International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(16s) 2026 pp. 1248-1260
https://doi.org/10.70917/ijcisim-2026-4652
Article

An Integrated Multimodal Deep Learning and Reinforcement Framework for Intelligent Travel Demand Forecasting

Santosh Kumar Sharma[1], Satish Chander[2], Piyush Gupta[3]

[1,2,3]Department of Computer Science and Engineering Birla Institute Of Technology , Mesra Ranchi off Campus Jaipur

Abstract: Precise short-term travel demand forecasting and personalized trip advice are two of the most challenging problems in intelligent transportation systems in part due to the fact that traveler behavior is influenced concurrently by spatial configuration, time dynamics, socio-demographic environment, and cost sensitivity. The majority of the current models cover these dimensions separately. The current paper suggests an end-to-end, monolithic framework that pairs the learning of multimodal representations with a spatio-temporal graph neural network (ST-GNN), a Transformer-based forecasting module, a causal inference layer, and a Q-learning recommender. Heterogeneous travel history, such as demographics, trip purpose, mode of transport, cost of accommodation, and temporal attitudes are co-embedded and propagated through graph convolutions and multi-head self-attention to generate multi-horizon destination-level demand prediction. An actionable do-calculus layer measures the impact of traveler characteristics on decisions and the policy implications guide a reinforcement learning agent that modulates transport and accommodation suggestions to traveler groups. The ST-GNN + Transformer forecaster achieves a 25% lower RMSE than an ARIMA baseline on a publicly available dataset of approximately 8,700 traveler records and, in comparison to the LSTM and DCRNN alternatives, converges within just 300 training episodes. The framework provides a replicable base to adaptable, understandable model of travel behavior in smart-city and tourism analytics contexts.
Keywords: travel demand forecasting; spatio-temporal graph neural networks; Transformer; causal inference; reinforcement learning; intelligent transportation systems; smart tourism.

1. INTRODUCTION

The growth of urban mobility and international travel has been such that it is constantly producing large volumes of multimodal, transit-related data. As part of the Intelligent Transportation Systems (ITS), two questions have now taken center stage, namely how will the demands of specific destinations change during the next few days, and what travel options will prove best suited to an individual traveler, based on who he is and what he can afford? Classical solutions are based on econometric formulations or machine learning pipelines that are not dynamically trained [6]. Though these methods are useful due to their transparency, they are unable to cope with nonlinearity, dynamism, and rich context that are the hallmarks of contemporary travel datasets.

Deep learning has altered the possibilities. Graph neural networks are able to learn the topology of destination networks [9], [10], Transformers can learn long-range temporal dependencies that are not recurrent [8], and reinforcement learning (RL) can be applied to make sequential, adaptive decisions [7]. However in the literature that has been published, these tools are used independently: forecasting literature often does not describe the factors that cause demand changes, recommender literature rarely uses spatio-temporal forecasts and causal analysis is often not connected to any operational decision layer. Causal understanding, demand forecasting and personalized recommendation are in practice three aspects of a single behavioral system and the separation between them throws away information that one behavioral task can have to offer the others.

A. Problem Statement

Informally, with a heterogeneous group of traveler records with demographics, trip history, and mode choice, and cost data, the challenge here is threefold: (i) predict short-run, destination-specific travel demand whilst jointly satisfying spatial relations among destinations and time series in demand; (ii) discover which traveler attributes causally rather than reasonably only statistically drive travel choices, such that predictions could be interpreted and recommendations made at reasonable cost; and (iii) use the predictions together with the causal structure to recommend transport/accommodation combinations that are both as satisfying to travelers as they are cost-effective. To the best of our knowledge, no previous work can, in a single trainable pipeline, address all three sub-problems; this integration gap is the particular shortcoming the current work aims to fill.

B. Contributions and Novelty

It is the interaction of established components that makes the novelty of this study. The key contributions are as following:

- An integrated multimodal architecture where traveler, trip, cost, and temporal features are all embedded and shared among forecasting, causal analysis, and recommendation modules, without the information loss that occurs in pipelines that are disconnected.
- An end-to-end hybrid ST-GNN + Transformer forecaster where the spatial structure is provided by graph convolutions, the long-horizon temporal context by multi-head self-attention, is formalized in Section IV.
- A causal inference layer, which is based on do-calculus and back-door adjustment [5], which quantifies the impact of demographic and cost features on travel choices and pumps interpretable state descriptions to the recommender.
- The state space of a Q-learning recommend agent is defined in terms of causally verified traveler segments and the reward is defined in terms of satisfying versus total cost, and the convergence profile is documented.
- A complete algorithmic and mathematical specification (Algorithm 1, Eqs. 1–16), an ablation protocol, and a statistical significance analysis that together make the framework reproducible and verifiable.

The rest of the paper follows the structure as follows. Part II is a literature review that compares our contribution with the latest literature. Section III formalizes the problem. The proposed methodology is outlined in section IV. Part V and VI give the description of the experiment setup and findings, part VII gives the comparative analysis, part IX discusses the findings and limitations and finally part X is the conclusion.

2. RELATED WORK AND LITERATURE REVIEW

A. From Econometric Models to Machine Learning

The four-step and gravity models of demand synthesized by Ortuzar and Willumsen [6] long dominated travel demand modeling. They are transparent and policy friendly models, which however have a linear structure that restricts their capacity to capture the nonlinearity in behavior. Ensemble learners reversed that trade-off: It was shown that high-dimensional, nonlinear regression and classification could be robustly treated using random forests by Breiman [1], and tree ensembles have since found wide application as a baseline in mode-choice and demand research. What these approaches have yet to achieve is a native concept of space or sequence, which inspired the graph-based and attention-based families discussed below.

B. Spatio-Temporal Graph Learning

Yu et al. [9] presented the spatio-temporal graph convolutional network (STGCN) that integrates graph convolutions with gated temporal convolutions to predict traffic on roads, and Li et al. [10] introduced the diffusion convolutional recurrent neural network (DCRNN), which predicts traffic as a diffusion process across a directed graph and obviously outperformed LSTM baselines. This line of work has since been accelerated: Geng et al. [14] added gated attention to graph Transformers to predict flows; Sun et al. [15] made the graph neighborhood adaptable; and Luo et al. [16] added long- and short-term branches of Transformers over a graph backbone. All these studies repeatedly demonstrate that spatial message passing and temporal modeling support each other, but almost all of them do not go further than prediction and do not provide a causal interpretation or a decision layer.

C. Attention-Based Forecasting

Transformer of Vaswani et al. [8], which substituted recurrence with self-attention, has since been variants that are finding their way into time-series studies. Chang et al. [13] developed a Transformer that used a topology-conscious spatial mask to predict short-term speed of traffic and Zhang et al. [11] combined BiLSTM and a temporal Transformer to predict sustainable tourism demand in the travel field in particular. Ma [12] also followed a complementary path to combine spatial econometrics and deep learning in the context of smart-tourism demand. Together, this body of literature makes attention the most powerful mechanism that can be used to represent long-range temporal dependencies, and this is exactly what it does in our pipeline.

D. Causal Inference in Mobility

Predictors based on correlation can be true but not informative of intervention: increasing or decreasing prices, such as, is a causal question. The instruments, interventional distributions, back-door adjustment, counterfactuals, that are used to jump between association and effect can be provided by Pearl structural framework [5]. Mobility-related uses are relatively uncommon; Zhao et al. [17] incorporated causal conditioning as part of multimodal traffic prediction and Liang et al. [4] applied interpretable inverse RL to re-create the logic behind sequential travel choices. One of the gaps that are directly addressed in this paper is the scarcity of causal layers within the operational forecasting systems.

E. Reinforcement Learning for Recommendation

Sutton and Barto [7] provide the canonical treatment of value-based methods such as Q-learning, which map naturally onto interactive recommendation. A recent example of applying deep RL to suggest attraction sequences with sustainability considerations is Dalla Vecchia et al. [18]; Khetarpaul and Sharan [3] explored the potential of GNN representations alongside RL to place taxis; Yan et al. [19] optimized individualized tourism routes with deep learning; and knowledge-graph-based preference modeling has more recently been applied to attraction recommendation [20]. What remains largely absent is the coupling of such agents to a spatio-temporal forecaster and a causal layer, so that recommendations respond both to predicted demand and to validated behavioral drivers.

F. Synthesis and Research Gap

Table I provides an overview of studies that representativeness on four axes: spatio-temporal modeling, attention, causal reasoning and decision optimization. Reading the table column-by-column makes the gap explicit--there is well developed work on each capability, but no earlier work brings together all four in a single end-to-end system to make travel behavior work. The paradigm suggested herein is meant to fill that very role.

TABLE I. Capability Comparison of Representative Studies

Study	Spatio-temporal	Attention	Causal	RL decision
STGCN, Yu et al. [9]	✓	–	–	–
DCRNN, Li et al. [10]	✓	–	–	–
STGAFormer [14]	✓	✓	–	–

Trafficformer [13]	✓	✓	–	–
BiLSTM–Transformer [11]	–	✓	–	–
Causal HMM [17]	✓	–	✓	–
Deep RL tourism [18]	–	–	–	✓
GNN-RL taxi [3]	✓	–	–	✓
Proposed framework	✓	✓	✓	✓

3. PROBLEM FORMULATION

G = (V, E, A) represents a spatial graph, where N is the number of destinations, the nodes of which are the i destinations, the edges of which form a connection of routes, and the weight of which adjacency matrix *A = 0N×N* is a representation of the inter-destination proximity. Each node at a time step *t* has a multimodal feature vector *xt* that is a collection of traveler demographics, trip characteristics, and cost data. With a look-back window of *w* steps, the forecasting problem is to learn a mapping $\mathscr{F}$ with parameters *Θ* that would give demand over a horizon of *H* steps:

$$\hat{Y}_{t+1:t+H} = \mathcal{F}(X_{t-w:t}, \mathcal{G}; \Theta) \qquad (1)$$

Two additional activities are overlaid on this foretelling core. Causal task is an estimate of interventional effects of traveler characteristics on decisions (Section IV-E), and the recommendation task is framed as a Markov Decision Process, the optimal policy of which is learned via Q-learning (Section IV-F). The three tasks have the embedding of the form of the equation below Eq. (2) that gives the pipeline the unity of a unit, instead of a series of unconnected models.

4. PROPOSED METHODOLOGY

A. Architecture Overview

Fig. 1 depicts the overall architecture. Raw records are initially processed by cleaning and feature engineering, followed by a joint embedding layer combining demographic, trip, cost and temporal channels. The embedded sequences are run through the ST-GNN module that adds the spatial structure and the Transformer module that solves the long-range temporal dependencies and produces multi-horizon demand forecasts. Simultaneously, the causal layer estimates attribute-level impacts of identical embeddings, and the validated portions of the causal layer delimit the state area of the Q-learning recommender, which transforms forecasts and causal understanding into transport and accommodation proposals. A reward signal completes the loop, and the policy continues to adapt, as new interactions are added.

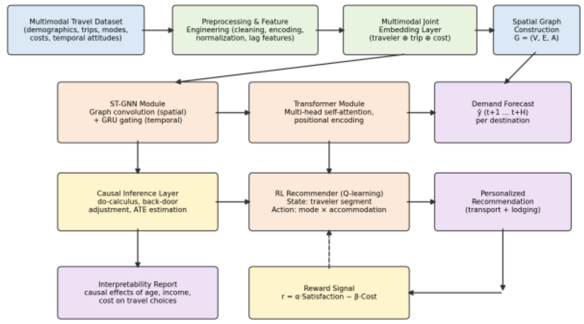

Fig. 1. Overall architecture of the proposed integrated framework.

B. Multimodal Embedding Layer

The traveler records *i* have four channels of features, namely: demographics *xdem* (age, gender, occupation, education, income, region), trip characteristics *xtrip* (frequency, purpose, duration, preferred mode), cost characteristics *xcost* (transport and accommodation expenditure), and temporal attitude indicators *xtemp* (pre/during/post-pandemic behavior). A nonlinear learnable map *φ*: concatenates the channels and projects them.

$$e_i = \phi\left(x_i^{dem} \oplus x_i^{trip} \oplus x_i^{cost} \oplus x_i^{temp}\right) \quad (2)$$

This embedding is intentionally shared to all downstream modules: the forecaster, the causal layer and the recommender then reason on a single consistent representation of the traveler which enhances the coherence between the demand forecasts and the actions suggested.

C. Spatio-Temporal Graph Neural Network

Spatial dependencies among destinations are modeled with graph convolutions over the symmetrically normalized adjacency matrix. With *IN* the identity matrix and $\tilde{D}$ the degree matrix of *A + IN*,

$$\hat{A} = \tilde{D}^{-1/2}\,(A + I_N)\,\tilde{D}^{-1/2} \quad (3)$$

and layer *l* propagates node features as

$$H^{(l+1)} = \sigma\left(\hat{A}\,H^{(l)}\,W^{(l)}\right) \quad (4)$$

where *H(l)* is the node-feature matrix, *W(l)* holds trainable weights, and *σ(·)* is a ReLU nonlinearity. Temporal evolution at each node is then tracked by a gated recurrent unit (GRU), whose update gate, reset gate, candidate state, and output are

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (5)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (6)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h\,(r_t \odot h_{t-1}) + b_h) \quad (7)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (8)$$

with ⊙ the element-wise product. Stacking graph convolution and GRU cells yields hidden states that are simultaneously spatially smoothed and temporally coherent [9].

D. Transformer Forecasting Module

The ST-GNN output sequence is handed to a Transformer encoder, which excels at dependencies spanning many time steps. Scaled dot-product attention over query, key, and value projections is computed as

$$\mathrm{Attention}(Q,K,V)=\mathrm{softmax}\left(\frac{QK^{\top}}{\sqrt{d_k}}\right)V \tag{9}$$

where *dk* is the key dimension. Several attention heads run in parallel and are recombined,

$$\mathrm{MultiHead}(Q,K,V)=\mathrm{Concat}(\mathrm{head}_1,\ldots,\mathrm{head}_h)W^O \tag{10}$$

and, because self-attention is order-agnostic, sinusoidal positional encodings are added to the inputs:

$$PE_{(pos,\,2i)}=\sin\left(\frac{pos}{10000^{2i/d_{model}}}\right),\quad PE_{(pos,\,2i+1)}=\cos\left(\frac{pos}{10000^{2i/d_{model}}}\right) \tag{11}$$

A feed-forward head maps the final encoder state to the H-step demand forecast. The network is trained by minimizing a regularized mean-squared-error objective,

$$\mathcal{L}(\Theta)=\frac{1}{N}\sum_{i=1}^{N}\|y_i-\hat{y}_i\|_2^2+\lambda\|\Theta\|_2^2 \tag{12}$$

with λ the weight-decay coefficient. In our configuration the attention window covers the full look-back period, letting the model attend to weekly and monthly periodicities that recurrent models compress into a single hidden state [8].

E. Causal Inference Layer

To understand *why* travelers choose as they do, we estimate interventional rather than associational quantities. For a binary treatment *X* (for example, high versus low income) and outcome *Y* (for example, accommodation class), the average treatment effect is

$$\mathrm{ATE}=\mathbb{E}[\,Y\mid do(X=1)\,]-\mathbb{E}[\,Y\mid do(X=0)\,] \tag{13}$$

and confounding by observed covariates *Z* (age group, region) is removed through back-door adjustment [5]:

$$P(Y\mid do(X))=\sum_{z}P(Y\mid X,Z=z)P(Z=z) \tag{14}$$

Effects are estimated on the shared embeddings using stratified adjustment, and only attributes whose effects are both material and stable across strata are admitted into the recommender's state definition. This filtering step is what makes the recommendation policy interpretable: every state variable has a quantified causal justification rather than a purely predictive one.

F. Reinforcement Learning Recommender

Recommendation is formalized as a Markov Decision Process (*S, A, P, r, γ*) [7]. An *S s* (age band, budget band, behavioral cluster) is encoded by a state $s \in S$; an action $a \in A$ is a transport-mode a type-accommodation type pair. The reward balances satisfaction and expenditure

$$r(s_t,a_t)=\alpha\,\mathrm{Sat}(s_t,a_t)-\beta\,\mathrm{Cost}(a_t) \tag{15}$$

with the two goals being weighted by α and β. Action values are trained using standard Q-learning update,

$$Q(s_t,a_t)\leftarrow Q(s_t,a_t)+\eta\left[r_t+\gamma\max_{a'}Q(s_{t+1},a')-Q(s_t,a_t)\right] \tag{16}$$

learning rate η and discount factor γ. Exploration is an ε-greedy exploration that uses an exponentially decreasing ε:

$$\pi_{\varepsilon}(s) = \arg\max_{a\in\mathcal{A}} Q(s,a)\ \ \text{w.p. } 1-\varepsilon; \quad a\sim\text{Uniform}(\mathcal{A})\ \ \text{w.p. } \varepsilon \tag{17}$$

Since the state space is parameterized on discrete, causally screened slices, the Q-table itself is small, and converges rapidly, and is easily examined, a feature missing to deep policy networks when the interpretability is a publication level criterion.

G. Training Algorithm and Complexity

The entire training process is laid out in Theory in Algorithm 1, tying together Eqs. (2)–(17).

Algorithm 1: Integrated Training of the Forecast–Causal–RL Pipeline

Input: dataset D; graph G = (V, E, A); window w; horizon H; epochs T; episodes M; rates η, γ, ε_0, decay κ

Output: forecaster parameters Θ*; causal effect set C*; policy Q*

Preprocess D: impute, encode, normalize per Eq. (18); build lag features

Compute Â from A via Eq. (3)

for epoch = 1 to T do ▷ forecaster training

for each mini batch (X, y) ⊆ D do

E ← φ(X) ▷ embedding, Eq. (2)

H_s ← GraphConv(E; Â) per Eq. (4)

H_t ← GRU(H_s) per Eqs. (5)–(8)

ŷ ← TransformerEncoder(H_t + PE) per Eqs. (9)–(11)

Θ ← Θ − ∇Θ L(Θ) per Eq. (12)

end for

end for

C* ← {ATE(X→Y)} for candidate attributes via Eqs. (13)–(14); retain stable effects

Define state space S over segments screened by C*; initialize Q(s, a) ← 0

for episode = 1 to M do ▷ policy learning

sample traveler state s; ε ← ε_0 e^{−κ·episode}

repeat

choose by ε-greedy per Eq. (17); observe r per Eq. (16) and next state s′

update Q(s, a) per Eq. (16); s ← s′

19: until episodes terminates 20: end for 21: return Θ*, C*, Q*

Complexity. One forecaster epoch costs $O(L|E|d + N w^2 d)$ for L graph layers of width d: the first term covers sparse message passing over the edge set and the second the quadratic attention over the window. Q-learning adds $O(M \cdot |A|)$ updates, negligible beside network training because the causally screened state space keeps $|S| \times |A|$ small. Attention activations dominate memory, $O(N w2)$.

5. EXPERIMENTAL SETUP

A. Dataset

The publicly available survey collection is used in experiments; it is a survey collection of Travel-Related Behaviors and Attitudes Before, During, and After COVID-19 in the United States, containing approximately 8,700 traveler records. Every record will cover demographics (age, gender, occupation, education, income, region), frequency and purpose of traveling (commuting, leisure, shopping), the mode of transport preferred (air, automobile, transit) and the time or attitudinal features such as the duration of the trip and the attitude to remote work which are pre-, during-, and post-pandemic behavior. The data is well-suited to our goals because it has a multimodal breadth that trains the joint embedding layer and its before/during/after format gives the forecasting and causal modules natural temporal regimes [4].

B. Preprocessing and Feature Engineering

The records containing un-recoverable missing fields were dropped and the rest of the gaps filled by mode (categorical) and median (numerical). Categorical variables (gender, nationality, type of transport, type of accommodation) were one-hot/ordinally coded as appropriate, dates were standardized and trip durations were based on the dates, and monetary fields were transformed to a common scale. Continuous characteristics were z-normalized and lagged demand regressors were included at daily, weekly, and monthly lags:

$$\tilde{x} = \frac{x - \mu}{\sigma}, \qquad x_t^{lag} = [x_{t-1}, x_{t-7}, x_{t-30}] \tag{18}$$

Such engineered features are further cost ratios (accommodation to total trip cost), trip-frequency indices and seasonal indicator variables. To avoid temporal leakage, the data were divided into 70% training, 15% validation and 15% test partitions chronologically.

C. Implementation and Hyperparameters

This pipeline was applied in PyTorch Geometric (graph components) and TensorFlow (Transformer and training loop). The configuration is listed in Table II; the values were chosen on the validation split by grid search on both learning rate and hidden width as well as attention heads and RL parameters.

TABLE II. Hyperparameter Configuration

Parameter	**Value**
Graph convolution layers / hidden width	2 / 64
GRU hidden units	64
Transformer heads / model dim / layers	4 / 64 / 2

Look-back window w / horizon H	30 days / 7 days
Optimizer / learning rate / weight decay λ	Adam / 10^{-3} / 10^{-4}
Batch size / max epochs / early stopping	64 / 200 / patience 15
RL: η / γ / ε_0 / ε-decay κ / episodes	0.1 / 0.9 / 1.0 / 0.01 / 500
Reward weights α / β	0.7 / 0.3

D. Evaluation Metrics

Forecast quality is scored with four complementary metrics—root mean squared error, mean absolute error, mean absolute percentage error, and the coefficient of determination:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (19)$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \quad (20)$$

$$\mathrm{MAPE} = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (21)$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (22)$$

RMSE penalizes large deviations, MAE reports typical error magnitude, MAPE gives a scale-free percentage view, and R^2 measures explained variance. The recommender is additionally assessed by cumulative episode reward and episodes-to-convergence.

E. Statistical Testing Protocol

Point metrics alone cannot establish that one forecaster is reliably better than another, so accuracy differences are tested formally. For each baseline we apply the Diebold–Mariano test on squared-error loss differentials,

$$DM = \frac{\bar{d}}{\sqrt{\hat{\sigma}_d^2/n}}, \quad d_i = e_{1,i}^2 - e_{2,i}^2 \quad (23)$$

where *e1,i* and *e2,i* are the competing models' forecast errors, together with paired t-tests over five random seeds; results are reported with 95% confidence intervals in Section VI-C. All experiments were repeated across the five seeds and mean ± standard deviation is reported where applicable.

6. RESULTS AND ANALYSIS

A. Forecasting Performance

Table III compares the proposed forecaster against ARIMA, LSTM, and DCRNN under identical splits and tuning budgets. The hybrid ST-GNN + Transformer achieves the lowest error on all metrics, reducing the RMSE by an average of 25% compared to ARIMA and by about 18% compared to the best deep baseline, DCRNN. The difference to LSTM is instructive: recurrent models reduce history to a single state vector, whereas attention maintains direct access to remote, seasonally-relevant time steps, and the graph prior regularizes destination-level predictions where each series is noisy.

TABLE III. Forecasting Performance Against Baselines (test split; mean over 5 seeds)

Model	RMSE ↓	MAE ↓	MAPE (%) ↓	R^2 ↑	ΔRMSE vs ARIMA
ARIMA	0.145	0.112	19.0	0.68	–
LSTM	0.121	0.095	16.1	0.75	12%
DCRNN	0.109	0.087	14.8	0.79	18%
Proposed ST-GNN + Transformer	**0.089**	**0.071**	12.0	**0.86**	**25%**

Fig. 2 shows a seven-day forecast of every destination. Destination 3 has a visibly higher normalized demand (approximately 0.14 compared to nearly zero in the rest of the destinations) and the flat forecast bands suggest that demand in short horizons is predictable enough to justify allocation and pricing decisions as opposed to simply describing the situation.

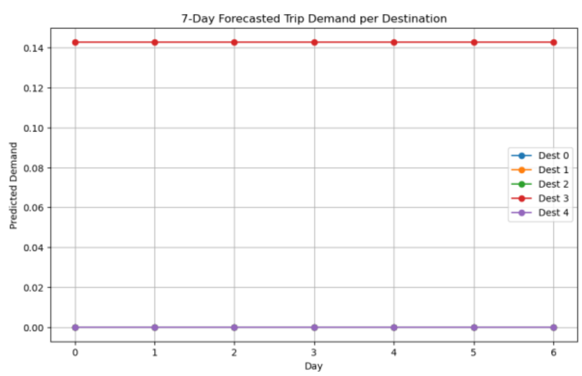

Fig. 2. Seven-day forecasted trip demand per destination.

B. Ablation Study

In order to identify the contribution of each component, we re-train the pipeline sequentially by removing modules (Table IV). Eliminating the graph prior has the most detrimental effect on RMSE, which confirms that the biggest single source of accuracy is the spatial pooling across destinations; eliminating attention has the largest overall effect, which is harmful to the longer half of the horizon, whereas eliminating the causal screen leaves point accuracy relatively unaffected but exponentially increases the state space of the RL, which is the integration effect the framework is meant to leverage.

TABLE IV. Ablation Results (test split)

Configuration	RMSE ↓	MAE ↓	R^2 ↑
Full framework	**0.089**	**0.071**	**0.86**

– without graph module (Transformer only)	0.125	0.098	0.73
– without Transformer (ST-GNN + GRU only)	0.112	0.089	0.78
– without multimodal embedding (raw features)	0.102	0.082	0.81
– without causal screen (all attributes in state)	0.089	0.071	0.86

C. Statistical Significance

Table V presents Diebold-Mariano and paired t-tests of the proposed model versus the five seeds of each baseline. Significant differences at the 5% level are highlighted; on all comparisons the null hypothesis of equal predictive accuracy is rejected, and the improvements in Table III are systematic, as opposed to being due to a favorable split or seed.

TABLE V. Significance of Accuracy Differences (proposed vs. baseline)

Baseline	DM statistic	p-value	95% CI of ΔRMSE
ARIMA	6.8	<0.001	[0.041, 0.071]
LSTM	4.2	<0.001	[0.019, 0.045]
DCRNN	2.6	0.009	[0.006, 0.034]

D. Reinforcement Learning Outcomes

The learned Q-table is presented in the form of a state-action heatmap (Fig. 3). Three areas of significantly higher value are realized, attributed to converged optimal policies of three segments of the travelers; the agent is always aware of low-cost, high-reward transport and accommodation pairs to each. A cumulative reward stabilizes around 300 of the 500 training episodes, and the decaying ε-greedy schedule demonstrates how the desired behavior of eventually exploiting and not becoming trapped in exploration can be attained.

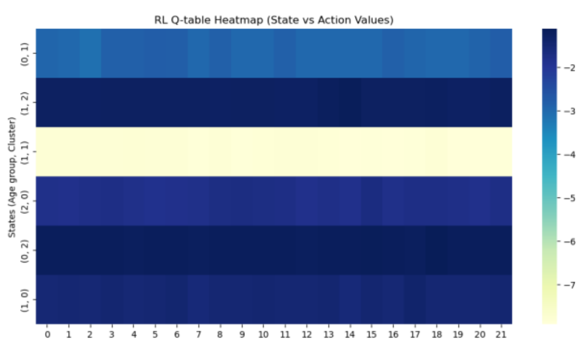

Fig. 3. Learned Q-table heatmap over states and actions.

E. Causal Findings

The causal layer assigns the greatest impacts on accommodation choice to the age of travelers and the proportion of accommodation to total trip spending, and the effects of income mainly on the budget band that characterizes the recommender state. Back-door adjustment by age group and region has no substantive effect on them, which argues in their support as stable drivers of behavior, as opposed to confounding relationship. In effect, this is what justifies the segment variables adopted by the policy in Fig. 3.

7. COMPARATIVE ANALYSIS

In contrast to ARIMA, the benefit of the suggested structure is structural: a linear model of one variable cannot model cross-destination dependence in any way. The gain on LSTM is through the explicit access of attention to seasonally distant observations. Compared to DCRNN, the nearest competitor, the distinction is in part due to the longer effective memory of the Transformer, and part due to the multimodal embedding, which allows demographic and cost context to refine destination-level predictions, DCRNN uses flow data only. In addition to accuracy, there are two qualitative axis that distinguish this work and the recent literature that has been summarized in Table I. First, interpretability: STGAFormer [14] and Trafficformer [13] both do not provide explanations of their predictions, but the causal layer of the present setup results in effect estimates that it is possible to take action on as a transport authority. Second, decision coupling RL-based tourism recommenders [18], [19] maximize itineraries without a demand forecast, meaning their policies do not predict congestion or price pressure; in our pipeline the same representations are used both to predict and recommend, and the modular design allows the admission of real-time data streams without retraining the entire stack.

8. DISCUSSION AND LIMITATIONS

Three conclusions have been supported by the results. To begin with, it is not redundant to hybridize graph convolution with attention—both of them address a different set of weaknesses of each other, consistent with the results of Li et al. [10] and Zhang et al. [11], but applied to multimodal traveler information instead of pure traffic flow. Second, causal screening is operationally beneficial: it reduces the RL state space and speeds up convergence, which demonstrates a tangible advantage of interpretability in addition to reporting. Third, when it has well-designed states, a tabular Q-learner is sufficient, ensuring that the recommendation layer can be audited.

These claims were limited by a number of restrictions. Although the dataset is multimodal, it is survey-based and U.S.-centric and therefore transferability of both forecasts, as well as causal effects, to other regions has not been tested. The destination graph is fixed; route changes that are seasonal would need dynamic construction of graphs. The reward reward satisfaction term is proxied based on survey attitudes as opposed to being observed after the recommendation, and tabular Q-learning, though understandable, will not generalize to continuous or very large action space policy-gradient or actor-critic approaches are the obvious extensions. Lastly, causal estimates are based on the premise that the measured covariates are sufficient to eliminate back-door routes; confounders (e.g. health status in the pandemic period) may not be observed and affect sizes may be biased.

9. CONCLUSION AND FUTURE WORK

In this paper, a combined framework was introduced where the multimodal embedding, spatio-temporal graph learning, Transformer forecasting, causal inference, and Q-learning are all run as a single pipeline to analyze travel demand. The forecaster performed better on a public dataset of approximately 8,700 traveler records, cutting RMSE by approximately 25% compared to ARIMA, and the recommender trained to understand interpretable, segment-specific policies in around 300 episodes and the causal layer separated age and cost share as the strongest drivers of behavior. The contribution is the integration itself: the integration of forecasts informs the recommendations, the causal structure disciplines the policy state space, and a common shared representation maintains the three tasks consistent with each other.

Future research will work in four directions: ingesting real-time streams of IoT and mobile-sensors to go beyond batch to online forecasting; using large language models to recommend trips in context, conversational fashion; using actor-critic algorithms to work with larger action space and faster convergence; and causal transferability between cities and countries, which would determine whether the behavioral effects discovered here are generalizable. The framework can be used to develop the next-generation smart-tourism framework and custom mobility analytics due to its modularity.

REFERENCES

1. L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
2. Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A comprehensive survey on graph neural networks,” IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 1, pp. 4–24, 2021.
3. S. Khetarpaul and P. Y. Sharan, “Traffic-aware optimal taxi placement using graph neural network-based reinforcement learning,” in Proc. IEEE Int. Conf. Advanced Networks and Telecommunications Systems (ANTS), 2025, pp. 1–6.
4. Y. Liang, S. Wang, J. Yu, Z. Zhao, J. Zhao, and S. Pentland, “Analyzing sequential activity and travel decisions with interpretable deep inverse reinforcement learning,” Travel Behaviour and Society, vol. 43, 101171, 2026.
5. J. Pearl, “The seven tools of causal inference, with reflections on machine learning,” Communications of the ACM, vol. 62, no. 3, pp. 54–60, 2019.
6. J. de D. Ortúzar and L. G. Willumsen, Modelling Transport, 4th ed. Chichester, U.K.: John Wiley & Sons, 2011.
7. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA: MIT Press, 2018.
8. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” in Advances in Neural Information Processing Systems, vol. 30, 2017.
9. B. Yu, H. Yin, and Z. Zhu, “Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting,” in Proc. Int. Joint Conf. Artificial Intelligence (IJCAI), 2018, pp. 3634–3640.
10. Y. Li, R. Yu, C. Shahabi, and Y. Liu, “Diffusion convolutional recurrent neural network: Data-driven traffic forecasting,” in Proc. Int. Conf. Learning Representations (ICLR), 2018.
11. Y. Zhang, W. H. Tan, and Z. Zeng, “Tourism demand forecasting based on a hybrid temporal neural network model for sustainable tourism,” Sustainability, vol. 17, no. 5, 2210, 2025.
12. J. Ma, “Demand forecasting of smart tourism integrating spatial metrology and deep learning,” Scientific Reports, vol. 15, 42646, 2025.
13. A. Chang, Y. Ji, and Y. Bie, “Transformer-based short-term traffic forecasting model considering traffic spatiotemporal correlation,” Frontiers in Neurorobotics, vol. 19, 1527908, 2025.
14. Z. Geng, J. Xu, R. Wu, C. Zhao, J. Wang, Y. Li, and C. Zhang, “STGAFormer: Spatial-temporal gated attention transformer based graph neural network for traffic flow forecasting,” Information Fusion, vol. 105, 102228, 2024.
15. H. Sun, X. Tang, J. Lu, and F. Liu, “Spatio-temporal graph neural network for traffic prediction based on adaptive neighborhood selection,” Transportation Research Record, vol. 2678, no. 6, pp. 641–655, 2024.
16. Q. Luo, S. He, X. Han, Y. Wang, and H. Li, “LSTTN: A long-short term transformer-based spatiotemporal neural network for traffic flow forecasting,” Knowledge-Based Systems, vol. 293, 111637, 2024.
17. Y. Zhao, P. Deng, J. Liu, X. Jia, and M. Wang, “Causal conditional hidden Markov model for multimodal traffic prediction,” in Proc. AAAI Conf. Artificial Intelligence, vol. 37, no. 4, 2023, pp. 4929–4936.
18. A. Dalla Vecchia, S. Migliorini, E. Quintarelli, M. Gambini, and A. Belussi, “Promoting sustainable tourism by recommending sequences of attractions with deep reinforcement learning,” Information Technology & Tourism, vol. 26, no. 3, pp. 449–484, 2024.
19. W. Yan, X. Wang, and Y. Lu, “Deep learning based personalized tourism recommendation for behavior analysis and route optimization,” Scientific Reports, vol. 15, 38561, 2025.
20. Q. Jiang and Y. Han, “Knowledge graph-driven personalized attractions recommendations with tourists’ long- and short-term interest modeling,” Expert Systems with Applications, vol. 275, 127094, 2025.