Flagged Vehicles Pattern Analysis

Date: December 24, 2024

Dataset: OBFCM 2021-2023 PHEV data (995,511 records)

Analysis Focus: Identifying patterns, correlations, and interesting connections in flagged vehicles

---

Table of Contents

Executive Summary

Analysis of flagged vehicles reveals several significant patterns:

Filter Step	Vehicles Flagged	Key Finding
Step 1 (CS Invalid)	33,348	Dominated by Stellantis group vehicles (Fiat, Jeep, Opel) and Mitsubishi Eclipse Cross
Step 2 (Missing RW_EC)	12,554	Strongly associated with Hyundai models (Tucson, Santa Fe) and Ford vehicles
Step 4 (RW_EC Zero)	18,779	Heavily dominated by Porsche Panamera and Cayenne E-Hybrid models
Step 5 (VFN Issue)	32,363	Volvo/Polestar models and Geely vehicles overrepresented

Key Insights

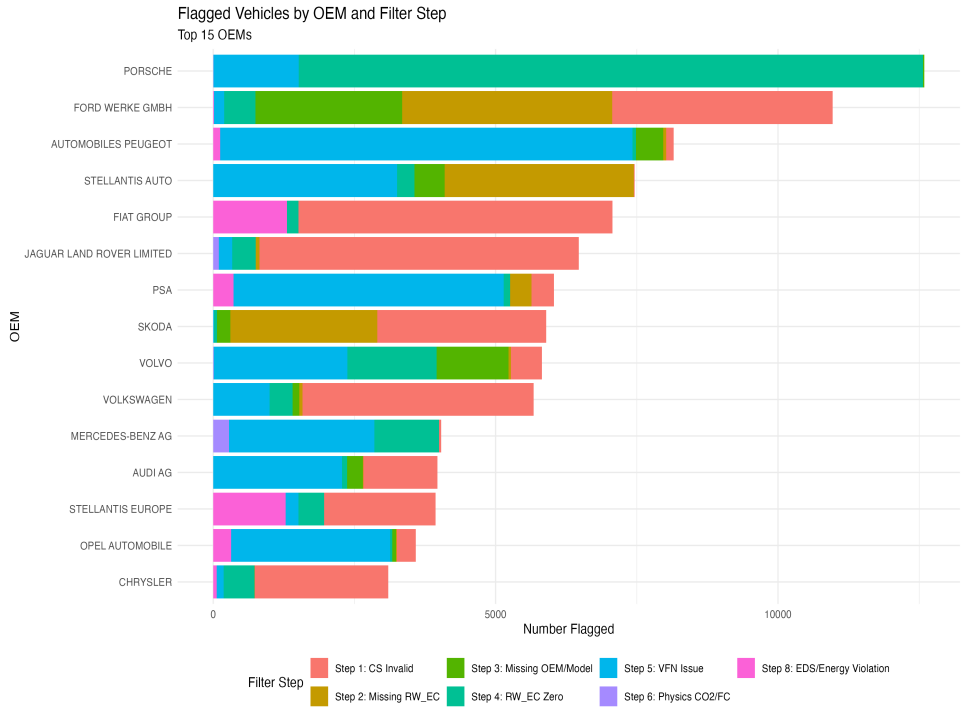

Figure: OEM breakdown by filtering step

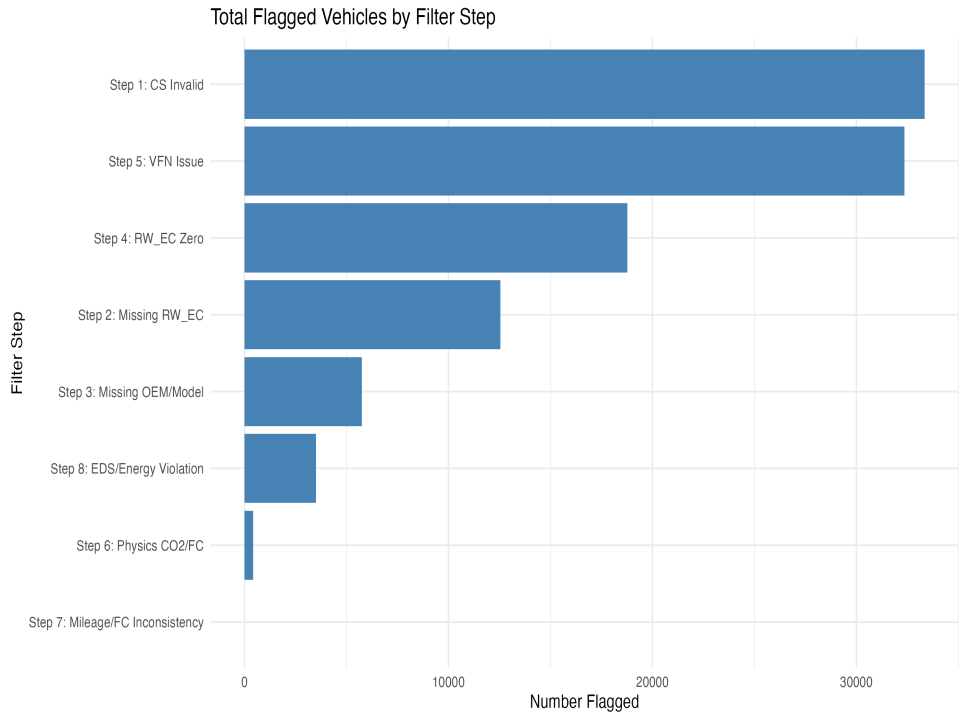

Figure: Total vehicles flagged per step

1. **Manufacturer Patterns** - Stellantis group shows consistent data quality issues across multiple steps

2. **Vehicle Characteristics** - Flagged vehicles show systematic differences in mass, engine power, and electric range

3. **Model-Specific Issues** - Certain models (Porsche Panamera, Hyundai Tucson, Mitsubishi Eclipse Cross) show extreme overrepresentation

4. **Data Quality Concerns** - Patterns suggest both genuine vehicle characteristics and potential data reporting issues

---

Detailed Findings by Filter Step

Step 1: CS Invalid (33,348 vehicles flagged)

What it means: Vehicles with invalid Charge-Sustaining (CS) mode data

Top Manufacturers

Manufacturer	Vehicles Flagged	Percentage of Flagged	Overrepresentation
Jaguar Land Rover Limited	5,654	16.95%	-
Fiat Group	5,561	16.68%	-
Volkswagen	4,091	12.27%	-
Ford Werke GmbH	3,903	11.70%	-
Skoda	2,991	8.97%	-

Top Models (by Overrepresentation)

Model	Overrepresentation
Mitsubishi Eclipse Cross	**19,345%**
Volkswagen Passat	**10,020%**
Jaguar E-PACE P300E R-Dynamic	**9,502%**
Opel Grandland X	**8,542%**

Vehicle Characteristics Comparison

Characteristic	Flagged Vehicles	Clean Vehicles	Difference	Direction
Mass	2,060 kg	2,121 kg	-2.89%	Lighter
TA_CO■	37.4 g/km	35.2 g/km	+6.10%	Higher
Electric Range	77.5 km	63.3 km	+22.32%	Longer
Engine Displacement	1,642 cc	1,863 cc	-11.84%	Smaller
Engine Power	123 kW	141 kW	-12.92%	Lower
Total Mileage	19,807 km	26,443 km	-25.10%	Lower

■ *Key Insight*

Vehicles flagged in Step 1 tend to be **lighter** and have **smaller engines**, but paradoxically show **higher CO■ emissions** and **longer electric range**. This suggests potential issues with:

• Charge-sustaining mode operation

• Data reporting in these specific models

• Possible calibration or sensor issues

■ *Online Research Context*

• **Jeep/Chrysler:** Research shows these brands have high complaint rates (1,488 complaints per car for Jeep) and recall rates (1.06 for Jeep/Chrysler), suggesting broader quality control issues that may extend to data reporting.

• Source: Carscoops - Safety Ratings Study

• **Mitsubishi Eclipse Cross PHEV:** Limited specific information found, but the extreme overrepresentation suggests model-specific issues with CS mode data collection or reporting.

Step 2: Missing RW_EC (12,554 vehicles flagged)

What it means: Vehicles missing Real-World Electric Consumption (RW_EC) data

Top Manufacturers

Manufacturer	Vehicles Flagged	Percentage of Flagged
Ford Werke GmbH	3,720	29.63%
Stellantis Auto	3,352	26.70%
Skoda	2,603	20.73%
BMW AG	597	4.76%
Hyundai Czech	590	4.70%

Top Models (by Overrepresentation)

Model	Overrepresentation
Hyundai Tucson/Tucson IX35	**216,430%**
Hyundai Santa Fe	**110,157%**

Vehicle Characteristics Comparison

Characteristic	Flagged Vehicles	Clean Vehicles	Difference	Direction
Mass	1,926 kg	2,122 kg	-9.21%	Lighter
TA_CO■	28.9 g/km	35.4 g/km	-18.20%	Lower
Engine Power	120 kW	141 kW	-14.85%	Lower
Total Mileage	18,907 km	26,314 km	-28.15%	Lower
FC_Tot (Total Fuel Consumption)	1,011 L	1,610 L	-37.18%	Lower

■ *Key Insight*

Missing RW_EC is **strongly associated** with Hyundai models (Tucson, Santa Fe) and Ford vehicles. These vehicles tend to be:

• **Lighter** than average

• Have **lower emissions**

• Have **lower total fuel consumption**

• May be **newer models** or have **different monitoring systems**

■ *Online Research Context*

• **Hyundai/Kia:** Recent recalls affecting millions of vehicles, though not specifically related to PHEV data reporting. The extreme overrepresentation of Hyundai models suggests potential issues with their OBFCM implementation or data transmission.

• Source: AP News - Hyundai/Kia Recalls

Step 4: RW_EC Zero (18,779 vehicles flagged)

What it means: Vehicles reporting zero Real-World Electric Consumption (indicating no electric mode usage)

Top Manufacturers (by Overrepresentation)

Manufacturer	Overrepresentation
Porsche	**225,325%**
Ferrari	420%
Suzuki	378%

Top Models (by Overrepresentation)

Model	Overrepresentation
Porsche Panamera 4S E-Hybrid	**4,426,839%**

Model	Overrepresentation
Porsche Panamera 4 E-Hybrid	**2,691,396%**
Porsche Panamera 4	**1,187,172%**
Porsche Cayenne E-Hybrid	**164,224%**

Vehicle Characteristics Comparison

Characteristic	Flagged Vehicles	Clean Vehicles	Difference	Direction
Mass	2,350 kg	2,115 kg	+11.11%	Heavier
TA_CO■	57.3 g/km	34.9 g/km	+64.36%	Much Higher
Electric Range	53.3 km	64.0 km	-16.75%	Shorter
Engine Displacement	2,513 cc	1,842 cc	+36.39%	Much Larger
Engine Power	213 kW	139 kW	+52.74%	Much Higher
FC_Tot (Total Fuel Consumption)	2,198 L	1,591 L	+38.20%	Higher

■ *Key Insight*

This is the **most striking pattern** - Porsche luxury PHEVs (Panamera, Cayenne) are reporting **zero electric consumption**. These are:

• **High-performance** vehicles

• **Heavy** vehicles (2,350 kg average)

• **Large engines** (2,513 cc average)

• **High power** (213 kW average)

Possible explanations for zero RW_EC:

1. ■ **Battery issues** preventing electric mode operation

2. ■ **Driver behavior** (not charging vehicles)

3. **Data reporting problems** specific to Porsche's OBFCM implementation

4. ■ **Design issues** where electric mode is rarely engaged

Step 5: VFN Issue (32,363 vehicles flagged)

What it means: Vehicle Family Name (VFN) validation failures

Top Manufacturers (by Overrepresentation)

Manufacturer	Overrepresentation
Geely	**529,641%**

Manufacturer	Overrepresentation
Ferrari	3,379%
Opel Automobile	1,619%

Top Models (by Overrepresentation)

Model	Overrepresentation
Volvo V60 T6 Twin Engine	**133,677%**
Polestar 1	**111,381%**
Volvo XC90 T8 Twin Engine	**92,057%**

■ *Key Insight*

VFN issues are **primarily** with:

• **Volvo/Polestar models** (owned by Geely)

• **Other Geely-owned brands**

This suggests:

• VFN whitelist may need updating

• Naming inconsistencies in Geely's reporting

• Corporate structure changes affecting data standardization

■ *Online Research Context*

• **Geely-Volvo Relationship:** Geely acquired Volvo Cars in 2010, explaining VFN naming inconsistencies between different reporting systems.

Cross-Step Patterns

Stellantis Group (Fiat, Opel, Peugeot, Chrysler, Jeep)

Consistent Issues Across Multiple Steps:

Filter Step	Issue Type	Affected Brands
Step 1	CS Invalid	Fiat, Opel, Jeep
Step 2	Missing RW_EC	Stellantis Auto
Step 3	Missing OEM/Model	Stellantis Auto, Peugeot
Step 8	EDS/Energy Violation	Stellantis Europe, Fiat, Opel, PSA

Analysis

The consistent data quality issues across Stellantis brands suggest:

1. **Common OBFCM implementation** across brands

2. **Shared data reporting infrastructure**

3. **Potential systemic issues** in their PHEV monitoring systems

■ *Online Research Context*

• **Stellantis Formation:** Stellantis was formed in 2021 from the merger of FCA (Fiat Chrysler Automobiles) and PSA (Peugeot Société Anonyme)

• The merger may have created integration challenges in data reporting systems

---

References

Online Research Sources

1. **Jeep/Chrysler Quality Issues:**

• Carscoops: Volvo, Subaru, Tesla Lead Study with Most 5-Star Safety Ratings

• Jeep: 1,488 complaints per car, recall rate of 1.06

2. **Hyundai/Kia Recalls:**

• AP News: Millions of recalled Hyundai and Kia vehicles with dangerous defect remain on road

3. **Automotive Recall Statistics:**

• AutoInsurance.com: Car Recall Facts & Statistics

• Ford: 94 recalls affecting 5.6M vehicles (2024-2025)

• Tesla: 20 recalls affecting 5.8M vehicles

4. **Stellantis Formation:**

• Stellantis formed in 2021 from merger of FCA (Fiat Chrysler Automobiles) and PSA (Peugeot Société Anonyme)

5. **Geely-Volvo Relationship:**

• Geely acquired Volvo Cars in 2010, explaining VFN naming inconsistencies

Data Sources

• **OBFCM dataset analysis** (2021-2023)

• **Pattern analysis** using statistical tests (t-tests, overrepresentation analysis)

• **Internal flagging statistics** from cleaning pipeline

Important Notes

> **Disclaimer:** This analysis is based on OBFCM data from 2021-2023. Patterns may reflect both genuine vehicle characteristics and data quality issues. Further investigation with manufacturers is recommended for flagged models.

Recommendations

1. **Contact manufacturers** for flagged models to verify data reporting accuracy
2. **Update VFN whitelist** to include Geely/Volvo naming variations
3. **Investigate Porsche OBFCM implementation** for zero RW_EC reporting
4. **Review Stellantis data pipeline** for systemic issues
5. **Monitor Hyundai models** for missing RW_EC data in future datasets

---

Document Version: 1.0

Last Updated: December 24, 2024

Author: Data Analysis Team

Flag Combination Visualizations

The following figures show the distribution of flag combinations across all vehicles. Flag codes represent which filtering steps were triggered:

• 1 = CS Invalid, 2 = Missing RW_EC, 3 = Missing OEM/Model, 4 = RW_EC Zero
• 5 = VFN Issue, 6 = Physics CO2/FC, 7 = Mileage/FC Inconsistency, 8 = EDS/Energy Violation
• 0 = Clean vehicles with no flags

How combinations work: Multi-digit codes combine steps. Examples: '5' = Step 5 only (VFN Issue), '12' = Steps 1+2 (CS Invalid + Missing RW_EC), '45' = Steps 4+5 (RW_EC Zero + VFN Issue), '58' = Steps 5+8 (VFN Issue + EDS/Energy Violation)

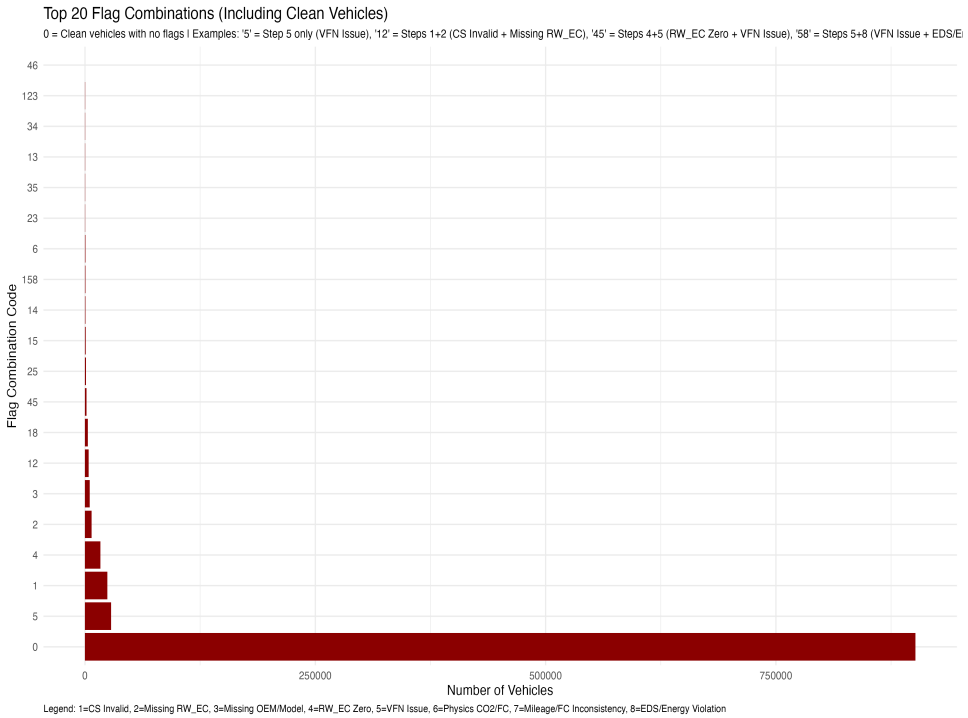

Figure: Top 20 flag combinations (including clean vehicles) - 0 represents clean vehicles. Examples shown in subtitle.

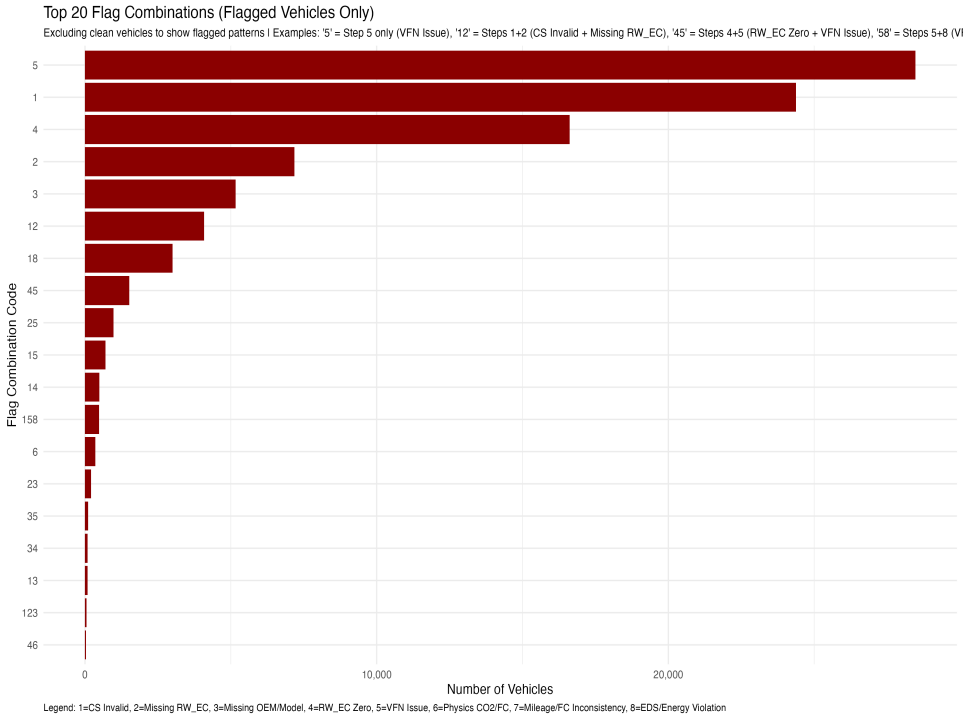

Figure: Top 20 flag combinations (flagged vehicles only) - better scale to see patterns. Examples shown in subtitle.

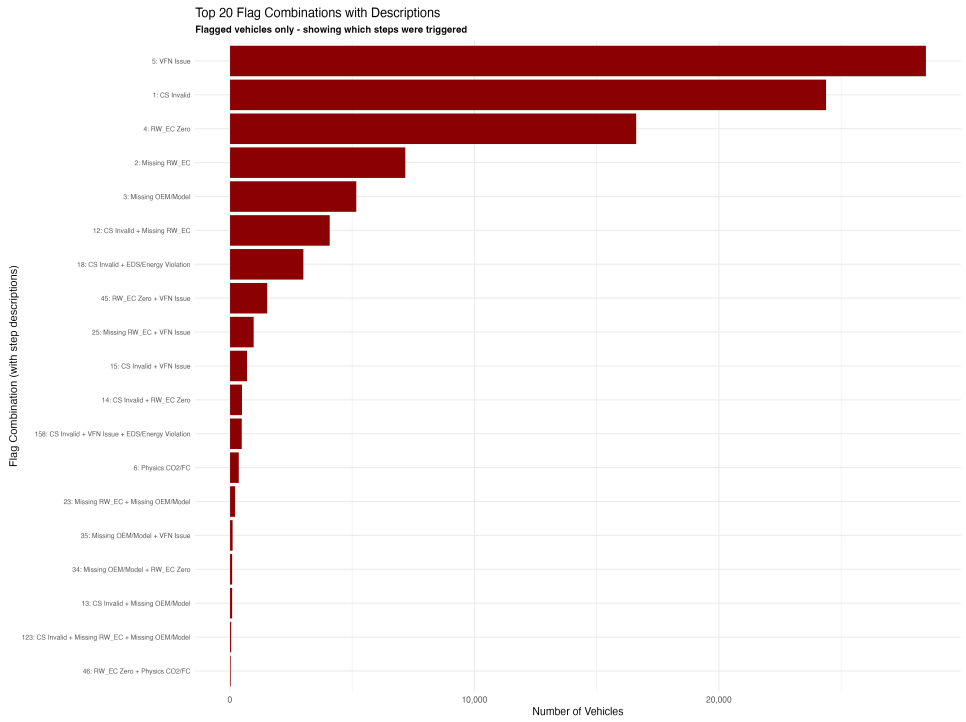

Figure: Top 20 flag combinations with full descriptions on y-axis

Additional Visualizations

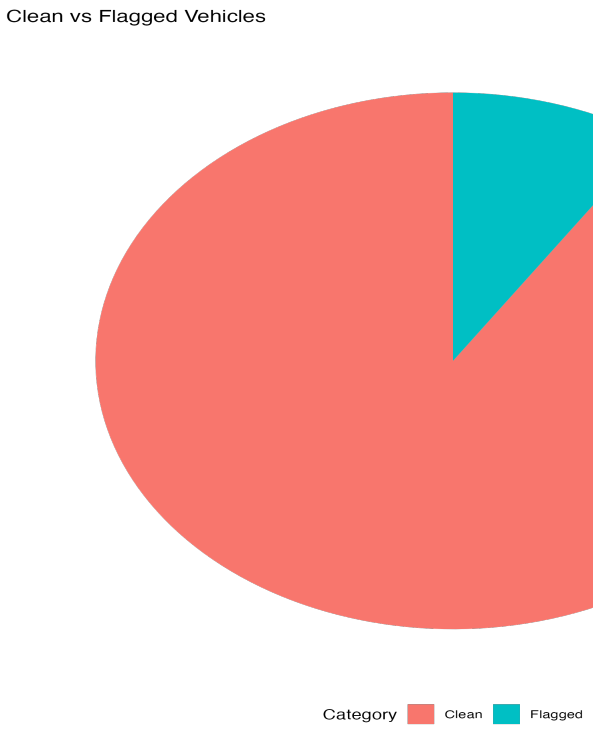

Figure: Clean vs flagged vehicle comparison