Long-distance line planning with endogenous demand

Zixuan Song*[1], Yongqiu Zhu[1], Shadi Sharif Azadeh[1], and Oded Cats[1]

[1]Transport & Planning Department, Delft University of Technology, Netherlands

Short summary

This study develops an integrated line planning framework for long-distance rail by treating passenger demand as endogenous, recognizing that both ridership level and distribution depend on service frequency and travel time. To address station heterogeneity, the methodology employs latent-class clustering to group stations into probabilistic typologies, which informs a gravity-based demand model. This is integrated into a mixed-integer linear program that builds upon hierarchical line generation and optimizes service planning while accounting for frequency-dependent transfer times. Applying this framework to the rail network of Great Britain, we conclude that demand-aware planning significantly enhances economic efficiency and service relevance compared to benchmark line generation methods and traditional fixed-demand models, respectively. Our findings demonstrate that capturing travellers' behavioural responses and structural roles of network stations leads to more realistic, cost-effective service designs that better align infrastructure supply with passenger needs.

Keywords: line planning, long-distance passenger transport, endogenous demand, mixed-integer program, transport network modelling.

1 Introduction

Long-distance public transport supports sustainable regional connectivity. Line planning is a critical component in determining routes, stopping patterns, and service frequencies. As is customary, we divide this comprehensive decision problem into three components: demand estimation, line pool generation, and optimization computation of the final line concept, each of them providing inputs for the subsequent one. In this study, we approach these stages in a conceptually consistent manner to obtain a practical line concept.

Despite a wealth of research and practice, line planning process is often hindered by a fundamental oversight: the use of exogenous (fixed) demand models. In reality, passenger demand is endogenous; travellers respond to service attributes like frequency and travel time. Ignoring this feedback loop can lead to suboptimal line concepts and biased ridership forecasts. While recent literature has begun acknowledging demand elasticity, focusing on either frequency (Hartleb et al., 2023) or travel time (Wang et al., 2025), existing models rarely integrate both or treat transfer time as a decision-dependent variable. This gap is critical for long-distance networks, where transfer penalties are substantial and line planning directly determines the interchange structure. Ignoring this interaction can therefore distort passenger cost estimation and network topology.

Research on line generation proposes heuristic construction methods (Lin & Tang, 2021; Grolle et al., 2024) and hierarchical approaches (Bagloee & Ceder, 2011). However, these methods generally rely on pre-estimated flows (Gattermann et al., 2017; Cipriani et al., 2012) given fixed OD demand, which introduces structural bias when service design influences ridership.

Furthermore, stations typically exhibit heterogeneous roles (e.g., hubs, connectors), each associated with different demand generation mechanisms. Incorporating this heterogeneity probabilistically improves behavioural fidelity but remains neglected.

To bridge these gaps, we propose an integrated framework that (i) incorporates probabilistic station typologies, (ii) models OD demand explicitly as a function of generalized travel time and service frequency, (iii) generates a structured line pool without pre-assigned flow assumptions, and (iv)

treats transfer time as frequency-dependent within the optimization model. The effectiveness of this workflow is demonstrated on an existing rail network.

2 Methodology

The framework (Figure 1) integrates station classification-based demand estimation, hierarchical line pool generation, and mixed-integer program (MIP) optimization to output line concept, demand flow, and system KPI results.

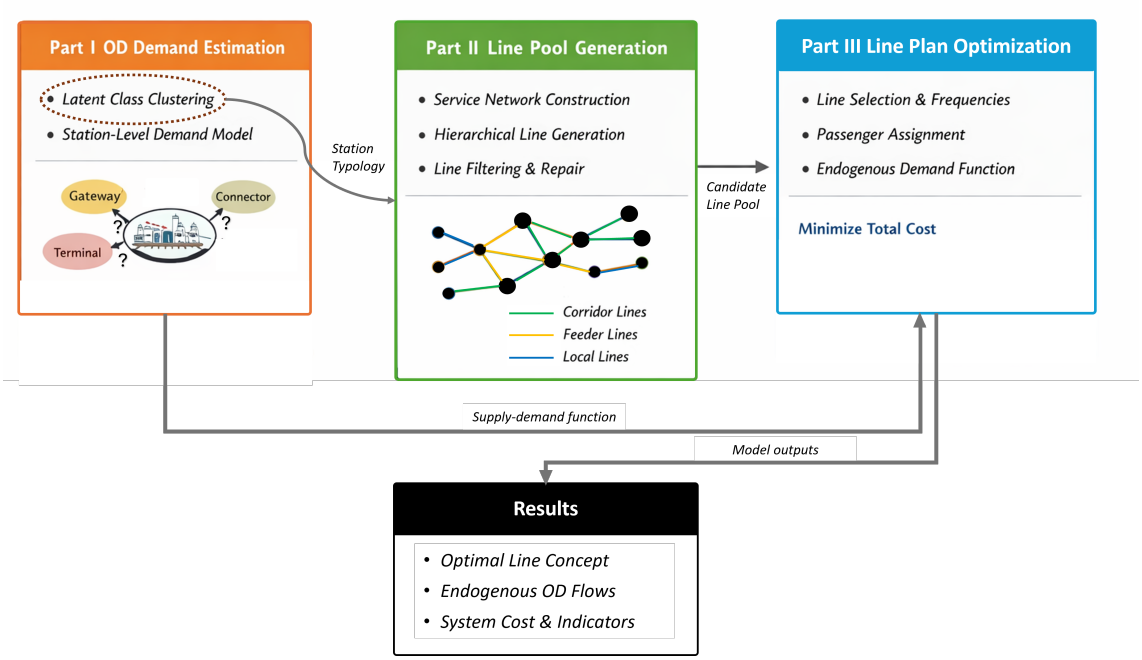


Figure 1: Methodology framework for line planning with endogenous demand

Passenger Demand Estimation

Gravity-type formulations remain a standard approach for intercity demand modelling de Dios Ortúzar & Willumsen (2024). We employ a Poisson-based gravity model incorporating probabilistic station typologies derived from latent-class clustering. Stations are grouped using socio-economic indicators (e.g., population, tourism) and network-topological metrics (e.g., centrality), with cluster counts optimized by the Bayesian Information Criterion.

The demand q_{vw} for OD pair (v, w) is a function of origin/destination potentials (O_v, D_w), generalized travel time (t_{vw}), and departure frequency (f_{vw}):

$$q_{vw} = O_v D_w \left[\alpha_1^{c(v),c(w)} (t_{vw})^{\alpha_2} + \alpha_3^{c(v),c(w)} (f_{vw}^{\text{depart}})^{\alpha_4}\right], \tag{1}$$

where potentials follow a multiplicative structure of explanatory variables X:

$$O_v = \prod_{p=1}^{P} (X_{vp})^{\beta_p^O}, \quad D_w = \prod_{p=1}^{P} (X_{wp})^{\beta_p^D}, \tag{2}$$

Class-pair parameters $\alpha^{c(v),c(w)}$ are weighted by membership probabilities:

$$\alpha_l^{c(v),c(w)} = \sum_{i,j} P(c(v) = i) P(c(w) = j) \alpha_l^{ij}, \quad l \in \{1, 3\} \tag{3}$$

Parameters $\theta = \{\alpha_{1,ij}, \alpha_2, \alpha_{3,ij}, \alpha_4, \beta\}$ are estimated via negative log-likelihood minimization using observed OD data.

For the MIP, endogenous demand is condensed to:

$$q_k = \gamma_1^k (t_k)^{\gamma_2} + \gamma_3^k (f_k^{\text{depart}})^{\gamma_4}, \tag{4}$$

where $\gamma_1^k = \alpha_1(k) \prod_{p=1}^{P} (X_{vp})^{\beta_p^O} \prod_{p=1}^{P} (X_{wp})^{\beta_p^D}$, $\gamma_3^k = \alpha_3(k) \prod_{p=1}^{P} (X_{vp})^{\beta_p^O} \prod_{p=1}^{P} (X_{wp})^{\beta_p^D}$, $\gamma_2 = \alpha_2$, $\gamma_4 = \alpha_4$. This formulation admits linearization by substituting t_k and f_k^{depart} with binary variables, allowing for direct integration into the MIP optimization.

Hierarchical Line Pool Generation

Candidate lines are generated hierarchically using station typologies: stations are categorized as primary and secondary, which can naturally be done based on the previously defined clusters.

First, a service network is constructed, where nodes represent stations, and edges (segments) comprise physical and "direct service" links between primary hubs (within a distance threshold) to allow skip-stop patterns.

Second, lines are routed via shortest paths subject to detour constraints and an upper bound on alternatives:

- **Corridor lines** connecting primary hubs;
- **Feeder lines** connecting secondary stations to the nearest primary hub;
- **Local lines** connecting stations whose connectivity falls below a threshold with their neighbours, ensuring minimum accessibility.

Third, the candidate set is pruned by (1) enforcing minimum length and stop counts, (2) applying overlap thresholds to eliminate near-duplicate lines, while (3) preserving node connectivity.

Finally, an enrichment procedure verifies that every OD pair is reachable with at most one transfer. For any failing OD pair, we traverse the shortest path and identify the point where a transfer becomes unavoidable. A new enrichment line between it and the destination is added if it fulfils the length requirements; otherwise an existing line is extended to the unserved destination.

Line Plan Optimization

The problem is a MIP that simultaneously determines line selection, frequencies and flows with endogenous OD demand. A compact notation is provided in Table 1.

Objective function The model minimizes net total societal cost by balancing operator expenses, passenger travel time and unserved demand, and ticket revenue:

$$\begin{aligned}
\min \quad & \underbrace{\sum_{l \in \mathcal{L}} \sum_{\phi \in \Phi} c_l \, \phi \, z_l^{\phi}}_{\text{Operating cost}} \\
& + \underbrace{\text{VoT} \sum_{k \in \mathcal{OD}} \sum_{r \in \mathcal{C}_k} \left(\bar{t}_r^{\text{in-veh}} + \kappa t_r^{\text{trans}} \right) w_r}_{\text{Passenger generalized travel cost}} \\
& + \underbrace{\sum_{k \in \mathcal{OD}} c_k u_k}_{\text{Penalties for unmet demand}} \\
& - \underbrace{\sum_{k \in \mathcal{OD}} \sum_{r \in \mathcal{C}_k} p_k w_r}_{\text{Ticket revenue from satisfied demand}} ,
\end{aligned} \tag{5}$$

Table 1: Notation

Sets	
$\mathcal{L}$	Set of candidate lines
Φ	Set of feasible frequency levels
$\mathcal{OD}$	Set of OD pairs
$\mathcal{C}_k$	Set of candidate routes for OD pair k
$\mathcal{E}$	Set of railway segments (edges)
Parameters	
c_l	Operating cost per vehicle of line l
Q_l	Vehicle capacity of line l
L_{ij}, U_{ij}	Lower/upper frequency bounds on segment (i,j)
$\bar{t}_r^{\text{in-veh}}$	In-vehicle time on route r
κ	Transfer time weight factor
VoT	Value of time
p_k	Ticket price per passenger per trip of OD pair k
c_k	Penalty per unserved passenger of OD pair k
T	Planning horizon
γ_1^k, γ_3^k	Demand scale parameters
γ_2, γ_4	Demand elasticities
Variables	
z_l^ϕ	binary, 1 if line l operates at frequency ϕ
x_l	binary, 1 if line l is selected
y_r	binary, 1 if route r is available
w_r	Passenger flow on route r
u_k	Unserved demand of OD pair k
η_r^k	binary, 1 if route r is shortest for OD pair k
q_k	Passenger demand of OD pair k

where z_l^ϕ is a binary variable determining line frequency, w_r corresponds to passenger route flow, and u_k unserved demand; $c_l, \text{VoT}, c_k, p_k$ are the corresponding cost and revenue parameters.

Constraints Constraints 6–8 govern line frequency selection (z_l^ϕ) and activation (x_l), enforcing edge-level bounds (L_{ij}, U_{ij}).

$$\sum_{\phi \in \Phi} z_l^\phi \leq 1, \quad \forall l \in \mathcal{L} \tag{6}$$

$$L_{ij} \leq \sum_{l \in \mathcal{P}_{ij}} \sum_{\phi \in \Phi} \phi z_l^\phi \leq U_{ij}, \quad \forall (i,j) \in \mathcal{E} \tag{7}$$

$$x_l = \sum_{\phi \in \Phi} z_l^\phi, \quad \forall l \in \mathcal{L} \tag{8}$$

Constraints 9–10 link route availability (y_r) to line activation, while Constraint 11 guarantees each passenger an available route. Constraint 12 affirms flows are allowed only on available routes, and flow conservation is enforced by Constraint 13. Constraint 14 ensures that passenger flows on each line segment cannot exceed vehicle capacity multiplied by frequency.

$$y_r \leq x_l, \quad \forall k \in \mathcal{OD}, r \in \mathcal{C}_k, l \in \mathcal{P}_r \tag{9}$$

$$y_r \geq \sum_{l \in \mathcal{P}_r} x_l - |\mathcal{P}_r| + 1, \quad \forall k \in \mathcal{OD}, r \in \mathcal{C}_k \tag{10}$$

$$\sum_{r \in \mathcal{C}_k} y_r \geq 1, \quad \forall k \in \mathcal{OD} \tag{11}$$

$$w_r \leq M y_r, \quad \forall k \in \mathcal{OD}, r \in \mathcal{C}_k \tag{12}$$

$$\sum_{r \in \mathcal{C}_k} w_r + u_k = q_k, \quad \forall k \in \mathcal{OD} \tag{13}$$

$$\sum_{k \in \mathcal{OD}} \sum_{r \in \mathcal{C}_k(l,i,j)} w_r \leq Q_l \sum_{\phi \in \Phi} \phi \, z_l^{\phi}, \quad \forall (i,j) \in \mathcal{E}, l \in \mathcal{P}_{ij} \tag{14}$$

As shown in Constraint 15, direct routes have zero transfer time. For single-transfer routes, transfer time depends on the frequency of the downstream line l_r^1 and is linearized into Constraint 16.

$$t_r^{\text{trans}} = 0, \quad \forall k \in \mathcal{OD}, r \in \mathcal{C}_k^{direct} \tag{15}$$

$$t_r^{\text{trans}} = \sum_{\phi \in \Phi} \frac{T}{2\phi} z_{l_r^1}^{\phi}, \quad \forall k \in \mathcal{OD}, r \in \mathcal{C}_k^{\text{trans}} \tag{16}$$

Using big-M Constraint 17 and indicator variable η_r^k in Constraints 18 and 19, the model identifies the shortest available route to determine the experienced travel time t_k and service frequency f_k^{depart} (Constraint 20, where l_r^0 represent the line serving route r).

$$\bar{t}_r^{\text{in-veh}} + \kappa t_r^{\text{trans}} - M(1 - \eta_r^k) \leq t_k \leq \bar{t}_r^{\text{in-veh}} + \kappa t_r^{\text{trans}} + M(1 - y_r), \quad \forall k \in \mathcal{OD}, r \in \mathcal{C}_k \tag{17}$$

$$\eta_r^k \leq y_r, \quad \forall k \in \mathcal{OD}, r \in \mathcal{C}_k \tag{18}$$

$$\sum_{r \in \mathcal{C}_k} \eta_r^k \geq 1, \quad \forall k \in \mathcal{OD} \tag{19}$$

$$f_k^{\text{depart}} = \sum_{r \in \mathcal{C}_k} \sum_{\phi \in \Phi} \phi \, \eta_r^k \, z_{l_r^0}^{\phi}, \quad \forall k \in \mathcal{OD} \tag{20}$$

As at most one transfer is allowed and frequencies are discrete, the non-linear demand q_k calculated via Equation 4 are reformulated into bi-linear expressions involving binary variables (Constraint 21), and subsequently handled using McCormick envelopes. This results in a solvable linear MIP.

$$\begin{aligned} q_k = & \gamma_1^k \sum_{r \in \mathcal{C}_k^{\text{direct}}} \eta_r^k \, (\bar{t}_r^{\text{in-veh}})^{\gamma_2} \\ & + \gamma_1^k \sum_{r \in \mathcal{C}_k^{\text{trans}}} \sum_{\phi \in \Phi} \eta_r^k \, z_{l_r^1}^{\phi} \, (\bar{t}_r^{\text{in-veh}} + \kappa \frac{T}{2\phi})^{\gamma_2} \\ & + \gamma_3^k \sum_{r \in \mathcal{C}_k} \sum_{\phi \in \Phi} \eta_r^k \, z_{l_r^0}^{\phi} \, (\phi)^{\gamma_4}, \quad \forall k \in \mathcal{OD} \end{aligned} \tag{21}$$

3 Results and discussion

We apply our model to 40 Great Britain rail stations selected for high inter-regional connectivity and volume using the 2023–2024 OD dataset (Office of Rail and Road, 2024). Figure 2 maps these nodes according to the three typologies derived along with the respective infrastructure links.

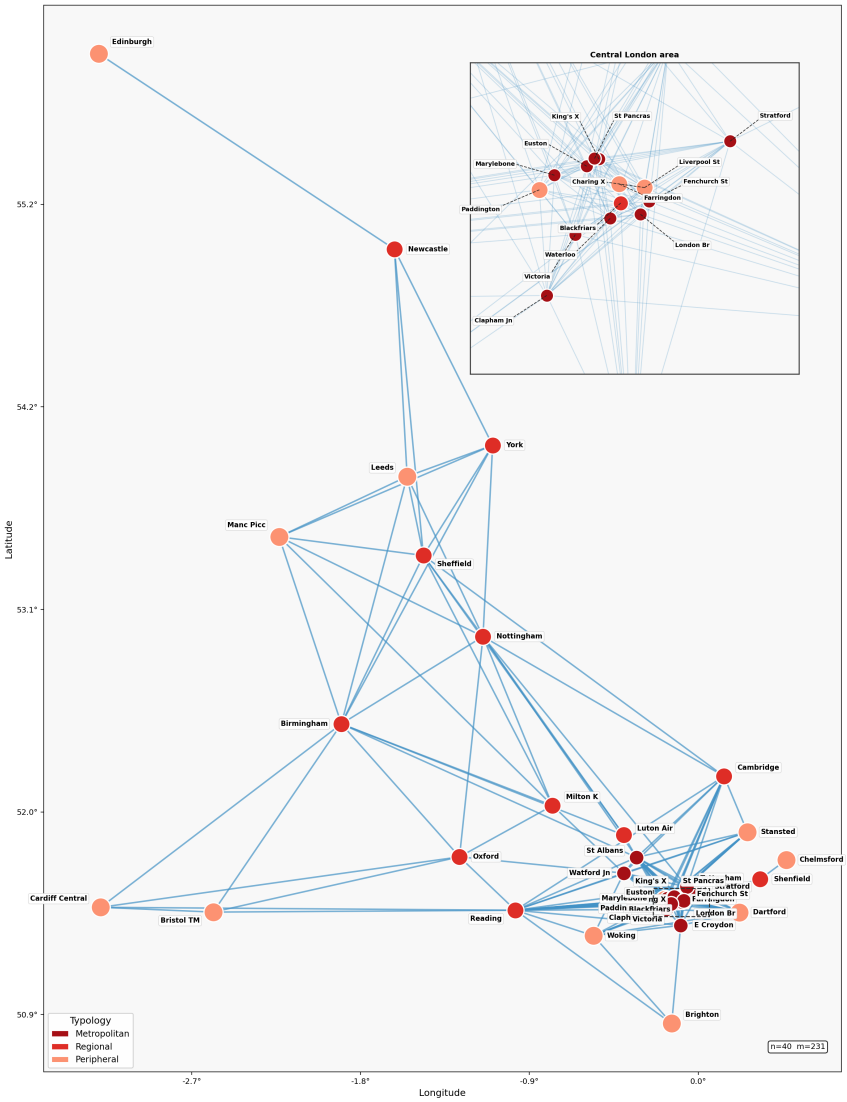


Figure 2: GB intercity rail network geographic layout

Demand Model Specifications

Using Gaussian mixture clustering, stations are grouped into three probabilistic typologies based on catchment characteristics and network centrality:

- **London metropolitan hubs** (15 stations; confidence: 0.998 ± 0.003) – massive catchment populations and high centrality.
- **Regional connectors** (12 stations; 1.000 ± 0.001) – moderate populations but high centrality, forming the national network backbone (e.g., Birmingham).
- **Peripheral/tourist stations** (13 stations; 0.986 ± 0.013) – above-average tourism activity, weaker airport competition, but lower centrality (e.g., Edinburgh).

The model achieves $R^2 = 0.60$ for OD demand prediction (Equation 1). Parameter estimates (Table 2) show behavioural consistency: demand decreases with travel time and increases with frequency. Station-level accessibility is the dominant structural driver, and airport proximity reflects intermodal competition. Travel-time scaling varies little across station-type combinations, whereas frequency scaling differs substantially, peaking for flows involving metropolitan hubs, indicating the importance of service intensity for high-tier connectivity. The typology-aware model significantly outperforms a uniform specification ($R^2 = 0.25$), which underestimates high-volume pairs.

Table 2: Estimated demand model parameters

Parameter	Value
Elasticities	
Travel-time exponent (α_2)	-0.9391
Frequency exponent (α_4)	0.7480
Station Potential Coefficients (β)	
PT accessibility	0.3669
Tourism	0.1168
Airport proximity	-0.1592
Intercept	3.7064
Travel-Time Scaling by Class Pair ($\alpha_1^{i,j}$)	
(Metropolitan, Metropolitan)	0.9141
(Metropolitan, Regional)	0.9086
(Metropolitan, Peripheral)	0.9126
(Regional, Regional)	0.9045
(Regional, Peripheral)	0.8990
(Peripheral, Peripheral)	0.9067
Frequency Scaling by Class Pair ($\alpha_3^{i,j}$)	
(Metropolitan, Metropolitan)	1.6837
(Metropolitan, Regional)	0.8926
(Metropolitan, Peripheral)	1.5304
(Regional, Regional)	0.1416
(Regional, Peripheral)	0.1905
(Peripheral, Peripheral)	0.3927

Line Pool Generation Results

The procedure was applied to the network of 40 nodes (1,360 OD pairs) and 231 infrastructure links, assuming symmetric service (i.e., same routing, stop pattern, and frequency for each line pair serving the same OD terminals but in opposite directions). *Regional connectors* were designated as primary hubs to anchor the inter-regional network, while London and peripheral stations served as secondary nodes.

The service network augmented the physical infrastructure with 44 direct hub-to-hub links. From an initial set of 426 lines by shortest-path algorithm, length and overlap filtering pruned the pool significantly, eliminating 165 lines. A final enrichment step ensures universal one-transfer connectivity. The final pool comprises 273 lines (213 corridor, 19 feeder, 29 local, and 12 enrichment). All lines feature three to seven stops. The average travel time per line is 156 minutes (approximately 312 km), ranging from 19 to 482 minutes. Corridor services are relatively long (averaging 164 minutes), forming the network backbone, while feeders are substantially shorter (averaging 60 minutes).

The spatial distribution of lines reveals a London-centric radial structure. This creates an "hourglass" geometry where cross-country capacity is heavily funnelled through southern hubs, concentrating the highest edge utilization around the capital.

Computational Experiments and Benchmarking

We tested the MIP model using an 8-hour planning horizon with discrete frequencies $\Phi = \{1, \ldots, 24\}$. Core parameters are listed in Table 3. Unserved penalty is set to a relatively large value to discourage unmet demand except when unavoidable. Line costs were modelled as $c_l = 800 + 25t_l$. Ticket prices are calibrated based on official ticketing information.

The model was solved using Gurobi 13 (TimeLimit: 24 hours, MIPGap: 0.01, CPU: 12th Gen Intel Core i7-12700H). It selects 174 selected lines, carrying 75% of the total demand. Average generalized passenger travel time is 109 minutes.

Based on the shortest passenger route identified for each OD pair, there is an average of 3.6 available departures per 8-hour window, with a mean generalized travel time of 192 minutes. The

Table 3: Core numerical parameters

Parameter	Symbol	Value	Unit
Planning horizon	T	480	min
Train capacity	Q_l	600	seats/train
Segment frequency bounds	L_e, U_e	0, 100	trains/period
Transfer weight (Department for Transport, 2025)	κ	2.0	–
Value of time	VoT	0.20	£/min
Unserved penalty	c_k	500	£/passenger

longest line is [Clapham → St Albans → Luton → Newcastle → Edinburgh] taking 482 minutes. In terms of the number of lines, London Blackfriars emerges as both the busiest terminus (52 lines terminate there) and the busiest interchange station (a total of 55 lines pass through it). Outside of London, Milton Keynes is the busiest hub with 42 lines passing through it. Several lines reach the maximum frequency of 24 departures, e.g., [Stansted → Tottenham Hale → Fenchurch Street] and [Stratford → St Albans → Luton → Newcastle → Edinburgh].

Next, we assess the impact of our proposed methodology by investigating two key factors: first, whether hierarchical line-pool generation yields a superior candidate set than the unstructured baseline; and second, how endogenous demand alters the optimal network configuration.

Effectiveness of hierarchical line pool generation We compare our *Hierarchical* pool (273 lines) against a *NoTypology* baseline (610 lines), which follows the same proposed generation procedure, but treats all stations as primary without service-role differentiation.

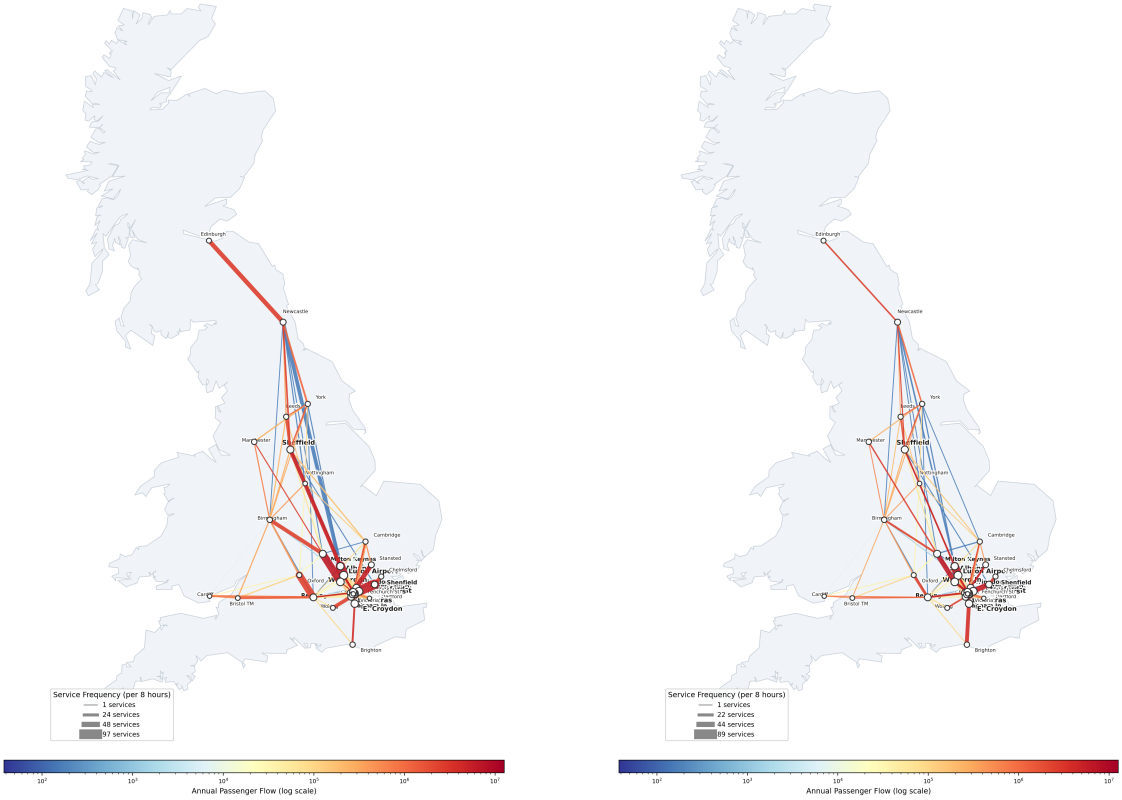
Table 4 reports the results when solving the line planning model using the proposed candidate line pool against the baseline version. Lines selected from the hierarchical pool are considerably fewer than those from the baseline while guaranteeing a sound level of service. The proposed method requires significantly less infrastructure than *NoTypology* and incurs lower operating cost, indicating a more economical network configuration, while improving unit efficiency by up to 9%. By distinguishing service roles, the hierarchical approach creates a more "surgical" candidate set.

Table 4: Line planning results under proposed and baseline line pools

Metric	Hierarchichal	NoTypology
Lines selected	174	425
Infrastructure length (km)	13,332	16,092
Operating cost (£)	6,152,261	9,999,429
Operating cost / Demand (£)	10.2	10.5
Operating cost / Pax-km (£)	0.0604	0.0665

Impact of endogenous demand integration To evaluate the impact of endogenous OD demand modelling, the proposed model was weighed against *Model-fixed* (fixed OD demand) and *Model-constantTransfer* (variable OD demand, but constant transfer times). In the assessment phase, the line concepts from the benchmarks are held fixed, while the OD demands and passenger route assignments are re-optimized to reflect the complete behavioural formulation. Figure 3 contrasts passenger flow and train frequency on service edges for the line concepts of Proposed and *Model-fixed*. The proposed line concept offers more intensive service and induces greater long-distance travel demand (e.g., more flow to/from Edinburgh).

Table 5 shows the proposed model achieves the best overall performance. While *Model-fixed* has the lowest operating cost, it severely underinvests, satisfying only 33% of the potential demand. This highlights the risk of using static OD matrices in long-distance planning. Compared to *Model-constantTransfer* assuming fixed transfer times, our model selects fewer lines and reduces operating costs by 8.5%. Furthermore, the proposed model outperforms *Model-constantTransfer* in service levels, satisfying 75.0% of demand compared to 67.5%, while simultaneously reducing passenger time costs by over £1 million. These results suggest treating transfer time as a variable allows the model to optimize connections more effectively.



(a) Proposed model line concept results.

(b) *Model-fixed* line concept results.

Figure 3: Realized passenger and train flows on edges based on computed line concept

Table 5: Performance of optimal line concepts under different demand formulations

Metric	Proposed	Model-fixed	Model-constantTransfer
Objective value (£)	101,580,968	323,945,037	111,838,936
Lines selected	174	144	182
Operating cost (£)	6,152,261	2,834,575	6,722,601
Passenger time cost (£)	13,189,242	7,225,665	14,253,204
Unserved penalty cost (£)	100,251,107	322,007,038	108,213,447
Ticket revenue (£)	18,220,701	8,122,241	17,350,316
Total demand	802,789	958,592	666,035
Demand satisfied	602,286	314,578	449,608
Served demand percentage	75.0%	32.8%	67.5%
Model solve time (second)	$>$ 86400 (gap 9.1%)	63	17075

4 Conclusions

This study proposes an integrated line planning framework that endogenizes demand and frequency-dependent transfers. It bridges demand modelling and line planning: the hierarchical line generation captures "good" lines, while the supply-demand interaction prevents the "underinvestment trap" typical of fixed-demand benchmarks. Validated on the Great Britain rail network, this market-responsive design simultaneously optimizes service attributes and passenger demand. Future work will address model scalability and generalized demand estimation.

Acknowledgements

This study is supported by the 3MARS project (grant number 101171152), funded by the European Research Council. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Bagloee, S. A., & Ceder, A. A. (2011). Transit-network design methodology for actual-size road networks. *Transportation Research Part B: Methodological*, *45*(10), 1787–1804.

Cipriani, E., Gori, S., & Petrelli, M. (2012). Transit network design: A procedure and an application to a large urban area. *Transportation Research Part C: Emerging Technologies*, *20*(1), 3–14.

de Dios Ortúzar, J., & Willumsen, L. G. (2024). *Modelling transport*. John wiley & sons.

Department for Transport. (2025). *Transport Analysis Guidance (WebTAG) — TAG Data Book*. GOV.UK. Retrieved from `https://www.gov.uk/government/publications/tag-data-book`

Gattermann, P., Harbering, J., & Schöbel, A. (2017). Line pool generation. *Public Transport*, *9*(1), 7–32.

Grolle, J., Donners, B., Annema, J. A., Duinkerken, M., & Cats, O. (2024). Service design and frequency setting for the european high-speed rail network. *Transportation Research Part A: Policy and Practice*, *179*, 103906.

Hartleb, J., Schmidt, M., Huisman, D., & Friedrich, M. (2023). Modeling and solving line planning with mode choice. *Transportation Science*, *57*(2), 336–350.

Lin, H., & Tang, C. (2021). Analysis and optimization of urban public transport lines based on multiobjective adaptive particle swarm optimization. *IEEE Transactions on Intelligent Transportation Systems*, *23*(9), 16786–16798.

Office of Rail and Road. (2024). *Origin and destination matrix (odm) 2023-24*. Rail Data Marketplace. Retrieved 2026-01-01, from `https://raildata.org.uk/dataProduct/P-420268d5-b082-469e-98e4-47d92f714bc6/overview`

Wang, P., Ye, Y., Nießen, N., D'Ariano, A., & Vansteenwegen, P. (2025). Fare system selection and line planning for new suburban rail lines considering the supply–demand interactions. *Transportation Research Part C: Emerging Technologies*, *178*, 105240.