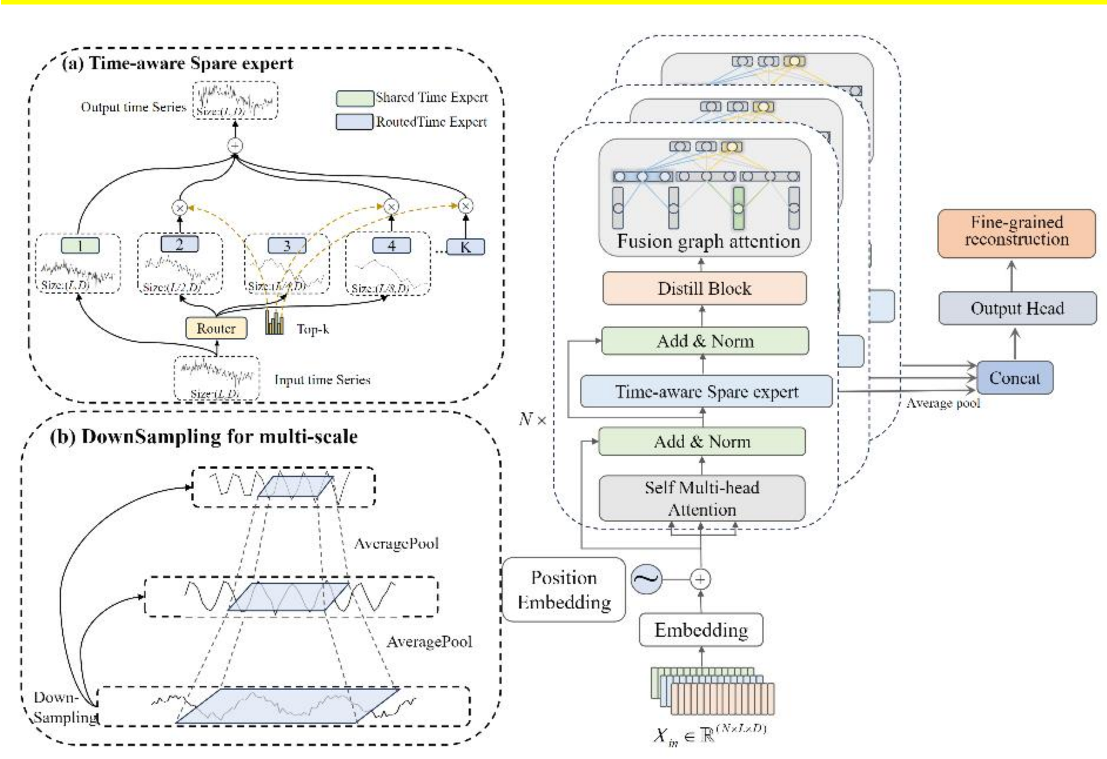
It would be beneficial to label the four main steps of the GAttention model in Figure 7 for clarity. Line 276 mentions, "Second, both the original and downsampled time series are input into their respective encoders," but the original time series is not labeled in Figure 7. Please ensure this is clarified.

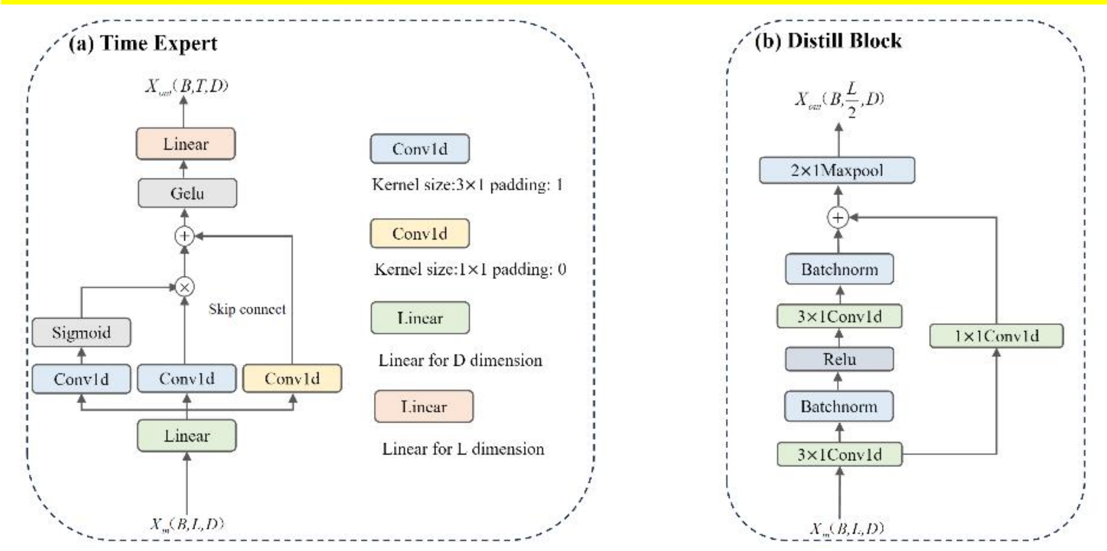
Authors' Reply:

Thank you very much for your valuable suggestions. Your feedback has significantly improved the logical structure and overall expression of this paper. Based on your input, we have further optimized the GAttention model, particularly by removing the original Decoder part and focusing solely on the Encoder. This adjustment not only simplifies the model structure but also effectively enhances its operational efficiency, making the entire system's computation process more efficient and streamlined.

In the new architectural design, we have provided detailed explanations for each module, aiming to clearly articulate the function and role of each part. This ensures that readers can fully understand the working principles of each module and their collaboration, making the operation of the GAttention model more intuitive and easier to grasp. Additionally, to help readers better understand and apply this model, we have included the complete model algorithm and provided a detailed description of the calculation process and implementation for each step. We believe this not only offers readers an opportunity to deeply understand the internal mechanisms of the GAttention model but also provides actionable algorithmic guidance for those who wish to apply, validate, or extend the model in their own research.

Each module's algorithm has been carefully designed to ensure that computational efficiency is maximized while maintaining the model's performance and accuracy. We highly value your input and recognize that these optimizations play a crucial role in enhancing the scientific rigor and applicability of the paper. Should you have any further questions or suggestions, we would be more than happy to continue the discussion and refine the model, ensuring that we provide even higher-quality research contributions to the academic community.





Time Expert processes input information at different scales and maps it to a unified temporal length. Specifically, it first applies a linear layer to transform the feature dimension, then uses gated 1D convolution and residual connections to condense and extract the temporal dimension. Finally, another linear layer adjusts the temporal dimension, projecting it to the required temporal length, preparing it for the subsequent fusion of each expert.

the model utilizes DistillBlock to reduce the temporal sequence length, employing techniques such as max-pooling to shorten the input sequence while preserving critical temporal information. This approach allows the model to retain important information while reducing computational complexity, thereby improving efficiency.

Comment 9:

In Figure 12, the three vertical plots in each column are not differentiated by labels. Additionally, the figure only provides the vehicle's dynamic response. Can the GAttention model also predict the bridge's dynamic response?

Authors' Reply:

We sincerely apologize for any misunderstanding. In the figure, we use 0-3 to represent the displacements and accelerations in the y and z directions, where the three rows correspond to three trains, and the four columns represent the displacements and accelerations in the y and z directions. We have already provided relevant explanations in the paper to ensure that readers can clearly understand.

Regarding the bridge response, it is indeed fully predictable. Since the number of output nodes is directly related to the training targets, the GAttention model we employed can accurately predict based on the features learned during training. However, considering that the focus of this study is on the dynamic response of the trains, the dynamic response of the bridge is not the primary research direction. Therefore, although predicting the bridge's response is feasible, we have not expanded on this aspect in the paper.

The study and prediction of the train's dynamic behavior are crucial for this research, so we focused on the dynamic response of the trains. If there is further research demand in the future, the dynamic response of the bridge can also be explored as a potential extension. We truly appreciate your attention and understanding and look forward to receiving any further valuable feedback and suggestions you may have.

Fig. \ref{fig:Fig10px}-\ref{fig:Fig12px} demonstrate the predictive capabilities of the GATtention model under the hybrid training approach, where 0-3 represent the acceleration and displacement of the train in the y and z directions, respectively. These figures highlight the model's accuracy in predicting the responses of different vehicle-bridge coupled systems. They also illustrate the model's ability to predict the behavior of vehicle-bridge systems with different structures, showcasing its structural generalization capability.

Comment 10:

The train model parameters, bridge model parameters, and track irregularity spectra are not introduced in the manuscript. It would be helpful to include these details for a clearer understanding of the experimental setup.

Authors' Reply:

Thank you very much for your valuable suggestions.

The core research direction of this paper focuses on modeling dynamic systems, particularly in the context of train-bridge coupling and traffic flow, and performing long-term sequence predictions. As a result, we have not provided an in-depth discussion of the design details of the train-bridge coupling system. At the same time, for specific parameter settings, you can refer to the detailed explanations in the code. Our primary focus is on the overall modeling framework and its application in long-term sequence prediction, aiming to demonstrate how precise modeling methods can address complex dynamic system prediction problems.

Nevertheless, to help readers better understand the model we used and its theoretical foundations, we have cited relevant literature in the introduction, providing a brief overview of the train-bridge coupling system and its development. These references provide solid theoretical support for our method selection and design, offering readers more detailed background information.

We recognize that the design details of the train-bridge coupling system are crucial for further optimizing the model and improving prediction accuracy. Therefore, in future research, we will consider exploring these details in greater greater depth to better advance.

[5] W. Zhai, Z. Han, Z. Chen, L. Ling, S. Zhu, Train – track – bridge dynamic801
interaction: a state-of-the-art review, Vehicle System Dynamics 57 (7)802
(2019) 984 – 1027. doi:10.1080/00423114.2019.1605085.803
URL https://doi.org/10.1080/00423114.2019.1605085804

[6] W. Zhai, K. Wang, J. Lin, Modelling and experiment of railway ballast vibrations, Journal of Sound and Vibration 270 (4) (2004) 673 – 683. doi:https://doi.org/10.1016/S0022-460X(03)00186-X. URL https://www.sciencedirect.com/science/article/pii/S0022460X0300186X
[7] W. Zhai, K. Wang, C. Cai, Fundamentals of vehicle – track coupled dynamics, Vehicle System Dynamics 47 (11) (2009) 1349 – 1376. doi:10.1080/00423110802621561. URL https://doi.org/10.1080/00423110802621561
[8] H. Xiang, W. Ren, C. Shang, J. Zhu, Y. Li, Dynamic response analysis of train vehicle bridge system under train-induced winds, International Journal of Structural Stability and Dynamics 23 (12) (2023) 2350132. doi:10.1142/S0219455423501328. URL https://doi.org/10.1142/S0219455423501328
[9] X. Tian, H. Xiang, Y. Li, Modeling and analyzing of high-speed maglev train-bridge systems considering centrifugal force induced by bridge vertical deformation, Structures 73 (2025) 108240. doi:https://doi.org/10.1016/j.istruc.2025.108240. URL https://www.sciencedirect.com/science/article/pii/S2352012425000542
[10] C. Shang, H. Xiang, Y. Bao, Y. Li, K. Luo, Mechanism of vortex-induced vibration in two lock-in regions for truss girder sections, Jour

<u>Comment 11:</u>

Formatting issues:

•In the abstract, TBCs should be spelled out in full on first mention.

•Table 1 contains two identical column headers; it is recommended to differentiate them. Similar issues are present in Tables 2 and 3. Additionally, please clarify the meanings of labels 3, 4, …, 11 in Table 2.

•Ensure consistency in the spelling of "GAttention" and "GATtention" throughout the manuscript.

Authors' Reply:

Thank you very much for your valuable feedback.

We have revised the title of the table based on your suggestion to ensure that it is clearer and more accurate, avoiding any potential ambiguity. The new title now clearly describes the content presented in the table, ensuring that readers can quickly understand its purpose and meaning.

Additionally, to further improve the clarity and comprehensibility of the table, we have provided a detailed explanation of the "3-11" time span in the text. This modification aims to help readers fully understand the specific significance and underlying logic of these time spans in the experiment, making the interpretation of the table data more accurate.

We have also corrected the inconsistency in the description of the GAttention model. The name and corresponding description of the GAttention model in the table are now unified, eliminating any potential confusion.

Once again, thank you for your careful feedback and suggestions. We greatly value your input and will continue to strive to improve and refine our research work.

Dataset	Bridge Span	Total Length	Time Interval	Total Length after downsampling	Time Interval after downsampling
TBC [26]	3	42654	5×10^5	2140	0.001
	4	50477	5×10^5	2530	0.001
	5	58300	5×10^5	2920	0.001
	6	66122	5×10^5	3310	0.001
	7	73945	5×10^5	3700	0.001
	8	79546	5×10^5	3970	0.001
	9	89591	5×10^5	4480	0.001
	10	97414	5×10^5	4880	0.001
	11	105236	5×10^5	5270	0.001

Model	Span							
	3	4	5	6	7	9	10	11
LSTM[27]	0.194	0.214	0.237	0.232	0.193	0.214	0.232	0.239
GNN[28]	0.302	0.326	0.380	0.376	0.302	0.336	0.367	0.381
GATs[22]	0.186	0.195	0.224	0.221	0.178	0.199	0.204	0.211
GATs + LSTM	0.182	0.193	0.222	0.220	0.180	0.194	0.221	0.198
GNBlock[14]	0.215	0.230	0.318	0.284	0.230	0.252	0.235	0.281
GAttention(Ours)	**0.087**	**0.092**	**0.103**	**0.114**	**0.111**	**0.117**	**0.116**	**0.112**

To evaluate the generalization capability of the GATtention model, it will be trained on TBCs datasets with 3 and 11 spans and then tested on a TBCs with 4-10 span bridges to demonstrate the model's generalization ability. Mean Absolute Error (MAE) will continue to be used as the evaluation metric, as shown in Table. \ref{tab:tab3}. The left side of the table represents models trained on a train-bridge coupled dataset with three spans, while the right side represents models trained on a train-bridge coupled dataset with eleven spans.

Reply to the reviewer's comment (Reviewer#2)

General comments:

This study proposes a hybrid deep learning approach for scalable predictions in Multibody Dynamics Systems. Specifically, it suggests a graph-based representation of dynamical systems and performs a multi-scale time series analysis, auxiliary learning strategies, and a hybrid training approach to improve long-term prediction accuracy and scalability. This research falls within the scope of the Nature Communications journal. However, after reading the manuscript carefully, I cannot agree to its publication. The reasons are stated in the following:

Authors' Reply:

I apologize for any confusion caused. In the initial draft, we did not clearly capture the main focus of the paper, and we failed to effectively articulate our research work. We have fully addressed the concerns you raised and made improvements based on your feedback in the revised version, aiming to better convey the core ideas of this paper.

The main research direction of this paper is the long-term time series prediction of dynamic systems, with the vehicle-bridge coupling system being a complex dynamic system that is the focus of our study. However, although there has been some progress with graph neural network (GNN) models in existing research, no literature has yet abstracted the vehicle-bridge coupling system into a graph structure using a regional graph representation. In this study, we are the first to propose an innovative method that transforms the vehicle-bridge coupling system into a graph structure using the regional graph representation, with seismic acceleration as the feature attribute of the bridge pier nodes, serving as the primary excitation for the entire graph structure. The innovation of this method lies in the fact that traditional graph representation methods are typically based on finite element node divisions, where each finite element node is treated as a graph node, and the complexity of the entire graph structure increases exponentially as the number of nodes increases. In contrast, we successfully overcome the computational bottleneck caused by large graph structures in finite element graph representations by adopting regional graph representation and fine-scale reconstruction methods. This is one of the key highlights of our research.

Additionally, traditional graph neural network-based models typically treat the graph structure as topological information and incorporate it into node features. However, most models fail to effectively account for different dynamic structures (e.g., changes in span number in vehicle-bridge coupling systems or changes in the number of train carriages) within the same model. In this regard, our research introduces an innovative hybrid training mode that allows the model to automatically recognize commonalities between different dynamic structures during the learning process. Furthermore, the graph attention mechanism is used to simulate "stiffness." This method effectively solves the issue where models fail to maintain stable predictive capabilities when handling different dynamic structures.

In terms of time series data processing, this paper also introduces a multi-scale analysis module—Time-aware Expert. This is a shared sparse mixture expert model that can automatically select appropriate signals based on the different periods of the time series data and integrate time series information at different scales. Different scales correspond to different periods and frequencies of data, which is particularly crucial for seismic wave analysis because the frequency components of seismic waves determine their impact on structural responses. By introducing this module, the model can adapt to time series signals at different frequencies and make more accurate predictions about the effects of seismic waves.

These innovative methods and techniques provide a solid theoretical foundation and practical guidance for this research. We believe these contributions not only advance the frontier of research in vehicle-bridge coupling system prediction but also have significant academic value, aligning well with the scope of Nature Communications.

We sincerely apologize for the shortcomings in the initial draft. Thank you for your valuable suggestions and feedback. We hope that, through this revision, we can better meet your expectations and work together to produce a high-quality academic paper.

Comment 1:

The manuscript's content, scope, and structure are not clear. The title and abstract suggest that a new deep learning approach for accurate and scalable predictions in Multibody Dynamics (MBD) Systems will be presented. Yet, there is no explanation as to what the term "Multibody Dynamics Systems" refers to. In fact, this term is only used in the abstract and conclusions of the manuscript. Then, the introduction of the manuscript stresses the current challenges of simulating MBD systems, such as high computational cost, but no analysis or comparison with existing algorithms is offered in this respect. Besides, the introduction presents primarily the work of the authors and not many other relevant works in the field.

Authors' Reply:

Thank you very much for your valuable feedback. In the initial draft, there was indeed an issue with the lack of clarity in the main theme. We have carefully revised and improved this aspect in the updated manuscript. Specifically, we have changed the term "Multibody Dynamics" to "dynamic systems" to more accurately reflect the content and objectives of our research. Our original intention was to find an effective modeling approach that could handle dynamic systems involving various complex phenomena, such as weather, traffic flow, and vehicle-bridge coupled systems. We focus particularly on using deep learning methods to model these systems and perform long-term time series prediction.

In the revised introduction, we first clearly define the concept of "dynamic systems" and introduce traffic networks and vehicle-bridge coupled systems as typical examples of dynamic systems. We further discuss some of the typical challenges these systems face, specifically: (1) a significant decrease in computational efficiency when the graph structure is excessively large; (2) difficulties in sharing information between nodes and propagating global information effectively in large-scale graph structures; and (3) the lack of a unified framework in existing methods, which makes effective generalization across different types of graph structures difficult. We emphasize that these issues limit the scalability and practical applicability of Graph Neural Networks (GNNs) in large-scale, complex systems.

To address these challenges, we have conducted corresponding experiments in the subsequent sections, aiming to demonstrate how the deep learning approach we propose effectively tackles these problems and enhances the computational efficiency, information propagation, and generalization ability across different graph structures. Through these revisions and experiments, we strive to present more clearly the significance and innovations of our research and hope to provide new ideas and solutions for related fields.

Comment 2:

The manuscript also focuses on the case of train-bridge coupled systems, which, however, is not mentioned either in the abstract or in the introduction, resulting in an abrupt transition to the topic in Section 2.2. That should have been mentioned earlier so that it is easier for the reader to follow. However, an application to solely one so-called MBD system, such as the train-bridge coupled system used herein, is not appropriate to prove the superiority of the proposed approach against existing algorithms and, more importantly, the applicability to different problems, which is in the scope of the Nature Communications journal.

Authors' Reply:

Thank you very much for your valuable comments. Indeed, there was a sudden transition in the original manuscript. To make the paper smoother and easier to understand, we have made significant improvements in the revised version. First, we have provided a detailed introduction to the concept of "dynamic systems," clearly explaining its meaning and the challenges such systems face in the real world. Dynamic systems refer to systems whose states evolve over time, influenced by initial conditions, external inputs, and governing laws. We have specifically highlighted the major issues encountered when dealing with complex, dynamic systems, such as bottlenecks in computational efficiency, difficulties in information propagation, and generalization issues in large-scale complex systems.

Next, we systematically introduced the research background and development history of train-bridge coupled systems, reviewing the work of relevant scholars in this field. In this section, we not only summarized the existing research on train-bridge coupled systems but also pointed out the current gaps and challenges. Based on these shortcomings, we have proposed the innovative aspects of this paper, particularly in improving existing methods and enhancing computational efficiency and information propagation within the model.

To better demonstrate the effectiveness of the deep learning approach we proposed, we conducted detailed experiments on train-bridge coupled systems, as well as tests on traffic flow prediction, which is another typical application of dynamic systems. These two cases effectively showcase the common characteristics of dynamic systems, namely, that their states evolve over time and are influenced by various factors. Through these experiments, we clearly validate the application and superiority of our model and algorithm in different types of dynamic systems, proving that our method not only applies to train-bridge coupled systems but also has the ability to handle other complex dynamic systems.

Moreover, we are aware that these cases represent only a portion of dynamic system applications. We are more than willing to supplement other types of cases as needed to further verify the generalizability and practical applicability of our model. We believe that through these additional experiments and cases, we can more comprehensively demonstrate the effectiveness of our approach and provide new ideas and solutions for related research fields.

Once again, thank you for your feedback. Your suggestions have helped us improve the structure and content of the paper, ensuring the research's breadth and depth. We look forward to presenting this revised manuscript, which better showcases the innovation and application prospects of our research, to the readers.

Comment 3:

The case study chosen for the demonstration of the proposed algorithm is not presented sufficiently. First of all, there is no description of the system used, i.e., the properties of the train and the bridge and what is the type of the bridge, or at least of the main properties of the system, for example, the frequencies of the train and the bridge. It is not clear why one node is chosen for each wagon of the train and three nodes for each span of the bridge. It is also unclear what is the feature vector of each node, what are the datasets of Table 1, what is the degree of required accuracy, and why only the number of spans of the bridge is used as a changing parameter for such a complicated coupled system. Changes in the speed of the train, the dynamic properties of the wagons, the physical properties of the bridge, or the contact characteristics between the train and bridge are more crucial for the dynamics of such a coupled system, especially in the case of long-term predictions, which the present study stresses as critical.

Authors' Reply:

Thank you very much for your valuable feedback. In response to the questions you raised, we will address each of them in detail and strengthen the corresponding content in the revised manuscript:

1. Explanation of Train and Bridge Properties:
    - This paper adopts an innovative regional graph representation approach, where different regions of the actual structure are abstracted as nodes. By doing so, we successfully reduce the size of the graph

structure, thus improving computational efficiency. Each node represents a structural region that may involve different physical properties in the real world and can adapt to various structural changes. Specifically, the properties of the bridge nodes include span length, material properties, and cross-sectional characteristics, which allow us to model different types of bridges using graph structures. Additionally, we can expand other features as needed based on practical requirements, but for simplicity and focus, we currently only use the aforementioned features as the basic characteristics of the bridge nodes.

2. Regionalization for Node Selection:
    - When choosing the number of nodes for the train cars and the bridge spans, we considered the balance between computational efficiency and accuracy. Assigning one node to each train car and three nodes to each bridge span is mainly aimed at optimizing the graph structure through regionalization, while avoiding unnecessary computational complexity. With this approach, we can significantly enhance computational efficiency without compromising model accuracy. In fact, in the revised version, we have further simplified this structure by merging the original three bridge nodes into one, further optimizing the graph structure size.
3. Detailed Description of Node Features:
    - In Section 2.2, we have already provided a detailed explanation of the main features of each node. The bridge node features include span length, material properties, and cross-sectional characteristics, while the train car node features are defined based on its dynamic characteristics, weight, and other factors. Additionally, the dataset in Table 1 is derived from Matlab simulation experiments, which include verified train-bridge coupling data used as the training, testing, and validation sets to ensure the model's efficiency and reliability.
4. Application of Graph Attention Mechanism:
    - As you mentioned, factors such as train speed, dynamic characteristics of the cars, physical properties of the bridge, and the contact characteristics between the train and bridge indeed have a significant impact on the dynamics of the train-bridge coupling system. To model

these complex interactions, we have employed a graph attention mechanism. This mechanism allows us to automatically learn and simulate the stiffness transmission process between various components in the graph structure. In the model, the dynamic characteristics of the train cars, physical properties of the bridge, and seismic inputs of the bridge piers are input as information to the respective nodes. By modeling this graph structure, we utilize the graph attention mechanism to adaptively learn and simulate the relationships between nodes, thus enhancing the model's generalization ability in various scenarios. The training process uses supervised learning to ensure that the model can quickly and autonomously handle different configurations of the train-bridge coupling system.

5. Focus on Bridge Span Length Variations:
    - The primary focus of this study is on the variation of the bridge span length, as from a practical perspective, the structure and span number of the bridge typically undergo more changes during the train's travel. Therefore, changes in the bridge span length have a more significant impact on the dynamics of the train-bridge coupling system. To ensure that these variations are adequately considered, we have emphasized the analysis of different combinations of bridge span lengths and their effects on the system in our experiments. We also recognize that the dynamic characteristics of the train cars and their variations are equally important to the system, and if needed, we can extend the study to include changes in the train car characteristics in future research to further validate and refine the model.

In conclusion, we greatly appreciate the valuable suggestions you have provided. These questions have allowed us to more clearly present the design ideas and innovations of the system in the revision. We believe that through these improvements and additions, our research will become more complete and provide valuable references for researchers in related fields.

However, when facing such complex dynamic systems, effectively modeling large and complicated systems remains a significant challenge. Many real-world systems consist of multiple coupled subsystems with intricate interactions, making traditional modeling methods difficult to apply. Additionally, predicting long-term sequences of dynamic systems poses another challenge. As time progresses, the system's state continuously changes, especially when external disturbances or internal parameters fluctuate, which further complicates prediction. Traditional forecasting methods may fail to capture such long-term changes, particularly in systems exhibiting strong nonlinearity or high dimensionality. Therefore, the development of algorithms capable of long-term forecasting, especially those that can adapt to system changes, is a crucial area of research. Furthermore, as dynamic systems evolve (such as changes in structure), ensuring that prediction accuracy is maintained in the face of ongoing changes is a key concern. For example, in vehicle-bridge coupled systems, different trains and bridge structures may lead to variations in system behavior. Similarly, in traffic flow, different urban planning strategies can induce changes in the system. In such cases, system behavior is often challenging to predict precisely. Improving the robustness of models to account for random fluctuations and external disturbances is essential to ensure effective prediction and control of dynamic systems.

For train-bridge dynamic systems, numerous advanced computational models have emerged for train-bridge dynamic systems. Zhai's foundational theory of train-bridge coupled \cite{zhai1, zhai2, zhai3} has provided a solid base for the calculation of train-bridge dynamics. Xiang introduced a high-speed train model under wind load \cite{Xianghuoyue1, Xianghuoyue2, Xianghuoyue3}, addressing a gap in the field. Elias G. Dimitrakopoulos proposed an improved method for decoupling the vehicle-bridge interaction under seismic effects \cite{Elias1, Elias2, Elias3}, thereby improving computational efficiency.. These optimized models and methods have significantly advanced the study of train-bridge coupling. However, the issue of computational efficiency in real-time applications remains a challenge. To address this, several deep learning models have been proposed for more efficient computation.The neural network models discussed above exhibit limitations in handling dynamic train-bridge systems, particularly when dealing with the continuously evolving bridge and train structures. These changes cite{Sun2024AUG6} make it challenging to accurately capture the impact of such variations on system responses. Consequently, graph-based neural networks have emerged as a promising research area, as they can more effectively represent the dynamic relationships between various components of the system.X. Peng et al. proposed the GN Block model \cite{xiang2024adaptive}, a highly adaptive graph neural network designed to address complex structural response prediction tasks, especially in dynamic train-bridge coupled systems. This model is capable of adjusting its network structure and learning strategy in real-time, thereby enhancing both adaptability and prediction accuracy.Similarly, P. Zhang et al. developed an adaptive graph neural network model \cite{zhang2024novel} that not only adapts to known structures but also autonomously predicts the responses of unknown structures. This adaptability allows the model to handle various types of train-bridge systems and deliver accurate predictions.These graph-based neural networks \cite{Elias4, 2020arXiv200511650W, 2017arXiv170904875Y, 2019arXiv190600121W} address the limitations of traditional neural networks, which struggle to fully leverage topological information. By aggregating information from nodes and edges, they capture the intricate relationships between different system components, thereby improving prediction accuracy and computational efficiency, especially in complex and dynamically changing structures. The application of these models has significantly advanced research in train-bridge dynamics, driving it towards greater efficiency and precision.

Although existing Graph Neural Networks (GNNs) have successfully integrated graph structure information as topological features into models, several significant challenges persist in practical applications. These challenges primarily include: (1) decreased computational efficiency when the graph structure is excessively large; (2) difficulties in sharing information between nodes and propagating global information effectively in large graph structures; and (3) a lack of a unified framework in existing methods, which hinders effective generalization across different types of graph structures. These issues limit the scalability and practical applicability of GNNs in large-scale, complex systems.

[5] W. Zhai, Z. Han, Z. Chen, L. Ling, S. Zhu, Train – track – bridge dynamic801
interaction: a state-of-the-art review, Vehicle System Dynamics 57 (7)802
(2019) 984 – 1027. doi:10.1080/00423114.2019.1605085.803
URL https://doi.org/10.1080/00423114.2019.1605085804
[6] W. Zhai, K. Wang, J. Lin, Modelling and experiment of railway ballast805
vibrations, Journal of Sound and Vibration 270 (4) (2004) 673 – 683.806
doi:https://doi.org/10.1016/S0022-460X(03)00186-X.807
URL https://www.sciencedirect.com/science/article/pii/808
S0022460X0300186X809
[7] W. Zhai, K. Wang, C. Cai, Fundamentals of vehicle – track coupled dy-810
namics, Vehicle System Dynamics 47 (11) (2009) 1349 – 1376. doi:811

10.1080/00423110802621561.812

URL https://doi.org/10.1080/00423110802621561813

Comment 4:

Finally, no detail is given as to how the other approaches that are used for comparison (LSTM, GNN, GATs, GATs+LSTM, GNBlock) are implemented, which is crucial to prove the superiority of the proposed method against existing approaches.

Authors' Reply:

Thank you very much for pointing out the issues. In the revised version, we have further improved the experimental section, where we provided detailed configurations and parameter settings for the other baseline models. We have carefully considered that differences in model configurations could significantly impact the results in comparative experiments. Therefore, to ensure fairness, we standardized the configurations across all baseline models (including LSTM, GNN, GAT, GAT+LSTM, and GNBlock) to ensure that each model is compared under identical conditions. Specifically, we standardized the hidden layer dimensions, encoding layers, and output layer structures for all models to avoid architectural differences that could introduce unnecessary biases in the experimental results.

Additionally, we paid particular attention to the graph structure representation capabilities of models based on Graph Neural Networks (GNN) and Graph Attention Networks (GAT). For these models, we ensured that the number of graph neural network layers matched that of the GAttention model, ensuring equal conditions for capturing graph structural information. This setup allows for a more accurate evaluation of the advantages of the GAttention model in terms of graph structure representation.

our goal is to demonstrate the superiority of the GAttention model in the field of graph neural networks through rigorous experimental design and fair comparison, ensuring the fairness and reliability of the evaluation results.

Reply to the reviewer's comment (Reviewer#3)

General comments:

Authors' Reply:

Your valuable comments have greatly helped to improve the quality of our article. We have done our best to reply to your comments and hope that we have not misunderstood your meaning. If there are any unclear points raised, we hope you can let us know during the next revision and we will carefully revise the article to your satisfaction. Thank you very much!

Comment 1:

Although the analysis of train-bridge systems may encounter computational problems (if a fined finite element model is applied), numerous previous studies have demonstrated that appropriate simplification and advanced modeling techniques are effective in lowering the computational cost while maintaining a satisfied prediction accuracy. For example, Homaei, H., Stoura, C. D., & Dimitrakopoulos, E. G. (2024). Extended Modified Bridge System (EMBS) method for decoupling seismic vehicle-bridge interaction. Earthquake Engineering & Structural Dynamics, 53(13), 4054-4075- Zeng, Q., Yang, Y. B., & Dimitrakopoulos, E. G. (2016). Dynamic response of high speed vehicles and sustaining curved bridges under conditions of resonance. Engineering Structures, 114, 61-74.- Du, X. T., Xu, Y. L., & Xia, H. (2012). Dynamic interaction of bridge–train system under non-uniform seismic ground motion. Earthquake Engineering & Structural Dynamics, 41(1), 139-157.

Authors' Reply:

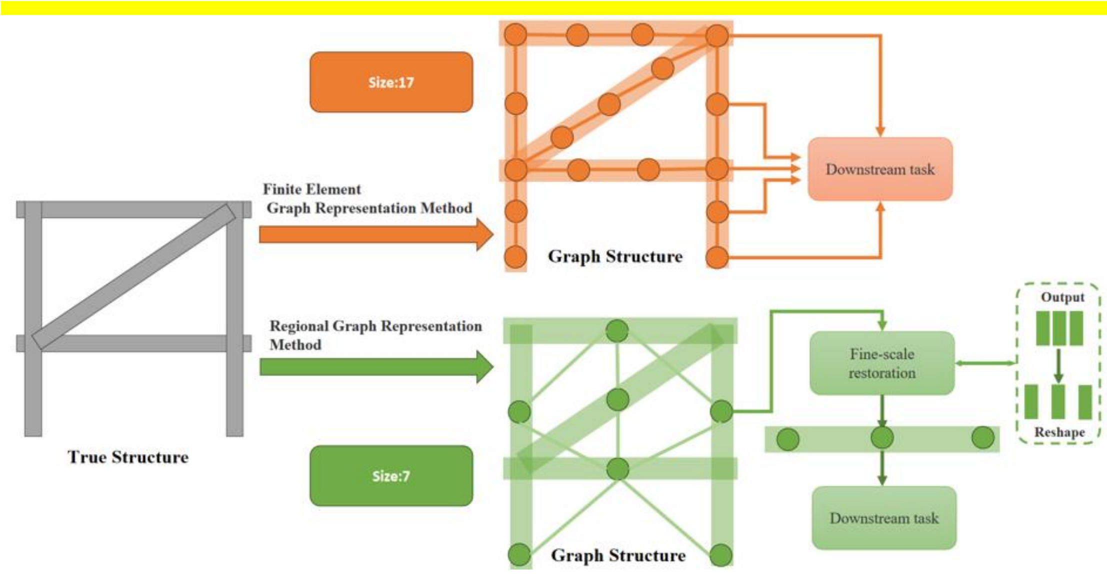
Thank you very much for your feedback. We carefully read the latest paper you provided and found it to be very valuable, so we have cited it and included it in the introduction section.

The train-bridge coupled system, as a complex coupled system, is the focus of this paper. The aim of this work is to use deep learning theories and methods to rapidly calculate the train response under seismic actions for train early warning. The early warning system requires the computation of the train's response for the next 2 to 3 seconds within 0.5 seconds, which traditional finite element methods cannot achieve in such a short time.

Unlike traditional graph neural network representations (e.g., treating each finite element grid node as a graph node), this paper adopts a regional graph representation method, where the entire region is abstracted as a single graph node. This method significantly simplifies the graph structure, speeds up the computation, and also allows for more detailed predictions of the entire region's response (whether to model the response of the entire region depends on specific needs, but this study focuses only on the wheel response). In brief, by using the regional graph structure strategy, this paper abstracts the complex train body as one node, abstracts a bridge span as three bridge nodes, and abstracts the pier as a single pier node. This simplified graph structure allows for high performance while achieving very efficient computational efficiency.

Once again, thank you for your valuable suggestions, and we look forward to your further feedback.

[11] Q. Zeng, Y. Yang, E. G. Dimitrakopoulos, Dynamic response858
of high speed vehicles and sustaining curved bridges under coURL
https://www.sciencedirect.com/science/article/pii/862
S0141029616000882863
[12] X. T. Du, Y. L. Xu, H. Xia, Dynamic interaction of bridge train864
system under non-uniform seismic ground motion, Earthquake En-865
gineering & Structural Dynamics 41 (1) (2012) 139 – 157. arXiv:866
https://onlinelibrary.wiley.com/doi/pdf/10.1002/eqe.1122,867
doi:https://doi.org/10.1002/eqe.1122.868
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.8691122870
[13] H. Homaei, C. D. Stoura, E. G. Dimitrakopoulos, Extended modified871
bridge system (embs) method for decoupling seismic vehicle-bridge872
interaction, Earthquake Engineering & Structural Dynamics 53 (13)873



Comment 2:

The utility of graph-based machine learning models to represent nonlinear structural systems to prompt response prediction is not new. For instance,

- Ref [8] in this manuscript.
- Chou, Y. T., Kuo, P. C., Li, K. Y., Chang, W. T., Huang, Y. N., & Chen, C. S. (2024). Inductive graph-based long short-term memory network for the prediction of nonlinear floor responses and member forces of steel buildings subjected to orthogonal horizontal ground motions. Earthquake Engineering & Structural Dynamics.

Authors' Reply:

Thank you very much for providing the paper on graph neural networks. We all agree that it represents the cutting-edge direction in this field and have cited it and included it in the introduction section. Graph neural networks have been widely applied to structural response prediction. However, as mentioned in the introduction of this paper, existing models (such as the one you provided) typically abstract the finite element model of the structure as a graph structure (this paper uses a regional graph structure representation method and fine-scale restoration) and embed certain structural properties into the graph nodes and edge feature vectors. By adjusting these features, they aim to predict the response of different structures. However, this method essentially does not alter the fundamental form of the graph structure, so it is only suitable for structures with fixed boundary conditions. For structures with varying boundary conditions, such as bridges with different spans, the input of seismic waves increases, and the traditional method of keeping the graph structure unchanged cannot handle this situation.

At the same time, most existing models (like the one you provided) adopt a step-by-step prediction approach, where the prediction for the next step depends on the result of the previous step. This is a common issue in recurrent neural networks, where errors can accumulate over time, and prediction efficiency is low. In contrast, this paper uses an encoder-only architecture that outputs all results at once, without the need for iterative prediction, which, in terms of efficiency, is far superior to the step-by-step approach.

We carefully studied the paper you provided, which embeds some unique structural properties (such as the first and second mode shapes) into the graph structure as inherent properties and generates a global graph vector through the graph aggregation process. This vector is then used as an auxiliary feature to help nodes with time-series prediction. The paper suggests using backpropagation to train this graph embedding. Firstly, we believe there is a more reasonable approach, namely using an autoencoder for unsupervised learning to obtain an embedded graph vector, which can support subsequent tasks effectively. Secondly, this approach almost ignores the coupling between different nodes in the time dimension, while the actual structural response fully depends on the response of the previous time step. This method overlooks the information exchange between nodes in the time dimension. In contrast, in this paper, temporal information and spatial structural information are exchanged simultaneously, similar to the coupling effects in physical structures, which is a significant difference from the existing methods.

.

Comment 3:

Given the stated dynamic systems on page 3, the last paragraph of the Introduction, the authors should clarify the definition of ‘complex dynamic systems’

and existing challenges more rationally.
- ‘Although the current GATtention model is primarily applied to structural response analysis in civil engineering, its efficient computational capabilities and strong nonlinear adaptability make it potentially expandable to other fields, such as weather forecasting and financial market analysis, which also require handling complex dynamic systems and large volumes of time series data.’.

Authors’ Reply:

Thank you very much for your valuable suggestions. Based on your feedback, we have further refined the paper. First, we have clearly added the definition of the concept of "dynamic system" in the article to ensure that readers can better understand the nature and characteristics of the system discussed in the paper. In addition, we have revised some of the wording in the paper, removing imprecise or unclear expressions, making the content more concise, clear, and easier to understand.

To further enhance the practical applicability and experimental nature of the paper, we have also added an experimental case on traffic flow prediction. This case effectively demonstrates the advantages and potential of the regional graph representation method we proposed. Through these modifications and additions, we aim to further improve the scientific, systematic, and practical value of the paper, and provide readers with a more comprehensive and clearer research outcome. Once again, thank you for your suggestions, which have provided important guidance and support for the improvement of the paper.

Dynamic systems \cite{Li2021,Katada2007JUN} refer to systems whose states evolve over time, influenced by initial conditions, external inputs, and governing laws. The behavior of such systems is typically described by specific mathematical models or physical laws that govern their evolution, often involving time-dependent variables.
.

<u>Comment 4:</u>

On page 3, the last paragraph of Introduction, the perspective in SHM is interesting, however, this paper did not present any outcomes.

- ‘The GATtention model can further develop its real-time data processing capabilities, achieving faster responses and higher real-time prediction accuracy, especially in applications such as structural health monitoring and immediate disaster response.’

Authors’ Reply:

Thank you very much for your sincere suggestions. The original design intent of GATtention was to handle dynamic systems, and health monitoring has become a highly regarded research area in recent years, especially in the fields of engineering and structural health. Health monitoring generally refers to the continuous monitoring and assessment of the health status of structures or equipment to ensure their safety, stability, and functionality. During the health monitoring process, various sensors, advanced technologies, and sophisticated algorithms are employed to track the operating status of the structure or system in real-time. This allows for the early detection of potential risks or faults, enabling timely preventive and corrective actions to avoid catastrophic accidents.

This process aligns closely with the design philosophy of GATtention. GATtention integrates sensor data from different regions to comprehensively assess the health status of the structure and identify potential issues. Essentially, this is a classification problem, and the GATtention model is fully capable of addressing this task, particularly in classifying and predicting the health state of structures. However, due to the lack of suitable publicly available datasets, this paper does not explore this direction in detail.

Nevertheless, this paper introduces a new experimental case—traffic flow prediction—to demonstrate the application of the model in a similar task. Traffic flow prediction shares certain similarities with health monitoring, especially in terms of handling dynamic data, pattern recognition, and early warning processes. Through the traffic flow prediction experiment, we are able to verify the effectiveness of the model in processing real-time dynamic system data and showcase its potential.

In the future, we plan to delve deeper into the topic of structural health monitoring in subsequent research, integrating more real-world application scenarios and datasets to further validate the applicability and advantages of the GATtention model in this domain. We believe that as the research progresses, GATtention will not only provide innovative solutions for health monitoring but also offer valuable technical insights for addressing other dynamic system problems.

Comment 5:

On page 5, lines 117-127, regarding the modeling object, it appears that the system and numerical model used in this study are simple cases of actual train-bridge systems and simplified models of this system, respectively. As a result, the computational cost seems not to be a critical problem, as mentioned in the Introduction.

- 'Each train carriage is abstracted as a train node, and each span of the bridge is abstracted into three bridge nodes'

Authors' Reply:

Thank you very much for your valuable suggestions and comments. The original intention of this paper was to address the long-term time series prediction of coupled responses in dynamic systems. However, the TBC system, being a complex system, despite the existence of some simplification methods, still falls short of meeting real-time computational requirements in practical applications. Therefore, this paper focuses on the issues of computational cost and time in long-term time series prediction for the TBC system, aiming to propose a solution that can ensure prediction accuracy while significantly reducing computational time and costs.

In this paper, the vehicle-bridge coupled system used is entirely based on actual finite element modeling. All response data were calculated using Matlab programs and compared with real experimental data from vehicle-bridge models, validating the feasibility of the data and the accuracy of the model. The finite element analysis process involves a large number of finite element divisions, which itself results in a huge computational load, and traditional computational methods are inadequate for meeting the demand for rapid real-time predictions. As a result, we propose a new

approach to optimize the computational process to achieve more efficient predictions. Specifically, the method mentioned in the paper—"each train carriage is abstracted as a train node, and each bridge span is abstracted as three bridge nodes"—is a way of abstracting the complex finite element structure through regional graph representation, converting the actual finite element structure into graph nodes within the graph structure. This abstraction method of the graph structure simplifies the complex structural system into a graphical representation, greatly improving computational speed. By using this approach, we can reduce the processing and division of finite element units during computation, while also improving computational efficiency, enabling faster predictions of structural responses without compromising accuracy.
The starting point of this method is to ensure a balance between prediction speed and computational accuracy, allowing for quicker responses to the long-term time series prediction requirements of complex systems in practical engineering applications. With the continuous advancement of technology, we believe that this regional graph representation method will have broad application potential in dynamic system prediction, providing an effective solution to the computational bottleneck problem of large-scale dynamic systems. In the future, we will continue to deepen research in this method, explore its application in other complex engineering problems, and optimize its performance to adapt to more complex and dynamically changing practical requirements.

Comment 6:

Do the generated earthquake ground motions based on the Cloug-Penzien model only include the y and z components, while the x component is excluded? What is the reason for neglecting this component?.

Authors' Reply:

It is my pleasure to clarify a misunderstanding for you. You may have confused the feature vectors for train nodes and pier nodes. In this paper, it is explicitly mentioned that "For pier nodes, which serve as input nodes for seismic accelerations, the feature vector $X_{pier} = \{x_{acc}, y_{acc}, z_{acc}\} \in \mathbb{R}^{N_{pier} \times 3}$ consists of seismic accelerations in the x, y, and z directions." We have considered the ground motion in the x, y, and z directions for the pier nodes. However, for the train nodes, since the x-direction has a minimal effect on derailment, it is not included in the scope of consideration.

Comment 7:

If the proposed machine learning model (or approach) holds the primary innovation, this paper can be considered to be submitted to a journal/conference with ML subjects.

Authors' Reply:

Thank you very much for your sincere suggestions and valuable feedback. Our research covers multiple fields and has broad application prospects, making it suitable for a variety of practical scenarios. In particular, the methods and models presented in this paper demonstrate strong potential and adaptability in areas such as engineering structures, dynamic systems analysis, and seismic response prediction. We believe that this work not only provides important theoretical support to the academic community, but also helps solve many pressing issues in practical applications.

Nature Communications, as a comprehensive journal, is one of the top journals in the field with significant academic influence and a wide readership. We believe that the research findings and innovations in this paper are well aligned with the publishing goals and subject areas of **Nature Communications**. By publishing in Nature Communications, we hope to share our research results with more scholars and engineers, and promote interdisciplinary discussion and collaboration. We are confident that this work will make valuable contributions to the academic community and drive further development in the field.

Comment 8:

In Figure 12, the maximum acceleration response is 6×10^(-1) m/s^2. Do I understand correctly that this represents a small earthquake event? In this context, the implication or the challenge of predicting structural/train responses is questionable.

Authors' Reply:

Thank you very much for your valuable suggestions. Regarding the issue you mentioned, where the maximum acceleration response is 6×10−1 m/s26×10−1m/s2, this value is indeed relatively small, typically reflecting a mild seismic event rather than a strong earthquake. However, when predicting structural or train responses, even though the acceleration is low, there are still certain challenges. Even minor vibrations can lead to small structural deformations, or cumulative effects over time, ultimately impacting the stability of the entire system.

In fact, even with a small acceleration value, the impact of earthquakes or other external disturbances on the structure and train should not be underestimated. Especially in complex dynamic coupled systems, even mild vibrations could lead to gradual degradation of system performance through repeated exposure, or even cause more serious structural instability or train derailments, which could pose safety risks. Such minor deformations, although difficult to detect directly, may accumulate over time and significantly affect the overall response of the system, thus impacting the safety and stability of train operations.

Therefore, the core objective of this paper is to identify and accurately predict train responses before the seismic event occurs, in order to assess whether a derailment or other abnormal situation might happen. The accuracy and consistency of the prediction model, especially in the case of such small accelerations, become particularly important. Even a lower acceleration response may induce complex interactions in a dynamic coupled system, so it is crucial to pay special attention to the performance of the model in this context, ensuring its stability and reliability under various external conditions.

In summary, although the predicted acceleration value is relatively small, in practical applications, especially in high-risk transportation environments, it is still necessary to consider the potential impact it could have. Ensuring that the prediction model maintains high accuracy and consistency in this context will be essential to effectively prevent accidents and ensure the safe operation of trains.

Reply to the reviewer's comment (Reviewer#4)

General comments:

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts. No code available for review for the paper.

Authors' Reply:

Reproducible code is a crucial component of research integrity, as it ensures that research results can be verified and disseminated. In response to the journal's requirements and to further enhance the transparency and credibility of our work, we have made the relevant code used in this paper publicly available for other researchers to reproduce and reference. We believe that open-sourcing code not only fosters collaboration within the academic community but also helps peers better understand and apply our models and methods, thus advancing the field.

Specifically, we have uploaded the core algorithms and experimental implementations from our paper to GitHub. The code repository provides detailed records of our modeling process, data processing methods, model training procedure, and evaluation metrics, ensuring that other researchers can replicate our experiments using the same setup to validate our results. This initiative is a significant step toward promoting research transparency and reproducibility.

You can access our open-source code repository through the following link: https://github.com/Amos-Yooupi/GAttention. On this page, you will find the complete code, documentation, and related dependencies, making it easier for you to conduct local tests and experiments. We also welcome any suggestions for improvements or feedback regarding the code, and we hope to collaborate with more researchers to explore and refine this work.

Thank you once again for your interest and support. We look forward to sharing our research findings with the broader academic community and contributing to the advancement of the field.

Responses to Reviewers' Comments

Dear Editor and Reviewers,

First of and foremost, we would like to express our sincere thanks and appreciation to the editors and reviewers for your professional handling and review of the manuscript. Your comments and constructive suggestions were very helpful in improving the quality of our manuscript and helped us identify several shortcomings in earlier versions of the manuscript. We have used these valuable suggestions as a guide to address these shortcomings where appropriate. The paper has now been revised following the reviewers' comments, changes in the manuscript have been marked in "blue". In the reply to comments, all the changes are marked in "yellow". Please check the following responses.

Thank you very much!

Best regards,
PingXiang

Reply to the reviewer's comment (Reviewer#1)

General comments:

Thank you very much for your recognition and valuable feedback on our paper. In this version, we have made comprehensive improvements to the paper, significantly enhancing its quality, as well as optimizing the expression and overall structure. We have more clearly articulated the core innovations of this paper—introducing a novel deep learning architecture designed to address the challenges and limitations of existing methods.

Our innovations are primarily reflected in the following aspects:

1.Regional Graph Representation: We introduce a novel regional graph representation method that significantly improves the computational efficiency of graph structures through region partitioning while maintaining the model's prediction accuracy.

2.Fusion Graph Convolutions: To address the issue of insufficient information propagation in large-scale graphs, we combine global and local information, aggregating different levels of data, which allows nodes to gain a global perspective, thereby enhancing the breadth and depth of information propagation.

3.Time-aware Sparse Expert: We propose a time-aware sparse expert mechanism that dynamically adjusts the selection of experts based on the characteristics of different time series. This mechanism significantly improves the model's flexibility and accuracy when handling long-term sequence tasks.

Additionally, we provide a detailed explanation of the motivation behind these innovations, analyze their theoretical foundations, and validate their effectiveness through extensive experiments. The experimental results show that the proposed architecture demonstrates superior performance across multiple tasks, particularly in handling complex dynamic systems, where it shows clear advantages.

Reply to the reviewer's comment (Reviewer#2)

General comments:

Authors' Reply:

Thank you sincerely for your recognition of our work and the valuable feedback on our manuscript. We fully agree that refining and enhancing a paper through reviewers' insights is crucial to elevating its quality, and we have taken your comments seriously by making substantial revisions to align with publication standards.

It is important to emphasize that the core objective of this study is to address key challenges in dynamic system modeling by proposing a dedicated deep learning framework. In the development of previous deep learning models, many approaches have often fallen into the dilemma of simultaneous surges in parameter scale and computational complexity: while excessive pursuit of large-scale parameters and high precision can enhance the model's expressive ability to a certain extent, it also leads to a sharp increase in computational complexity—especially when dealing with large-scale dynamic systems and long-term sequence data, where computational costs and storage requirements become prohibitive.

To tackle this, we propose the Regional Graph Representation (RGR) method: it achieves efficient simplification of graph structures through minimal structural adjustments, aggregating connected nodes into a single region represented by a new node. Connections between regions are defined by the original links of cross-region nodes, significantly reducing the scale of the graph structure while preserving core adjacency relationships. By adaptively adjusting the key parameter of "region order", this method can flexibly balance computational efficiency and modeling accuracy; combined with fine-grained reconstruction technology, it can reversely restore node features within regions to the original graph structure, providing feasibility for accurate modeling of large-scale dynamic systems.

To address the insufficient information interaction between nodes in large-scale graphs, we further propose the Fusion Graph Convolution (FGC) module: by balancing local neighborhood details and global system trends, it effectively expands the information perception range of nodes and improves the comprehensiveness and precision of information flow transmission. It is worth noting that although the graph representation method has reduced the structural scale, the traditional graph attention mechanism may still increase computational overhead. Therefore, redundant graph attention mechanisms are removed through module optimization in our framework, further improving operational efficiency.

In time series prediction tasks, especially in long-term prediction scenarios, the prediction accuracy of the Gattention model is significantly improved after integrating the Time-Aware Sparse Expert (TASE) module, which fully verifies the core value of multi-scale time analysis. This module captures diverse temporal patterns through sparse expert layers and dynamically optimizes module selection strategies with the help of an adaptive routing mechanism, effectively enhancing the ability to model complex time series.

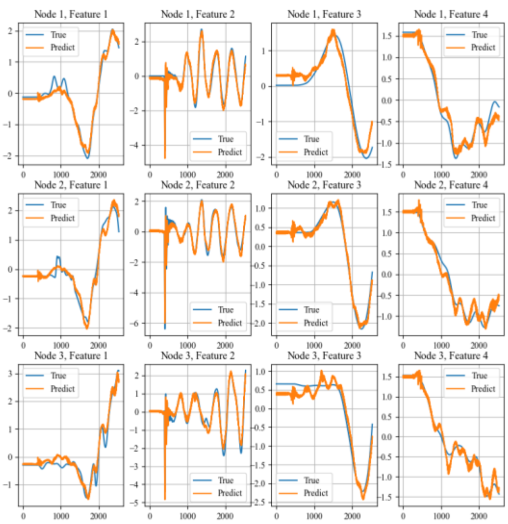
We remain committed to refining these points further based on your feedback and ensuring our contributions are rigorously articulated. Thank you again for guiding us to strengthen the clarity and impact of our work.

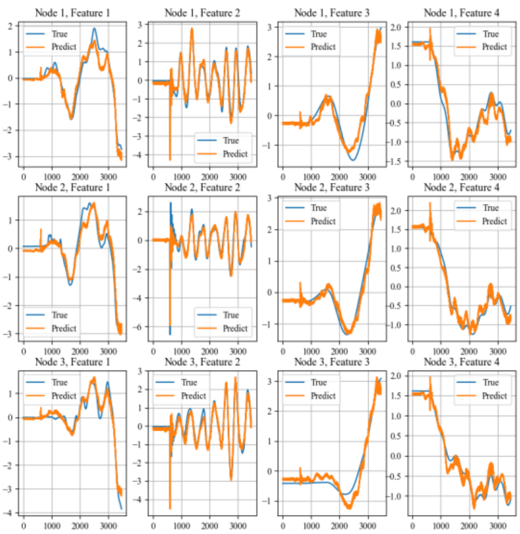
Comment 1:

In lines 305-317, the features used for the train, bridge, and pier nodes in the GNN are presented. However, the rationale behind the chosen features is not clearly discussed. It is mentioned that real-time responses from the bridge are unavailable, while such responses from the train are available, which is reasonable depending on the application. Nonetheless, the assumption that displacements of the train, on top of accelerations, are available in real time may not always be practical. Thus, further justification or clarification is needed.

Authors' Reply:

Thank you very much for your valuable comments; your perspective is crucial to us. In this paper, the train node setting indeed takes displacement into account, but we realize that, in practical situations, displacement may not always be a feasible option. Generally, the state of a train under seismic influence is mainly determined by the seismic effects and the train's acceleration. Therefore, even if displacement is used as a feature vector, its impact on the final prediction results is relatively small. At the same time, the core purpose of this paper is to validate the generalization capability of the Gattention framework through the vehicle-bridge coupling example, to adapt it to different types of bridge structures. Hence, the choice of feature vectors has no significant impact on the problem under study. We have further clarified the selection of feature vectors and sincerely appreciate your valuable suggestions, which have been instrumental in enhancing our paper. The following prediction figures shows the results without considering the train displacement.





The code of **DataIterTBC** is modified as follows:

```python
class DataIterTBC(DataIter):
    def __init__(self, data_file_path, split_ratio: list):
        super().__init__()
        assert sum(split_ratio) == 1, "Split ratio should sum up to 1"
        # 读取数据
        self.vehicle, self.bridge, self.pier = read_file_TBC(data_file_path, time_freq=10)
        # 划分数据集
        dataset_dict = self.split_data( *original_data: self.vehicle,
                                        self.bridge, self.pier,
                                        split_ratio=split_ratio)
        # 归一化数据
        self.dataset_dict = self.norm_dict(dataset_dict)

        # 设置参数
        self.adj, self.num_train, self.num_bridge = None, None, None
        self.get_adj()
        # 只预测列车节点的响应
        self.num_task = self.vehicle.shape[0]

        self.set_train_dis_to_zero()

    1 usage
    def set_train_dis_to_zero(self):
        self.vehicle[:, :, 2:] = 0
```

The new definition for the feature vector of the train node has been provided, and the selection criteria for displacement are discussed:

For train nodes, in practical situations, displacement may not always be a feasible option. Generally, the state of a train under seismic influence is mainly determined by the seismic effects and the train's acceleration. Therefore, even if displacement is used as a feature vector, its impact on the final prediction results is relatively small.so their feature vector consists of displacements and accelerations in the y and z directions.

<u>Comment 2:</u>

Section 4.1 is somewhat unclear and would benefit from a more structured explanation. The authors state that the dataset is "derived from MATLAB computations" but do not specify what this entails. Details about the models used for the vehicle, bridge, and piers—including their physical parameters and assumptions—are missing. Additionally, the statement that "experimental verification shows that the discrepancies between these results and those obtained from structural simulations are negligible" lacks supporting evidence and should be substantiated with references or quantitative comparisons.There is further confusion regarding the description of seismic waveform data preprocessing. While the manuscript refers to "collected" seismic data and discusses preprocessing steps such as denoising and baseline correction, it appears that no actual experimental signals are involved in this study. The authors should clarify this part and avoid including parts that are out of context in the manuscript. Finally, the last paragraph of this section seems unnecessarily long and somewhat out of place, with a disproportionate focus on dataset downsampling that could be more concisely stated.

Authors' Reply:

Thank you very much for your detailed review and valuable suggestions on our paper. We fully understand your concerns regarding Section 4.1, and we have clearly explained in the revised manuscript that the dataset used in this paper is generated based on the MATLAB structural dynamics calculation framework developed by our research team in previous work. This framework integrates a train-bridge coupled vibration model, and the specific calculation process and validation methods are detailed in reference [40, 41, 42]. We have added the relevant citation to Section 4.1 and included the following explanation.

Our previous research has thoroughly validated the vehicle-bridge coupling system. In our laboratory, we have conducted extensive experiments, and for more details, please refer to reference [42].

FIGURE REDACTED

FIGURE REDACTED

Thank you very much for your valuable suggestions. We fully understand your concerns regarding the seismic waveform data preprocessing section, and we have clarified this part in the revised manuscript. We have reviewed the relevant content and removed sections that are not directly related to the objectives of this study, ensuring that the focus is more on the core aspects of the research. Finally, regarding the last paragraph of this section, which discusses dataset downsampling, we have also discarded it, further streamlining the content to ensure that the manuscript stays focused on the key elements of the study.

The TBC dataset, obtained from Matlab computations \cite{ZHAO2023105244, zhao2022seismic}, exhibits negligible discrepancies when compared to structural simulations. Therefore, future research will prioritize data from numerical computations \cite{PANG2018233}, using seismic waveforms based on the Clough-Penzien model \cite{zhao2022seismic} and recorded by high-precision sensors during actual earthquake events.

<u>Comment 3:</u>

The manuscript tends to repeat claims about the model's effectiveness, especially in Section 4.2, where long-term prediction capabilities are mentioned multiple times without accompanying detail. Statements such as "the significant performance improvement of the Gattention model is primarily attributed to its precise simulation of real-world data structures and its high adaptability" are not well supported by quantitative analysis or specific evidence. Such claims should be either substantiated or more conservatively stated.

Authors' Reply:

Thank you for your valuable feedback. We fully understand your concerns. Indeed, in Section 4.2, the description of the model's validity contained redundancy and lacked specificity, which may have caused confusion for readers. In the revised manuscript, we have significantly removed the redundant parts, ensuring that the content is more concise and focused. We have paid particular attention to strengthening the clear presentation of the methods and experimental conclusions, making each conclusion more rigorous and precise.

In this revision, we have also enhanced the quantitative analysis and empirical support for the long-term prediction capability of the model. We conducted more detailed data analysis regarding the model's prediction performance. Specifically, we provided more intuitive experimental results and validated the model's long-term prediction effectiveness through various perspectives and comparative experiments. These enhanced empirical supports provide a solid foundation for the model's validity, ensuring that the conclusions of the paper are more reliable.

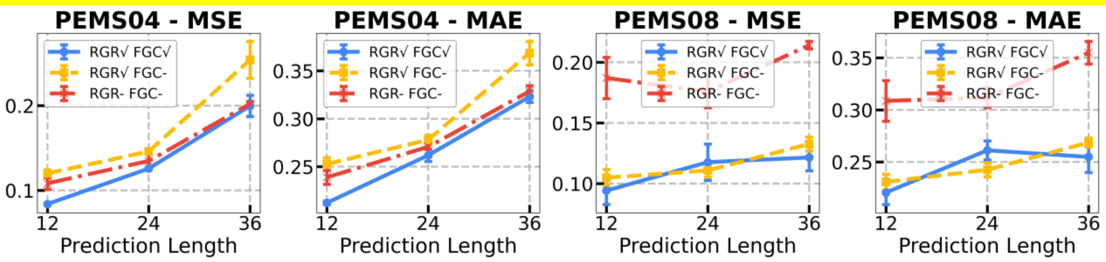
Thank you for your feedback, which has provided us with valuable insights and helped us further improve the rigor, logic, and readability of the paper. We believe that this revision has significantly enhanced the overall quality of the paper, enabling us to better convey our research findings.

The detail experiment for efficiency is as follows:

As shown in Table \ref{tab:region}, the graph structure utilizing regional graph representation (RGR) achieves significant reductions in node count, computation time, and memory usage compared to the original graph structure. For the PEMS08 dataset, with a region order of 3, the total number of nodes drops from 170 in the original graph to 22, marking an 87\% reduction. Concurrently, the computation speed is substantially improved: the processing time decreases from 1.646 seconds in the original graph to 0.036 seconds, representing an approximately 45.7-fold acceleration. Memory consumption also sees a notable decrease of 87\%, falling from 23.15 MB to 2.99 MB. In the PEMS04 dataset, similar performance gains are observed. When the region order is set to 3, the total node number reduces from 307 to 64, a 79\% decrease. The computation speed is enhanced by about 22.4 times, with the processing time shortening from 3.316 seconds in the original graph to 0.148 seconds. Additionally, memory usage is reduced by 79\%, declining from 39.79 MB to 8.29 MB.

Dataset	Graph Structure	Region Order	Total Node Number	Speed	Memory
PEMS08	Original Graph	–	170	1.646 s	23.15 MB
	RGR	1	64	0.115 s	8.71 MB
		2	33	0.055 s	4.49 MB
		3	22	0.036 s	2.99 MB
PEMS04	Original Graph	–	307	3.316 s	39.79 MB
	RGR	1	129	0.498 s	16.72 MB
		2	78	0.152s s	10.11 MB
		3	64	0.148 s	8.29 MB

Table 2 Comparison of Graph Structures with Different Region Orders



<u>Comment 4:</u>

The ablation study of Section 4.3 aims to assess the contribution of individual modules to the model's performance. However, the methodology is not described in sufficient detail. For instance, it is unclear whether only the graph structure was altered or if other components, such as the fusion graph convolution or time-aware experts, were also varied. Likewise, the formation of regions is not clearly defined. How is the central node selected, what determines region size, and do regions overlap?

Moreover, in Figure 13, it is unclear whether the comparison models (e.g., LSTM, GNN) were applied to the full graph or to the regionalized version. A more comprehensive analysis, including a combined ablation, e.g., multiple modules removed simultaneously, would offer deeper insight into which components are most critical to performance. The study lacks a clear, quantitative summary of which modules drive improvement.

Authors' Reply:

Thank you for your valuable feedback. We have indeed noticed that the description of the ablation study in Section 4.3 lacks clarity in some areas. In the revised version, we will provide a more detailed explanation of the following aspects:

1. **Experiment Setup and Component Replacement**: We have revised the ablation section to provide a more detailed experimental setup and explicitly describe the components that are enabled or replaced in the experiments.

The objective of our experiment is to predict the traffic flow of all nodes within the largest region. Through this task setup, we aim to validate the effectiveness of the fine-grained reconstruction method.

Ablation experiments will be conducted on the Traffic dynamic systems to thoroughly investigate the contributions and performance variations of each model component. Specifically, the Fusion Graph Convolution (FGC) and Regional Graph Representation (RGR) will be simplified. In these experiments, the improvement in system efficiency following the introduction of RGR will be assessed, while also testing whether the model's predictive accuracy decreases when RGR is excluded. The objective is to determine whether the incorporation of RGR enhances computational efficiency without compromising accuracy, thereby providing valuable insights for optimizing the Traffic dynamic systems.

Furthermore, the Time-aware expert module will be replaced with a basic Feed Forward module to evaluate the effectiveness of the Time-aware expert. This experiment will help assess whether the Time-aware expert offers a distinct performance advantage and its overall impact on system performance. Additionally, the effect of different Region Orders on model predictions will be examined, with the goal of optimizing regional representation processing and improving both model performance and generalization ability. Through these ablation experiments, more practical and targeted optimization strategies for the various modules and components within the Traffic dynamic systems will be offered.

2. **Ablation Experiments on PEMS04 and PEMS08 Datasets**: The ablation experiments were conducted on the PEMS04 and PEMS08 datasets, which have different graph structures. The component replacement refers to the swapping of modules within the Gattention framework.

Ablation experiments will be conducted on the Traffic dynamic systems to thoroughly investigate the contributions and performance variations of each model component. Specifically, the Fusion Graph Convolution (FGC) and Regional Graph Representation (RGR) will be simplified. In these experiments, the improvement in system efficiency following the introduction of RGR will be assessed, while also testing whether the model's predictive accuracy decreases when RGR is excluded. The objective is to determine whether the incorporation of RGR enhances computational efficiency without compromising accuracy, thereby providing valuable insights for optimizing the Traffic dynamic systems.

3. **Revised Definition of Regional Graph Representation**: We have redefined the regional graph representation and provided the definition of graph region order. Specifically, the selection of the central node is based on its degree, and we define the node with the highest number of connections as the central node. The nodes that are directly connected to this central node form a region. Once a node is assigned to a region, it will not be included in other region formations, thus preventing overlapping regions. Additionally, we control the size of the region by defining the k-hop neighbors of each node. For further details, please refer to Section Method.

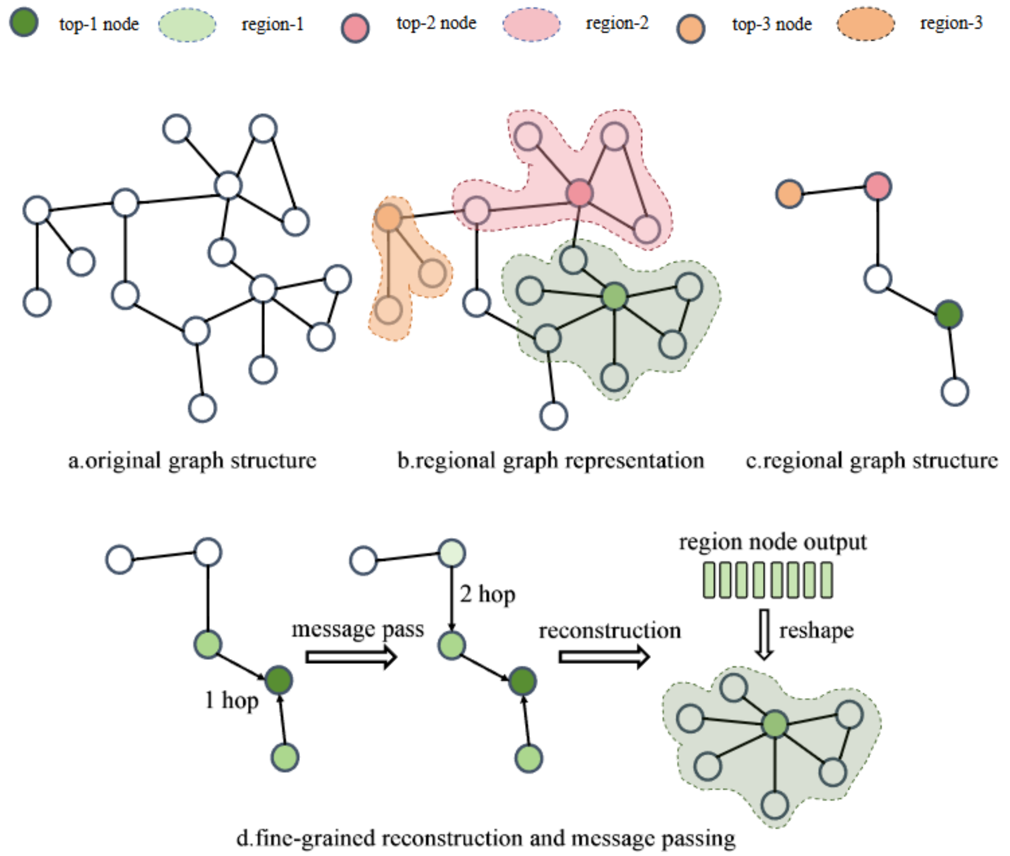


Figure 2: Regional graph structure representation.a.Original graph structure;b.Visualization of regional graph representation;c.Graph structure after regional graph representation;d.Specific details of fine-grained reconstruction, by reorganizing the output results and assigning them to the nodes in region.

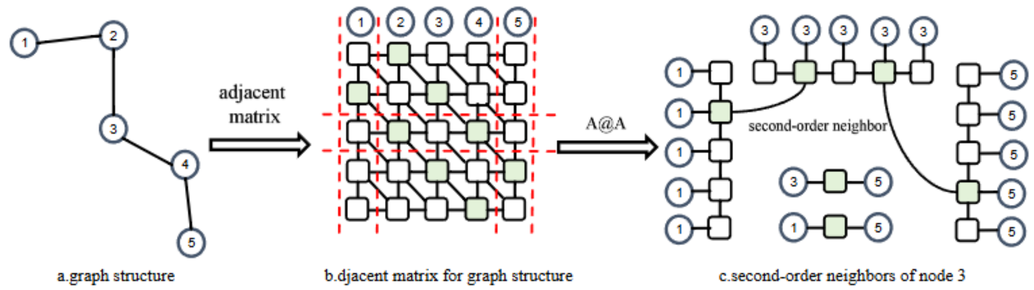


Figure 3: Explanation for region order.a.Original graph structure;b.Adjacency matrix of the graph structure, where bright colors indicate a connection between two nodes;c.The process of finding the second-order neighbors of node 3.

Algorithm 1 Regional Graph Representation Algorithm

Input:
  Original adjacency matrix $A \in \mathbb{R}^{n \times n}$
  Region order $k \in \mathbb{N}$
  Feature matrix $X_{\text{graph}} \in \mathbb{R}^{n \times m}$ (node features)
Output:
  Adjacency matrix $A' \in \mathbb{R}^{n' \times n'}$ of the regional graph representation
  Region feature matrix $X_{\text{graph, region}} \in \mathbb{R}^{n' \times m}$ (node features)
Initialize A_k (from equation 4)
while $\text{Sum}(A_k) \neq 0$ **do**
  Find the node v with the highest degree:

$$v = \arg\max_i \left(\sum_j A_{ij} \right)$$

  Select node v as the center of the region
  Find the k-hop neighbors of node v:

$$N_k(v) = \{ i : d(i, v) \leq k \}$$

  where $d(i, v)$ represents the number of hops between nodes i and v, with no regard to the actual path length.
  Compute the feature vector for the region:

$$\mathrm{X}_{\text{region},i} = \frac{1}{|N_k(v)|} \sum_{i \in N_k(v)} X_{\text{graph},i}$$

  where $X_{\text{graph},i}$ is the feature vector of node i
  Remove the region nodes from A_k:

$$A_k = A_k - A_k[N_k(v), N_k(v)]$$

  Update A_k
end while
Establish connection for region graph structure A_k based on the feature vectors
Output: Final adjacency matrix A', Region feature matrix $X_{\text{graph, region}}$

4. **Application of Comparison Models** : In our study, we present some more complex demonstration cases. Specifically, we conducted experiments using coupled train-bridge structures with varying numbers of train carriages and bridge spans, which represent more complex transportation scenarios in reality. These intricate structures allow us to delve deeper into the model's performance and advantages in handling dynamic and complex temporal data.Additionally, we compared our model with several classic graph neural network temporal models, including STGNN and GraphWaveNet. These classic models have significant application value in time-series prediction tasks. By comparing with them, we were able to thoroughly evaluate the performance improvements of our approach in practical applications. Regarding the application of comparison models LSTM and GNN, it is important to clarify that our task involves multi-target prediction, i.e., predicting multiple nodes simultaneously. For such tasks, LSTM processes the data by outputting high-dimensional data and then reshaping it to achieve multi-target prediction. In contrast, for GNN, we extract the features of the

corresponding nodes directly for prediction. A more detailed explanation can be found in Section Result.

Model	Metric	4 trains			5 trains			6 trains			7 trains		
		5 spans	6 spans	7 spans	5 spans	6 spans	7 spans	5 spans	6 spans	7 spans	5 spans	6 spans	7 spans
GAT [35]	MSE	0.418	0.295	0.242	0.353	0.358	0.344	0.350	0.413	0.312	0.392	0.354	0.333
	MAE	0.468	0.376	0.344	0.427	0.423	0.411	0.415	0.447	0.385	0.441	0.421	0.413
STGNN[36]	MSE	0.335	0.226	0.184	0.273	0.254	0.263	0.261	0.363	0.229	0.297	0.278	0.317
	MAE	0.407	0.348	0.305	0.373	0.348	0.361	0.342	0.400	0.316	0.374	0.371	0.378
GraphWaveNet[37]	MSE	0.138	0.103	0.111	0.124	0.140	0.104	0.133	0.159	0.109	0.147	0.154	0.135
	MAE	0.273	0.235	0.235	0.252	0.267	0.233	0.251	0.282	0.232	0.267	0.268	0.255
GAttention	MSE	**0.088**	**0.080**	**0.072**	**0.098**	**0.0991**	**0.084**	**0.092**	**0.122**	**0.070**	**0.098**	**0.107**	**0.095**
	MAE	**0.221**	**0.204**	**0.189**	**0.227**	**0.224**	**0.209**	**0.205**	**0.242**	**0.186**	**0.217**	**0.223**	**0.211**

Table 1 Comparison of different models under varying numbers of trains and spans.The predicted length is 24, with the best result highlighted.

For graph-based models such as GAT (Graph Attention Network) and GCN (Graph Convolutional Network), graph convolution operations are applied directly, with the output of the feature vectors from each node after convolution used as the final prediction. All graph-based models are trained using a hybrid approach, which incorporates both the spatial information of the graph structure and the temporal characteristics of the train's dynamic response. This ensures that the relationships between graph nodes and the temporal evolution of the train's response can be learned by the models.

We provide comprehensive ablation experiments to validate the effectiveness of the framework proposed in the paper.

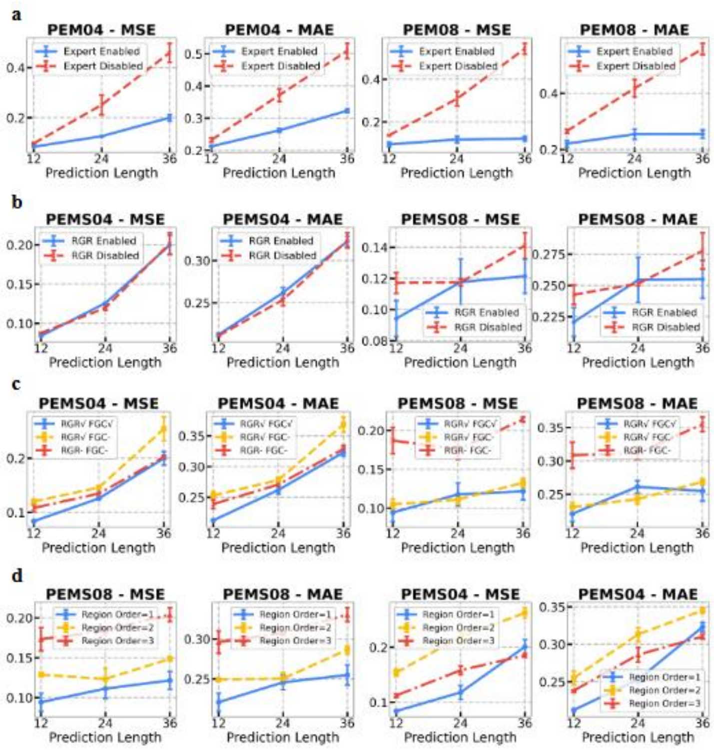


Figure 11: The result of ablation experiemnt.a.Time-aware expert Module impact on prediction accuracy and stability;b.Regional graph representation impact on prediction accuracy and stability;c.Fusion graph convolution impact on prediction accuracy and stability;d.Region order impact on prediction accuracy and stability.

<u>Comment 5:</u>

Inconsistencies in terminology and repetitive phrasing make the manuscript harder to follow. For example, in line 727, Table 3 is said to demonstrate the Fusion Graph Convolution (FGC), but this table does not directly relate to that topic. Abbreviations such as FGC are also not consistently defined upon first mention. Additionally, lines 704-718 repeat ideas previously stated, reducing the overall clarity and conciseness. These issues should be revised for improved readability.

Authors' Reply:

Thank you very much for your suggestions. We have addressed the redundant information in the paper and unified the abbreviations' explanations. This makes the paper more concise and clear, ensuring that readers can better understand the content. We appreciate your advice, as it has been very helpful in improving the quality and readability of the paper.

The statement for the abbreviation has been revised, and we have ensured that it is clear this time:

This experiment as presented in Fig. \ref{fig:ablation} aims to predict the traffic flow of all nodes within the largest region to evaluate the effectiveness of the fine-grained reconstruction method. Ablation experiments are conducted to examine the contributions and performance variations of each model component. Specifically, we simplified the Fusion Graph Convolution and Regional Graph Representation to FGC and RGR, respectively. These experiments will assess improvements in system efficiency upon introducing RGR and determine whether excluding RGR affects the model's predictive accuracy. The objective is to establish whether RGR enhances computational efficiency without compromising accuracy, providing valuable insights for optimizing the traffic dynamic system.

Table 7: Ablation Study on RGR(Best Results Highlighted)

Dataset	Predict Length	RGR	FGC	Time-aware expert	Region Order	MSE	MAE
PEMS04	36	✓			1	**0.1996 ± 0.0124**	**0.3230 ± 0.0059**
		-			1	0.2016 ± 0.0135	0.3239 ± 0.0090
	24	✓	✓	✓	1	0.1257 ± 0.0009	0.2619 ± 0.0063
		-			1	**0.1195 ± 0.0026**	**0.2535 ± 0.0067**
	12	✓			1	**0.0839 ± 0.0006**	0.2123 ± 0.0018
		-			1	0.0877 ± 0.0003	**0.2109 ± 0.0008**
PEMS08	36	✓			1	**0.1215 ± 0.0110**	**0.2549 ± 0.0150**
		-			1	0.1411 ± 0.0082	0.2776 ± 0.0144
	24	✓	✓	✓	1	0.1176 ± 0.0150	0.2544 ± 0.0180
		-			1	**0.1175 ± 0.0023**	**0.2514 ± 0.0015**
	12	✓			1	**0.0942 ± 0.0115**	**0.2207 ± 0.0115**
		-			1	0.1171 ± 0.0067	0.2426 ± 0.0078

The repeated content has been revised. The new content for lines 704-718 is as follows:

As shown in Table \ref{tab:region}, the graph structure utilizing regional graph representation (RGR) achieves significant reductions in node count, computation time, and memory usage compared to the original graph structure. For the PEMS08 dataset, with a region order of 3, the total number of nodes drops from 170 in the original graph to 22, marking an 87\% reduction. Concurrently, the computation speed is substantially improved: the processing time decreases from 1.646 seconds in the original graph to 0.036 seconds, representing an approximately 45.7-fold acceleration. Memory consumption also sees a notable decrease of 87\%, falling from 23.15 MB to 2.99 MB. In the PEMS04 dataset, similar performance gains are observed. When the region order is set to 3, the total node number reduces from 307 to 64, a 79\% decrease. The computation speed is enhanced by about 22.4 times, with the processing time shortening from 3.316 seconds in the original graph to 0.148 seconds. Additionally, memory usage is reduced by 79\%, declining from 39.79 MB to 8.29 MB.

Reply to the reviewer's comment (Reviewer#3)

General comments:

Authors' Reply:

Thank you for your valuable comments; your insights are crucial to refining our research.After revising the paper in light of your suggestions, we have further clarified the core contributions and innovations of this study: To address the challenges of multi-task modeling in long-term time series prediction for dynamic systems, we have proposed a comprehensive deep learning framework. This framework systematically elaborates on dynamic system simplification strategies, model construction logic, and training schemes, with its applicability validated across multiple fields such as civil engineering, traffic management, and meteorological forecasting, rather than being limited to a single disciplinary scenario. The innovatively proposed Regional Graph Representation (RGR) method effectively tackles the computational efficiency bottleneck of large-scale graph structures by dynamically aggregating associated nodes and preserving core topological relationships. Complementing this method is the fine-grained reconstruction mechanism, which serves as a key component—it constructs a reverse mapping path from regional global features to node-level local features, enabling accurate recovery of the original features of individual nodes within regions on the basis of graph structure simplification. This not only avoids the common issue of detail loss in traditional simplification methods but also ensures the accuracy of large-scale system modeling. Such a "simplification-recovery" collaborative design has not been reported in existing literature, providing a new technical pathway for efficient and accurate modeling of graph structures. Meanwhile, the designed Fusion Graph Convolution (FGC) module compensates for the insufficiency of information transmission in traditional graph models under large-scale scenarios by balancing information interaction between local neighborhoods and the global system, with its effectiveness fully verified by experimental data. Additionally, to address the multi-periodic characteristics of time series, we pioneered the Time-Aware Sparse Expert module, which achieves precise modeling of the time dimension by dynamically matching different periodic patterns through an adaptive routing mechanism—this design is the first of its kind in this field. Currently, we have comprehensively revised the paper, focusing on enhancing the rigor of research logic and the precision of expression, striving to meet the publication standards of Nature Communications. Should you have any further revision suggestions in the subsequent review process, we will promptly make adjustments in response. We are well aware that academic research requires continuous refinement and will spare no effort to ensure that the quality of our work is highly consistent with the journal's publication standards. Thank you again for your careful guidance!

Comment 1:

The knowledge contribution and novelty are unclear. From the perspective of AI, hybrid training, multi-scale time series analysis and graph representation are well-established concepts. Their integration in this work lacks substantial novelty. From the civil/mechanical engineering standpoint, the train-bridge coupled behaviors and related analyses have also been extensively studied. The manuscript fails to clearly identify existing knowledge gaps or explicitly clarify its unique contributions to either field.

Authors' Reply:

Thank you for your valuable feedback; your perspective is very important to us. After carefully revising the paper, we have more clearly articulated the contributions and novelty of our work. We acknowledge that concepts such as hybrid training, multi-scale time series analysis, and graph representation are well-established techniques in the AI field. However, we propose a novel deep learning framework for predicting long-term time series of dynamic systems, which addresses a complex multi-task prediction problem. This framework spans multiple domains, including civil engineering, traffic, and weather, thus not being limited to the field of civil engineering.

Additionally, this paper introduces a regional graph representation method, which significantly improves the computational efficiency of large-scale graph structures, a method that has not been explored in existing research. Alongside this, we introduce the Fusion Graph Convolution module, which addresses the issue of insufficient information exchange between nodes in large-scale graphs. Experimental results validate the effectiveness of this module.

Moreover, to tackle the multi-periodicity problem of time series data, we propose the Time-aware Sparse Expert module, which autonomously selects different period segments for precise modeling of the time dimension—an innovation that has not been introduced in the field before.

We have diligently revised and optimized our paper, and should you have any further suggestions or areas for improvement during the review process, we will make timely adjustments based on your feedback.

We have identified and highlighted the gap in current research, and our study aims to bridge this gap by addressing the limitations found in existing methods. The revised introduction is as follows:

The technology for modeling dynamic systems \cite{5405072} has matured considerably, with the prevailing approach representing dynamic systems as graph structures or grid structure \cite{Cheng2022, Meller2023, xiang2024adaptive}. In this framework, the system's components are treated as nodes, and spatial topological information is incorporated through convolution to model the spatial characteristics. This approach has yielded results across various scenarios. However, as the size of the graph structure increases \cite{Batarfi2015, ZHAO2021358}, its limitations become more apparent. Specifically, when the graph structure becomes too large \cite{Elias4}, computational and information propagation efficiency decline, particularly in large-scale dynamic systems. Although existing deep learning models have proposed some new convolution methods \cite{2017Inductive, 2015LINE} to address the challenges posed by large-scale graph convolutions, these methods typically focus solely on reducing the number of convolutions and lowering computational complexity. While these improvements alleviate computational pressure to some extent, they still do not fundamentally overcome the bottlenecks in handling long time-series data. In particular, in the long-term prediction tasks of complex dynamic systems, relying on large-scale convolution operations still results in excessively long training times and consumes substantial hardware resources. Therefore, how to effectively balance the trade-off between computational efficiency and prediction accuracy during the expansion of graph structures is crucial.

Currently, two primary temporal modeling techniques are widely utilized in graph neural networks \cite{Peng2023, 10636792, LAI2025107298}. One technique involves one-dimensional spatiotemporal convolution with a gating mechanism \cite{9661093}, while the other utilizes an attention mechanism for temporal information extraction. Both methods effectively address various challenges within the field. However, for dynamic systems characterized by strong periodicity and frequent information exchange, modeling across multiple time scales is particularly important. For instance, in weather systems, temperature fluctuates on both daily and annual cycles. This suggests that analyzing the same temporal data across different time scales can uncover information at different levels.

Comment 2:

The study cases are of limited impact. Despite the broad range of dynamic system problems presented in the introduction, the actual scenarios studied—train-bridge systems under earthquakes and traffic flow systems—are common and well-researched examples. The train-bridge system studied here is relatively simple, failing to adequately demonstrate the method's potential or general applicability.

Authors' Reply:

We sincerely apologize for any confusion caused. We acknowledge that exclusive reliance on the train-bridge coupled system fails to comprehensively validate our Gattention framework. Therefore, in the revised version, we have included more complex dynamic system cases, such as traffic flow and weather datasets, instead of limiting ourselves to the train-bridge coupled system. Specifically, we first tested the generalization capability of Gattention on the train-bridge coupled system to verify whether it could adapt to different train-bridge coupled structures. Then, on the traffic flow dataset, we decoupled the Gattention framework and conducted a series of ablation experiments to demonstrate the effectiveness of its individual modules and validate its computational efficiency. On the weather dataset, we tested Gattention's adaptability to another grid structure, which also presents a significant and complex forecasting challenge. The experimental results show that the framework's average temperature error on the weather data is only 0.8086 degrees, proving its broad adaptability and efficiency.

In addition, regarding the issue you raised about the train-bridge coupled system being too simple, we have expanded the dataset by introducing variations in both the number of trains and bridge spans. This makes the experimental scenarios more complex than the previous train-bridge coupled structures with only variations in bridge spans.

Model	Metric	4 trains			5 trains			6 trains			7 trains		
		5 spans	6 spans	7 spans	5 spans	6 spans	7 spans	5 spans	6 spans	7 spans	5 spans	6 spans	7 spans
GAT [35]	MSE	0.418	0.295	0.242	0.353	0.358	0.344	0.350	0.413	0.312	0.392	0.354	0.333
	MAE	0.468	0.376	0.344	0.427	0.423	0.411	0.415	0.447	0.385	0.441	0.421	0.413
STGNN[36]	MSE	0.335	0.226	0.184	0.273	0.254	0.263	0.261	0.363	0.229	0.297	0.278	0.317
	MAE	0.407	0.348	0.305	0.373	0.348	0.361	0.342	0.400	0.316	0.374	0.371	0.378
GraphWaveNet[37]	MSE	0.138	0.103	0.111	0.124	0.140	0.104	0.133	0.159	0.109	0.147	0.154	0.135
	MAE	0.273	0.235	0.235	0.252	0.267	0.233	0.251	0.282	0.232	0.267	0.268	0.255
GAttention	MSE	**0.088**	**0.080**	**0.072**	**0.098**	**0.0991**	**0.084**	**0.092**	**0.122**	**0.070**	**0.098**	**0.107**	**0.095**
	MAE	**0.221**	**0.204**	**0.189**	**0.227**	**0.224**	**0.209**	**0.205**	**0.242**	**0.186**	**0.217**	**0.223**	**0.211**

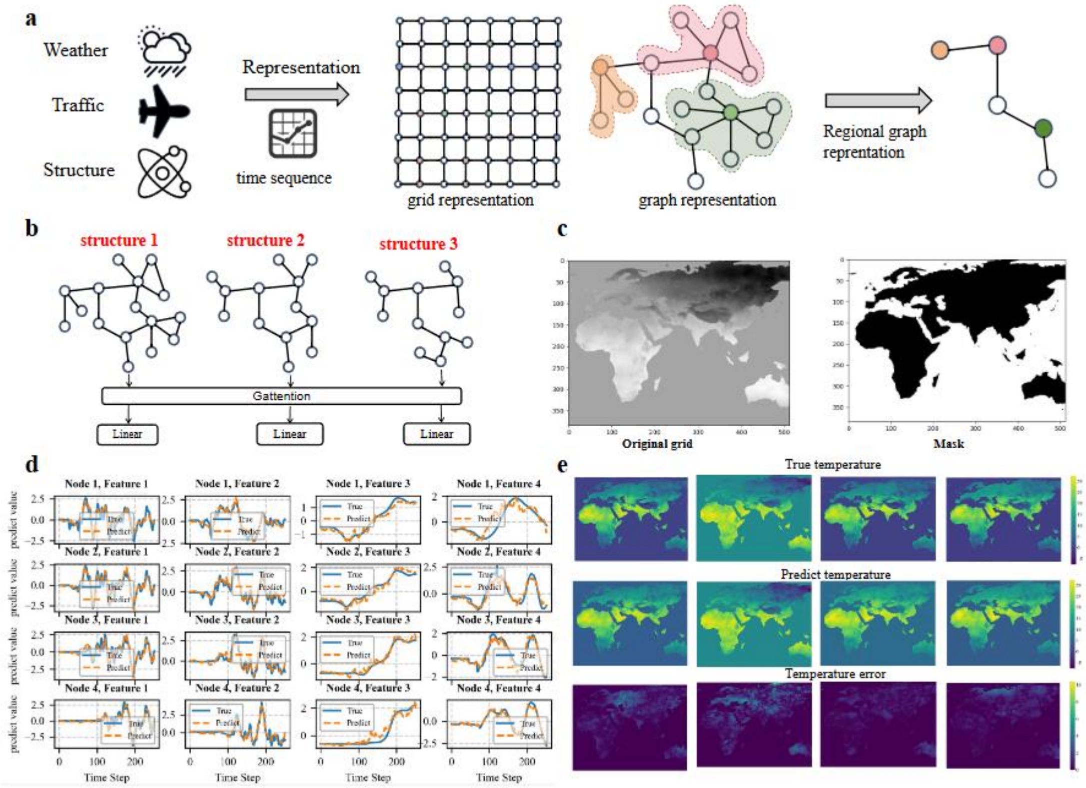
Table 1 Comparison of different models under varying numbers of trains and spans.The predicted length is 24, with the best result highlighted.

In this our paper, a performance evaluation of the Gattention model is conducted using three datasets: the train-bridge coupled dataset, the traffic flow dataset, and the global weather dataset. The objectives of the tests will vary for each dataset to evaluate the model's performance.

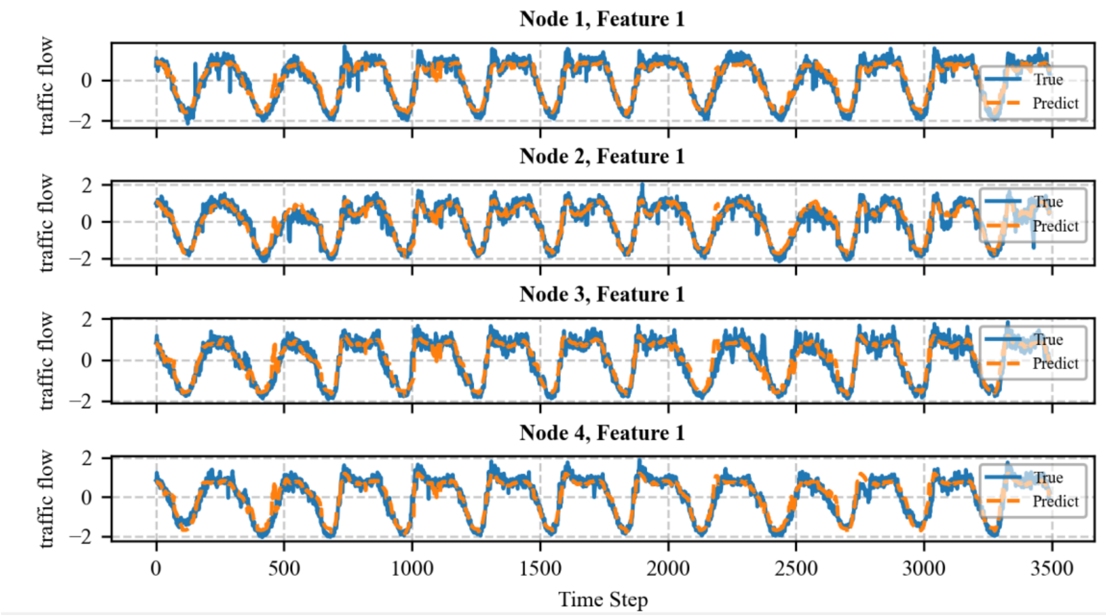
First, on the train-bridge coupled system dataset, the generalization ability of the Gattention model will be assessed. The model's performance stability and adaptability in different bridge-coupling scenarios will be evaluated to determine whether it can learn patterns from historical data and apply them to new similar scenarios.Next, on the traffic flow dataset, ablation experiments will be performed on the individual modules of the Gattention model. By removing or adjusting key components, the contribution of each module to the prediction performance will be analyzed.Finally, on the global weather dataset, the long-term prediction capability of the Gattention model for grid-structured data will be tested. The model's ability to handle and predict weather changes over extended periods will be assessed.

Through testing on these three datasets, a comprehensive evaluation of the Gattention model will be conducted, providing insights for optimizing its practical applications.

The detail prediction result of response for trains and global temperature is as illustrated below:



The prediction result of traffic flow is as illustrated below:



Comment 3:

The improvement of the proposed approach appears largely incremental. According to Tables 3–5, the proposed method improves accuracy by approximately 10% compared to existing methods. Such gains suggest that traditional AI/ML techniques remain competitive. The authors should further justify the practical value of these incremental improvements.

Authors' Reply:

I would be glad to answer your question. The improvement in model accuracy is indeed not significant compared to existing models. The main reason is that current models are already relatively mature in terms of graph neural network temporal dynamics, leaving limited space for further accuracy improvements. It is important to emphasize that the core focus of this paper is not solely on enhancing accuracy, but rather on improving the computational speed and efficiency of the model. As mentioned in the abstract, our primary goal is to strike a balance between accuracy and efficiency. In order to achieve faster computation, we may, in certain cases, accept some loss in accuracy as long as it remains within an acceptable range. Therefore, despite the limited improvement in accuracy, the optimization in efficiency holds significant value in terms of real-time performance and computational resource savings in practical applications.

The detail improvement of speed for different dataset is as follows:

Dataset	Graph Structure	Region Order	Total Node Number	Speed	Memory
PEMS08	Original Graph	–	170	1.646 s	23.15 MB
	RGR	1	64	0.115 s	8.71 MB
		2	33	0.055 s	4.49 MB
		3	22	0.036 s	2.99 MB
PEMS04	Original Graph	–	307	3.316 s	39.79 MB
	RGR	1	129	0.498 s	16.72 MB
		2	78	0.152s s	10.11 MB
		3	64	0.148 s	8.29 MB

Table 2 Comparison of Graph Structures with Different Region Orders

The balance between speed and accuracy is presented as follows:

Dataset	Predict Length	RGR	FGC	Time-aware Expert	Region Order	MSE (Mean±Std)	MAE (Mean±Std)
PEMS08	36	✓	✓	✓	1	**0.1215 ± 0.0113**	**0.2549 ± 0.0124**
					2	0.1485 ± 0.0025	0.2866 ± 0.0052
					3	0.2041 ± 0.0085	0.3306 ± 0.0086
	24				1	**0.1112 ± 0.0124**	**0.2456 ± 0.0092**
					2	0.1233 ± 0.0136	0.2504 ± 0.0076
					3	0.1838 ± 0.0034	0.3079 ± 0.0027
	12				1	**0.0942 ± 0.0115**	**0.2207 ± 0.0115**
					2	0.1287 ± 0.0026	0.2490 ± 0.0011
					3	0.1733 ± 0.0146	0.2962 ± 0.0140
PEMS04	36	✓	✓	✓	1	0.2009 ± 0.0132	0.3230 ± 0.0060
					2	0.2624 ± 0.0092	0.3451 ± 0.0035
					3	**0.1850 ± 0.0032**	**0.3103 ± 0.0032**
	24				1	**0.1178 ± 0.0126**	**0.2529 ± 0.0050**
					2	0.2190 ± 0.0087	0.3132 ± 0.0092
					3	0.1581 ± 0.0080	0.2860 ± 0.0097
	12				1	**0.0839 ± 0.0005**	**0.2123 ± 0.0018**
					2	0.1537 ± 0.0064	0.2551 ± 0.0090
					3	0.1122 ± 0.0031	0.2382 ± 0.0007

Table 6 Ablation Study on Region Order Impact (Best Results Highlighted)

Comment 4:

Validation based solely on simulation data may not convincingly demonstrate practical effectiveness. Authors should strengthen validation with laboratory or field test data and explicitly evaluate performance under various measurement noise levels.

Authors' Reply:

Thank you for your valuable feedback. It is true that while simulation data can provide useful initial insights, they often fail to fully capture the complex scenarios encountered in real-world applications. Therefore, in the revised manuscript, we have incorporated data from various dynamic systems such as traffic and weather, ensuring that our research findings possess higher credibility and practical relevance.

Furthermore, regarding the simulation data of the train-bridge coupled system, we have already conducted comprehensive and thorough validation in our previous work—our team systematically verified the dynamic response characteristics of the train-bridge coupled system through scaled model experiments and on-site tests in a laboratory environment. The relevant experimental designs included vibration tests under different load conditions and track irregularity parameters. These validation results have been elaborated in corresponding studies; specifically, the research in references [41, 42, 43] provides multi-dimensional validation evidence for the train-bridge coupled model, further confirming the reliability and accuracy of the simulation calculations. Thus, we believe that based on such simulation data supported by laboratory empirical evidence and literature, we can more comprehensively evaluate the model's performance in practical applications.

The description of the TBC dataset is as follows:

The TBC dataset, obtained from Matlab computations \cite{ZHAO2023105244, zhao2022seismic}, exhibits negligible discrepancies when compared to structural simulations. Therefore, future research will prioritize data from numerical computations \cite{PANG2018233}, using seismic waveforms based on the Clough-Penzien model \cite{zhao2022seismic} and recorded by high-precision sensors during actual earthquake events.the specific calculation process and validation methods are detailed in reference [31, 32, 33]

Comment 5:

Mean Absolute Error (MAE) alone is insufficient as a metric for performance evaluation. It is strongly recommended to benchmark the proposed method against advanced AI/ML techniques used for general time-series prediction beyond the limited scope of bridge engineering.

Authors' Reply:

Thank you very much for pointing out this issue. Based on your suggestion, we have introduced other more advanced models as comparison models in the revised manuscript. Additionally, we have included the error values of two commonly used evaluation metrics, MSE and MAE.

In the field of time-series prediction, MSE and MAE are widely used evaluation metrics. MSE places more emphasis on the model's ability to predict extreme values, effectively measuring its performance in handling outliers. On the other hand, MAE focuses on the average error across the overall prediction process, which helps assess the model's stability and generalizability. By combining these two metrics, we can comprehensively evaluate the performance of the model.

These revisions are detailed in Section Method-[Training setting] of the revised manuscript, and we welcome further feedback to refine our analysis. Thank you again for helping us enhance the validity and generalizability of our work.

The detailed adjustments to the compared models and loss metrics are as follows:

Model	Metric	4 trains			5 trains			6 trains			7 trains		
		5 spans	6 spans	7 spans	5 spans	6 spans	7 spans	5 spans	6 spans	7 spans	5 spans	6 spans	7 spans
GAT [35]	MSE	0.418	0.295	0.242	0.353	0.358	0.344	0.350	0.413	0.312	0.392	0.354	0.333
	MAE	0.468	0.376	0.344	0.427	0.423	0.411	0.415	0.447	0.385	0.441	0.421	0.413
STGNN[36]	MSE	0.335	0.226	0.184	0.273	0.254	0.263	0.261	0.363	0.229	0.297	0.278	0.317
	MAE	0.407	0.348	0.305	0.373	0.348	0.361	0.342	0.400	0.316	0.374	0.371	0.378
GraphWaveNet[37]	MSE	0.138	0.103	0.111	0.124	0.140	0.104	0.133	0.159	0.109	0.147	0.154	0.135
	MAE	0.273	0.235	0.235	0.252	0.267	0.233	0.251	0.282	0.232	0.267	0.268	0.255
GAttention	MSE	**0.088**	**0.080**	**0.072**	**0.098**	**0.0991**	**0.084**	**0.092**	**0.122**	**0.070**	**0.098**	**0.107**	**0.095**
	MAE	**0.221**	**0.204**	**0.189**	**0.227**	**0.224**	**0.209**	**0.205**	**0.242**	**0.186**	**0.217**	**0.223**	**0.211**

Table 1 Comparison of different models under varying numbers of trains and spans.The predicted length is 24, with the best result highlighted.

Dataset	Predict Length	RGR	FGC	Time-aware expert	Region Order	MSE	MAE
PEMS04	36	✓	✓	✓	1	**0.1996 ± 0.0124**	**0.3230 ± 0.0059**
				-	1	0.4587 ± 0.0378	0.5089 ± 0.0237
	24			✓	1	**0.1257 ± 0.0009**	**0.2619 ± 0.0063**
				-	1	0.2510 ± 0.0397	0.3716 ± 0.0198
	12			✓	1	**0.0839 ± 0.0006**	**0.2123 ± 0.0018**
				-	1	0.0971 ± 0.0035	0.2321 ± 0.0065
PEMS08	36	✓	✓	✓	1	**0.1215 ± 0.011**	**0.2549 ± 0.015**
				-	1	0.5432 ± 0.025	0.5581 ± 0.020
	24			✓	1	**0.1176 ± 0.015**	**0.2544 ± 0.018**
				-	1	0.3084 ± 0.034	0.4183 ± 0.031
	12			✓	1	**0.0942 ± 0.0115**	**0.2207 ± 0.0115**
				-	1	0.1372 ± 0.0005	0.2655 ± 0.007

Table 3 Ablation study on time-aware expert ,the time-aware expert module is replaced with a basic Feed Forward module to assess its effectiveness. (Best results highlighted)

Dataset	Predict Length	RGR	FGC	Time-aware expert	Region Order	MSE	MAE
PEMS04	36	✓	✓	✓	1	**0.1996 ± 0.0124**	**0.3230 ± 0.0059**
		-			1	0.2016 ± 0.0135	0.3239 ± 0.0090
	24	✓			1	0.1257 ± 0.0009	0.2619 ± 0.0063
		-			1	**0.1195 ± 0.0026**	**0.2535 ± 0.0067**
	12	✓			1	**0.0839 ± 0.0006**	0.2123 ± 0.0018
		-			1	0.0877 ± 0.0003	**0.2109 ± 0.0008**
PEMS08	36	✓	✓	✓	1	**0.1215 ± 0.0110**	**0.2549 ± 0.0150**
		-			1	0.1411 ± 0.0082	0.2776 ± 0.0144
	24	✓			1	0.1176 ± 0.0150	0.2544 ± 0.0180
		-			1	**0.1175 ± 0.0023**	**0.2514 ± 0.0015**
	12	✓			1	**0.0942 ± 0.0115**	**0.2207 ± 0.0115**
		-			1	0.1171 ± 0.0067	0.2426 ± 0.0078

Table 4 Ablation study on RGR. The regions are restored to the original number of nodes through fine-grained reconstruction. (Best results highlighted)

Dataset	Predict Length	RGR	FGC	Time-aware expert	Region Order	MSE	MAE
PEMS04	36	✓	-	✓	1	0.2538 ± 0.0217	0.3685 ± 0.0121
		-	-			0.2020 ± 0.0028	0.3286 ± 0.0059
		✓	✓			**0.1996 ± 0.0124**	**0.3230 ± 0.0059**
	24	✓	-			0.1459 ± 0.0031	0.2780 ± 0.0049
		-	-			0.1347 ± 0.0079	0.2707 ± 0.0101
		✓	✓			**0.1257 ± 0.0009**	**0.2619 ± 0.0063**
	12	✓	-			0.1200 ± 0.0042	0.2528 ± 0.0066
		-	-			0.1083 ± 0.0069	0.2389 ± 0.0073
		✓	✓			**0.0839 ± 0.0006**	**0.2123 ± 0.0018**
PEMS08	36	✓	-	✓	1	0.1326 ± 0.0057	0.2687 ± 0.0043
		-	-			0.2140 ± 0.0031	0.3548 ± 0.0108
		✓	✓			**0.1215 ± 0.0110**	**0.2549 ± 0.0150**
	24	✓	-			0.1108 ± 0.0055	0.2426 ± 0.0068
		-	-			0.1756 ± 0.0131	0.3117 ± 0.0094
		✓	✓			**0.1176 ± 0.0150**	**0.2612 ± 0.0092**
	12	✓	-			0.1048 ± 0.0069	0.2309 ± 0.0075
		-	-			0.1870 ± 0.0170	0.3086 ± 0.0195
		✓	✓			**0.0942 ± 0.0115**	**0.2207 ± 0.0115**

Table 5 Ablation study on RGR and FGC. We replace FGC with a standard Graph Convolution layer for comparison purposes, which aims to validate the ability of FGC to model large-scale graph structures effectively. (Best results highlighted)

Comment 6:

It is recommended to include at least one more complex scenario to demonstrate broader applicability. Specifically, address multi-target prediction scenarios (e.g., simultaneous multiple train crossings), generalization to different structural configurations, and explicitly discuss computational efficiency. Although such capabilities are claimed, validation and discussion are currently inadequate.

Authors' Reply:

Thank you very much for your valuable suggestions. Based on your feedback, we include more complex scenarios in the revised manuscript, specifically considering the simultaneous prediction of the responses of all trains in a multi-bridge coupled system. This is a complex multi-target prediction scenario involving multiple dynamically changing prediction targets. Specifically, we need to address the fact that the number of prediction targets will change over time, and we also need to handle the continuous changes in the graph structure, ensuring that the model can adapt flexibly to this complex temporal data structure and multi-target task.

In addition, we supplement the revised manuscript with a detailed analysis of computational efficiency, particularly focusing on the trade-offs between the model's running speed and accuracy under different region orders. We will provide detailed experimental results under various conditions and conduct an in-depth analysis of these results to demonstrate the balance between computational efficiency and predictive accuracy in our method. With these additions, we aim to more comprehensively validate the practical applicability of our approach, especially in complex scenarios and large-scale data.

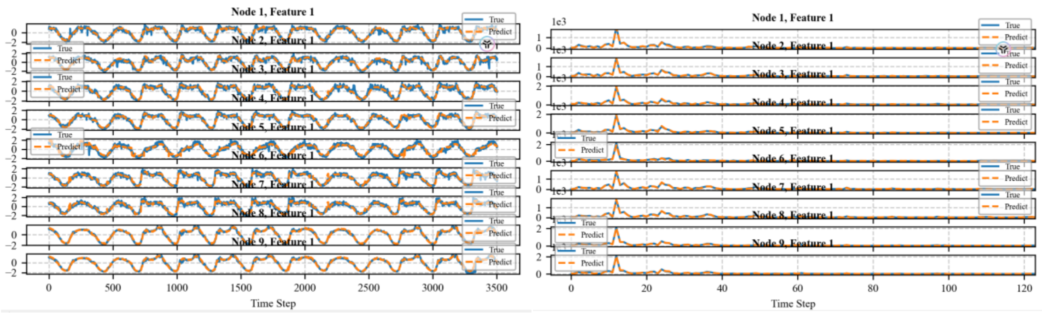
<u>Comment 6:</u>

To fully demonstrate prediction accuracy, comparison of time-series predictions in the frequency domain is highly recommended. Additionally, authors should clearly state the method’ s performance limitations and underlying assumptions.

Authors’ Reply:

Thank you very much for your valuable suggestion. Based on your advice, we have added a frequency-domain comparison of time-series predictions in the code to provide a more comprehensive demonstration of prediction accuracy. Additionally, we have included a more detailed explanation of the limitations and underlying assumptions of the method in the revised manuscript.

The prediction for traffic flow and the frequency domain comparison for both the true values and predictions are as follows:




```
2 usages
def fourier_loss(true, predict):
    """compare in freq domain"""
    true_fft = torch.fft.fft(true, dim=0)
    predict_fft = torch.fft.fft(predict, dim=0)
    loss = torch.mean(torch.abs(true_fft - predict_fft))
    return loss.item()
```

The assumptions and limitations of our method are clarified in detail in the revised version:

Graph convolution operations inherently introduce an inductive bias, assuming that a node's features are primarily influenced by its neighboring nodes. Based on this assumption, we introduce a novel strategy: the neighbors of a node are pre-aggregated, with the node itself serving as the central hub of a region, We found that in many cases, information exchange between tightly connected nodes is frequent and redundant, which results in unnecessary computational overhead. To address this issue, we propose a strategy that pre-combines tightly connected nodes into a single region, effectively performing an average pooling operation at the region.

we introduced an intuitive definition to determine the relationships between different regions by assessing whether there are connections between nodes across regions. Through this approach, we are able to preserve most of the topological information in the graph structure.For grid structures

Although the framework excels in enhancing efficiency, it still has certain limitations. When predicting nodes across multiple regions, simultaneous information recovery for these regions introduces redundant computational overhead, undermining overall model performance. To mitigate this issue, we propose an enhanced solution: instead of considering all nodes within a region, we focus exclusively on those that are relevant to the prediction targets. Specifically, we map the dimensions of regional nodes directly to those of the target nodes, thereby eliminating unnecessary computations for irrelevant nodes. This approach resolves the redundancy problem and improves the computational efficiency and resource utilization of the model, particularly in cross-region prediction tasks involving large-scale graph.

Although the regional graph representation reduces the size of the graph structure, the graph attention mechanism introduced by Fusion Graph Convolution still leads to some time-consuming operations. Furthermore, the running efficiency is quadratic with respect to the number of nodes. If necessary, the graph attention mechanism can be removed to improve computational efficiency.

Comment 7:

Page 8, Line 223: Verify whether "Figure 4" should be correctly cited as "Figure 2".In Figure 3, clarify "GATtension." Similarly, on Line 400, verify if "Gattentionas" should be corrected to "Gattention."Equations (15) and (16) lack explicit definitions of variables "y" and "z".Figure 12 appears in the manuscript without explicit citation or discussion in the main text.Figures 9 – 11 lack detailed explanations for each subplot, making it difficult to interpret the response and behavior of the train-bridge systems under earthquake scenarios. Additional clarification is required.

Authors' Reply:

We truly appreciate your careful identification of the errors in the manuscript. In the revised version, we have addressed all similar issues and made the necessary corrections. Specifically, we have standardized the naming of the framework to "Gattention" throughout the manuscript to ensure consistency. Your detailed feedback has been invaluable in improving the clarity and accuracy of the paper.

The definition for “y” and “z” as follows:

Therefore, even if displacement is used as a feature vector, its impact on the final prediction results is relatively small.so their feature vector consists of displacements and accelerations in the y and z directions.

For TBC, the x-direction is the driving direction of the train, while y and z are the lateral and vertical directions, respectively. For train nodes, in practical situations, displacement may not always be a feasible option. Generally, the state of a train under seismic influence is mainly determined by the seismic effects and the train's acceleration. Therefore, even if displacement is used as a feature vector, its impact on the final prediction results is relatively small.so their feature vector consists of displacements and accelerations in the y and z directions.

$$X_{\text{train}} = \{y_{\text{acc}}, z_{\text{acc}}, y_{\text{dis}}, z_{\text{dis}}\} \in \mathbb{R}^{N_{\text{train}} \times 4}$$

For bridge nodes, real-time responses are not available, so their feature vector

$$X_{\text{bridge}} = \{L_{\text{span}}, E_{\text{material}}, I_z\} \in \mathbb{R}^{N_{\text{bridge}} \times 3}$$

where L_{span} is the length of each span; E_{material} is the elastic modulus of the material, and I_z is the moment of inertia of the section with respect to the z-axis.

For pier nodes, which serve as input nodes for seismic accelerations, the feature vector

$$X_{\text{pier}} = \{x_{\text{acc}}, y_{\text{acc}}, z_{\text{acc}}\} \in \mathbb{R}^{N_{\text{pier}} \times 3}$$

consists of seismic accelerations in the x, y, and z directions.

Thus, the graph structure of the TBC is established. The subscripts "acc" and "dis" represent acceleration and displacement, respectively.

We have removed Figures 9 – 11 and restructured the illustrations. Additionally, we have clarified the specific prediction targets in the revised version.

we used train-bridge coupled system with varying bridge spans and train numbers to predict the train's response under seismic excitation.

The detail explanations for each subplot as follows:

The prediction task in the graph structure focuses on the response of train nodes in the y and z directions within a train-bridge coupled system under seismic action. node 1 to 4 corresponding to the 1st to 4th carriages respectively, feature 1-4 represent the acceleration of the carriage in the y and z direction.

Fig. 2 Overview of long-term prediction results for dynamic systems. **a** Illustration of dynamic systems represented as graph and grid structures. **b** Hybrid training method to enable the model's applicability to different graph structures. **c** Mask training approach designed for adapting the model to various grid structures. **d** The prediction task in the graph structure focuses on the response of train nodes in the y and z directions within a train-bridge coupled system under seismic action. node 1 to 4 corresponding to the 1st to 4th carriages respectively, feature 1-4 represent the acceleration of the carriage in the y and z direction. **e** Global monthly average temperature prediction, corresponding to the prediction task in a grid structure.

Reply to the reviewer's comment (Reviewer#4)

General comments:

Authors' Reply:

We sincerely appreciate your joint contribution to the peer review of this manuscript. In response to your valuable feedback, we have made necessary revisions to the article and provided more detailed code in the revised version. Your input has significantly enhanced the quality of the manuscript, and we are grateful for your support in this collaborative process. Thank you again for your effort and expertise in reviewing this work.

Responses to Reviewers' Comments

Dear Editor and Reviewers,

First of and foremost, we would like to express our sincere thanks and appreciation to the editors and reviewers for your professional handling and review of the manuscript. Your comments and constructive suggestions were very helpful in improving the quality of our manuscript and helped us identify several shortcomings in earlier versions of the manuscript. We have used these valuable suggestions as a guide to address these shortcomings where appropriate. The paper has now been revised following the reviewers' comments, changes in the manuscript have been marked in "blue". In the reply to comments, all the changes are marked in "yellow". Please check the following responses.

Thank you very much!

Best regards,
PingXiang

Reply to the reviewer's comment (Reviewer#2)

Comment 1:

The comments to the authors have been adequately addressed. I do not have any more comments.

Authors' Reply:

Thank you very much for your efforts and your recognition of the work presented in this paper.

Reply to the reviewer's comment (Reviewer#3)

Comment 1:

The manuscript has been revised to a satisfactory level for publication.

Authors' Reply:

We greatly appreciate your support for the work presented in this paper, as well as your efforts in reviewing it.

Reply to the reviewer's comment (Reviewer#4)

Comment 1:

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts..

Authors' Reply:

Thank you very much for your efforts in reviewing this paper.