ACCSCIENCE PUBLISHING

Asian Journal of Water, Environment and Pollution

ORIGINAL RESEARCH ARTICLE

Big data-driven carbon footprint tracking and game-theoretic incentive design in green logistics supply chains

Supplementary file

Table S1. Variable definitions and data sources (variable dictionary)

Variable (symbol)	Level/stream	Definition/construction	Unit (original) / scale used in modeling	Primary source
y	Micro (TMS+OBD/GPS)	Carbon emission intensity computed from activity data and calibrated/predicted emission factor; used as the prediction target	kg CO_2 per ton km; reported as standardized *z*-score in modeling	Enterprise TMS/OBD/GPS + model-calibrated emission factor (Section 3.1)
Load	Micro (TMS)	Load factor/utilization proxy reflecting payload relative to capacity for a shipment/route	Ratio (0–1) or %; standardized z-score in modeling	Enterprise TMS orders (internal)
Speed	Micro (OBD/GPS)	Average operating speed over the trip/route segment	km/h; standardized z-score in modeling	Vehicle OBD/GPS records (internal)
Distance	Micro (TMS)	Transport distance for the shipment/route	km; standardized z-score in modeling	Enterprise TMS / routing records (internal)
StopStart	Micro (OBD/GPS)	Stop-and-go intensity proxy (e.g., stop-start events or idling/stop frequency)	Index/count rate; standardized z-score in modeling	Vehicle OBD/GPS records (internal)
ColdChain	Micro (TMS)	Cold-chain service indicator/business-line flag (1=cold-chain; 0=otherwise)	Binary; encoded and used in modeling	Enterprise TMS orders (internal)
FuelPrice	Macro/exogenous	Fuel price level/index used as an exogenous driver and scenario input	RMB/L or price index (as provided); standardized when used in ML	NBS/National Data and transport statistics portals (Section 4.1)
Pandemic	Macro/exogenous	Pandemic regime indicator for COVID-19 disruption periods used for interaction checks	Binary regime indicator	Constructed from calendar/time period definition (Section 4.1–4.2)
S	Policy lever	Normalized subsidy intensity used in response surface simulations and evolutionary game scenarios	Normalized to [0,1]	Policy scenario design (Section 3.2.1)
P	Policy lever	Normalized policy stringency/penalty intensity used in response surface simulations and evolutionary game scenarios	Normalized to [0,1]	Policy scenario design (Section 3.2.1)
c_i	Network	Node importance/centrality multiplier for targeted regulation on high-centrality nodes (super-nodes).	Centrality score (normalized where required)	Constructed from a carbon-flow linkage network (Section 3.1)

Note: For confidentiality and comparability, micro-level variables are summarized and modeled using standardized (*z*-score) values.
Abbreviations: GPS: Global positioning system; ML: Machine learning; NBS: National Bureau of Statistics; OBD: On-board diagnostics; TMS: Transport management system.

Table S2. Evolutionary game parameter settings and simulation details

Parameter	Symbol	Baseline value	Range (sensitivity)	Calibration/source
Subsidy intensity (normalized)	S	0.50	$S \in [0,1]$	Policy scenario design; normalized from the unit subsidy coefficient in Equation 2
Penalty intensity (normalized)	P	0.50	$P \in [0,1]$	Policy scenario design; normalized from unit penalty coefficient in Equation 3
Initial government strategy share	$x0$	0.30	[0,1]	Set to reflect baseline policy strictness
Initial enterprise low-carbon share	$y0$	0.10	[0,1]	Scenario design to reflect low initial adoption; sensitivity tested
Initial supplier compliance share	$z0$	0.10	[0,1]	Scenario design to reflect low initial compliance; sensitivity tested
Baseline emission intensity	y_base	1.00 (unit-normalized)	[0.5, 1.5] (normalized)	Normalized baseline from emission model outputs; robustness over scaling range
Emission reduction effectiveness	η	0.20	[0.05, 0.50]	Scenario-based effectiveness; calibrated within plausible ranges; sensitivity tested
Abatement cost parameter	c	0.30	[0.10, 0.80]	Scenario-based cost parameter; calibrated within plausible ranges; sensitivity tested

Note: We evaluated (S, P) on a uniform grid $S, P \in \{0.0, 0.1, \ldots, 1.0\}$. For each grid point, replicator dynamics were simulated for $T = 5{,}000$ steps ($\Delta t = 0.01$). Convergence was declared when $\max\{|dx/dt|, |dy/dt|, |dz/dt|\} < 1e{-}6$ for 200 consecutive steps. Heterogeneous initial conditions reflect low initial penetration: $x0 \in \{0.1, 0.3\}$, $y0 \in \{0.05, 0.2\}$, $z0 \in \{0.05, 0.2\}$.

Table S3. Sample construction and preprocessing rules

Step	Operation	Applied to	Details/rationale
Data sourcing	Collect and align multi-stream data	Macro + enterprise micro data	Macro statistics compiled at province-year/city-year level (2018–2023). Micro operational data included TMS orders/route records and OBD/GPS traces from participating logistics enterprises
Unit harmonization	Standardize measurement units	All streams	Harmonize metric units (e.g., distance, weight, time) to ensure consistent feature construction and comparability across years/regions
Record linkage (matching)	Link TMS orders to OBD/GPS traces	Micro data	Match by vehicle/route identifiers and time windows; map each trip to origin–destination regions for network aggregation
Deduplication & plausibility filters	Remove duplicates and invalid records	Micro data	Drop duplicate orders/trips; remove records with missing key fields and physically implausible values (e.g., non-positive distance/time)
Missing data treatment	Imputation after cleaning	Micro + exogenous variables	Impute missing values using conservative rules (e.g., within-route medians or time-aware imputation) to avoid leakage; record the missing rate after cleaning (summarized in Table 5 of the main text)
Outlier handling	Detect and cap/flag extremes	Key continuous variables	Identify outliers via rule-based and distribution-based checks; cap extreme values where appropriate to stabilize model fitting and interpretation.
Winsorization (robustness)	Tail capping sensitivity variant	Heavy-tailed variables	As an additional robustness option, apply symmetric tail capping (winsorization) and confirm that key conclusions are stable (reported in Table S5)
Temporal alignment	Lag expansion and normalization	Time-series variables	Construct lagged features and normalize time-series inputs using information available up to the forecast origin to prevent look-ahead bias
Privacy protection	De-identification and aggregation	Micro data	Micro records are processed in de-identified form and reported at aggregated/standardized levels to protect confidentiality while enabling external-validity assessment

Abbreviations: GPS: Global positioning system; OBD: On-board diagnostics; TMS: Transport management system.

Table S4. Feature engineering procedures (standardization, lagging, aggregation, encoding)

Feature type	Procedure	Examples	Implementation notes
Numeric features	Standardization (z-score)	Load, Speed, Distance, StopStart, FuelPrice	StandardScaler applied within the modeling pipeline; statistics computed on the training window to prevent leakage
Categorical features	One-hot encoding	Business line, energy route, cold-chain flag	Categorical variables encoded to enable nonlinear models and linear baselines
Temporal features	Lag expansion	Lagged fuel price/demand proxies	Lagged features were constructed using information available up to the forecast origin
Aggregation	Trip/order aggregation to region-year for network construction	Edge weight as aggregated carbon flow	Shipment-level emissions aggregated to origin–destination corridors to build annual directed networks (2018–2023)
Normalization (network)	Row-normalization for transition-based measures	PageRank on a directed network	Normalized adjacency used for random-walk measures; robustness checked with alternative normalizations

Note: All preprocessing steps follow the unified pipeline described in the main text (Sections 4.1–4.2).

Table S5. Additional robustness checks and supplementary analyses

Check / analysis	Design	Outcome/reporting	Purpose
Out-of-time split	Train 2018–2021; validate 2022; test 2023	Main text Table 3; split dates documented here	Leakage-free temporal evaluation
Rolling-window backtesting	36-month train, 6-month validation, 1-month test step	Distribution of errors across time (summary)	Temporal stability under drifting environments
COVID-19 structural break	Pandemic regime indicator and interactions (Jan 2020–Dec 2021/2020–2022)	Pre-/during-/post-COVID performance comparison	Shock robustness and generalization after disruption
Corrected significance testing	Nadeau–Bengio-corrected paired *t*-test on fold-level errors	Main text Table 4	Avoid inflated significance under temporal dependence
Seed stability	Repeat stochastic learner training under 10 seeds (0–9)	Stability summary (mean ± SD) and rank invariance	Confirm that results are not driven by a single random initialization
Outlier/winsor sensitivity	Tail-capping variant on heavy-tailed inputs	Qualitative stability of key effects	Robustness to extreme observations
Network concentration validation	Bootstrap CI (B = 2,000) + degree-preserving null model (R = 1,000)	Observed share, CI, null distribution, p-value	Validate “super-node” concentration beyond chance
Policy interval reporting	Effective P interval (P_min, P_max) accompanying Figure 9	Supplementary reporting of interval endpoints	Policy calibration and guardrails
Top 10% node contribution share (weighted degree)	0.58		

Note: Numeric results are already presented in the main text; this table documents the design choices, reproducibility settings, and the location of each robustness result.
Abbreviations: CI: Confidence interval; *P*: Penalty; SD: Standard deviation.

Table S6. Model hyperparameters and training settings (seeds, epochs/iterations, early stopping, grid)

Item	Setting/selected value	Applies to	Notes
Temporal split	Train: 2018–2021; validation: 2022; test: 2023	All models	Out-of-time evaluation to avoid temporal leakage (main text Section 4.2)
Time-series CV	Expanding-window CV, K = 5 folds; horizon H = 1; gap G = 1	Hyperparameter selection	Validation RMSE used for selection; fold-level errors used for corrected significance tests
Random seeds	10 seeds (0–9) for stochastic learners; deterministic learners unchanged	Tree ensembles / stochastic learners	Main tables report median across seeds; mean ± SD stability can be reported
Stopping criterion	Iteration-based with validation monitoring; avoid overfitting	GBDT/LightGBM	Early stopping used the 2022 validation year; specific patience was configured in the training pipeline
Search approach	Grid search guided by empirical intervals and prior knowledge	All models	Search ranges documented in evaluation code

Note: RMSE is reported in kg CO_2 per ton km (original unit). Abbreviations: CV: Cross-validation; GBDT: Gradient-boosted decision trees; GBM: Gradient boosting machine; RMSE: Root mean square error; SD: Standard deviation; SVR: Support vector regression.

Table S7. Network construction and centrality computation details

Component	Definition	Formula/parameterization	Notes
Nodes	Regions (province/city) or aggregated enterprise clusters	Node i represents an origin/destination region in the logistics OD system	Analysis conducted on annual networks (2018–2023) and a pooled network for robustness
Edges	Directed OD corridors carrying carbon flows	Edge $i \to j$ exists if shipments from i to j occur within the time window t	Directionality retained for corridor analysis
Shipment emissions	Shipment-level CO_2 emissions	Emissions_n = Q_n (tons) × D_n (km) × EF_n (kg CO_2 per ton km)	EF_n is predicted/calibrated by the machine learning model using operational features (Section 3.1)
Edge weight (linkage strength)	Aggregated carbon flow on $i \to j$	W_ij, t = Σ_n∈($i \to j$,t) Emissions_n	Used to form a weighted-directed adjacency matrix W_t
Normalization	Row-normalized adjacency for transition-based metrics	Ŵ_ij = W_ij / Σ_j W_ij (row-normalization)	Required for PageRank/random-walk measures; robustness also checked with alternative normalizations
Centrality families	Multiple notions of node importance	Weighted degree, betweenness, eigenvector, Katz, PageRank, collective influence	Robustness assessed via Spearman rank correlations and top k overlap (e.g., Jaccard)
Community detection	Modularity-based clustering	Louvain method on weighted graph; 4–5 stable communities	Used to interpret corridor groupings and regional structures.

Note: For PageRank and other transition-based measures, the normalized network is used.
Abbreviation: OD: Origin–destination.