Preprints of the 23rd IFAC World Congress
Busan, Republic of Korea, August 23-28, 2026

Unsupervised Driving Regime Discovery using Interpretable TCN Transformer Autoencoders with Spectral Clustering

Amirreza Yasami * **Mohamadali Tofigh** * **Mahdi Shahbakhti** *
Charles Robert Koch *

* *Mechanical Engineering Department, University of Alberta, AB, CA (e-mail: ayasami, tofigh, mahdi, ckoch@ualberta.ca).*

Abstract: Understanding and quantifying driving behavior is essential for enhancing fuel efficiency, reducing emissions, and improving fleet-level operational efficiency. This paper presents an interpretable deep learning framework for unsupervised driving regime discovery, the Temporal Convolutional Transformer with Spectral Clustering (TCTSC). The model integrates a Temporal Convolutional Network (TCN) for short-term vehicle dynamics, a Transformer encoder for long-range temporal dependencies, and spectral regularization with a differentiable K-Means module for compact and balanced clustering. The framework is trained and validated on large-scale Controller Area Network (CAN) data from Edmonton Transit Service diesel buses operating under diverse traffic and weather conditions. TCTSC achieves over 75% lower reconstruction errors than DenseAE and more than 85% lower than PCA, preserving both steady-state and transient behaviors. Outcome-grounded evaluation using fuel consumption shows that TCTSC explains 59.7% of the variance (R^2), achieves a Spearman correlation of 0.55, and attains the lowest MAE and RMSE among all models. The identified regimes, idling, deceleration, efficient cruising, high-load cruising, and aggressive acceleration, provide interpretable, physically meaningful insights into real-world bus operations and a robust foundation for data-driven eco-driving analysis and driver assessment.

Keywords: Artificial Intelligence in Transportation, Convolutional Networks, Transformer, Unsupervised Learning.

1. INTRODUCTION

The transportation sector is a major contributor to greenhouse gas (GHG) emissions and urban air pollution. In 2024 it accounted for approximately 8416 Mt of CO_2-equivalent emissions, or 15.8% of global GHG output (Crippa et al., 2025). Canadian transportation emissions have grown 36% since 1990 to 195 Mt CO_2 eq in 2023 ($\sim$28% of national GHG), and Canadian road transport contributed 38% of NO_x and 54% of CO emissions in 2023 (Environment and Climate Change Canada, 2025). While advances in powertrain and aftertreatment technology have helped, real-world variability driven by *driver behavior* remains a dominant and under-exploited source of efficiency gains: eco-driving strategies have been shown to reduce fuel use by 5-10% and GHG emissions by up to 10% (Martin et al., 2012), and driver-assistance and eco-feedback systems offer further safety and environmental benefits (Wu et al., 2017).

Characterizing driving styles from in-vehicle data has therefore become an active research area. Early studies relied on physics-based models and classical machine learning such as SVMs, HMMs, and decision trees (Liang et al., 2007; Hallac et al., 2016; Miyajima et al., 2007). Deep learning is now dominant, with convolutional networks Tofigh et al. (2024), autoencoders Fugiglando et al. (2018), and multimodal frameworks Choi et al. (2016) delivering strong representation quality. More recently, Temporal Convolutional Networks (TCNs) within autoencoders capture long-range temporal patterns with stable, causal training (Thornton et al., 2025), while Transformer-based autoencoders use self-attention for global context (Li et al., 2025; Yasami et al., 2025; Zerveas and et al., 2021). Hybrid TCN-Transformer architectures show further improvements in time-series and behavioral modeling (Wang and Chen, 2025; Isreal and Alonge, 2025; Yasami et al., 2026).

Despite these advances, existing studies are limited by (i) reliance on simulator or proprietary, small-scale, labeled data; (ii) weak coupling between discovered behavior types and operational outcomes such as fuel use; and (iii) limited interpretability in unsupervised settings. To address these gaps, we propose the *Temporal Convolutional Transformer with Spectral Clustering* (TCTSC), an unsupervised, interpretable framework for driving regime discovery from high-resolution (1 Hz) real-world Controller Area Network (CAN) data. The key contributions are: (i) joint integration of convolutional and attention-based temporal modeling; (ii) spectral regularization combined with a differentiable K-Means head, which yields compact and balanced clusters that align with operational outcomes; and (iii) SHAP-based feature attribution to map each cluster to physically meaningful CAN signals. The framework

©2026 the authors. Accepted by IFAC for publication under a Creative Commons Licence CC-BY-NC-ND

provides a unified basis for objective driver assessment under diverse traffic, terrain, and weather conditions.

2. EXPERIMENTS AND DATA COLLECTION

This study uses a large real-world dataset from Edmonton Transit Service (ETS) diesel buses operating on Route 9, a high-frequency urban corridor in Edmonton, Canada. Data span the 2024 calendar year, covering over 120 full service days (∼9 h per day) under diverse traffic, roadway, and seasonal weather conditions. The test platform is a 2020 model-year, EPA 2020-compliant diesel transit bus (GVWR 19,724 kg; 9 L inline-6 Cummins engine). Full instrumentation, calibration, and validation details are provided in our prior work (Yasami et al., 2025).

3. METHODOLOGY: TCT WITH SPECTRAL CLUSTERING

The proposed TCTSC framework learns compact, discriminative latent representations from multivariate vehicular time-series data. A TCN encoder captures fine-grained local dynamics; a Transformer encoder models long-range contextual relationships; a spectral regularizer aligns the latent space with the dominant eigenstructure of the affinity graph; and a differentiable K-Means head produces compact, balanced clusters. SHAP attribution then links each cluster to physically meaningful CAN signals. The complete workflow is summarized in Fig. 1.

3.1 Inputs and Notation

Let $\mathbf{X} \in \mathbb{R}^{B\times F\times W}$ denote a batch of input driving windows (B batch size, F features, W window length) which is denated as input data in Fig. 1. Each sample is a short driving segment of CAN signals. Although the CAN database contains many signals, this study uses a curated subset of $F = 8$ physically meaningful features capturing vehicle kinematics, powertrain state, and load conditions (Table 1). The objective is to encode $\mathbf{X}$ into a low-dimensional latent representation $\mathbf{Z} \in \mathbb{R}^{B\times D}$ that captures meaningful temporal dynamics and supports the discovery of distinct driving regimes.

Table 1. CAN-based input features used in the proposed framework.

Feature	**Physical meaning**
Vehicle speed	Longitudinal velocity (m/s)
Longitudinal acceleration	Time derivative of speed (m/s^2)
Engine percent load	Normalized engine load at current speed (%)
Engine speed	Crankshaft rotational speed (RPM)
Engine fuel rate	Instantaneous fuel consumption rate (g/s)
Intake manifold pressure	Intake air pressure, load proxy (kPa)
Aftertreatment outlet NO_x	NO_x downstream of aftertreatment (ppm)
Vehicle specific power	Power demand per unit mass (kW/ton)

3.2 Temporal Convolutional Encoder

The TCN operates entirely through one-dimensional dilated causal convolutions, ensuring that each output at time t depends only on current and past inputs and enabling efficient parallel training over long horizons. Each layer is composed of residual blocks; for the l-th layer and filter m, the activation at time t is Yasami et al. (2025):

$$Z_m^{[l]}(t) = \sum_{j=1}^{C^{[l-1]}} \sum_{p=0}^{K-1} w_{j,m}^{[l]}(p)\, Y_j^{[l-1]}(t - p\, d_l) + b_m^{[l]}, \quad (1)$$

with kernel size K and dilation d_l. A ReLU nonlinearity, normalization, dropout, and a skip connection produce the temporally rich embedding $\mathbf{H} = \mathrm{TCN}(\mathbf{X}) \in \mathbb{R}^{B\times F\times W}$.

3.3 Transformer Temporal Abstraction

Long-term dependencies are captured by a Transformer encoder (Vaswani et al., 2017). The TCN output is reshaped to sequence-major form $\mathbf{H}' \in \mathbb{R}^{W\times B\times F}$ and augmented with sinusoidal positional encodings:

$$\mathrm{PE}_{(t,2i)} = \sin\Big(\tfrac{t}{10000^{2i/d}}\Big), \quad \mathrm{PE}_{(t,2i+1)} = \cos\Big(\tfrac{t}{10000^{2i/d}}\Big). \quad (2)$$

Multi-head self-attention then relates each time step to all others:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\Big(\tfrac{QK^\top}{\sqrt{d_k}}\Big)\, V, \quad (3)$$

where Q, K, V are linear projections of $\mathbf{H}'$. The Transformer output $\mathbf{M} \in \mathbb{R}^{W\times B\times F}$ is mean-pooled over time:

$$\mathbf{z}_b = \frac{1}{W}\sum_{t=1}^{W} \mathbf{M}_{t,b}, \quad b = 1, \ldots, B, \quad (4)$$

yielding compact window-level embeddings $\mathbf{z}_b \in \mathbb{R}^D$.

3.4 Spectral Regularization

The latent embeddings $\mathbf{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_B]^\top$ are regularized to align with K orthogonal directions corresponding to the dominant cluster structure. With Gram matrix $\mathbf{G} = \mathbf{Z}\mathbf{Z}^\top$, the spectral objective seeks an orthonormal $\mathbf{F} \in \mathbb{R}^{B\times K}$ ($\mathbf{F}^\top\mathbf{F} = \mathbf{I}_K$) that maximizes the projected variance (Ng et al., 2001):

$$\max_{\mathbf{F}} \ \mathrm{tr}\big(\mathbf{F}^\top \mathbf{G}\mathbf{F}\big), \quad (5)$$

solved via a truncated SVD of $\mathbf{Z}$ at each iteration. The normalized spectral loss is

$$\mathcal{L}_{\mathrm{spec}}(\mathbf{Z}, \mathbf{F}) = \lambda_s \frac{\|\mathbf{Z}\|_F^2 - \|\mathbf{Z}^\top\mathbf{F}\|_F^2}{\|\mathbf{Z}\mathbf{Z}^\top\|_F + \varepsilon}, \quad (6)$$

with weighting λ_s and stabilizer ε. Minimizing $\mathcal{L}_{\mathrm{spec}}$ encourages well-separated and cluster-consistent latent embeddings.

3.5 Differentiable K-Means

The latent space is jointly refined by a differentiable K-Means head with trainable centroids $\{\mu_k\}_{k=1}^K$ and soft assignments (Bauckhage, 2015):

$$p_{ik} = \frac{\exp(-\|z_i - \mu_k\|^2/\tau_{\mathrm{km}})}{\sum_{r=1}^{K}\exp(-\|z_i - \mu_r\|^2/\tau_{\mathrm{km}})}, \quad (7)$$

where τ_{km} controls assignment sharpness. The compactness loss is

$$\mathcal{L}_{\text{km-exp}} = \frac{1}{B}\sum_{i=1}^{B}\sum_{k=1}^{K} p_{ik}\, \|z_i - \mu_k\|^2. \quad (8)$$

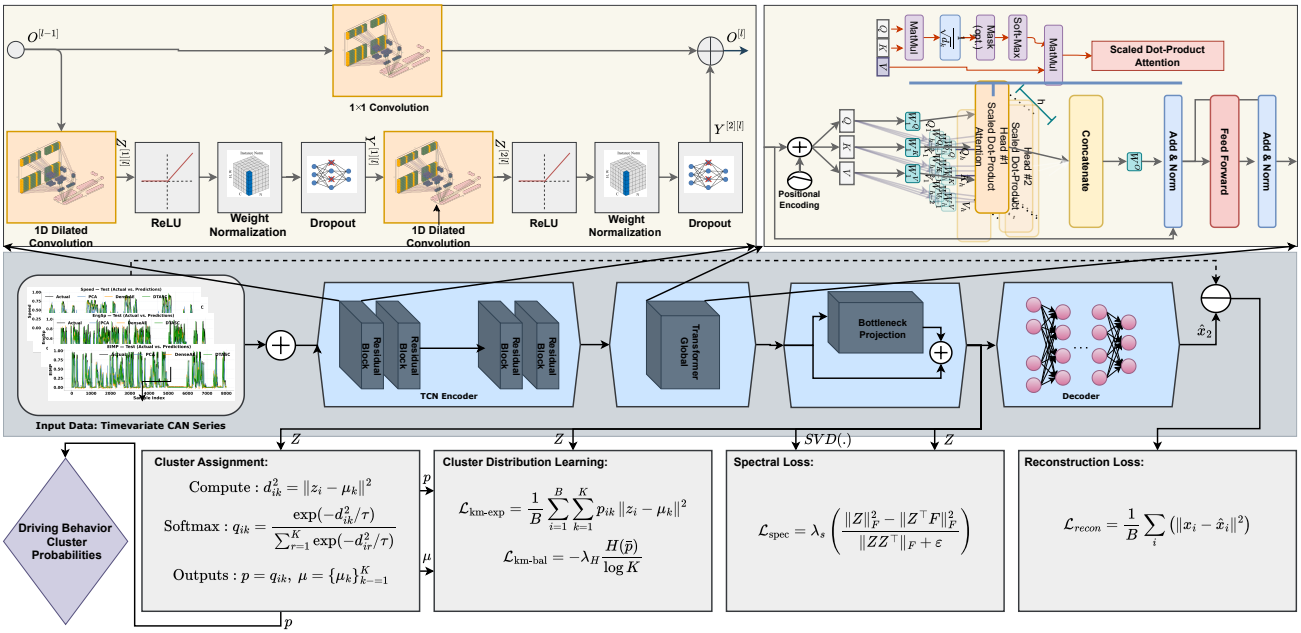


Fig. 1. Overall workflow of the proposed TCTSC framework. TCN and Transformer encoders are jointly trained with spectral regularization and a differentiable K-Means head to learn interpretable behavioral regimes from multivariate CAN time-series.

To prevent cluster collapse, a Shannon-entropy regularizer ($H(.)$) (Bromiley et al., 2004) on the average assignment $\bar{p} = \frac{1}{B}\sum_i p_i$ is added:

$$\mathcal{L}_{\text{km-bal}} = -\lambda_H \frac{H(\bar{p})}{\log K}, \quad H(\bar{p}) = -\sum_k \bar{p}_k \log \bar{p}_k. \tag{9}$$

The differentiable K-Means objective is $\mathcal{L}_{\text{km}} = \mathcal{L}_{\text{km-exp}} + \mathcal{L}_{\text{km-bal}}$.

3.6 Training Objective

Encoder weights and centroids are trained end-to-end by minimizing

$$\mathcal{L}_{\text{total}} = w_{\text{rec}}\,\mathcal{L}_{\text{rec}} + w_{\text{spec}}\,\mathcal{L}_{\text{spec}} + w_{\text{km}}\,\mathcal{L}_{\text{km}}, \tag{10}$$

where $\mathcal{L}_{\text{rec}} = \|\widehat{\mathbf{X}} - \mathbf{X}\|_2^2$ stabilizes encoding–decoding consistency, and $w_{\text{rec}}, w_{\text{spec}}, w_{\text{km}}$ are scalar weights. Optimization uses Adam Kingma (2014) with adaptive learning rates and validation-based early stopping.

4. RESULTS AND DISCUSSION

Thesis unsupervised models are compared, each with a soft K-Means head for fairness: (1) the proposed **TCTSC**; (2) a **TCN** autoencoder, isolating the contribution of attention and spectral learning; (3) a **DenseAE** (Mrabah et al., 2020), a nonlinear non-temporal baseline; and (4) **PCA** (Pham et al., 2022), a linear reference. All architectures and hyperparameters are tuned via Bayesian optimization (Wu et al., 2019). PCA preserves variance under linear orthogonality; DenseAE adds nonlinearity without temporal structure; TCN adds temporal structure but lacks attention and spectral constraints; TCTSC adds both. This staged comparison provides a principled ablation of the proposed framework.

4.1 Reconstruction Performance

Reconstruction quality reflects how well the latent space preserves both steady-state behavior and transient dynamics. Quality is quantified using MAE, RMSE, and symmetric MAPE on a held-out test set; transient fidelity is visualized as reconstruction-error time series (Fig. 2). Results for four representative CAN signals are summarized in Table 2. TCTSC achieves the lowest error on every metric and feature, with order-of-magnitude reductions over PCA and 3–5× reductions over DenseAE for transient-sensitive variables (acceleration, engine load, fuel rate). The TCN baseline improves substantially over PCA/DenseAE but still trails TCTSC, confirming that temporal convolution alone cannot fully capture long-range dependencies and that spectral regularization further stabilizes the latent representations during transients.

4.2 Outcome-Grounded Clustering Evaluation

To assess real-world relevance, clusters are scored by how well they explain variation in fuel consumption. With cluster assignments z_i and outcome y_i, predictions are taken as cluster means $\hat{y}_i = \mu_{z_i}$, where cluster means are estimated on the training split. Four metrics are reported: explained variance (R^2_{cluster}), MAE, RMSE, and Spearman rank correlation ρ_s. As shown in Table 3, TCTSC explains nearly 60% of fuel-consumption variance, attains the lowest absolute and squared errors, and produces the strongest monotonic agreement with observed fuel use, indicating that its learned representations capture meaningful real-world structure that PCA and DenseAE largely miss.

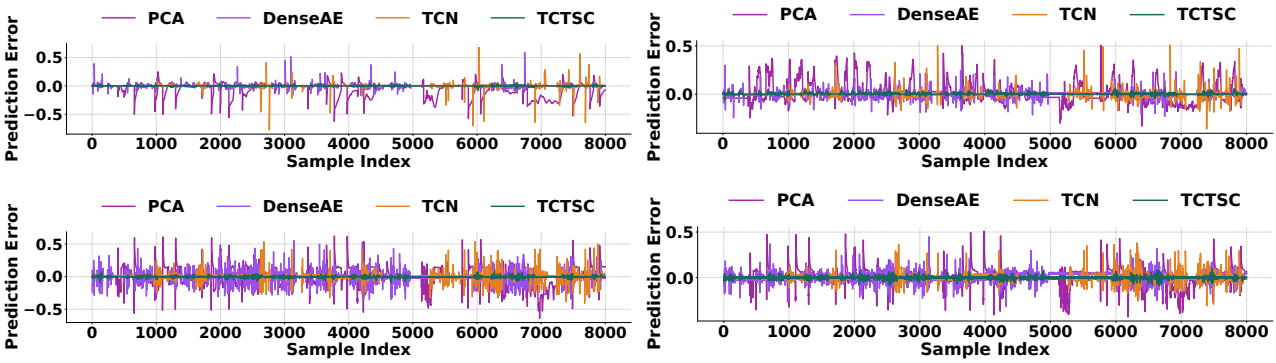


Fig. 2. Test-split reconstruction error time series ($\hat{y} - y$) for PCA, DenseAE, and TCTSC across representative CAN features (first 8,000 samples). TCTSC shows visibly lower transient deviations.

Table 2. Reconstruction performance on the test set. Lower is better.

Feature	MAE				RMSE				sMAPE (%)			
	PCA	DenseAE	TCN	TCTSC	PCA	DenseAE	TCN	TCTSC	PCA	DenseAE	TCN	TCTSC
Speed	0.044	0.018	0.014	**0.005**	0.086	0.035	0.049	**0.009**	149.86	93.33	139.35	**87.69**
Acc	0.060	0.028	0.018	**0.007**	0.089	0.042	0.042	**0.012**	13.68	7.73	4.55	**1.65**
EngLoad	0.083	0.041	0.032	**0.008**	0.124	0.069	0.055	**0.013**	46.48	43.92	40.13	**35.98**
FR	0.047	0.038	0.024	**0.008**	0.085	0.050	0.040	**0.014**	48.67	78.49	50.70	**39.39**

Table 3. Outcome-grounded clustering evaluation using fuel consumption as the target.

Model	R^2_{cluster}	MAE	RMSE	Spearman ρ_s
PCA	−0.003	0.165	0.214	0.186
DenseAE	0.067	0.144	0.197	0.069
TCN	0.312	0.118	0.172	0.341
TCTSC	**0.597**	**0.086**	**0.136**	**0.553**

Table 4. Intrinsic clustering evaluation. Higher Silhouette and CH and lower DB are better.

Model	Split	Silhouette	CH	DB
PCA	Train	0.513	4,381.93	1.726
	Test	0.601	2,305.71	1.186
DenseAE	Train	0.642	9,845.27	0.932
	Test	0.714	6,412.68	0.743
TCN	Train	0.668	18,972.44	0.812
	Test	0.732	8,906.15	0.681
TCTSC	Train	**0.692**	**34,917.45**	**0.722**
	Test	**0.751**	**12,406.88**	**0.584**

4.3 Intrinsic Clustering Evaluation

To assess structural cluster quality independent of any downstream target, three intrinsic metrics are computed in the latent space: the Silhouette coefficient (Rousseeuw, 1987), the Calinski-Harabasz (CH) index (Caliński and Harabasz, 1974), and the Davies-Bouldin (DB) index (Davies and Bouldin, 1979). As summarized in Table 4, a clear and monotonic progression PCA → DenseAE → TCN → TCTSC is observed across all metrics and on both train and test splits, with TCTSC achieving the highest Silhouette and CH values and the lowest DB index. The substantial CH gain (over 3× relative to TCN on training data) directly reflects the effect of spectral regularization, which enforces approximately orthogonal latent subspaces and suppresses overlapping cluster geometry. Some overlap between neighboring clusters persists, which is physically meaningful: real-world driving evolves continuously rather than switching abruptly between discrete states.

4.4 Cluster Interpretation

Cluster-level statistics in Table 5 reveal five distinctive operating regimes characteristic of urban bus driving. All features are standardized prior to clustering, so values reflect relative magnitudes rather than absolute physical quantities, and the interpretation relies on the joint behavior of multiple features. *Cluster 0* (high-load cruising) exhibits elevated Speed, Engine Load, Fuel Rate, and Vehicle Specific Power (VSP), consistent with sustained motion under moderate-to-high propulsion demand. *Cluster 1* (deceleration / braking transitions) shows reduced Acceleration (< 0.5), low Fuel Rate, low VSP, and low Engine Load, characteristic of coasting and braking phases. *Cluster 2* (idling / standstill) is dominated by near-zero Speed and minimal Fuel Rate and VSP, capturing dwell at stops and signals. *Cluster 3* (aggressive acceleration) achieves the highest values of Acceleration, Engine Load, Fuel Rate, and VSP, reflecting transient high-power demand that disproportionately drives fuel use and emissions. *Cluster 4* (efficient steady driving) combines moderate Speed and Fuel Rate with low acceleration magnitude, representing efficient operation under steady traffic. The dominant separating features across all clusters are Speed, Acceleration, Engine Load, Fuel Rate, and VSP, which yields a clear and physically meaningful mapping between learned clusters and real-world bus dynamics.

4.5 Explainability via SHAP

To explain cluster assignments in terms of physically interpretable signals, SHAP attribution (Lundberg and Lee, 2017; Lundberg et al., 2020) is applied to the clustering

Table 5. Statistical summary (mean ± std (median)) of selected features across clusters.

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Acc	0.62±0.09 (0.61)	0.47±0.07 (0.48)	0.49±0.08 (0.50)	0.73±0.12 (0.74)	0.28±0.13 (0.30)
EIMP	0.50±0.24 (0.44)	0.17±0.07 (0.16)	0.01±0.02 (0.00)	0.52±0.25 (0.49)	0.12±0.07 (0.12)
ERP	0.55±0.17 (0.54)	0.46±0.09 (0.45)	0.44±0.06 (0.43)	0.68±0.14 (0.68)	0.47±0.09 (0.46)
EngLoad	0.54±0.36 (0.58)	0.17±0.22 (0.01)	0.27±0.12 (0.23)	0.69±0.25 (0.70)	0.13±0.17 (0.00)
EngSp	0.50±0.27 (0.46)	0.24±0.18 (0.19)	0.15±0.07 (0.13)	0.66±0.23 (0.64)	0.28±0.12 (0.25)
FR	0.59±0.32 (0.58)	0.19±0.14 (0.17)	0.08±0.03 (0.08)	0.74±0.22 (0.77)	0.29±0.09 (0.29)
Speed	0.51±0.27 (0.47)	0.23±0.18 (0.18)	0.14±0.06 (0.12)	0.67±0.23 (0.66)	0.27±0.11 (0.24)
VSP	0.61±0.37 (0.59)	0.18±0.14 (0.15)	0.07±0.03 (0.06)	0.79±0.23 (0.82)	0.29±0.11 (0.27)

head of TCTSC. SHAP values are computed with respect to the logit of the soft cluster-assignment probability so that attributions are additive and comparable across regimes; gradient SHAP is used to approximate contributions efficiently (Lundberg et al., 2020). Per-cluster importances are aggregated as mean absolute SHAP values across assigned samples. As shown in Table 6, intake manifold pressure, vehicle speed, and acceleration consistently dominate, confirming their central role in defining transient and load-intensive regimes. Clusters with elevated EIMP, acceleration, and fuel-related proxies correspond to aggressive or high-load conditions, while clusters dominated by stable speed and moderate engine load reflect smoother, more efficient driving. SHAP therefore links the learned latent structure directly to interpretable vehicle signals and supports actionable insights for eco-driving feedback and driver coaching.

Table 6. Top features per cluster by mean absolute SHAP value. EIMP: engine intake manifold pressure; VSP: vehicle specific power; EngSp: engine speed; Acc: vehicle acceleration; EngLoad: engine load.

Feature	C0	C1	C2	C3	C4
EIMP	2.63	1.78	2.41	3.69	0.67
Speed	1.39	2.49	2.87	0.68	0.53
VSP	0.86	0.10	0.63	0.61	1.16
EngSp	0.63	0.19	0.19	0.19	0.44
Acc	0.26	0.56	0.59	0.88	1.29
EngLoad	0.11	0.09	0.08	0.14	0.10

5. CONCLUSION

This paper introduced TCTSC, an interpretable unsupervised framework for driving regime discovery from real-world transit-bus CAN data. By combining temporal convolutional encoding, Transformer-based long-range attention, spectral regularization, and a differentiable K-Means head, TCTSC learns compact and globally structured latent representations of driving behavior. On a large transit bus ETS dataset, TCTSC consistently outperforms PCA, DenseAE, and TCN baselines, achieving substantially lower reconstruction errors and superior preservation of transient dynamics. Outcome-grounded evaluation against fuel consumption shows that TCTSC explains nearly 60% of outcome variance with the lowest prediction errors and highest rank correlation among all models. Intrinsic clustering metrics confirm that spectral regularization yields well-separated yet softly bounded behavioral regimes consistent with the continuous nature of real-world driving, and SHAP-based analysis ties each regime to physically meaningful CAN driving signals. Together these results establish TCTSC as a robust and interpretable foundation for eco-driving analysis, driver assessment, and fleet-level performance monitoring. Future work will extend the framework with contrastive learning and multi-output emission modeling to tighten the link between regimes and pollutant outcomes.

FUNDING AND ACKNOWLEDGMENT

This research was supported by the Climate Action and Awareness Fund through Environment and Climate Change Canada and by Emissions Reduction Alberta (ERA) under the AZEHT project. The authors also acknowledge Bahram Bahri for his dedicated efforts in conducting the data collection.

REFERENCES

Bauckhage, C. (2015). Lecture notes on data science: Soft k-means clustering. Technical report, B-IT, University of Bonn. URL https://www.researchgate.net/publication/283327156_Lecture_Notes_on_Data_Science_Soft_k-Means_Clustering. Technical note describing soft k-means clustering with fuzzy/soft membership assignments, where points can belong to multiple clusters with degrees of membership.

Bromiley, P., Thacker, N., and Bouhova-Thacker, E. (2004). Shannon entropy, renyi entropy, and information. *Statistics and Inf. Series (2004-004)*, 9(2004), 2–8.

Caliński, T. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods*, 3(1), 1–27. doi:10.1080/03610927408827101.

Choi, S., Kim, J., Kwak, D., Angkititrakul, P., and Hansen, J.H.L. (2016). Analysis and classification of driver distraction using in-vehicle can-bus data. In *Proceedings of the 2016 Biennial Workshop on DSP for In-Vehicle and Mobile Systems*.

Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., and et al. (2025). Ghg emissions of all world countries - 2025 report. https://edgar.jrc.ec.europa.eu/report_2025. Accessed October 31, 2025.

Davies, D.L. and Bouldin, D.W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2), 224–227. doi:10.1109/TPAMI.1979.4766909.

Environment and Climate Change Canada (2025). Greenhouse gas sources and sinks in canada: executive summary 2025. https://www.canada.ca/en/environment-climate-change/services/climate-change/gree

nhouse-gas-emissions/sources-sinks-executive-summary-2025.html. Accessed October 31, 2025.

Fugiglando, U., Massaro, E., Santi, P., Milardo, S., Abida, K., Stahlmann, R., Netter, F., and Ratti, C. (2018). Driving behavior analysis through can bus data in an uncontrolled environment. *IEEE Transactions on Intelligent Transportation Systems*, 20(2), 737–748.

Hallac, D., Nystrup, P., and Boyd, S. (2016). Driver identification using automobile sensor data from a single turn. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 953–958. IEEE. doi:10.1109/ITSC.2016.7795665.

Isreal, O. and Alonge, M. (2025). Combining temporal convolutional networks and transformer models for time series forecasting. Preprint, ResearchGate. Hybrid TCN–Transformer framework with enhanced forecasting accuracy over standalone models; available online at: https://www.researchgate.net/publication/390494107_Combining_Temporal_Convolutional_Networks_and_Transformer_Models_for_Time_Series_Forecasting.

Kingma, D.P. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Li, C., Kiat, Y.C., Jing, J., and Long, C. (2025). T-vae: Transformer-based variational autoencoder for perceiving anomalies in multivariate time series data. *Expert Systems*. doi:10.1111/exsy.70078. Transformer-based variational autoencoder leveraging self-attention to capture global dependencies in multivariate time series.

Liang, Y., Reyes, M.L., and Lee, J.D. (2007). Real-time detection of driver cognitive distraction using support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, 8(2), 340–350. doi:10.1109/TITS.2007.895298.

Lundberg, S.M., Erion, G.G., Chen, H., DeGrave, A., Prutkin, J.T., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.I. (2020). From local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence*, 2(1), 56–67. doi:10.1038/s42256-019-0138-9.

Lundberg, S.M. and Lee, S.I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 30, 4765–4774. URL https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.

Martin, E.W., Chan, N.D., and Shaheen, S.A. (2012). How public education on ecodriving can reduce both fuel use and greenhouse gas emissions. *Transportation research record*, 2287(1), 163–173.

Miyajima, C., Nishiwaki, Y., Ozawa, K., Wakita, T., Itou, K., Takeda, K., and Itakura, F. (2007). Driver modeling based on driving behavior and its evaluation in driver identification. *Proceedings of the IEEE*, 95(2), 427–437.

Mrabah, N., Khan, N.M., Ksantini, R., and Lachiri, Z. (2020). Deep clustering with a dynamic autoencoder: From reconstruction towards centroids construction. *Pattern Recognition*, 106, 107412.

Ng, A.Y., Jordan, M.I., and Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Pham, C.H., Ladjal, S., and Newson, A. (2022). Pca-ae: Principal component analysis autoencoder for organising the latent space of generative networks. *Journal of Mathematical Imaging and Vision*, 64(5), 569–585.

Rousseeuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. doi:10.1016/0377-0427(87)90125-7.

Thornton, C.E., Sloop, J., Brown, S., Orndorff, A., Headley, W.C., and Young, S. (2025). Temporal convolutional autoencoder for interference mitigation in fmcw radar altimeters. *arXiv preprint*. URL https://arxiv.org/abs/2505.22783. Shows a Temporal Convolutional Network autoencoder that exploits temporal correlations in sequential data, learning effective representations of temporal patterns.

Tofigh, M., Kharazmi, A., Smith, D.J., Koch, C.R., and Shahbakhti, M. (2024). Temporal dilated convolution and nonlinear autoregressive network for predicting solid oxide fuel cell performance. *Engineering Applications of Artificial Intelligence*, 136, 108994.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

Wang, Y. and Chen, P. (2025). Network traffic prediction based on transformer and temporal convolutional network. *PLOS ONE*, 20(4), e0320368. doi:10.1371/journal.pone.0320368. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0320368. Hybrid model combining TCN and Transformer for network traffic prediction; captures local and global temporal dependencies.

Wu, J., Chen, X.Y., Zhang, H., Xiong, L.D., Lei, H., and Deng, S.H. (2019). Hyperparameter optimization for machine learning models based on bayesian optimization. *Journal of Electronic Science and Technology*, 17(1), 26–40.

Wu, Y., Zhao, X., Chen, C., Rong, J., and Zhang, Y. (2017). Development and application of an ecodriving support platform based on internet+: Case study in beijing taxicabs. *Transportation Research Record*, 2645(1), 57–66.

Yasami, A., Tofigh, M., Bahri, B., Koch, C.R., and Shahbakhti, M. (2025). Intelligent ensemble architecture for capturing transient nitrogen oxides emission spikes in real-world driving conditions. *Engineering Applications of Artificial Intelligence*, 158, 111359.

Yasami, A., Tofigh, M., Shahbakhti, M., and Koch, C.R. (2026). Dual-stream transformer autoencoding with contrastive k-means and spectral learning for explainable unsupervised driving-behavior clustering. *Advanced Engineering Informatics*, 74, 104739.

Zerveas, G. and et al. (2021). A transformer-based framework for multivariate time series representation learning. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*. Demonstrates how Transformer encoders capture global contextual relationships across multivariate sequences using self-attention.