Larsson *et al. Journal of Artificial Intelligence for Healthcare Systems and Analytics* *(2026) 1:79*
https://doi.org/10.51847/j185378437

ORIGINAL RESEARCH **Open access**

Spatiotemporal Graph Neural Network for Forecasting Daily Hospital Bed Demand at Regional Level Using Influenza-Like Illness Surveillance and Environmental Data

Sven Larsson[1*], Erik Johansson[1], Anna Nilsson[2]

Abstract

Seasonal influenza and respiratory viruses cause recurring surges in hospital bed demand, often resulting in overcrowding under reactive capacity management. ILI surveillance data and environmental factors such as temperature and humidity provide early signals of transmission, making them valuable for forecasting healthcare demand. Traditional forecasting methods rely on simple time series models or historical averages and fail to capture spatial disease spread or environmental influences, leading to inaccurate predictions and poor resource allocation. We propose a spatiotemporal graph neural network (ST-GNN) that integrates ILI surveillance and environmental data for regional daily hospital bed demand forecasting. The model represents regions as graph nodes connected by population flow, enabling joint spatial-temporal modeling of disease dynamics. The framework uses regional graph construction, ILI data from healthcare visits, and environmental variables such as temperature, humidity, and air quality. These inputs are processed by the ST-GNN to predict daily bed demand. The approach captures spatial disease propagation and improves early detection of demand surges, supporting proactive healthcare planning. The ST-GNN provides a scalable, data-driven framework for improving hospital bed demand forecasting and enhancing preparedness during seasonal epidemics.

Keywords Graph neural networks, Spatiotemporal graph neural networks, Hospital bed demand forecasting, Influenza-like illness surveillance, Environmental data, Regional healthcare capacity planning

*Correspondence:
Sven Larsson
sven.larsson@gmail.com
[1] Department of Intelligent Healthcare Systems, Karolinska Institute, Stockholm, Sweden
[2] Department of AI Medical Analytics, Lund University, Lund, Sweden

Introduction

Seasonal influenza and respiratory viruses cause predictable annual surges in hospital bed demand. The COVID-19 pandemic demonstrated the vulnerability of healthcare systems to unexpected surges from novel pathogens. Bed shortages frequently result in diverted ambulances, delayed patient care, and increased excess mortality rates [1, 2]. Hospitals must therefore develop robust forecasting tools to anticipate these demands. Proactive planning can substantially improve patient outcomes and operational efficiency.

Current hospital capacity planning is often reactive, responding only to current occupancy levels rather than anticipating future needs. Many facilities rely on simple time series models based on historical averages that fail to capture dynamic changes. These methods ignore the spatial spread of diseases as they move between regions through travel and commuting patterns [3, 4]. Such limitations lead to inefficient resource utilization during peak periods. A shift toward spatially aware models is critical for modern healthcare management.

ILI surveillance systems operated by organizations such as the CDC and local health departments provide valuable early warning signals of respiratory disease activity. These systems track outpatient visits, emergency department presentations, and laboratory confirmed cases in real time. Environmental data including temperature, humidity, and air quality measurements further influence transmission dynamics and disease severity [5, 6]. Integrating these data sources can enhance the predictive power of forecasting models. This multi-modal approach aligns with recent advances in public health informatics.

This article proposes an ST-GNN framework for daily bed demand forecasting at the regional level by integrating ILI surveillance and environmental data alongside spatial connectivity information. The model offers a comprehensive roadmap for implementation in healthcare systems. It addresses key gaps in current forecasting methodologies by incorporating spatiotemporal dependencies [7]. The framework emphasizes practical applications for capacity planning. Ultimately, it aims to support better decision-making during seasonal and emergent health challenges.

Figure 1 presents the end-to-end conceptual architecture through which regional graph construction, multi-source surveillance integration, spatiotemporal graph learning, multi-horizon forecasting, uncertainty estimation, and operational decision support are linked for daily hospital bed demand forecasting.

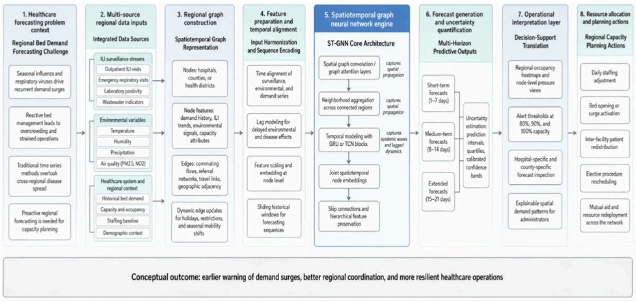

Figure 1. Conceptual architecture of the spatiotemporal graph neural network framework for regional daily hospital bed demand forecasting using influenza-like illness surveillance, environmental data, and spatial connectivity

Background

Hospital bed demand and capacity

Planning Hospital bed demand is primarily driven by infectious diseases such as seasonal influenza along with other factors including elective procedures and public health emergencies. Seasonal patterns create predictable peaks that strain capacity, while unexpected events like pandemics exacerbate the situation. Under-capacity leads to overcrowding, increased wait times, and compromised quality of care [8]. Effective capacity planning requires accurate demand predictions to optimize resource allocation. Without reliable forecasts, healthcare systems face significant operational challenges.

Consequences of inadequate bed capacity planning include ambulance diversions and postponed elective surgeries. These issues contribute to higher mortality rates and reduced patient satisfaction. Recent studies highlight the need for advanced predictive models in hospital operations [2, 9]. Machine learning approaches have shown promise in addressing these forecasting challenges. The integration of diverse data sources can improve overall system resilience.

Influenza-like illness surveillance

Influenza-like illness surveillance relies on networks such as the CDC ILINet, which monitors outpatient visits for influenza-like symptoms. Emergency department visits for respiratory issues and laboratory confirmed cases provide additional granularity to these datasets. Wastewater surveillance has emerged as a complementary tool for early detection of community transmission [10]. These data streams offer real-time insights into disease activity levels. Their timely availability makes them ideal for integration into forecasting models.

Surveillance data enables tracking of ILI trends at regional and local levels. It captures variations in disease burden across different geographic areas. Historical ILI records from 2017 to 2023 demonstrate consistent seasonal patterns that correlate with hospital admissions [1, 11]. Combining multiple surveillance modalities enhances the robustness of predictions. This comprehensive monitoring supports proactive public health responses.

Environmental drivers of respiratory disease

Temperature plays a critical role in respiratory disease transmission, with colder and drier conditions often increasing viral stability and spread. Low absolute humidity has been associated with higher rates of influenza-like illness incidence. Air quality parameters such as PM2.5 and NO2 concentrations can exacerbate respiratory symptoms and influence hospitalization rates [12]. These environmental factors exhibit both direct and lagged effects on disease dynamics. Understanding their impact is essential for accurate modeling.

Environmental data provides contextual information that complements traditional epidemiological surveillance. Variations in weather patterns can predict shifts in ILI activity weeks in advance. Studies have incorporated these variables to improve the precision of infectious disease forecasts [13]. The spatiotemporal nature of environmental influences aligns well with graph-based modeling approaches. This integration allows for more nuanced predictions of healthcare demand.

Spatiotemporal Graph Neural Networks Spatiotemporal graph neural networks such as ST-GCN, DCRNN, and Graph WaveNet combine graph convolutions for spatial dependencies with temporal modeling components like RNNs or TCNs. These architectures effectively capture how information propagates across connected nodes over time. Applications in healthcare have demonstrated their utility for epidemic-related predictions [4, 14]. The dual focus on space and time makes them suitable for regional forecasting tasks. Recent advancements have extended their capabilities to multi-source data integration.

Graph neural networks model complex relational structures where nodes represent entities and edges denote interactions. In spatiotemporal variants, temporal dynamics are incorporated through recurrent or convolutional layers applied sequentially. This enables the modeling of evolving patterns such as disease spread waves [15, 16]. The framework leverages these strengths for hospital demand forecasting. Such models outperform traditional methods in capturing heterogeneous spatial influences.

Framework Overview

High-level architecture

The high-level architecture of the proposed ST-GNN framework begins with a region graph constructed from hospitals and counties as nodes. ILI surveillance time series and environmental data serve as input features to the model. The spatiotemporal graph neural network processes these inputs to generate forecasts of daily hospital bed demand for the next 7 to 14 days [7]. This integrated pipeline ensures that both spatial connectivity and temporal evolution are considered. The output supports operational planning in healthcare facilities.

Data flows through the system starting from multi-source collection to graph representation and finally to prediction layers. The architecture maintains modularity to allow for easy incorporation of additional data streams. Core components include spatial convolution layers followed by temporal modeling blocks [4]. This design facilitates scalable application across different regional contexts. Overall, the framework provides a cohesive solution for demand forecasting.

Core assumptions

The framework assumes the availability of historical ILI and bed demand data from previous seasons within the 2017-2023 period. Regions are defined with clear hospital catchment areas to facilitate accurate graph construction. Population flow patterns are considered relatively stable under normal conditions [17]. These assumptions enable the model to learn meaningful spatiotemporal relationships. Validation against real-world scenarios would confirm their applicability.

Data quality and completeness are presupposed for effective training and inference. The model relies on consistent reporting from ILI surveillance systems and environmental monitoring stations. Underlying disease dynamics are assumed to follow patterns observed in historical records [18]. External disruptions such as policy changes are accounted for through dynamic adjustments. These foundational assumptions ground the conceptual design in practical realities.

Design principles

Spatial propagation awareness is a key design principle, allowing the model to capture how ILI spreads from one area to adjacent regions via the graph structure. Multi-horizon forecasting ensures predictions are available at various time scales for different decision needs. Uncertainty quantification is incorporated to provide confidence intervals around bed demand estimates [14]. These principles enhance the framework's utility for real-world applications. The overall design prioritizes interpretability and actionability.

The principles emphasize integration of heterogeneous data sources for comprehensive modeling. Graph-based representations promote scalability and flexibility in handling regional variations. Temporal components are designed to model epidemic waves and seasonal trends effectively [15]. This principled approach distinguishes the ST-GNN from conventional forecasting tools. It aligns with best practices in AI for healthcare systems.

Graph Construction

Node definition

Each node in the regional graph represents either an individual hospital or a county or health district. Node features include static attributes such as bed capacity, baseline occupancy rates, and staffing levels. Historical demand patterns are also encoded as time-varying features for each node [9]. This rich representation allows the model to account for local variations in healthcare infrastructure. The node definition forms the foundation of the spatiotemporal modeling process.

Nodes are selected based on geographic and administrative boundaries to ensure comprehensive regional coverage. Additional features may incorporate demographic information relevant to disease susceptibility. The inclusion of these attributes enables personalized forecasting at the facility level [8]. Careful node definition ensures that the graph accurately reflects real-world healthcare networks. This step is critical for subsequent convolution operations.

Edge definition

Edges between nodes are defined based on population flow metrics including commuting patterns, patient referral networks, and general travel behaviors. Geographic adjacency serves as a baseline for connectivity, with edge weights reflecting the strength of interactions. These connections model the potential pathways for disease transmission across the region [4]. Dynamic weighting allows for more accurate representation of spatial dependencies. The edge structure is essential for propagating information in the GNN layers.

Edge definitions can incorporate multiple types of relationships to capture complex interactions. Data sources such as mobility traces or hospital transfer records inform the weighting scheme. This approach ensures that the graph reflects realistic connectivity [16]. Variable influence between nodes is handled through attention mechanisms in later stages. Proper edge construction enhances the model’s ability to simulate spatial spread.

Dynamic graph updates

The graph is designed to support dynamic updates where edges can change over time due to factors such as travel restrictions or seasonal holidays. Time-varying graph structures allow the model to adapt to evolving connectivity patterns. This capability is important for capturing transient effects on disease propagation [14]. Updates are performed periodically based on available data. The dynamic nature improves forecasting accuracy during anomalous periods.

Mechanisms for graph evolution include addition or removal of edges as conditions change. Historical patterns inform the prediction of future graph states where data is limited. Integration with temporal modeling ensures seamless handling of these dynamics [15]. This feature distinguishes the framework from static graph approaches. Overall, dynamic updates contribute to the robustness of the ST-GNN.

Spatiotemporal GNN Architecture

Spatial graph convolution

Spatial graph convolution layers aggregate information from neighboring nodes to model the spread of disease from high-ILI areas to connected regions. Graph attention mechanisms assign variable influence weights to different neighbors based on their relevance. This allows the framework to focus on the most impactful spatial relationships [7, 16]. The convolution operation updates node embeddings by incorporating contextual information from the graph. It effectively captures heterogeneous spatial dependencies in the regional network.

Multiple layers of graph convolution enable hierarchical feature extraction across different scales of connectivity. The spatial component processes the graph structure at each time step independently before temporal integration. Attention-based variants improve interpretability by highlighting key transmission pathways [4]. This design principle aligns with the needs of epidemic modeling in healthcare contexts. Spatial convolutions form the backbone of the ST-GNN's predictive power.

Temporal modeling

Temporal modeling is achieved through gated recurrent units or temporal convolutional networks applied to the sequence of node embeddings over time. These layers capture the progression of epidemic waves and seasonal trends in bed demand. The temporal component processes historical ILI and environmental data to predict future states [15]. Recurrent architectures maintain memory of past events for improved long-range dependencies. This integration ensures that forecasts reflect evolving disease dynamics.

Temporal convolutions offer an efficient alternative to recurrent units for handling long sequences without vanishing gradients. The model learns patterns such as lagged effects from environmental factors on hospitalization rates. By combining with spatial convolutions, it produces spatiotemporal representations [14]. This dual modeling approach outperforms purely temporal or spatial methods. Temporal layers are crucial for multi-step forecasting horizons.

Multi-horizon output

The decoder component of the ST-GNN is designed to predict bed demand at multiple horizons including 7, 14, and 21 days ahead. Direct multi-step forecasting is employed rather than iterative single-step predictions to avoid error accumulation. This strategy provides comprehensive outputs for both short and medium-term planning [9]. Output layers are tailored to each forecasting horizon independently. The multi-horizon capability supports diverse operational needs in hospital management.

Specialized output heads generate predictions for different future time windows simultaneously. The architecture includes skip connections to preserve information across horizons. This design facilitates accurate long-range forecasts by leveraging the full spatiotemporal context [16]. Multi-horizon outputs enable strategic decision-making at various scales. The approach enhances the practical value of the forecasting framework.

Uncertainty quantification

Uncertainty quantification is incorporated through probabilistic outputs that provide both mean predictions and variance estimates for bed demand. Quantile regression techniques allow the model to output prediction intervals at specified confidence levels. This feature helps decision-makers assess risk in capacity planning [14]. Conformal prediction methods can further calibrate the uncertainty estimates post-training. Reliable uncertainty information is vital for robust operational use.

The framework outputs distributional forecasts rather than point estimates to better reflect inherent variability in disease dynamics. Ensemble techniques or Bayesian approximations support the generation of uncertainty metrics. Integration of these elements ensures that forecasts are accompanied by measures of confidence [15]. This probabilistic perspective improves trust in the model's recommendations. Uncertainty quantification completes the comprehensive architecture of the ST-GNN.

Multi-Source Data Integration

ILI surveillance data

Influenza-like illness surveillance data forms a foundational input for the ST-GNN framework through weekly proportions of outpatient visits exhibiting ILI symptoms. Real-time emergency department records for respiratory complaints and laboratory positivity rates add granularity to capture emerging trends at the regional level. These streams enable the model to detect early signals of demand surges before they manifest in hospital admissions [10, 11]. Integration occurs by aligning surveillance timestamps with graph node features for seamless processing. Such multi-modal surveillance enhances the framework's responsiveness to dynamic disease activity.

Additional layers of ILI data including wastewater monitoring provide complementary community-level insights that complement traditional reporting. The framework aggregates these sources to create enriched temporal sequences for each node in the regional graph. Historical patterns from 2017 to 2023 inform the learning of seasonal and anomalous behaviors [1, 6]. This comprehensive approach ensures that the ST-GNN captures subtle shifts in transmission intensity. Overall, ILI surveillance integration strengthens the predictive foundation of daily bed demand forecasts.

Environmental data

Environmental data integration incorporates daily measurements of minimum, maximum, and average temperature alongside relative humidity and precipitation levels. Air quality index values including PM2.5 and NO2 concentrations are fused as node-level features to account for their influence on respiratory disease severity. Lagged effects spanning one to two weeks are explicitly modeled to reflect delayed impacts on hospitalization rates [12, 13]. The ST-GNN processes these variables through dedicated embedding layers before spatiotemporal propagation. This step enriches the graph representations with contextual drivers of ILI transmission.

Precipitation and broader meteorological indices further refine the model's understanding of environmental modulators across the region. Data alignment with ILI streams allows the framework to learn joint spatiotemporal patterns observed in prior seasons [19, 20]. Dynamic scaling of environmental inputs accommodates variations in regional climate profiles. Such integration aligns with established approaches to respiratory disease modeling and supports robust multi-source forecasting [21]. The result is a more holistic representation of factors driving hospital bed demand.

Table 1 clarifies how each major data domain is translated into graph-level representations and forecast functions within the proposed regional ST-GNN architecture.

Table 1. Conceptual mapping of data domains, graph representations, and forecast functions in the proposed ST-GNN framework

Framework domain	Data elements	Representation within the ST-GNN	Prim modelii
Regional healthcare infrastructure	Hospital identity, county or district membership, bed capacity, baseline occupancy, staffing level, catchment characteristics	Static node attributes and slowly varying node-level covariates	Anchor health contex struc heterog acros gra
Historical utilization dynamics	Daily bed demand history, prior seasonal peaks, recent occupancy trajectories	Time-varying node feature sequences	Provide tempora fo autoreg dem struc
ILI outpatient surveillance	Proportion of visits meeting ILI criteria, clinic-level or county-level trend intensity	Temporal node features aligned to daily or interpolated forecasting windows	Capture comm respir disease befo hospital pea
Emergency respiratory presentations	Respiratory-related emergency visits and acute care presentations	High-frequency node-level surveillance features	Dete accelera clinic meani disease
Laboratory positivity and confirmatory surveillance	Virologic confirmation rates, subtype prevalence, confirmed respiratory pathogen activity	Temporal node features or auxiliary surveillance channels	Impro epidemi specifi disease sigr
Wastewater and community early warning indicators	Wastewater viral concentration and analogous community transmission proxies	Regional exogenous temporal feature streams mapped to nodes or subregions	Extends detec earlier transm cyc
Environmental exposures	Temperature, humidity, precipitation, PM2.5, NO2, air quality indices	Node-level exogenous features with lagged embeddings	Moc transmi modulat seve modul environ condi

Population mobility and referral connectivity	Commuting flows, patient transfers, travel behavior, geographic adjacency	Weighted graph edges and dynamic edge updates	Encodes regio propag pathwa intera stren
Spatial graph learning layer	Neighbor aggregation, graph convolution, graph attention weighting	Learned spatial embeddings	Identifies conne regions stron influen each
Temporal learning layer	GRU, TCN, or equivalent sequence model over node embeddings	Learned temporal embeddings	Captu epide progre season and structur tim
Forecast decoder	Horizon-specific prediction heads for 1–7, 8–14, and 15–21 day outputs	Multi-horizon output layer	Produ operati differen dem estim
Uncertainty layer	Quantile estimates, variance outputs, conformal calibration, prediction intervals	Probabilistic forecast wrapper	Quan confiden ris surrou predicte dem

Forecasting Horizons

Short-term (1-7 days)

Short-term forecasting horizons of one to seven days prioritize operational decisions such as daily staffing adjustments and immediate bed opening or closing protocols. The ST-GNN leverages recent ILI and environmental inputs to generate precise daily bed demand estimates that inform real-time diversion planning across the regional network. These forecasts enable hospitals to respond swiftly to localized surges without overcommitting resources [2, 22]. The architecture's temporal layers ensure that short-horizon outputs remain sensitive to rapid changes in disease activity. This capability directly supports frontline operational efficiency in healthcare facilities.

Integration of high-frequency surveillance data allows the model to update short-term predictions continuously as new information arrives. Edge definitions based on patient flow further refine spatial awareness for near-term resource needs [23]. Decision thresholds can trigger automated alerts for bed reallocation within the graph structure. The framework thus bridges immediate data inputs with actionable short-term outputs. Hospitals benefit from reduced uncertainty in day-to-day capacity management.

Medium-term (8-14 days)

Medium-term forecasting horizons extending from eight to fourteen days support strategic decisions including elective procedure rescheduling and activation of temporary surge capacity. The ST-GNN's multi-horizon decoder produces these extended predictions by propagating spatiotemporal patterns learned from historical ILI and environmental sequences. Regional connectivity encoded in the graph enables anticipation of propagating demand across hospital catchments [15, 17]. Such foresight allows administrators to coordinate inter-facility support well in advance of peak loads. The design principles of uncertainty quantification further inform risk-adjusted planning at this scale.

By incorporating lagged environmental effects, the model captures slower-evolving influences on medium-term bed requirements. Multi-source integration ensures that forecasts remain aligned with evolving epidemic trajectories [8]. Hospitals can therefore initiate preparatory measures such as supply stockpiling or staff training programs. This horizon balances the need for proactive action with sufficient lead time for implementation. The framework ultimately enhances strategic resilience in regional healthcare systems.

Operational Decision Support

Dashboard visualization

Dashboard visualization translates ST-GNN outputs into intuitive regional heatmaps that overlay forecasted bed demand against current capacity metrics. Color-coded alert thresholds at 80 percent, 90 percent, and 100 percent occupancy highlight impending pressure points across the graph nodes. Interactive elements allow users to drill down into individual hospital or county projections derived from ILI and environmental inputs [7, 24]. The interface incorporates uncertainty bands to communicate prediction reliability visually. This design facilitates rapid comprehension by hospital administrators and public health officials.

Real-time synchronization with incoming surveillance data keeps the dashboard current and actionable. Spatial graph representations enable zoomable views of disease propagation pathways. Threshold-based notifications can be configured to trigger automated reports or escalation protocols [4]. The visualization layer thus serves as the primary interface between the model and end users. Effective dashboard implementation ensures that complex forecasts translate into practical operational insights.

Resource allocation

Recommendations Resource allocation recommendations generated by the ST-GNN include predictions for staff redeployment across connected hospitals based on forecasted demand imbalances. Mutual aid requests between facilities are suggested when graph edges indicate strong patient flow linkages and localized capacity shortfalls. Elective procedure postponement triggers are derived from medium-term horizons to preserve surge beds for infectious cases [16]. These recommendations are accompanied by confidence intervals to guide risk-informed choices. The framework prioritizes equitable distribution of resources throughout the region.

Integration of dynamic graph updates allows recommendations to adapt to changing conditions such as holiday travel or policy shifts. Environmental data influences the prioritization of certain allocation strategies during high-transmission periods [14]. Hospital networks benefit from coordinated responses that minimize overall system strain. The recommendation engine embeds domain knowledge to align outputs with operational constraints. This support layer transforms raw forecasts into concrete, implementable actions for healthcare leaders.

Table 2 translates the model's multi-horizon outputs into a structured decision framework linking forecast timing, alert intensity, and corresponding regional capacity actions.

Table 2. Decision-oriented interpretation of forecast horizons, alert conditions, and regional response actions for ST-GNN-enabled hospital bed demand planning

Forecast horizon and decision layer	Typical trigger signal	Primary operational question	Recommen actions
Immediate situational awareness (1–3 days)	Sharp recent rise in ILI indicators, ED respiratory activity, or node-specific occupancy pressure	Is a localized surge imminent within the next few days?	Activate da bed manageme review, adju nurse staffi prepare overflow be intensify discharge coordinatio
Short-term response planning (4–7 days)	Persistent growth in forecasted demand with narrowing uncertainty intervals	Which hospitals are likely to exceed safe capacity this week?	Rebalance staffing, pre position supplies, rev rota assignment open contingenc beds, initiat intra-netwo referral planning
Tactical regional coordination (8–14 days)	Sustained upward multi-node forecasts, regional clustering of high-demand predictions, widening transmission footprint	Should the region initiate coordinated surge preparation?	Reschedul elective procedure prepare mut aid agreement plan transfe pathways review surg staffing poo escalate command

			center monitoring
Extended preparedness (15–21 days)	Medium-confidence forecasts indicating broad regional pressure or recurrent seasonal escalation	Are longer-lead preparations justified despite forecast uncertainty?	Review stockpiles refresh respiratory surge protoc plan tempora unit conversi schedule cro facility staf training
High-risk alert state	Forecast occupancy exceeds 90% with upper uncertainty band approaching or exceeding full capacity	Which areas require immediate escalation and contingency activation?	Trigger executive alerting, def non-urgen admissions where appropriate deploy surg beds, activa mutual aid prioritize hig risk nodes
Moderate-risk alert state	Forecast occupancy between 80% and 90% or rapidly rising demand slope	Which hospitals should enter heightened readiness?	Intensify monitoring preserve flexible capacity, prepare staff redeployme review electi scheduling thresholds
Uncertainty-sensitive planning state	Wide forecast intervals despite rising expected demand	How should administrators act when risk is plausible but not fully certain?	Use precautiona staged actio combine forecasts wi expert revie emphasize reversible intervention increase da refresh frequency
Cross-regional spillover management	Demand forecast rises first in one subregion but edge structure indicates probable transmission to neighbors	Where will the next wave of bed pressure emerge?	Pre-allocat transfers, reinforce adjacent hospitals, anticipate secondary hotspots, ali ambulance a referral routi
Post-forecast review and learning	Forecasted surge resolved or failed to materialize	How should the system refine future actions?	Compare forecast intervals wi realized demand, au decisions taken, recalibrate thresholds refine grap update rule

Evaluation Strategy

Forecasting metrics

Forecasting metrics for the ST-GNN framework emphasize mean absolute error, root mean square error, and mean absolute percentage error applied to daily bed demand predictions. Probabilistic outputs are assessed through calibration scores that evaluate the reliability of uncertainty intervals across multiple horizons. Baselines including ARIMA, LSTM, and non-spatial graph neural networks provide comparative benchmarks to highlight the added value of spatiotemporal modeling [3, 5, 18]. These metrics ensure that the conceptual framework is designed for rigorous quantitative validation. The focus remains on practical utility rather than isolated performance claims.

Additional evaluation includes coverage probability for prediction intervals to confirm that uncertainty quantification meets operational needs. Multi-horizon specific metrics allow differentiated assessment of short-term versus medium-term accuracy [25, 26]. The strategy incorporates sensitivity analyses on environmental and ILI input variations to test robustness. Such comprehensive metrics align with standards in healthcare forecasting literature.

The evaluation design supports iterative refinement of the ST-GNN architecture.

Validation protocols

Validation protocols employ temporal cross-validation by training on three to five years of historical data from 2017 to 2023 and testing on subsequent seasons to simulate real-world deployment. This approach preserves the chronological integrity of epidemic patterns and environmental influences. Geographic transfer learning is incorporated by training on one region's graph and evaluating generalization to another with similar connectivity structures [27, 28]. The protocols account for potential distribution shifts across seasons or locales. Such rigorous validation ensures the framework's conceptual soundness prior to implementation.

Cross-regional experiments further assess the model's ability to handle varying population densities and hospital networks. Hold-out periods aligned with peak ILI seasons provide targeted stress testing of surge prediction capabilities [29]. The strategy includes ablation studies on individual data sources to quantify their contribution within the ST-GNN pipeline. These protocols collectively establish a pathway for safe and effective translation into healthcare operations. The validation framework is designed to build confidence in the overall conceptual approach.

Conclusion

The ST-GNN framework offers a comprehensive conceptual approach for regional daily hospital bed demand forecasting by integrating influenza-like illness surveillance data with environmental variables and spatial connectivity. It constructs a dynamic graph where hospitals and counties serve as nodes to capture disease propagation patterns across time. The architecture processes multi-source inputs through spatial convolutions and temporal modeling layers to produce multi-horizon predictions. This design addresses longstanding gaps in traditional capacity planning methodologies. The result is a scalable model tailored to the complexities of healthcare operations.

Key advantages include explicit modeling of spatial disease spread, seamless fusion of heterogeneous data streams, and built-in support for uncertainty quantification. Multi-horizon outputs align directly with both operational and strategic decision needs in hospital networks. The framework's principled design promotes interpretability and adaptability across diverse regional contexts. These strengths position the ST-GNN as a forward-looking tool for epidemic-informed resource management. Its conceptual foundation draws on established advances in graph neural networks for healthcare applications.

Limitations of the framework stem from its reliance on comprehensive regional data sharing agreements and assumptions regarding relatively stable population flow patterns. Novel pandemic scenarios may require additional adaptation layers to handle unprecedented transmission dynamics. Environmental data quality and reporting consistency across jurisdictions could also influence overall model performance in certain settings. Addressing these challenges through collaborative infrastructure development remains an important consideration for future extensions. The conceptual nature of the framework acknowledges these dependencies while outlining mitigation strategies.

Implementation of the ST-GNN on public datasets such as CDC ILINet, HHS Protect, and integrated weather archives is encouraged to accelerate adoption in real-world healthcare systems. Open-source reference implementations would facilitate broader validation and customization by regional authorities. Continued refinement through community contributions can enhance the framework's applicability to emerging respiratory threats. This pathway supports the broader goal of data-driven regional capacity planning. Ultimately, the proposed model advances the integration of artificial intelligence into proactive healthcare resource management.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 13 Mar 2023 Revised: 08 Jun 2023 Accepted: 10 Aug 2023

Rights and permissions

References

Novaes de Amorim A, Deardon R, Saini V. A stacked ensemble method for forecasting influenza-like illness visit volumes at emergency departments. PLoS One. 2021;16(3):e0241725.

Tello M, Reich ES, Puckey J, Maff R, Garcia-Arce A, et al. Machine learning based forecast for the prediction of inpatient bed demand. BMC Med Inform Decis Mak. 2022;22(1):55.

Zhang J, Nawata K. Multi-step prediction for influenza outbreak by an adjusted long short-term memory. Epidemiol Infect. 2018;146(7):809-16. https://doi.org/10.1017/S0950268818000449.

Croft VM, van Iersel SC, Della Santina C. Forecasting infections with spatio-temporal graph neural networks: a case study of the Dutch SARS-CoV-2 spread. Front Phys. 2023;11:1277052. https://doi.org/10.3389/fphy.2023.1277052.

Kutafina E, Bechtold I, Kabino K, Jonas SM. Recursive neural networks in hospital bed occupancy forecasting. BMC Med Inform Decis Mak. 2019;19(1):39. https://doi.org/10.1186/s12911-019-0766-5.

Morbey RA, Charlett A, Lake I, Mapstone J, Pebody R, et al. Can syndromic surveillance help forecast winter hospital bed pressures in England? PLoS One. 2020;15(2):e0228804. https://doi.org/10.1371/journal.pone.0228804.

Wang Z, Guo R, Hong L, Wang C, Chen L. Demand-responsive windows scheduling in tertiary hospital leveraging spatiotemporal neural networks. In: International Conference on Green, Pervasive, and Cloud Computing. Cham: Springer; 2020. p. 231-43. https://doi.org/10.1007/978-3-030-56036-7_19.

Deschepper M, Eeckloo K, Malfait S, Benoit D, Callens S, et al. Prediction of hospital bed capacity during the COVID-19 pandemic. BMC Health Serv Res. 2021;21(1):468. https://doi.org/10.1186/s12913-021-06447-3.

Mahmoudian Y, Nemati A, Safaei AS. A forecasting approach for hospital bed capacity planning using machine learning and deep learning with application to public hospitals. Healthc Anal. 2023;4:100245. https://doi.org/10.1016/j.health.2023.100245.

Spreco A, Eriksson O, Dahlström Ö, Cowling BJ, Timpka T. Integrated detection and prediction of influenza activity for real-time surveillance: algorithm design. J Med Internet Res. 2017;19(6):e211. https://doi.org/10.2196/jmir.7101.

Ben-Nun M, Riley P, Turtle J, Bacon DP, Riley S. Forecasting national and regional influenza-like illness for the USA. PLoS Comput Biol. 2019;15(5):e1007013. https://doi.org/10.1371/journal.pcbi.1007013.

Mansour M, Donmez TB, Kutlu MÇ, Freeman C. Respiratory diseases prediction from a novel chaotic system. Chaos Theory Appl. 2023;5(1):20-6. https://doi.org/10.51537/chaos.1248645.

Taylor G. Animal models of respiratory syncytial virus infection. Vaccine. 2017;35(3):469-80. https://doi.org/10.1016/j.vaccine.2016.11.054.

Cini A, Marisca I, Bianchi FM, Alippi C. Scalable spatiotemporal graph neural networks. In: Proc AAAI Conf Artif Intell. 2023;37(6):7218-26. https://doi.org/10.1609/aaai.v37i6.25877.

Qiu M, Tan Z, Bao BK. MSGNN: Multi-scale spatio-temporal graph neural network for epidemic forecasting. arXiv. 2023;arXiv:2308.15840.

Tang S, Tariq A, Dunnmon JA, Sharma U, Elugunti P, et al. Predicting 30-day all-cause hospital readmission using multimodal spatiotemporal graph neural networks. IEEE J Biomed Health Inform. 2023;27(4):2071-82. https://doi.org/10.1109/JBHI.2023.3241236.

Latruwe T, Van der Wee M, Vanleenhove P, Devriese J, Verbrugge S, et al. A long-term forecasting and simulation model for strategic planning of hospital bed capacity. Oper Res Health Care. 2023;36:100375. https://doi.org/10.1016/j.orhc.2022.100375.

Tsan YT, Chen DY, Liu PY, Kristiani E, Nguyen KL, et al. The prediction of influenza-like illness and respiratory disease using LSTM and ARIMA. Int J Environ Res Public Health. 2022;19(3):1858. https://doi.org/10.3390/ijerph19031858.

Aveyard P, Gao M, Lindson N, Hartmann-Boyce J, Watkinson P, et al. Association between pre-existing respiratory disease and its treatment, and severe COVID-19: a population cohort study. Lancet Respir Med. 2021;9(8):909-23. https://doi.org/10.1016/S2213-2600(21)00095-3.

Lutz CS, Huynh MP, Schroeder M, Anyatonwu S, Dahlgren FS, et al. Applying infectious disease forecasting to public health: a path forward using influenza forecasting examples. BMC Public Health. 2019;19(1):1659. https://doi.org/10.1186/s12889-019-7966-8.

Lau YC, Shan S, Wang D, Chen D, Du Z, et al. Forecasting of influenza activity and associated hospital admission burden and estimating the impact of COVID-19 pandemic on 2019/20 winter season in Hong Kong. PLoS Comput Biol. 2024;20(7):e1012311. https://doi.org/10.1371/journal.pcbi.1012311.

Chou SC, Chang YS, Chen PC, Schuur JD, Weiner SG. Hospital occupancy and its effect on emergency department evaluation. Ann Emerg Med. 2022;79(2):172-81. https://doi.org/10.1016/j.annemergmed.2021.08.429.

Pieters AJ, Schlobach S. Combining process mining and time series forecasting to predict hospital bed occupancy. In: Int Conf Health Inf Sci. Cham: Springer; 2022. p. 76-87. https://doi.org/10.1007/978-3-031-16266-9_7.

Venna SR, Tavanaei A, Gottumukkala RN, Raghavan VV, Maida AS, et al. A novel data-driven model for real-time influenza forecasting. IEEE Access. 2018;7:7691-701. https://doi.org/10.1109/ACCESS.2018.2888592.

Viboud C, Vespignani A. The future of influenza forecasts. Proc Natl Acad Sci U S A. 2019;116(8):2802-4. https://doi.org/10.1073/pnas.1822167116.

Dal Negro RW, Zanasi A, Turco P, Povero M. Influenza and influenza-like syndromes: the subjects’ beliefs, the attitude to prevention and treatment, and the impact in Italian general population. Multidiscip Respir Med. 2018;13(1):7. https://doi.org/10.1186/s40248-018-0117-9.

Nasrullah A, Gangu K, Garg I, Javed A, Shuja H, et al. Trends in hospitalization and mortality for influenza and other respiratory viruses during the COVID-19 pandemic in the United States. Vaccines (Basel). 2023;11(2):412. https://doi.org/10.3390/vaccines11020412.

Hy TS, Nguyen VB, Tran-Thanh L, Kondor R. Temporal multiresolution graph neural networks for epidemic prediction. In: Workshop on Healthcare AI and COVID-19. PMLR; 2022. p. 21-32.

Schaum A, Bernal-Jaquez R, Sánchez-González G. Model-based monitoring of dengue spreading. IEEE Access. 2022;10:126892-8. https://doi.org/10.1109/ACCESS.2022.3225379.