DATATHON 2025

TEAM CLESMOR: LYRA, CARMEN, CALVIN, SHANE, THOMAS

TABLE OF CONTENTS

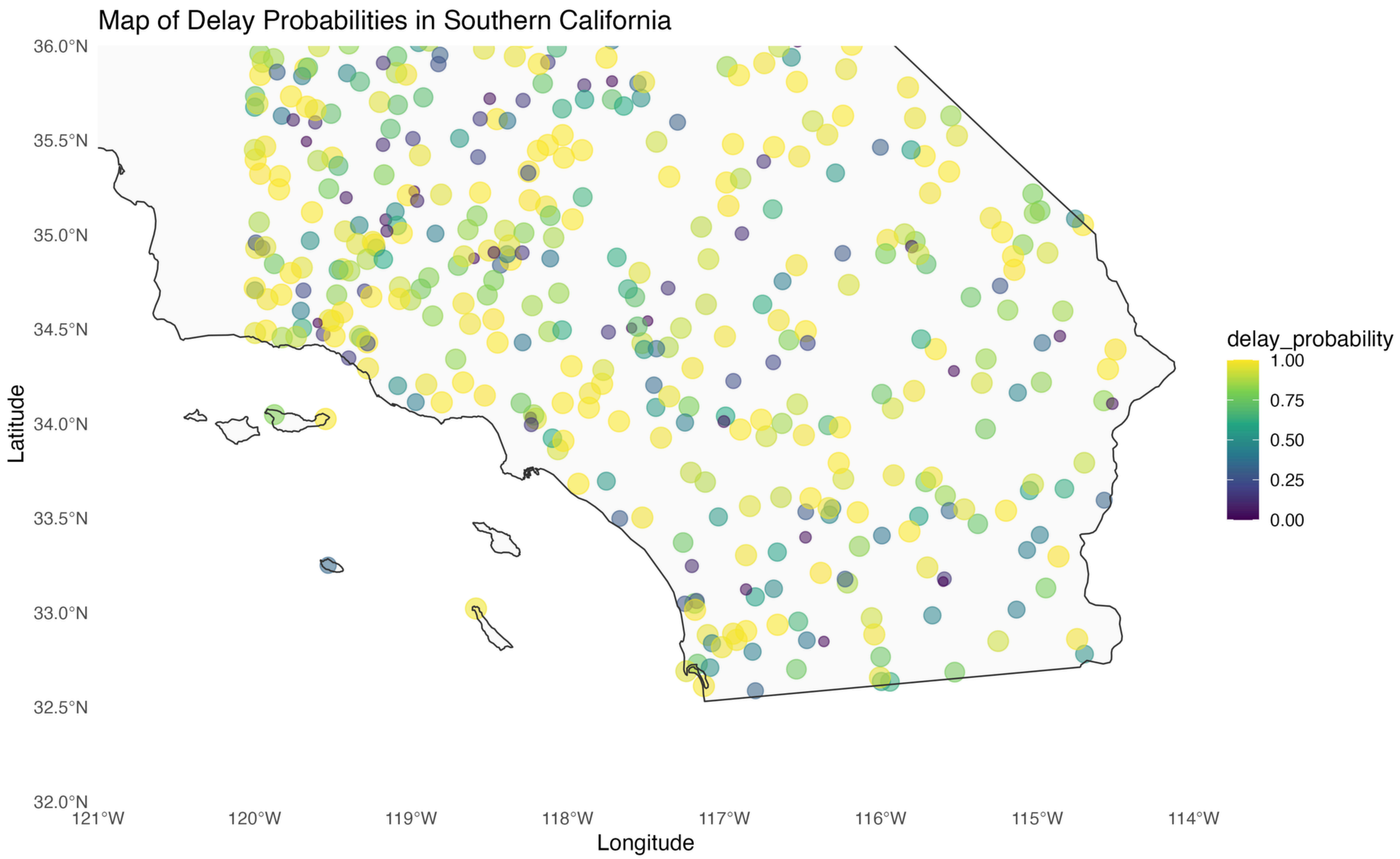
PROBLEM STATEMENT

***"THE CURRENT LOGISTICS NETWORK IS ROUTED SPARSELY;
THIS IS THE KEY DRIVER OF UNOPTIMISED SHIPPING COST, DELIVERY DELAY,
AND CARBON EMISSIONS."***

Logistics network is built on **raw GPS points**, with **NO** structured network design:

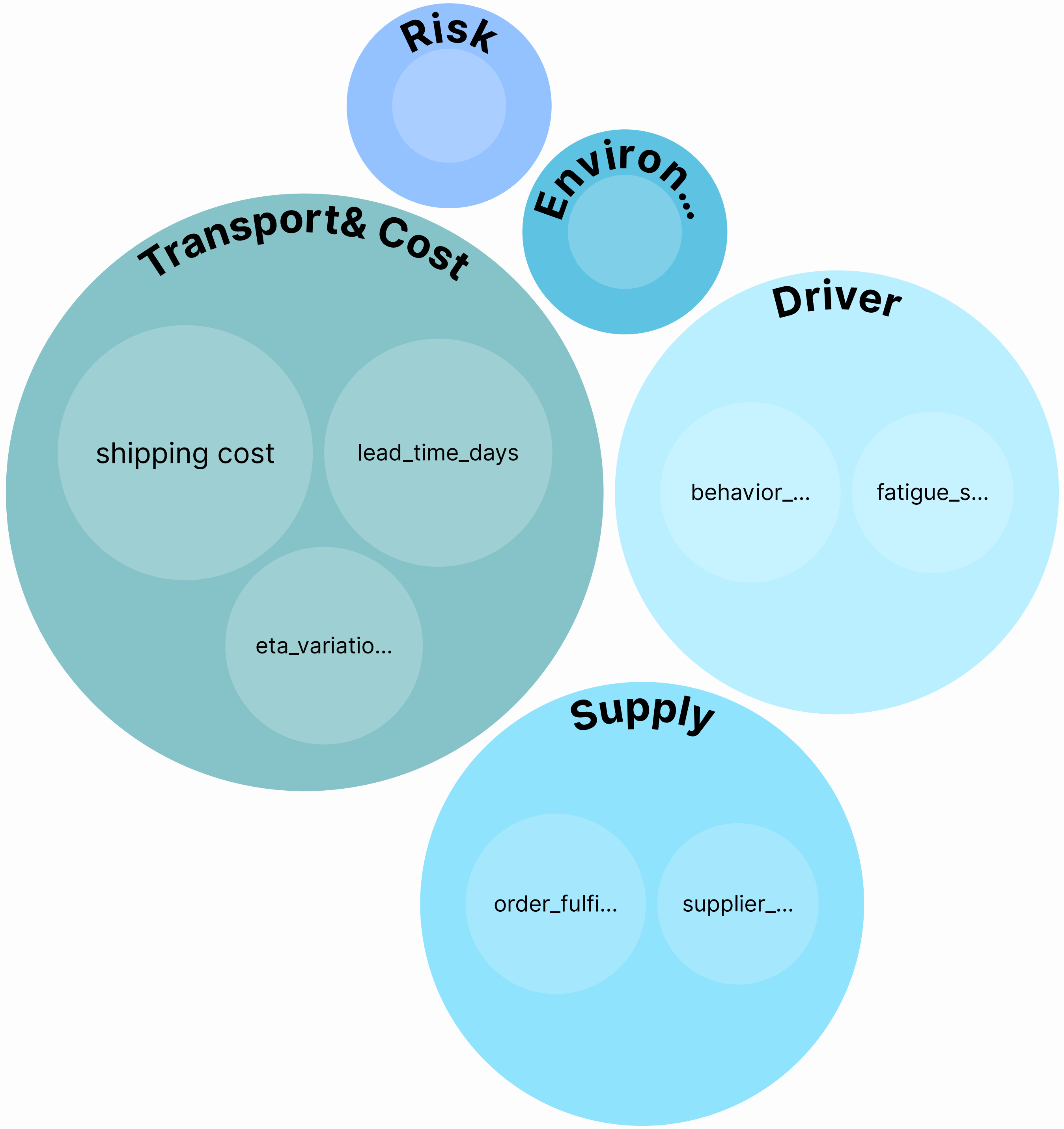
- **Total Shipping Cost**
  \$14.729M
- **Avg. Delivery time Deviation**
  5h 10m
- **Avg. CO_2 emissions**
  0.812M kg

EDA: IDENTIFYING RISK REGIONS

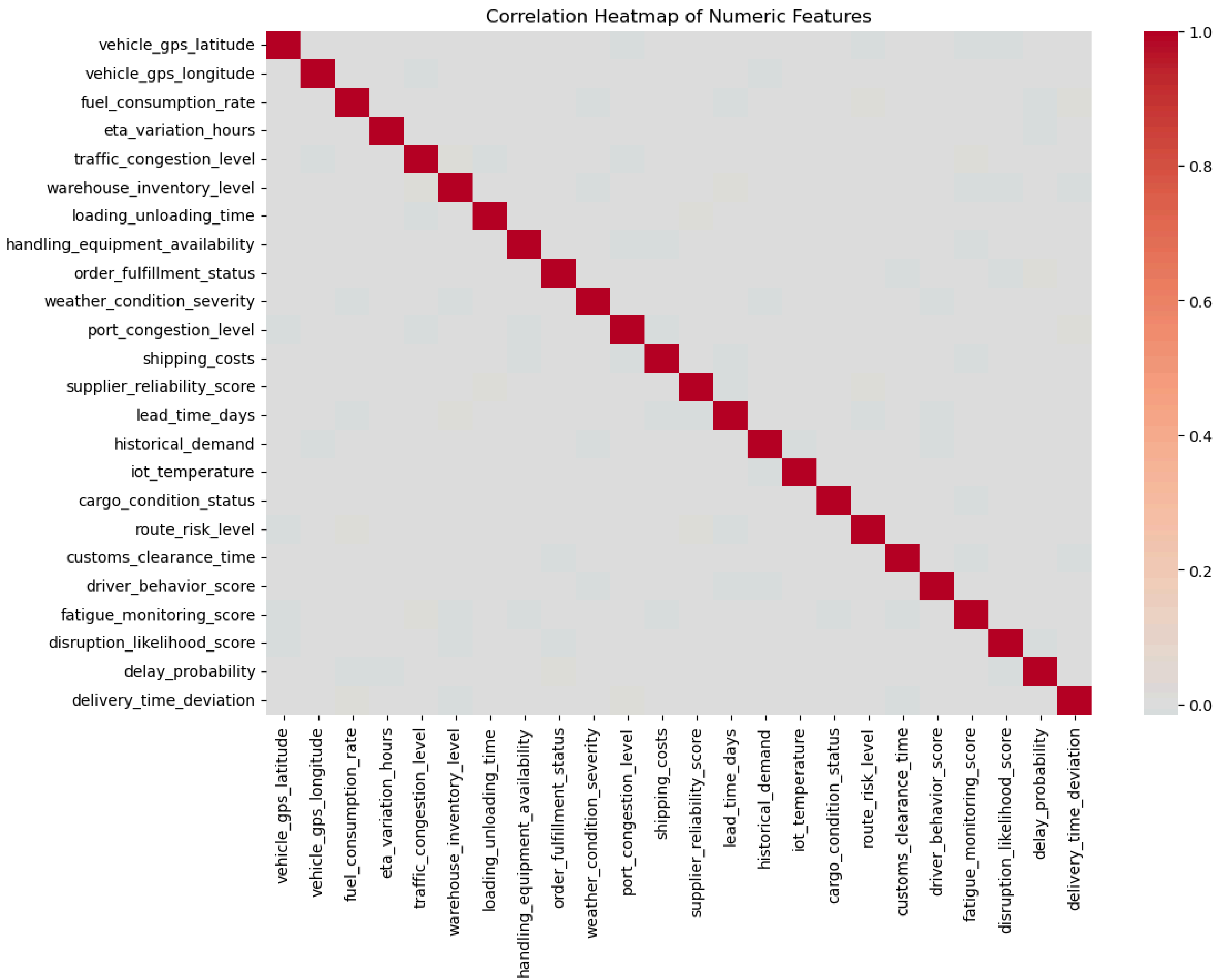


FROM **DATA** TO **INSIGHTS**

- **Large internal dataset accumulated over many years**
- **Covers multiple operational dimensions**
- **Reflects operational efficiency signals**

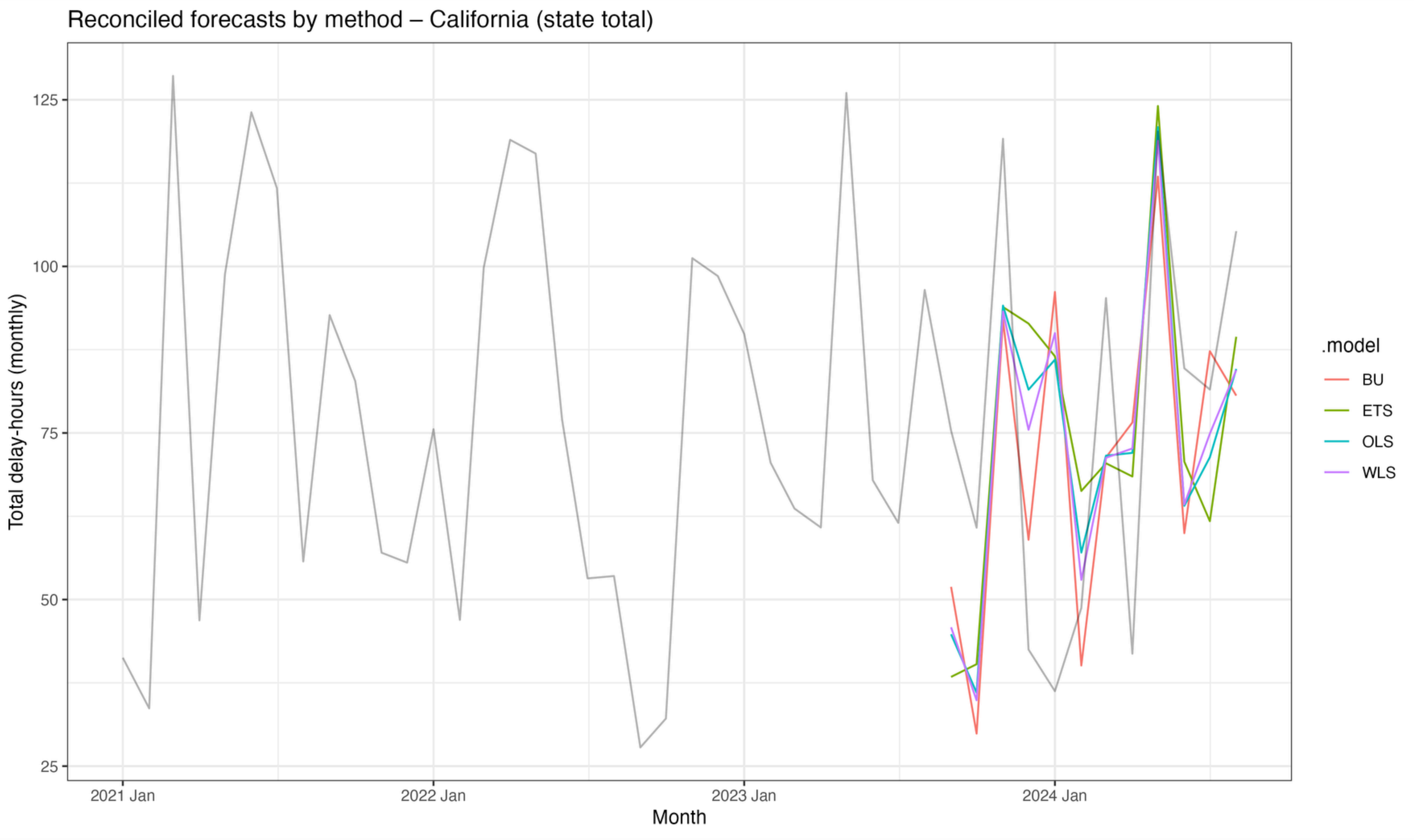


FROM **DATA** TO **INSIGHTS**



- **Low correlations**
- **Singular factors have little impact**
- **Highest correlation is only 0.01**

FROM **DATA** TO **INSIGHTS** - **FORECASTING**



ETS Time Series Model

- **ETS(A, N, A)**
- **Using features as x-regressors**
- **Reconciled with BU, OLS, WLS**

Almost no predictive power

- **Delays mainly caused by unexpected events**
- **ARIMA model not predictive**

MODEL BUILDING

OBJECTIVE SETTING

- Forecasting gives poor results
- Decision-making model
- Leverages existing routes, manpower and resources
- Using optimisation to balance cost, time, and sustainability

MODEL BUILDING

Step 1. Estimate Stable Parameters

- **Use light ML / statistics to get reliable inputs**
- *Examples: cost uplifts (by bucket), delay probability, expected deviation hours, emission factors*

Step 2. Optimization Model (Gurobi)

- **Objective: minimize (cost, delay risk, emissions)**
- *Feed those parameters into an optimization model (Gurobi) that selects routes/allocations*
- **Constraints: capacity, reliability, emissions caps**

Step 3. Business Trade-offs & Decisions

- **Outputs: allocation plans, system-level KPIs**
- *Show trade-offs clearly: cost vs. on-time vs. sustainability*

MODEL BUILDING - DETAIL

Hub Construction

- Input: raw latitude/longitude shipment records.
- Method: **KMeans (K=12)** (picked via elbow/silhouette and runtime trade-off).
- Each record is mapped to the nearest hub; **hub "count"** ≈ demand weight.

Corridor Design

- Each hub connects to its **3 nearest** hubs → realistic connectivity with low complexity.
- Arc parameters (from hub stats): **unit_cost (c_{ij}), expected delay hours (d_{ij}), emissions factor (e_{ij}).**
- We **drop delay_probability** and use **expected delay hours** to avoid conflicting targets.

Optimization Model

$$\boxed{\min \sum_{(i,j)} \left(c_{ij} + \lambda_{\text{sla}}\, d_{ij} + \lambda_{\text{co2}}\, e_{ij}\right) x_{ij}}$$

Subject to:

1. **Supply** at source: $\sum_j x_{sj} = \text{TOTAL\_SUPPLY}$
2. **Demand** at each destination k: $\sum_u x_{uk} - \sum_v x_{kv} = \text{demand}_k$
3. **Transshipment**: inflow = outflow for intermediate hubs
4. **Capacity**: $x_{ij} \le \text{cap}_{ij}$ (set high unless constrained)

Assumptions

- *Total supply = 10,000 units.*
- *Demand is allocated by hub weights from clustering.*
- *Road corridors only;*
- *Fuel → CO_2 proxy = fuel × 3.16 kg CO_2e/litre.*
- *etc (find Appendix)*

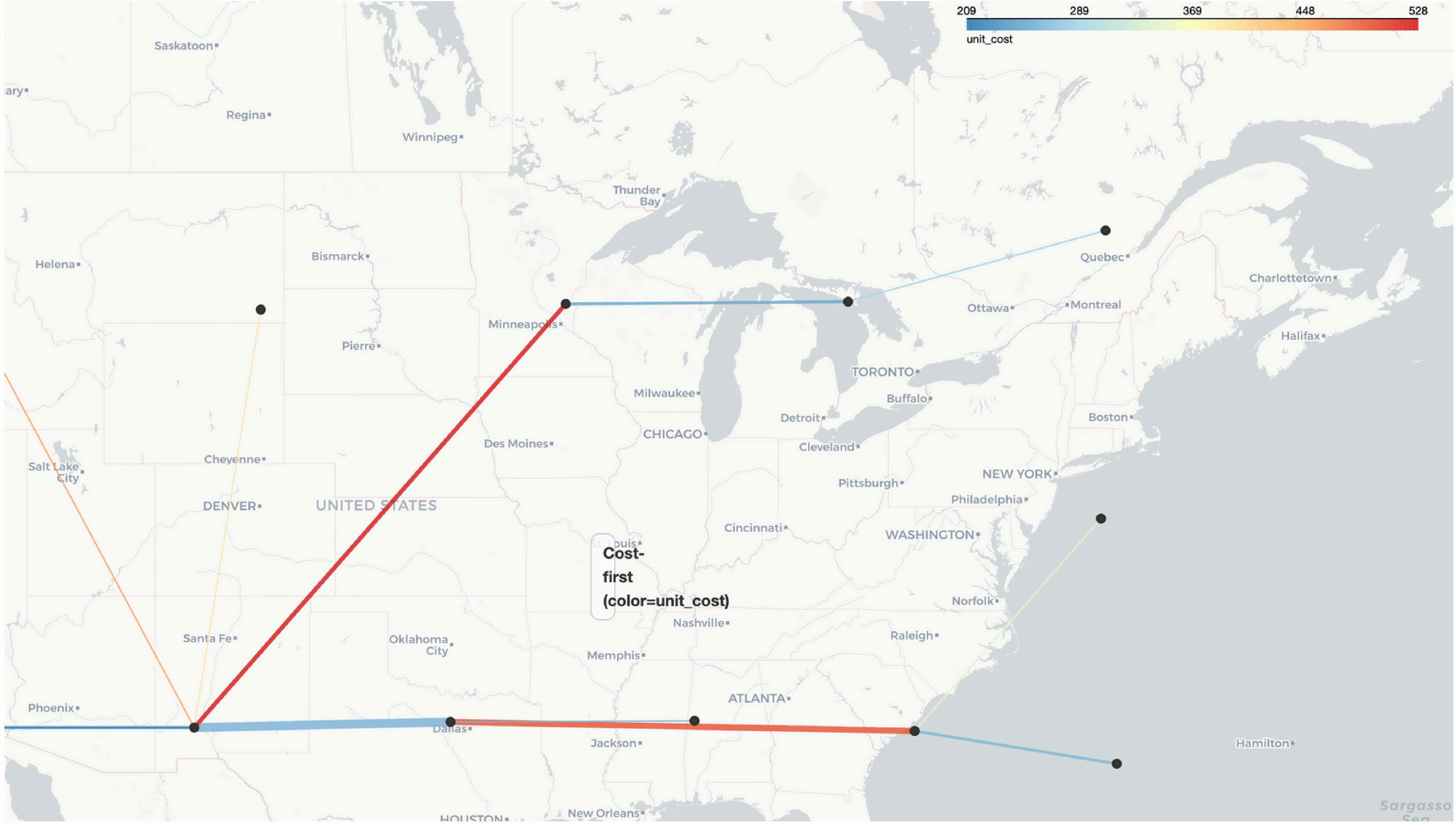
We simulate three policies by adjusting λ:

- **Cost-first**:
  $\lambda_{\text{sla}} = 0,\ \lambda_{\text{co2}} = 0$
- **On-time-first**:
  $\lambda_{\text{sla}} = 50,\ \lambda_{\text{co2}} = 0$
- **Low-carbon-first**:
  $\lambda_{\text{sla}} = 0,\ \lambda_{\text{co2}} = 10$

On-time-first: pay +$11.3k to save −893 h and −11.6k CO_2 proxy units.
Low-carbon-first: pay +$56.9k to save −16.5k CO_2 proxy units, similar SLA.

Strategy	Total Cost	vs Cost-first	Deviation Hours	vs Cost-first	CO_2 Proxy	vs Cost-first
Cost-first	$6,751,660	—	109,354 h	—	6,933,047	—
On-time-first	$6,763,006	**0.17%**	**108,461 h**	**−0.82%**	**6,921,415**	**−0.17%**
Low-carbon-first	$6,808,575	0.84%	109,084 h	−0.25%	**6,916,506**	**−0.24%**

IMPLEMENTATION & EXECUTION



1. Cost-first (cost prioritisation)

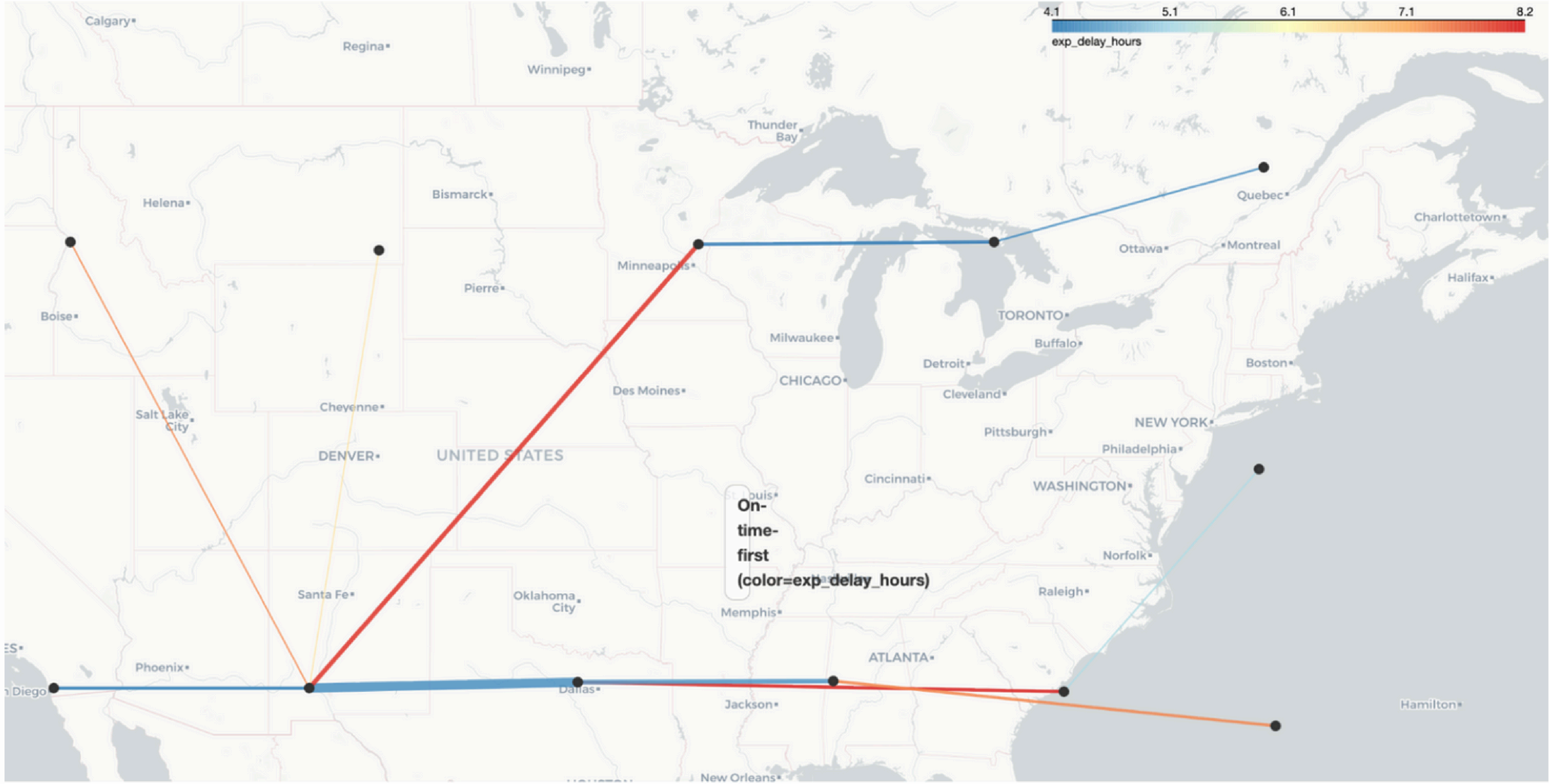
- Flows concentrate on the **southern trunk**, minimizing cost but sacrificing reliability and carbon.
- Most volume runs on a single long, low-cost corridor; expensive diagonal links are rarely used.
- **Business meaning:** best for budget, but higher risk of delays and CO_2.

IMPLEMENTATION & EXECUTION

2. Low-carbon-first (emission prioritisation)

- Flows shift off **carbon-heavy corridors**, **paying a small cost premium for lower CO_2** without hurting reliability.
- Traffic is spread across multiple low-emission paths instead of one polluting trunk.
- **Trade-off: ~+0.8% cost for ~-0.24% CO_2, delay rate stable.**
- **Business meaning:** best when environmental targets or carbon budgets are binding.

IMPLEMENTATION & EXECUTION



3. On-time-first (reliability prioritisation) ⭐

- **A slight cost increase** brings **fewer delays** and **a small CO_2 improvement—often the most balanced choice.**
- Flows shift from the southern trunk to more reliable connectors; unreliable arcs shrink or disappear.
- **Trade-off: ~+0.17% cost, −0.82% delay hours, ~−0.17% CO_2.**
- **Business meaning:** balanced option, best for customer experience .

COMPARISONS TO THE BASELINE

Strategy	Total Cost	vs Cost-first	Deviation Hours	vs Cost-first	CO_2 Proxy	vs Cost-first
Cost-first	$6,751,660	—	109,354 h	—	6,933,047	—
On-time-first	$6,763,006	**0.17%**	**108,461 h**	**−0.82%**	**6,921,415**	**−0.17%**
Low-carbon-first	$6,808,575	0.84%	109,084 h	−0.25%	**6,916,506**	**−0.24%**

Baseline Statistics

Total Shipping Cost
$14.729M

Avg. Delivery time Deviation
5h 10m

Avg. CO_2 emissions
0.812M kg

On-time-first Model Statistics

Total Shipping Cost
$6.763M (54% decrease)

Avg. Delivery time Deviation
3h 22m (35% decrease)

Avg. CO_2 emissions
0.692M kg (15% decrease)

BUSINESS IMPACT ⭐

In the current network, the three strategies yield modest differences, showing a relatively balanced system:

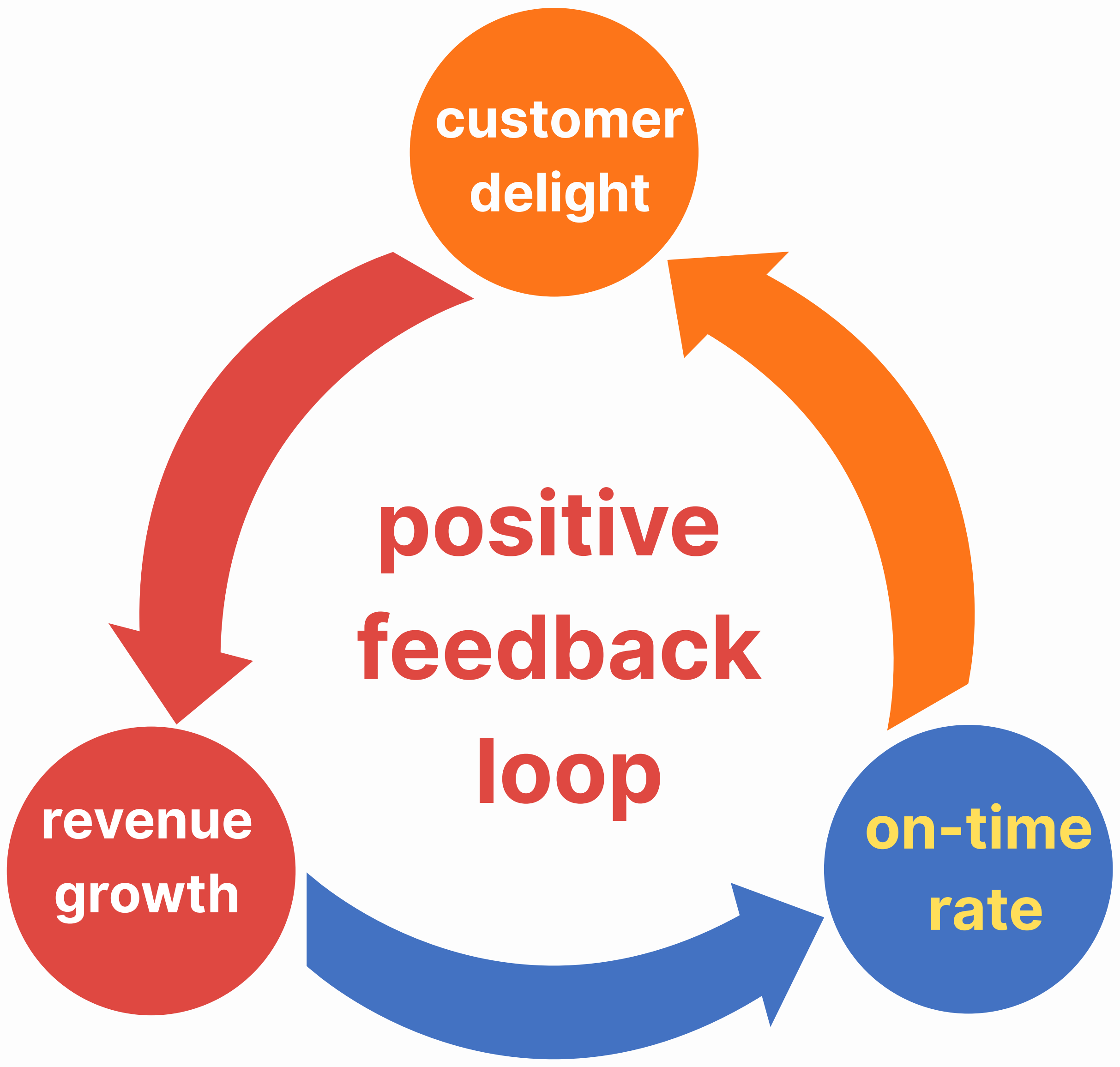
- **Cost-first** minimises expense but risks SLA and emissions.
- **Low-carbon-first** fits when environmental targets are binding.
- **On-time-first**⭐ offers the most balanced trade-off and improves customer experience.

Given revenue pressures , we recommend implementing the on-time-first strategy. This will improve the delivery rate, thus enhancing *customer loyalty* and drive *revenue growth*, offsetting short-term cost increases and creating a positive cycle for the company long-term growth.

Cost vertex → lower total cost

Deviation (On-time) vertex → lower total deviation hours

CO_2 vertex → lower total emissions proxy



THANK YOU

QUESTIONS?