Title

Remote DCFC Reliability and Downtime Detection Tool: Detecting EV Charging Failures That Standard Reliability Protocols Cannot Detect

Authors

Karanam, Vaishnavi
Tal, Gil
Garas, Dahlia

Publication Date

2026-08-12

DOI

10.7922/G2SQ8XS6

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report Number	2. Government Association Number	3. Recipient's Catalog Number
UCD-ITS-RR-26-42	N/A	N/A

4. Title and Subtitle	5. Report Date
Remote DCFC Reliability and Downtime Detection Tool: Detecting EV Charging Failures That Standard Reliability Protocols Cannot Detect	07/08/2026
	6. Performing Organization Code UC Davis

7. Author(s)	8. Performing Organization Report Number
Vaishnavi Karanam (ORCID: 0000-0002-4647-5250) Gil Tal (ORCID: 0000-0001-7843-3664) Dahlia Garas (ORCID: 0000-0002-6222-9690)	UCD-ITS-RR-26-42

9. Performing Organization Name and Address	10. Work Unit Number
UC Davis Electric Vehicle Research Center One Shields Avenue Davis, CA 95616	N/A
	11. Contract or Grant Number 65A1188

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
California Department of Transportation Division of Research, Innovation, and System Information PO Box 942873 Sacramento, CA 94273-0001	Research Report 6/28/2024 - 2/25/2026
	14. Sponsoring Agency Code Caltrans

15. Supplementary Notes

doi:10.7922/G2SQ8XS6

16. Abstract

The Caltrans ZEV 30-30 project is established to support California's goal of deploying five million ZEVs on the road by 2030, focusing on filling gaps in the State Highway System's corridor ZEV charging network. The reliability of DC Fast Charging (DCFC) infrastructure along these corridors is essential. Unreliable chargers erode driver confidence and undermine the transition to zero-emission transportation. Charging Station Operators (CSOs) currently rely on conventional monitoring protocols, primarily the Open Charge Point Protocol (OCPP), to detect charger failures. While OCPP-based monitoring is effective for detecting most electrical and software failures, it cannot identify a broad class of faults arising from mechanical damage, physical obstruction, network communication outages, or logistical barriers. [1, 2] These hidden issues persist until an EV driver encounters the faulty charger and reports the problem, leading to delayed fault resolution and degraded consumer experience.

This report presents a predictive anomaly detection tool developed under Caltrans Agreement No. 65A1188 that enables Charging Station Operators to detect hidden charging faults by analyzing habitual EV driver usage patterns. The tool incorporates two anomaly detection models: a naïve probability distribution-based technique and a Long Short-Term Memory (LSTM) autoencoder.

17. Key Words	18. Distribution Statement
Electric Vehicle, DCFC, Charging system operator, charger failure, fault, reliability, remote, anomaly detection	No restrictions

19. Security Classification (of this report)	20. Number of Pages	21. Cost of Report
Unclassified	34	$ 0.00

DOT F 1700.7 (8.72) *Reproduction of completed page authorized.*

DISCLAIMER STATEMENT

This document is disseminated in the interest of information exchange. The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the Federal Highway Administration. This publication does not constitute a standard, specification or regulation. This report does not constitute an endorsement by the Department of any product described herein.

For individuals with sensory disabilities, this document is available in alternate formats. For information, call (916) 654-8899, TTY 711, or write to California Department of Transportation, Division of Research, Innovation and System Information, MS-83, P.O. Box 942873, Sacramento, CA 94273-0001.

STATE OF CALIFORNIA • DEPARTMENT OF TRANSPORTATION

FINAL RESEARCH REPORT

Remote DCFC Reliability and Downtime Detection Tool:

Detecting EV Charging Failures That Standard Reliability Protocols Cannot Detect

Prepared for:

California Department of Transportation (Caltrans)

Division of Research, Innovation and System Information

ZEV 30-30 Program

Prepared by:

Vaishnavi Karanam (Postdoctoral Researcher)

Gil Tal, Ph.D. (Principal Investigator)

Dahlia Garas (Project Manager)

Electric Vehicle Research Center, UC Davis Institute of Transportation Studies

Agreement No. 65A1188 | UCD #A24-3062-001 | Contract Period: June 28, 2024 – February 25, 2026

February 2026

EXECUTIVE SUMMARY

The Caltrans ZEV 30-30 project is established to support California's goal of deploying five million ZEVs on the road by 2030, focusing on filling gaps in the State Highway System's corridor ZEV charging network. The reliability of DC Fast Charging (DCFC) infrastructure along these corridors is essential. Unreliable chargers erode driver confidence and undermine the transition to zero-emission transportation.

Charging Station Operators (CSOs) currently rely on conventional monitoring protocols, primarily the Open Charge Point Protocol (OCPP), to detect charger failures. While OCPP-based monitoring is effective for detecting most electrical and software failures, it cannot identify a broad class of faults arising from mechanical damage, physical obstruction, network communication outages, or logistical barriers. [1, 2] These hidden issues persist until an EV driver encounters the faulty charger and reports the problem, leading to delayed fault resolution and degraded consumer experience.

This report presents a predictive anomaly detection tool developed under Caltrans Agreement No. 65A1188 that enables CSOs to detect hidden charging faults by analyzing habitual EV driver usage patterns. The tool incorporates two anomaly detection models: a naïve probability distribution-based technique and a Long Short-Term Memory (LSTM) autoencoder.

Key Findings

- Detected anomalies resulted in uptime reductions of 6%–38% (mean 16%) across 12 California chargers, representing reliability deficits invisible to standard OCPP monitoring.
- At 50% confidence, CSOs could detect faults up to 3× faster (naïve) and 2.4× faster (LSTM), reducing total gap hours by 62% and 68% respectively.
- At 75% confidence, CSOs could detect faults 2× faster (naïve) and 1.8× faster (LSTM), reducing gap hours by 43% and 50%.
- At 90% confidence, both methods enabled ~1.5× faster detection, reducing gap hours by 28% (naïve) and 46% (LSTM).
- Average gap duration was reduced by 34%–66% depending on method and confidence threshold.

Regulatory Relevance

The California Energy Commission (CEC) is developing uptime recordkeeping and reporting standards for EV charging stations that received public funding. [36] The tool developed in this project directly supports CSOs in meeting these standards by detecting fault categories outside the scope of conventional OCPP compliance monitoring, including network communication failures, blocked access, and physically damaged equipment.

Field Validation Status and Next Steps

Despite repeated outreach to ZEV 30-30 stakeholders, the research team has been unable to establish consistent contact and remains uncertain whether a dedicated team exists to support model evaluation through on-site testing. As a result, formal field validation using Caltrans-owned DCFC station data could not be carried out as originally scoped. To partially bridge this gap, approximately three months of publicly available live charger status data were scraped to generate utilization-over-time feeds for preliminary validation. A minimum of one year of continuous data per station is needed to adequately capture utilization patterns; the team will continue collecting this data and iteratively calibrating the models as the dataset grows. Renewed stakeholder engagement to

enable on-site validation and access to internal fault records remains a priority for the next project phase.

1. INTRODUCTION

The Caltrans ZEV 30-30 project was established to support California's goal of deploying five million zero-emission vehicles on the road by 2030. It aims to fill gaps in California's corridor ZEV charging network, focusing on key routes within the State Highway System. Reliable DC Fast Charging (DCFC) infrastructure along these corridors is essential for enabling long-distance EV travel and maintaining public confidence in battery-electric vehicles.

Charging Station Operators (CSOs) are the primary stakeholders responsible for ensuring the ongoing operational performance of their charging stations, including managing backend technologies and communications between backend systems and the charging hardware. CSOs typically detect most electrical and software failures through conventional monitoring protocols, primarily the Open Charge Point Protocol (OCPP). [1] However, a critical class of faults arising from mechanical damage, physical obstruction, network communication outages, and logistical barriers remains invisible to CSOs until encountered by a driver.

This visibility gap creates a systematic delay in fault detection. A broken charging cable, a blocked station, or a communication network failure may persist for hours or days before a driver reports the issue, eroding consumer experience and contributing to overestimation of true uptime. [2] This project, funded under Caltrans Agreement No. 65A1188 (PI: Dr. Gil Tal, UC Davis EV Research Center), develops and demonstrates a tool that addresses this gap by analyzing habitual EV charging usage patterns to detect anomalous gaps indicative of hidden faults.

The tool incorporates two complementary anomaly detection approaches: a naïve probability distribution model and an LSTM autoencoder. Both were demonstrated on 12 real-world chargers across California. The tool can subsequently be calibrated and validated against data from approximately 37 Caltrans-owned DCFC stations under the ZEV 30-30 project.

1.1 Background and Motivation

Figure 1 illustrates the dichotomy between CSO-measured uptime and consumer-experienced uptime. While CSOs can detect most electrical and software failures via OCPP, a class of remotely unobservable failures remains hidden. **Figure 2** enumerates the timeline of a charging attempt and the failure modes at each stage, separated by CSO visibility level. [1, 2]

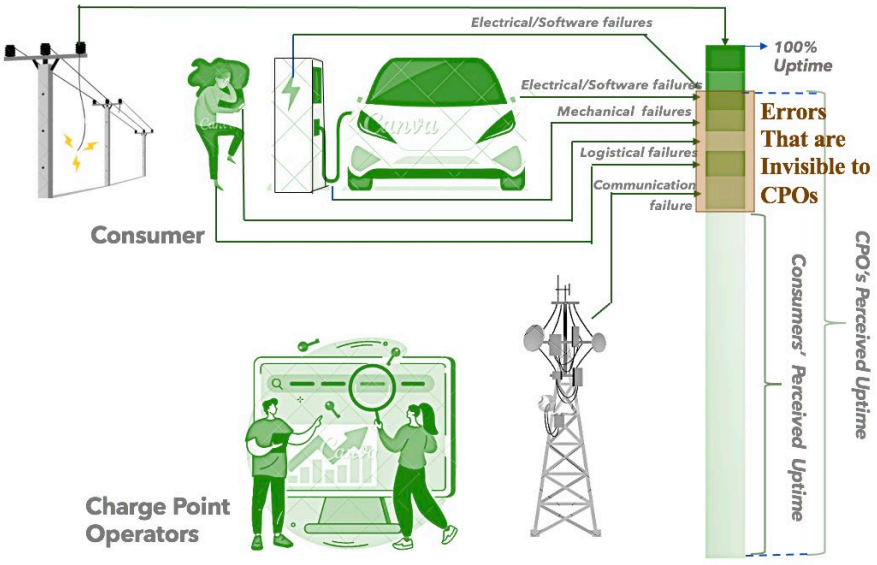


Figure 1: *Consumer Uptime vs. CSO-Measured Uptime*

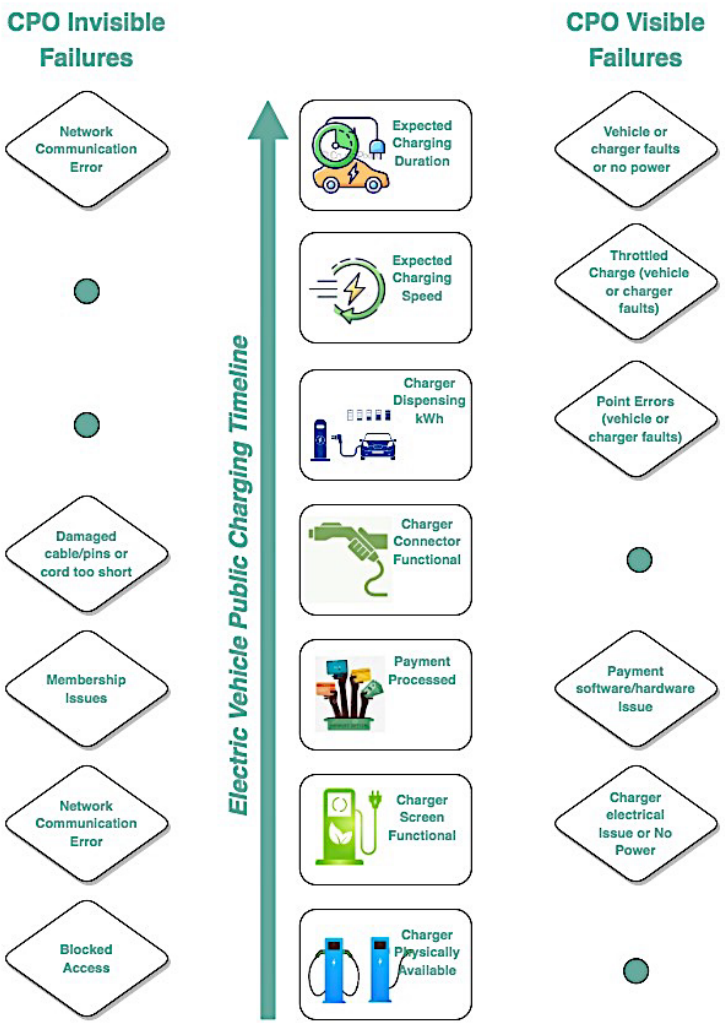


Figure 2: *EV Charging Attempt Timeline and Associated Failure Modes by CSO Visibility*

A failure invisible to CSOs may persist until an EV driver encounters the faulty charger and files a report. This reactive model exacerbates the reliability problem and is particularly acute for corridor DCFC stations, which are the primary focus of the Caltrans ZEV 30-30 program, where driver stranding risk is highest. [37]

1.2 Project Scope and Organization

This report constitutes the Task 6 Final Report deliverable under Caltrans Agreement No. 65A1188. It is organized as follows: Section 2 describes the taxonomy of EV charging failures. Section 3 reviews time series anomaly detection literature (Task 1 deliverable). Section 4 presents the full methodology including data preparation and model architecture. Section 5 reports results from demonstration on California charging data (Tasks 2–3). Section 6 discusses policy implications and integration recommendations (Tasks 4–5). Section 7 provides conclusions and future calibration directions.

2. TYPES OF EV CHARGING FAILURES

Rempel et al. [1] identified six types of charge failures encountered in a study of open DCFC reliability in the California Bay Area: broken connectors, non-responsive screens, error messages on screens, connection errors, payment system failures, and charge initiation failures. **Table 1** abstracts and extends these failure types to more comprehensively capture the obstacles consumers encounter while attempting to charge their EVs. [1, 2]

Failures are classified by observability to CSOs:

2.1 Remotely Observable Failures

CSOs using OCPP-based monitoring can typically detect the following failure types in near real-time: [1]

***Table 1**:. Remotely Observable EV Charging Failures [1, 2]*

Failure Type	Description
Charger-to-Vehicle Communication Failure	Malfunction in the EV charging port or station connector; issue with the communication protocol between the EV and the charging station.
Connector / Cable Issue	Charging cable improperly placed into the vehicle charging port; poor conductivity due to corrosion.
Electrical Insulation / Safety Issue	Electrical system of charger may be overheating; insulation may need to be inspected.
Payment Errors	Technical issues with the payment system, compatibility issues with the payment method, or user error during the payment process.
Vehicle Errors	Software or hardware malfunctions in the vehicle; charging port incompatibility; battery issues.
Charger Equipment Errors	Software or hardware malfunctions in the charger; power supply issues.
Power Outage	A power outage can cause an EV charger to shut down or interrupt an ongoing charging process.

2.2 Remotely Unobservable Failures

The following failure types cannot be detected through standard OCPP-based remote monitoring and persist until physically encountered or reported by a driver. [1, 2] These are the primary target of the tool developed in this project:

***Table 2:**. Remotely Unobservable EV Charging Failures [1, 2]*

Failure Type	Description
Blocked Access to Station	Physical obstruction by non-EV vehicles, fencing, snow, flood water, or other barriers preventing access to the charging port.
Physically Damaged Equipment	External EVSE components damaged by environmental factors, vandalism, or vehicle collisions, undetectable via OCPP.
Logistical and Interoperability Issues	Membership requirements, payment incompatibility, equipment incompatibility, poor cell/Wi-Fi coverage, and difficult EVSE operations.
Network Communication Failure	Configuration errors, line damage, power loss, or hardware failure in the communication network that prevents OCPP reporting entirely.

This failure taxonomy defines the core opportunity: because remotely unobservable failures do not generate OCPP fault codes, CSOs have no direct mechanism to detect them.[37] The tool developed in this project exploits unexpected gaps in charger usage as behavioral signals and an indirect proxy for these hidden faults. This is particularly critical for remote corridor DCFC stations under the ZEV 30-30 program, where extended downtime poses the greatest risk of stranding drivers.

3. TIME SERIES ANOMALY DETECTION: LITERATURE REVIEW

Time series anomaly detection is a process of identifying abnormal patterns within a sequence of data points collected over time. [3] Techniques span six categories: forecasting, reconstruction, encoding, distribution, distance, and isolation tree approaches. [3] All six were evaluated for suitability in detecting anomalous EV charging usage patterns indicative of CSO-invisible charger faults.

3.1 Forecasting Methods

Forecasting-based anomaly detection trains a model to predict future values from a current data window; deviations between predicted and actual values identify anomalies. [3] Models are typically trained in a semi-supervised manner on normal (non-anomalous) data. Traditional statistical forecasting methods include Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and Triple Exponential Smoothing. [4, 5, 6, 7] Deep learning forecasting methods include LSTM autoencoders, Convolutional Neural Networks, Graph Attention Networks, and Echo State Networks. [8, 9, 10, 11, 12]

3.2 Reconstruction Methods

Reconstruction methods encode subsequences of normal training data into a low-dimensional space, then measure reconstruction error on test data; a high error signals an anomaly. [3] Traditional approaches include Principal Component Isolation (PCI). [13] Machine learning approaches include Principal Component Analysis (PCA) and Principal Component Classifier (PCC). [14, 15] Deep learning approaches include autoencoders, variational autoencoders (VAE), LSTM-VAE, Recurrent Neural Networks (RNN), Spectral Residual, and Generative Adversarial Networks (GAN). [16, 17]

3.3 Encoding Methods

Encoding methods transform subsequences into a low-dimensional latent space and compute anomaly scores from the encoded representations. Methods include grammar-based compression, graph-based compression, suffix trees, time series bitmaps, Hidden Markov Models (HMM), and Dynamic Bayesian Networks (DBN). [18, 19] Subsequences that are difficult to compress or have low frequency are considered anomalous.

3.4 Distribution Methods

Distribution methods estimate the probability distribution of normal data, then flag observations with low likelihood. [20] Anomalous points are judged by frequency rather than distance. Approaches include extreme value theory, [20] copula-based outlier detection, [21] discrete wavelet transforms and maximum likelihood estimation, [22] Histogram-based Outlier Scores (HBOS), [23] and Normalizing Flows (NF). [24]

3.5 Distance Methods

Distance-based methods compare subsequences using specialized distance metrics, treating anomalies as those with unusually large distances to neighbors. [3] These methods are typically unsupervised. Traditional approaches include density- and cluster-based local outlier detection (CBLOF, COF, LOF). [25, 26, 27] Machine learning methods include k-means, K-nearest neighbors (KNN), and Support Vector Machines (SVM). [28, 29, 30] Deep learning methods include hybrid KNNs. [31]

3.6 Isolation Tree Methods

Isolation tree methods build random trees that partition test time series samples; anomalous samples are closer to tree roots and have shorter path lengths. [32] Representative algorithms include Isolation Forest (iForest), Extended Isolation Forest (EIF), Hybrid Isolation Forest (HIF), Sub-IF, and IF-LOF. [32, 33, 34, 35]

3.7 Comparison and Method Selection

Table 3 presents the advantages and disadvantages of all anomaly detection techniques considered for the development of this tool, reproducing Table 21 from the underlying research. [3]

Table 3: *Advantages and Disadvantages of Anomaly Detection Techniques (adapted from dissertation Table 21) [3]*

Method Family	Method Class	Method(s)	Advantages	Disadvantages
Forecasting	Deep Learning	AD-LTI, DeepAnt, DeepNap, HealthESN, LSTM-AD, MTAD-GAT, Telemanom, Torsk	Automated feature learning, ability to capture non-linear relationships, consideration of temporal context, and adaptability to changing data distributions	Data requirements, computational complexity, black-box nature, susceptibility to overfitting, and the need for careful hyperparameter tuning, which can limit interpretability and performance in scenarios with limited labeled data
Forecasting	Traditional Statistics	ARIMA, MedianMethod, SARIMA, Triple ES	Interpretability, less data requirements, computational efficiency, and robustness to outliers	Assuming linearity in data, limited temporal context, manual feature engineering, limited adaptability to changing data distributions, and sensitivity to certain data patterns, which may hinder performance in complex and dynamic anomaly detection scenarios
Reconstruction	Deep Learning	Autoencoder, Bagel, Donut, EncDec-AD, LSTM-VAE, MSCRED, OmniAnomaly, Spectral Residual, TAnoGAN	Unsupervised learning, feature learning, and the ability to capture non-linear relationships and temporal context	Data requirements, computational intensity, hyperparameter tuning, reconstruction ambiguity, limited interpretability, and susceptibility to overfitting
Reconstruction	Machine Learning	PCC, RobustPCA	Interpretability, computational efficiency, and robustness to outliers	Assuming linearity in data, limited feature learning capabilities, reduced consideration of temporal context, and the lack of detailed anomaly localization compared to deep learning reconstruction methods
Reconstruction	Traditional Statistics	PCI	Interpretability, computational efficiency,	Assuming linearity in data, limited feature learning

Method Family	Method Class	Method(s)	Advantages	Disadvantages
			and dimensionality reduction	capabilities, sensitivity to data distribution, challenges in selecting an appropriate anomaly detection threshold, and limited adaptability to complex and dynamic anomaly detection scenarios
Encoding	Data Mining	Ensemble GI, GrammarViz, PST, Series2Graph, TARZAN, TSBitmap	Graph representation, data abstraction, flexibility, and temporal context	Complexity, hyperparameter tuning, interpretability, data transformation, data requirements
Encoding	Stochastic Learning	LazerDBN, MultiHMM	Probabilistic modeling, sequential data handling, representation learning, scalability, and adaptability	Complexity, data requirements, hyperparameter tuning, interpretability, initialization sensitivity, and data distribution assumptions
Distance	Data Mining	HOT SAX, NormA-SJ, SSA	Computational efficiency, interpretability, and scalability	Parameter sensitivity, feature engineering requirements, lack of full temporal context, sensitivity to data scaling, limited representation learning, and dependence on data distribution
Distance	Deep Learning	Hybrid KNN	Feature learning, modeling non-linear relationships, data abstraction, and adaptability	Computational intensity, data requirements, hyperparameter tuning, interpretability, overfitting, and sensitivity to data characteristics
Distance	Machine Learning	k-means, KNN, PS-SVM	Interpretability, simplicity, scalability, and parameter tuning	Limitations in feature learning capabilities, sensitivity to data scaling, curse of dimensionality, data requirements, computation time, and hyperparameter tuning
Distance	Traditional Statistics	CBLOF, COF, LOF	Interpretability, scalability (CBLOF), and robustness to noise (COF and LOF)	Limitations in feature learning, sensitivity to data scaling and distribution, curse of dimensionality (LOF), computation time (COF and LOF), data requirements, and hyperparameter tuning
Distribution	Deep Learning	Normalizing Flows (NF)	Flexible distribution modeling, generative	Computational intensity, data requirements,

Method Family	Method Class	Method(s)	Advantages	Disadvantages
			capabilities, scalability, and non-linear relationship modeling	hyperparameter tuning, data distribution assumptions, interpretability, and data scaling
Distribution	Machine Learning	HBOS	Computational efficiency, interpretability, scalability, and non-linear relationship modeling	Limitations in feature learning, sensitivity to data distribution assumptions, data scaling, data requirements, anomaly localization, and susceptibility to overfitting
Distribution	Signal Analysis	DWT-MLEAD	Signal processing expertise, noise robustness, multiscale analysis, and non-linear relationship modeling	Limitations in data requirements, interpretability, parameter tuning, computation time, feature learning capabilities, and sensitivity to data scaling
Distribution	Traditional Statistics	COPOD, S-H-ESD, DSPOT	Interpretability, robustness to outliers (S-H-ESD), data distribution modeling, and anomaly localization (DSPOT)	Limitations in data scaling, data requirements, computation time, data distribution assumptions, hyperparameter tuning, feature learning capabilities, and non-linear relationship modeling
Isolation Trees	Traditional Statistics	EIF, HIF, Isolation Forest–Local Outlier Factor (IF-LOF), iForest	Outlier-focused design, model simplicity, scalability, and non-linear relationship modeling	Data scaling sensitivity, data requirements, hyperparameter tuning (IF-LOF), anomaly localization, limited feature learning capabilities, curse of dimensionality, and data distribution assumptions

The tradeoff between deep learning and traditional statistical methods for anomaly detection revolves around complexity, interpretability, feature learning, data requirements, and computational resources. [3] Deep learning excels at capturing complex non-linear patterns but requires substantial labeled training data and computational resources. Traditional statistical methods are more interpretable and data-efficient but may lack automatic feature learning capability.

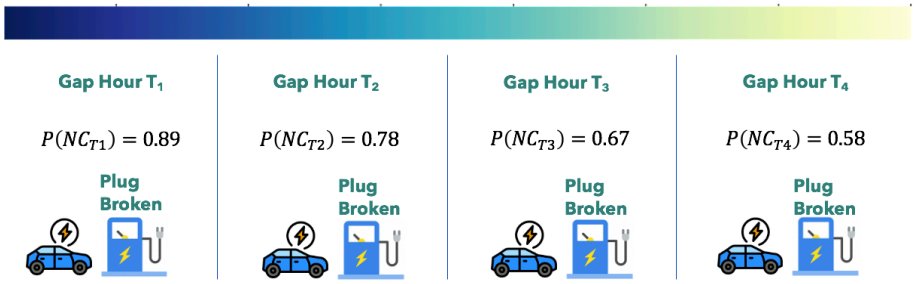
To harness the complementary strengths of both paradigms, this project implements both a traditional statistical method (a naïve probability distribution technique) and a deep learning method utilizing LSTM networks. This dual-model design supports performance comparison across different charger deployment contexts and enables CSOs with varying computational resources to deploy the tool.

4. METHODOLOGY

4.1 Methodological Intuition

EV drivers are likely to charge their EVs at the same public charging locations along their regular travel routes.[37] Any sudden gap within the usage pattern of a given EVSE could reveal a technical or logistical failure that standard monitoring protocols fail to capture. **Figure 3** provides an intuitive demonstration.

Consider a charger with a broken plug. The heatmap defines the hourly probabilities of a charging session occurring on a typical summer Saturday. At hour T_0, a vehicle that usually charges at this station attempts to do so but fails; the probability of no charging in that specific hour is 89%. At hours T_1, T_2, and T_3, three more vehicles similarly fail, with individual no-charge probabilities of 78%, 67%, and 58%. Individually, each no-charging event appears plausible. But the joint probability of the entire four-hour sequence of non-use is only 26% - a meaningful signal of a potential charger fault.



$$P(NC_{T1}, NC_{T2}, NC_{T3}, NC_{T4},) = 0.89 * 0.78 * 0.67 * 0.58 = 0.26$$

Figure 3: *Charging Anomaly Detection Intuition*

4.2 Methodology Framework

The methodology consists of three major stages, illustrated in **Figure 4**:[3]

1. Data Collection and Preparation: EVSE charging session data are collected, cleaned, and transformed into hourly time traces. Data are partitioned into training (one year) and testing (one month) sets.
2. Model Training: Both anomaly detection models are trained on the cleaned training data to learn regular charging patterns, enabling them to distinguish anomalies from normal behavior.
3. Anomaly Detection: Trained models are applied to the testing dataset to flag instances that deviate significantly from learned normal patterns.

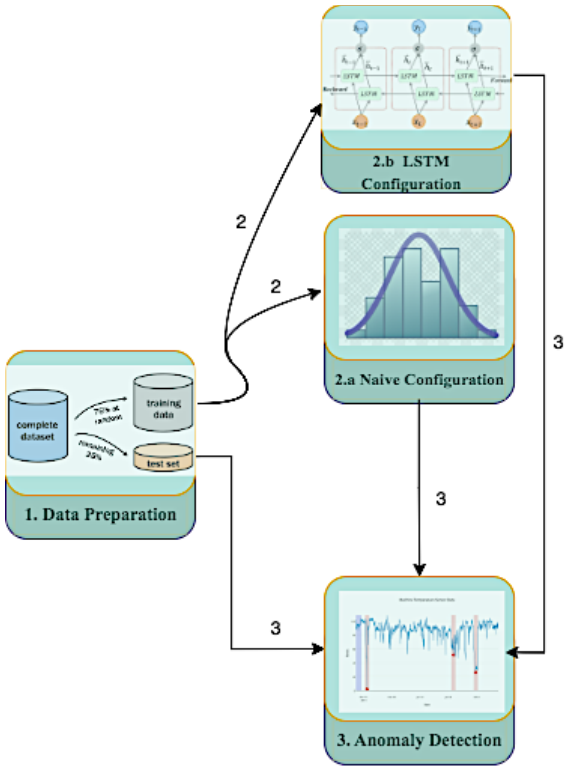


Figure 4: *Charging Anomaly Detection Methodology Framework*

4.3 Data Overview and Preparation

Charging session data from 12 EVSE units across four California location types were used to develop and demonstrate the tool. The dataset spans the range of deployment contexts relevant to the ZEV 30-30 program: residential Level 2 (Oakland), commercial Level 2 (San Francisco), workplace Level 2 (San Diego), and corridor DC Fast Chargers (Northern California highway corridor). Session details were obtained from the corresponding installation companies; identities have been anonymized to protect confidentiality.

Table 4 summarizes the dataset. A total of 12,006 Level 2 sessions are logged over 700+ days (April 2021–January 2023) and 2,341 DCFC sessions over 263 days (January–October 2021). To construct hourly time series traces, each charger's logging period was iterated hour by hour, recording the energy dispensed and number of sessions in each hour. Training data consisted of one full year per charger; the immediately following month served as the test dataset.

Table 4: *Charging Session Dataset Overview*

Charger Type	ID	# Sessions	Total kWh	Days Logged	Utilization Rate
Residential Level 2 (Oakland)	1	1,742	34,466	1,055	32.9%
Residential Level 2 (Oakland)	2	1,786	36,577	1,055	38.6%
Commercial Level 2 (San Francisco)	1	2,841	18,729	731	30.3%
Commercial Level 2 (San Francisco)	2	1,098	12,672	731	14.7%

Charger Type	ID	# Sessions	Total kWh	Days Logged	Utilization Rate
Commercial Level 2 (San Francisco)	3	2,079	29,703	731	42.1%
Workplace Level 2 (San Diego)	1	592	4,529	575	28.9%
Workplace Level 2 (San Diego)	2	717	2,744	575	34.9%
Workplace Level 2 (San Diego)	3	772	3,293	575	34.5%
Workplace Level 2 (San Diego)	4	479	3,078	575	25.3%
Corridor DCFC (N. California)	1	736	9,959	263	4.5%
Corridor DCFC (N. California)	2	396	6,790	263	3.0%
Corridor DCFC (N. California)	3	627	12,684	263	5.2%
Corridor DCFC (N. California)	4	582	6,407	263	16.4%

4.4 Naïve Probability Distribution Anomaly Detection Model

The naïve probability distribution model detects charging usage gaps through the following algorithm:

Input: Charging behavior data for EV chargers, including hourly time traces of charging sessions, day type (weekday or weekend), and holiday information.

4. Calculate Daily Charging Probability Distribution: For each charger, calculate the hourly probability of a charging session occurring, segmented by day type (weekday/weekend) and holiday status. This models the expected charging behavior for each day and time slot.
5. Iterate Through Time Traces by the Hour: Traverse each charger's hourly time trace sequentially.
6. Detect No-Charging Sessions in an Hour: When no charging session is recorded in a given hour, proceed to compute the probability of no charging in that hour.
7. Calculate Joint Probability of No Charging Sessions: Using the daily charging probability distribution from Step 1, compute the joint probability of no charging sessions for consecutive no-charge hours by multiplying the individual no-charge probabilities. For example, if hours T_0 and T_1 both show no activity, the joint probability is p(no charge at T_0) × p(no charge at T_1). This continues for longer uninterrupted gaps.
8. Identify Anomalies: If the joint probability falls below the threshold (default: 0.5), mark the window as an anomalous gap and record the charger ID, joint probability, and time window.

Output: Identified anomalies including charger ID, joint probability, and corresponding time window(s) where charging behavior deviates from the expected distribution.

4.5 LSTM Autoencoder Anomaly Detection Model

The LSTM autoencoder is a deep learning approach for detecting anomalous events within time series data, based on Long Short-Term Memory networks, a type of Recurrent Neural Network (RNN) well-suited for modeling sequential data. [8] The architecture was implemented in Keras.

4.5.1 Architecture Overview

The autoencoder consists of an encoder and a decoder. The encoder LSTM reduces the dimensionality of the input sequence into a compact representation; the decoder LSTM reconstructs the original sequence from that representation. The reconstruction error, measured using Mean Squared Error (MSE), serves as the anomaly score. [12] **Figure 5** and **Figure 6** illustrate the architecture for one of the Level 2 chargers in the study.

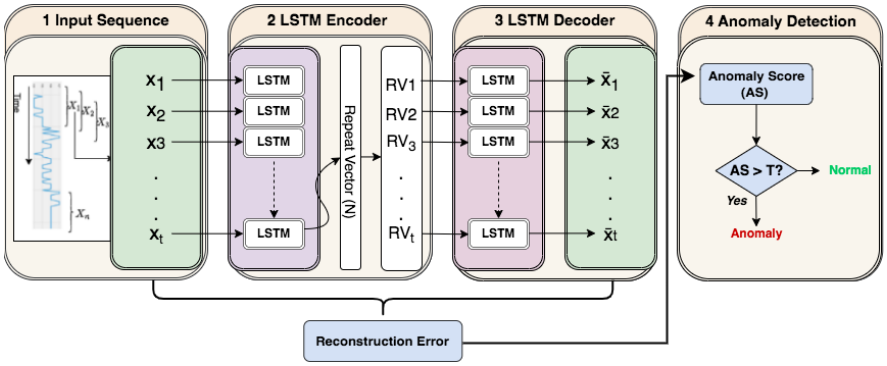


***Figure 5**: LSTM Autoencoder Architecture*

```
Training shape: (4151, 42, 1)
Testing shape: (145, 42, 1)
Model: "sequential_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_6 (LSTM)                (None, 128)               66560
_________________________________________________________________
dropout_6 (Dropout)          (None, 128)               0
_________________________________________________________________
repeat_vector_3 (RepeatVecto (None, 42, 128)           0
_________________________________________________________________
lstm_7 (LSTM)                (None, 42, 128)           131584
_________________________________________________________________
dropout_7 (Dropout)          (None, 42, 128)           0
_________________________________________________________________
time_distributed_3 (TimeDist (None, 42, 1)             129
=================================================================
Total params: 198,273
Trainable params: 198,273
Non-trainable params: 0
```

***Figure 6:** LSTM Autoencoder Architecture Detail*

4.5.2 Input Data Sequence

The dataset is structured as a succession of time sequences $[X_1, X_2, X_3, ..., X_n]$. Each sequence X_i consists of a T-length window $[x_1, x_2, ..., x_t]$, where $x_t \in \mathbb{R}^m$ denotes m input features at time t. This is transformed into a 2D array (samples × time steps). The input in this application is the hourly charger energy output, configured as a 42×1 vector (42-time step windows).

4.5.3 LSTM Encoder

The encoder functions as a sequential folding layer, transforming features into batches of time-dependent feature sequences. As shown in **Figure 7**, Layer 1 is an LSTM network with 42 cells that

sequentially process individual samples.[8] Each LSTM cell decides whether to retain or discard the preceding sample's information via gating mechanisms. The final (42nd) cell accumulates all pertinent sample information, producing a 1×128 encoded feature vector. Layer 2 is a RepeatVector layer that replicates this vector across all 42 time steps, generating 42 replicas as input to the decoder.

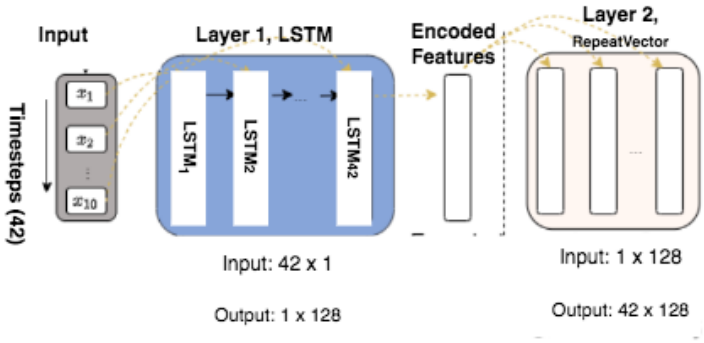


Figure 7: *LSTM-AD Encoding Layer*

4.5.4 LSTM Decoder

The decoder operates as a sequential unfolding layer, reinstating the original sequence arrangement. As shown in **Figure 8**, Layer 3 encompasses 42 LSTM units; each processes a 1×128 encoded feature set. A TimeDistributed dense layer multiplies the LSTM output (42×128) by a weight matrix (128×1) to produce a 42×1 reconstructed sequence matching the original input dimensions.[12]

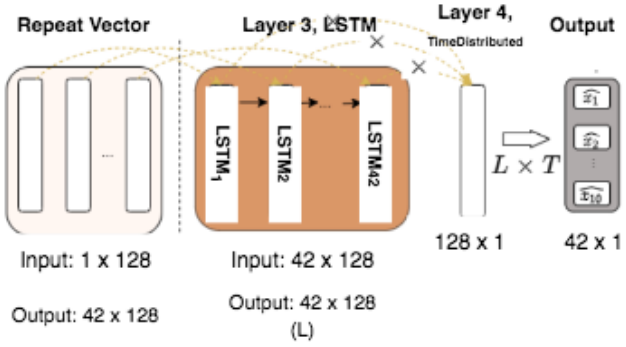


Figure 8: *LSTM-AD Decoding Layer*

4.5.5 Anomaly Detection Threshold

The model is trained on one year of per-charger data, enabling computation of baseline reconstruction error rates for normal energy output. After training, the 90th percentile MSE across training samples is set as the anomaly threshold. During testing, sequences whose reconstruction error exceeds this threshold are flagged as anomalous.

4.5.6 Training and Testing Protocol

Each charger's model was trained separately using backpropagation with the ADAM optimizer[8], for 100 epochs with a batch size of 32, with early stopping to prevent overfitting. The MSE formula is given in **Equation 1**.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\left|\widehat{Y}_i - Y_i\right|\right)^2$$

Equation 1*: Mean Squared Error Formula*

Where N is the total number of observations, i is the index of the output layer sample, $\hat{Y}_i$ is the desired vector of sample i, and Y_i is the output vector of sample i. Model performance was evaluated using precision, recall, F1-score, and AUC-ROC.

5. RESULTS

5.1 Detected Charging Usage Gaps

Figures 9–12 present detected charging usage gaps for the testing dataset of all chargers in each location. The base time series represents hourly energy usage. Colored markers indicate detected gaps; darker pink/red shades indicate higher-confidence anomalies, lighter yellow shades indicate lower confidence. Both the naïve probability distribution method and the LSTM method were applied, assuming both models had effectively learned usage patterns from the prior year of training data.

5.1.1 Residential Chargers (Oakland)

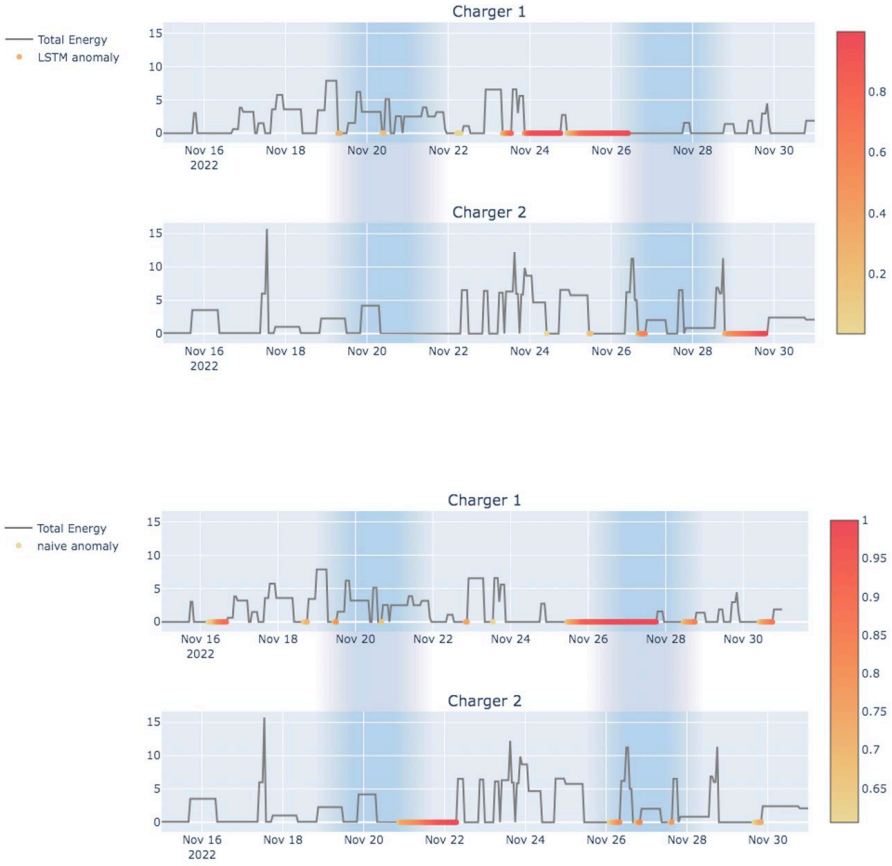


Figure 9: *Residential Charger Charging Usage Gaps*

Figure 9 illustrates detected gaps for the residential location. Charger 1 exhibited 94 detected gap hours (13.6% uptime reduction); Charger 2 exhibited 52 gap hours (7.2% reduction). Average gap length for both chargers was approximately 10 hours, with maximum gaps exceeding 30 hours.

5.1.2 Commercial Chargers (San Francisco)

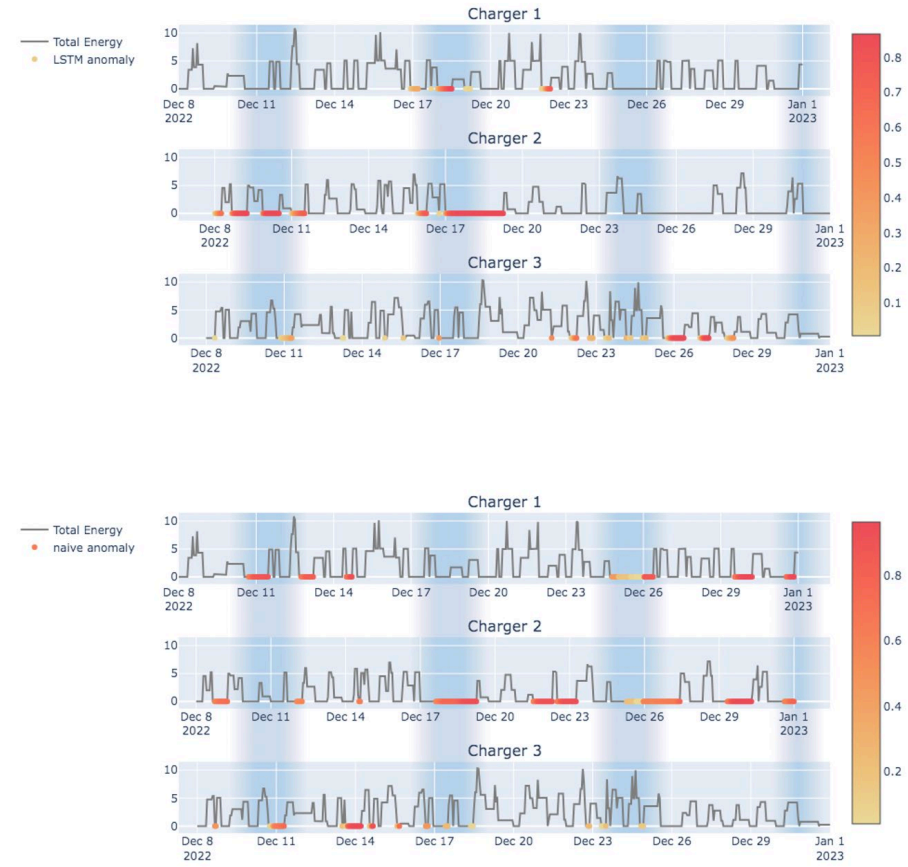


Figure 10: *Commercial Charger Charging Usage Gaps*

Figure 10 depicts detected gaps for the commercial location. Charger 1 had 101 gap hours (13.6% uptime reduction); Charger 2 had 185 hours (25% reduction); Charger 3 had 57 hours (7.7% reduction). Average gap lengths ranged from ~4 hours (Charger 3) to over 15 hours (Chargers 1 and 2), with Chargers 1 and 2 experiencing maximum gaps exceeding 30 hours.

5.1.3 Workplace Chargers (San Diego)

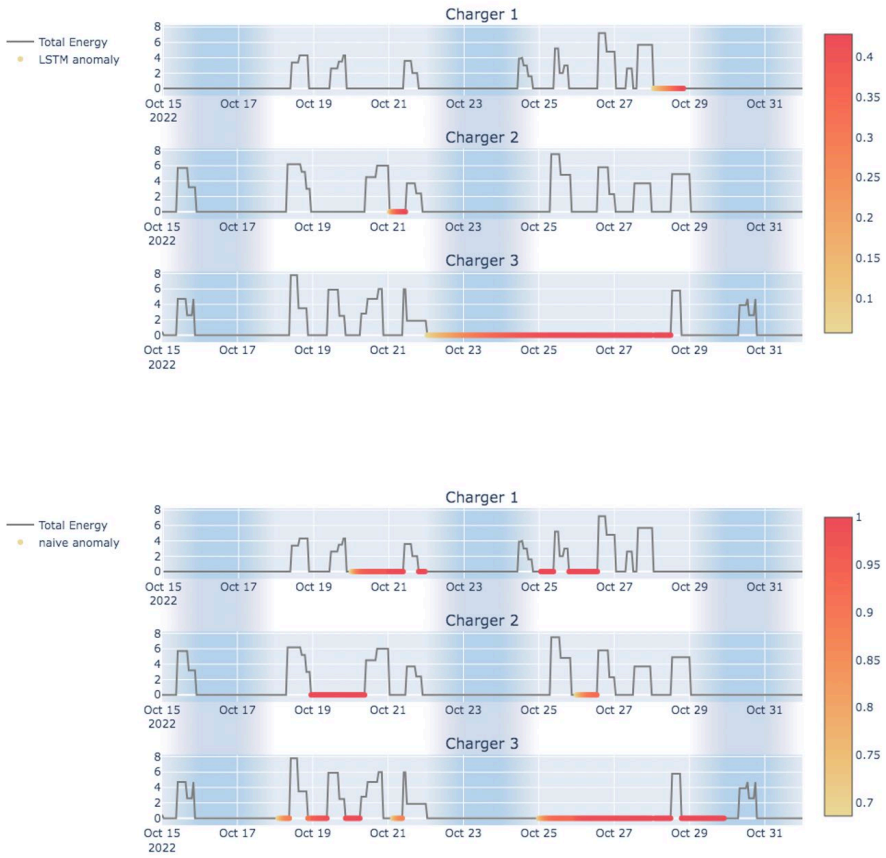


Figure 11: *Workplace Charger Charging Usage Gaps*

Figure 11 illustrates detected gaps for the workplace location. Charger 1 had 67 gap hours (9% uptime reduction); Charger 2 had 49 hours (6.6% reduction); Charger 3 had 159 hours (29% reduction). Average gap lengths ranged from ~15 hours (Charger 1) to over 20 hours (Chargers 2 and 3), with all three chargers experiencing maximum gaps exceeding 30 hours.

5.1.4 Corridor DC Fast Chargers (Northern California)

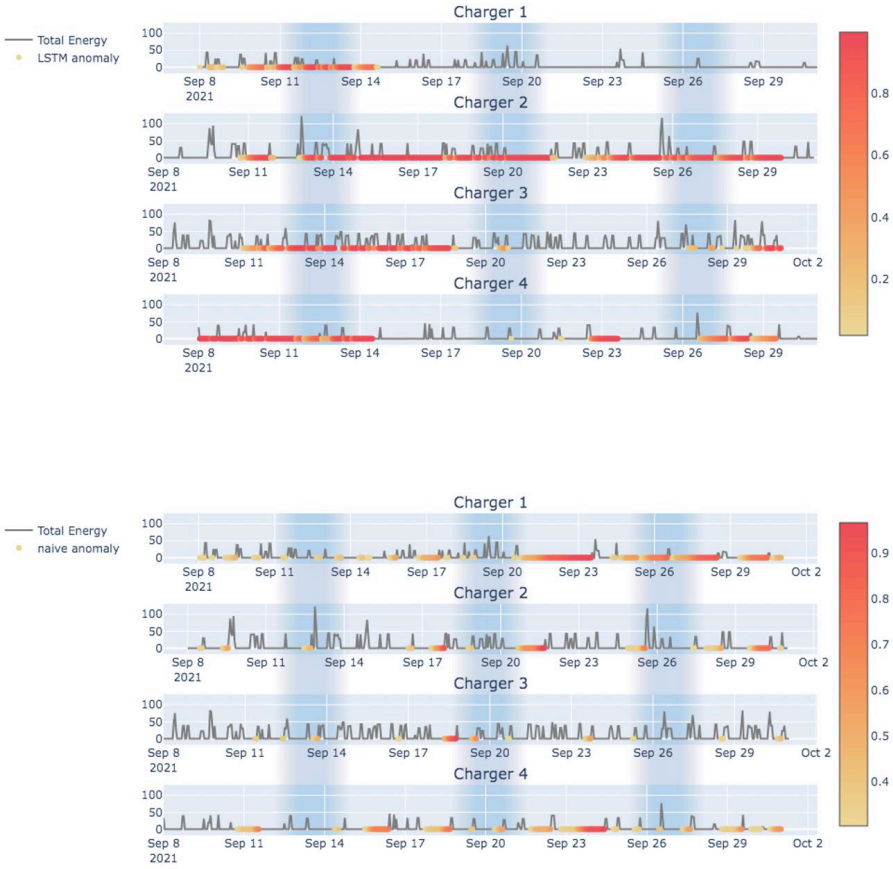


Figure 12: *Corridor DCFC Charging Usage Gaps*

Figure 12 depicts detected gaps for the corridor DCFC location. Chargers 1, 2, and 4 each exceeded 100 detected gap hours, with uptime reductions ranging from 16% to 38%. Average gap lengths remained under 15 hours, but maximum gap length varied significantly from 11 to 68 hours - underscoring the particular risk of extended outages at corridor stations where driver stranding is most likely.

5.2 Potential Benefits of Tool Deployment

Figure 13 illustrates a specific detected gap for a residential charger, with estimated repair time savings at three confidence levels, assuming a CSO responds within one hour of receiving a tool alert.

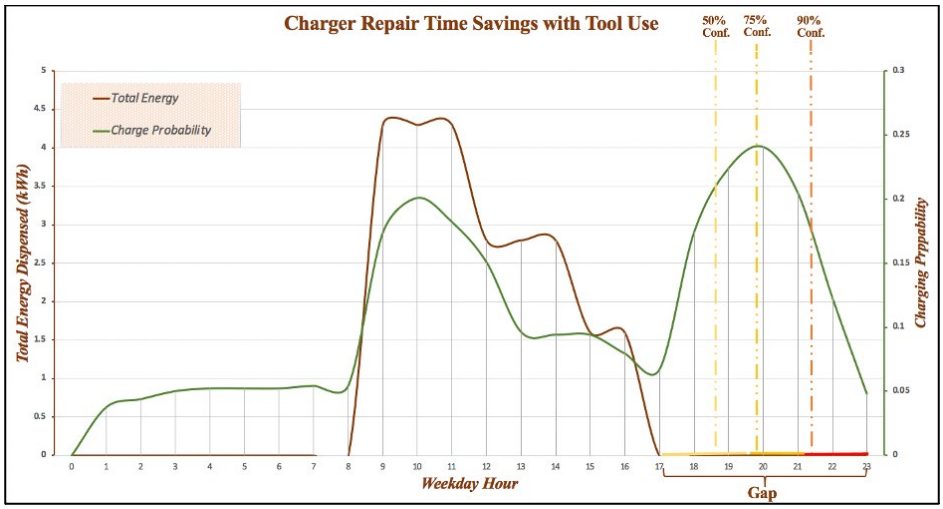


Figure 13: *Charger Repair Time Savings with Tool Use*

For this charger, a gap ran from 5 PM to 11 PM. At a 50% confidence level (joint probability > 0.5), the tool would have notified the CSO at 6:40 PM, one hour and 40 minutes into the gap, potentially reducing the total gap by 4 hours 20 minutes (72%). At 75% confidence, notification at 7:55 PM would reduce the gap by 3 hours 5 minutes (51%). At 90% confidence, notification at 9:22 PM would reduce the gap by 1 hour 38 minutes (27%).

Figure 14 shows total gap hours across all chargers by location, with and without tool use at three confidence levels. **Figure 15** shows average gap duration under the same conditions. Both analyses assume the CSO repairs the faulty charger within one hour of a tool alert.

Figure 14: *Total Detected Gap Hours With and Without Tool Use*

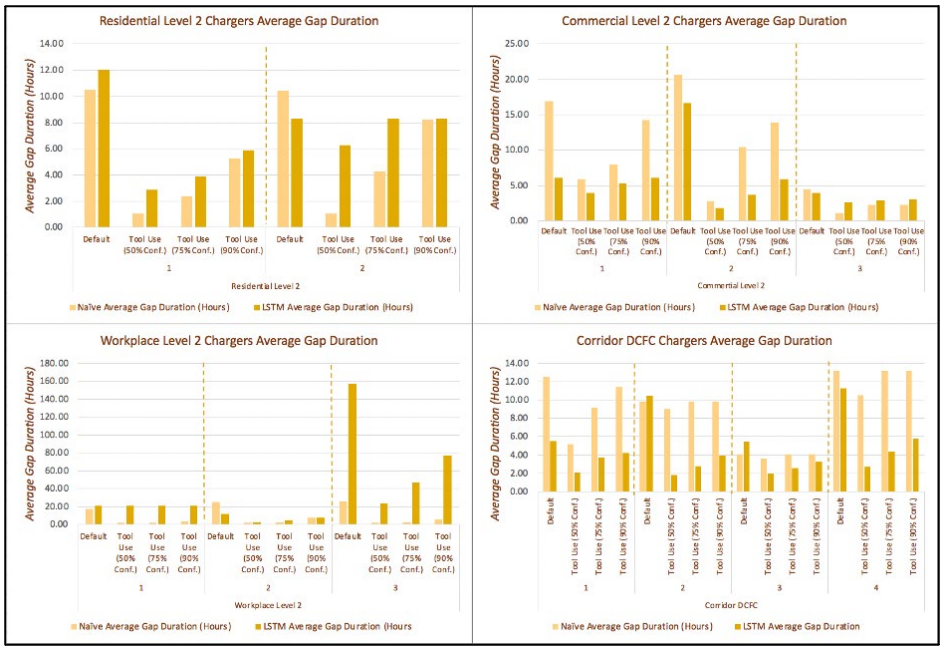


Figure 15: *Average Gap Duration With and Without Tool Use*

Table 5: *Tool Performance Summary by Confidence Level*

Metric	50% Naïve	50% LSTM	75% Naïve	75% LSTM	90% Naïve	90% LSTM
Total gap hours reduced	62%	68%	43%	50%	28%	46%
Average gap duration reduced	66%	59%	50%	45%	35%	34%
Fault detection speed-up	3×	2.4×	2×	1.8×	1.5×	1.5×

6. DISCUSSION

6.1 Tool Effectiveness

The results demonstrate that behavioral pattern analysis can meaningfully complement OCPP-based monitoring.[1, 2] The reduction in uptime resulting from detected anomalies ranges from 6% to 38%, with a mean of 16% for both methods. These are reliability deficits that would not have been captured by standard fault detection protocols.

By deploying the tool in real-time and responding to alerts within one hour, At 50% confidence, CSOs could detect gaps three times faster using the naïve method and 2.4 times faster using the LSTM method. At 75% confidence, detection speed improved to two times and 1.8 times faster, respectively. At 90% confidence, both methods enabled approximately 1.5 times faster detection. The naïve method offers superior interpretability, as CSOs can directly understand why a gap was flagged. The LSTM method offers higher gap-hour reduction at equivalent confidence levels, particularly at 90% confidence (46% vs. 28%).

6.2 Regulatory Alignment with CEC Standards

The CEC is developing uptime recordkeeping and reporting standards for EV charging stations that received public funding.[36] The standards will address charger interoperability and payment system failures prior to installation, while charger and network failures will be addressed through performance standards and monitoring.[36] Remote and physical monitoring options under consideration include operative status monitoring, random field inspections, and preventive maintenance requirements.

The tool developed in this project directly supports CSOs in meeting these emerging standards by detecting fault categories outside conventional OCPP monitoring scope, including network communication failures, blocked access, and physically broken cables.[36] The tool can be configured to send automated alerts upon anomaly detection, enabling immediate corrective action and providing an auditable record of fault detection events for compliance purposes.

6.3 Field Validation Challenges and Interim Data Approach

The original project scope called for calibration and validation of the tool using data from approximately 37 Caltrans-owned DCFC stations under the ZEV 30-30 program. However, despite repeated outreach, the research team has been unable to establish consistent contact with a dedicated team to support model evaluation through on-site testing. It remains unclear whether such a team is currently in place to facilitate this collaboration.

As an interim measure, the research team scraped approximately three months of publicly available live charger status data to generate utilization-over-time feeds. This data provides a limited basis for approximating real-time operational patterns and conducting preliminary validation of the tool's anomaly detection behavior against live charger activity. While this represents a meaningful step toward real-world validation, it is not a substitute for the formal field testing and fault-record comparison originally planned.

A minimum of approximately one year of continuous per-station data is required to adequately capture the seasonal and weekly utilization variation that the probability distribution and LSTM models depend on for reliable baseline construction. The research team will continue collecting this data on an ongoing basis and will iteratively refine model thresholds as the dataset matures. Once stakeholder access is established, the tool will be integrated with existing CSO monitoring protocols

for real-time and near-real-time data exchange, with a user-friendly interface enabling CSOs to visualize anomalies, configure alert thresholds, and log corrective actions.

6.4 Structural and Policy Recommendations

Effective fault detection alone is insufficient to resolve all charging reliability concerns. [37] Lack of stakeholder profit incentives, unclear accountability structures, and absence of performance monitoring can significantly delay fault resolution even when faults are detected. [37, 38] The following structural measures are recommended to complement the technical tool:

- Require CSOs to provide data access to hosts and third-party service providers to facilitate independent fault diagnostic and performance monitoring platforms. [36]
- Mandate roaming SIM network providers and standardized communication restoration protocols to reduce the frequency and duration of lost OCPP connectivity. [38]
- Ensure public EV charging tariff structures provide sufficient revenue to fund timely maintenance and servicing. [38]
- Require plug-and-charge capability (eliminating app or RFID card requirements) to reduce logistical barriers contributing to invisible failure modes. [38]
- Invest in training and expanding the accredited EV charging repair workforce to reduce time-to-repair for detected faults. [37]

7. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

7.1 Conclusions

Reliable and functional EV chargers are crucial for the widespread adoption of EVs and for the success of the Caltrans ZEV 30-30 program. While uptime is the most commonly used reliability metric, it fails to capture all technological and logistical challenges that determine true consumer-facing reliability.[1, 2] This study developed and demonstrated a tool that leverages habitual EV charger usage patterns to identify potential faults invisible to standard monitoring protocols.

Key conclusions:

- Both the naïve probability distribution model and LSTM autoencoder successfully detected anomalous usage gaps across all 12 California chargers, with detected anomalies corresponding to uptime reductions of 6%–38% (mean 16%).
- Real-time deployment could reduce total gap hours by 28%–68% and accelerate fault detection by 1.5–3×, depending on confidence threshold and model choice.
- Corridor DCFC chargers, the primary target of the ZEV 30-30 program, exhibited the largest maximum gaps (up to 68 hours), underscoring the value of the tool for this deployment context.
- The tool is directly applicable to emerging CEC uptime monitoring standards and supports compliance reporting for publicly funded ZEV charging infrastructure.[36]

7.2 Limitations and Field Validation Status

The anomaly detection thresholds used in this study were set heuristically rather than empirically calibrated against confirmed fault records from the ZEV 30-30 network. This may produce false positives (non-fault gaps flagged as anomalies) or false negatives (faults not flagged due to insufficient deviation from baseline). Calibration against labeled fault data remains a critical next step.

With respect to field validation, despite repeated outreach to the relevant ZEV 30-30 stakeholders, the research team has been unable to establish consistent contact and remains uncertain whether a dedicated team exists to support model evaluation through on-site testing. As a result, the formal field validation phase anticipated under the project scope of work could not be carried out as expected.

To partially address this gap, the research team scraped a limited sample of approximately three months of publicly available live charger status data to generate utilization-over-time feeds for a subset of relevant stations. This data allows for an approximation of real-time operational patterns and provides a preliminary basis for validation purposes. However, three months of data is insufficient to adequately capture the seasonal and temporal variation in utilization rates that the models rely on; a minimum of approximately one year of continuous data is needed to establish robust usage pattern baselines for each charger.

Additionally, the tool's performance on very low-utilization chargers (e.g., corridor DCFC stations with 3%–5% utilization rates) may be more variable, as sparse usage patterns yield less stable probability distributions and reduce the LSTM model's ability to learn reliable baselines.

7.3 Future Research Directions

Several enhancement and continuation directions are identified for future project phases:

- Continued live data collection and calibration: The research team will continue collecting publicly available live charger status data over an extended period. The goal is at least one year of continuous data per station to enable robust empirical calibration of detection thresholds against real operational patterns. Model parameters and anomaly thresholds will be iteratively refined as data accumulates.
- Renewed stakeholder engagement: The team will continue efforts to establish working contact with a dedicated ZEV 30-30 CSO team to enable on-site testing, access to internal fault records, and collaborative validation of tool detections against confirmed charger failure events.
- Supplementary data integration: Incorporating weather forecasts, traffic patterns, and real-time power outage maps could enable proactive identification of fault risks before they manifest in usage patterns.
- Direct hardware integration: Charging stations equipped with cable damage sensors and port malfunction detectors could feed signals directly into the tool, complementing the behavioral anomaly detection approach.
- Automated maintenance triggering: Bidirectional communication between the tool and charging infrastructure could initiate automated repair workflows upon detection, reducing the time between alert and resolution.
- Network-level anomaly detection: Extending the tool to detect correlated anomalies across multiple ZEV 30-30 corridor stations could identify systemic failures, such as grid outages or CSO network provider failures, affecting entire charging corridors simultaneously.

REFERENCES

[1] Rempel, D., Kurani, N., Turrentine, T., & Burke, A. (2022). Reliability of Open-Access, DC Fast Charging Infrastructure in the San Francisco Bay Area. UC Berkeley Transportation Sustainability Research Center.

[2] Rempel, D., et al. (2021). Electric Vehicle DC Fast Charger Reliability: A Consumer Perspective. SAE Technical Paper 2021-01-0182.

[3] Bontempi, G., Ben Taieb, S., & Le Borgne, Y.-A. (2012). Machine Learning Strategies for Time Series Forecasting. European Business Intelligence Summer School (eBISS).

[4] Brockwell, P. J., & Davis, R. A. (2002). Introduction to Time Series and Forecasting (2nd ed.). Springer.

[5] Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and Practice (3rd ed.). OTexts.

[6] Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting, 20(1), 5–10.

[7] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. Management Science, 6(3), 324–342.

[8] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

[9] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521, 436–444.

[10] Velickovic, P., et al. (2018). Graph attention networks. Proceedings of ICLR.

[11] Jaeger, H. (2001). The echo state approach to analysing and training recurrent neural networks. GMD Technical Report 148.

[12] Malhotra, P., et al. (2016). LSTM-based encoder-decoder for multi-sensor anomaly detection. Proceedings of ICML Anomaly Detection Workshop.

[13] Jolliffe, I. T. (2002). Principal Component Analysis (2nd ed.). Springer.

[14] Bouwmans, T., El Baf, F., & Vachon, B. (2008). Background modeling using mixture of Gaussians for foreground detection. Recent Patents on Computer Science, 1(3), 219–237.

[15] Shyu, M., et al. (2003). A novel anomaly detection scheme based on principal component classifier. Proceedings of ICDM.

[16] Kingma, D. P., & Welling, M. (2014). Auto-encoding variational Bayes. Proceedings of ICLR.

[17] An, J., & Cho, S. (2015). Variational autoencoder based anomaly detection using reconstruction probability. Proceedings of ICDM Workshop.

[18] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257–286.

[19] Murphy, K. P. (2002). Dynamic Bayesian networks: Representation, inference and learning. Ph.D. dissertation, UC Berkeley.

[20] Pickands, J. (1975). Statistical inference using extreme order statistics. Annals of Statistics, 3(1), 119–131.

[21] Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris, 8, 229–231.

[22] Mallat, S. (1989). A theory for multiresolution signal decomposition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), 674–693.

[23] Goldstein, M., & Dengel, A. (2012). Histogram-based outlier score (HBOS). Proceedings of German AI Conference (KI).

[24] Papamakarios, G., et al. (2021). Normalizing flows for probabilistic modeling and inference. Journal of Machine Learning Research, 22(57), 1–64.

[25] Breunig, M. M., et al. (2000). LOF: Identifying density-based local outliers. Proceedings of ACM SIGMOD.

[26] He, Z., Xu, X., & Deng, S. (2003). Discovering cluster-based local outliers. Pattern Recognition Letters, 24(9–10), 1641–1650.

[27] Tang, J., Chen, Z., Fu, A. W., & Cheung, D. W. (2002). Enhancing effectiveness of outlier detections for low density patterns. Proceedings of PAKDD.

[28] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning (2nd ed.). Springer.

[29] Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27.

[30] Schölkopf, B., et al. (2001). Estimating the support of a high-dimensional distribution. Neural Computation, 13(7), 1443–1471.

[31] Zhang, J., et al. (2019). HIFI: Anomaly detection for multivariate time series using hybrid isolation forest. Proceedings of ICDE.

[32] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. Proceedings of IEEE ICDM, 413–422.

[33] Hariri, S., Kind, M. C., & Brunner, R. J. (2019). Extended isolation forest. IEEE Transactions on Knowledge and Data Engineering.

[34] Lesouple, J., et al. (2021). Generalized isolation forest for anomaly detection. Pattern Recognition Letters, 149, 109–119.

[35] Xu, H., et al. (2017). Improving the isolation forest algorithm for anomaly detection. Proceedings of ICDE.

[36] California Energy Commission. (2023). Electric Vehicle Charging Station Uptime Recordkeeping and Reporting Standards. CEC Staff Report, Sacramento, CA.

[37] Hardman, S., et al. (2018). A review of consumer preferences of and interactions with electric vehicle charging infrastructure. Transportation Research Part D, 62, 508–523.

[38] ChargeUK / OZEV. (2022). Public EV Charging Infrastructure: Reliability Best Practices. UK Government Report.