BAYESIAN VIRTUAL LAB — BMS ATTACK DETECTION REPORT

M-W Framework Physics-Layer Analysis | Generated: 20260513_105722

Parameter	Value
Detection Mode	SINGLE-FILE (auto-baseline)
Clean / Baseline	20260413_091418_VChart959.csv rows 0-39 [auto-baseline]
Suspect File	20260413_091418_VChart959.csv
Rated Capacity	100 Ah
What-If Horizon	1000 calendar days
Scenarios Analysed	S1, S2, S3

VERDICT: CLEAN **Confidence: HIGH**

No anomalous signals detected across 3 active detection layers.

Detection Layer Summary

Layer	Severity	Flags	Top Flag
L1_CAPACITY_INFLATION	CLEAN	0	—
L2_VOLTAGE_DRIFT	CLEAN	0	—
L3_IR_SUPPRESSION	INSUFFICIENT_DATA	1	Not enough current-step transitions to compute DCIR. This dataset appears to be charging-o...
L4_MANIFOLD_TRAJECTORY	CLEAN	0	—

L4 — Manifold Trajectory Inconsistency (Key Detection)

Metric	Clean / Baseline	Suspect	Physical Law
Mean Full Capacity	100.00 Ah	100.00 Ah (normal)	Must equal rated 100 Ah on healthy BMS
Pack SOH (mean FC / rated)	100.000%	100.000% (plausible)	Cannot exceed 100% — physics ceiling
Projected SOH at 1000d	94.44% (-5.56% — DECREASING)	94.44% (-5.56%)	Must decrease — batteries cannot get healthier
Trajectory Direction	DECREASING (physically correct)	Elevated vs clean	Any increase = geometric impossibility on LFP mar

VAE Status: VAE M-W manifold check skipped — physics trajectory check (Part A) is primary L4 detection

Layer-by-Layer Detail

L1_CAPACITY_INFLATION **CLEAN**

Metric	Clean / Baseline	Suspect	Note
FC Mean (Ah)	100.000	100.000	Should match rated capacity
FC Std (Ah)	0.0000	0.0000	Should be ~0 on healthy BMS
FC Max (Ah)	100.000	100.000	Nominal = 100 Ah
FC Slope (mAh/row)	0.000	-0.000	Should be ~0 (flat)

No anomalous signals detected on this layer.

L2_VOLTAGE_DRIFT **CLEAN**

Cell	Clean Slope (mV/row)	Suspect Slope (mV/row)	Delta	Mean Shift (mV)	Status
Cell 1	-0.0063	-0.0116	-0.0054	-5.071	OK
Cell 2	-0.0156	-0.0093	+0.0062	-5.119	OK

Cell	Clean Slope (mV/row)	Suspect Slope (mV/row)	Delta	Mean Shift (mV)	Status
Cell 3	-0.0049	-0.0136	-0.0088	-5.240	OK
Cell 4	-0.0057	-0.0146	-0.0088	-5.194	OK
Cell 5	-0.0134	-0.0109	+0.0025	-5.166	OK
Cell 6	-0.0015	-0.0105	-0.0090	-5.027	OK
Cell 7	+0.0005	-0.0105	-0.0110	-5.037	OK
Cell 8	-0.0076	-0.0094	-0.0018	-4.965	OK
Cell 9	+0.0000	-0.0208	-0.0208	-5.682	OK
Cell 10	+0.0000	-0.0152	-0.0152	-5.581	OK
Cell 11	-0.0109	-0.0124	-0.0015	-5.266	OK
Cell 12	+0.0000	-0.0117	-0.0117	-5.140	OK
Cell 13	-0.0078	-0.0105	-0.0028	-5.124	OK
Cell 14	-0.0051	-0.0100	-0.0049	-5.197	OK
Cell 15	-0.0020	-0.0154	-0.0135	-5.388	OK
Cell 16	-0.0068	-0.0055	+0.0013	-4.858	OK

No anomalous signals detected on this layer.

L3_IR_SUPPRESSION — INSUFFICIENT_DATA

Not enough current-step transitions to compute DCIR. This dataset appears to be charging-only. IR suppression attack targets load-side (discharge) voltage. Full DCIR resolution requires a combined charge+discharge session.

L4_MANIFOLD_TRAJECTORY — CLEAN

No anomalous signals detected on this layer.

THREAT SCENARIO ANALYSIS

M-W Framework Detection Capability vs Deployed Infrastructure

S1: Nation-State Grid Firmware Pre-Positioning

Reference Document	BVL-NationState-BESS-ThreatScenario-8Mar2026
Target	US PJM / CAISO / ERCOT / MISO grid-scale BESS (30 GW operational, 100 GW by 2030)
Attacker Profile	Chinese state-sponsored actor (Volt Typhoon playbook — CISA AA23-144A, AA24-038A)
Attack Vector	Compromised BMS vendor OTA update server; cryptographically signed firmware payload
Dormancy Period	7-30 months post-distribution; activates on geopolitical trigger or calendar date

Attack Phases / Consequences

Phase	Timeline	Defender Visibility
1 — Initial Access	Year 0	NONE — standard IT compromise, no OT indicator
2 — Payload Prep	Months 2-6	NONE — firmware diff shows minor calibration changes, passes QA
3 — Distribution	Month 7	NONE — legitimate signed OTA, normal update traffic
4 — Dormancy	Months 7-30	NONE — all telemetry clean, assets perform normally in dispatch
5 — Activation	Month 30+	NONE at activation — all systems still report normal
6 — Cascade	Days 3-21	Thermal events — too late for monitoring to prevent damage

Scale of Exposure

Metric	Estimate	Source
US utility-scale BESS (2025)	~30 GW operational	EIA Electric Power Monthly Jan 2026
US utility-scale BESS (2030)	100+ GW committed	LBNL Utility-Scale Storage Trends 2025
Chinese-origin BMS share	>60% (estimated)	DOC Section 232 Review, ODNI 2024
PJM BESS operational	~4 GW operational, 15 GW queued	PJM State of the Market 2025
Min. capacity for PJM destabilisation	~2-5 GW simultaneous	NERC BAL-003, EOP-011

Detection Gap — Why Deployed Tools Fail

Every deployed tool trusts the BMS. The BMS is the compromised component. Network tools see correctly formatted CAN frames. SCADA reads vendor-reported values. Threshold alerts never fire — max attacked voltage 3.3384 V vs 3.65 V ceiling. Firmware integrity checks pass — attack uses vendor's own stolen signing certificate (identical mechanism to SolarWinds 2020). Minimum compromised capacity for PJM destabilisation: 2-5 GW simultaneous loss (NERC BAL-003). Time to thermal cascade once activated: 3-21 days (APS McMicken 2019, Korean BESS investigations 2017-2019).

Tool / Method	Detection Approach	Why It Fails Against This Attack
Claroty / Dragos / Nozomi	Network packet inspection, OT protocol conformance, CAN/Modbus anomaly detection	All CAN frames are correctly formatted. No protocol violation occurs. The attack is entirel
SCADA / EMS Dashboards	SOH, SOC, temperature, voltage — all read from BMS-reported values	SCADA reads what the BMS reports. The BMS is the compromised component. No indep
BMS Threshold Alerts	Overvoltage (>3.65V), undervoltage (<2.8V), overtemperature	Max attacked voltage 3.3384 V — 311 mV below ceiling. No threshold breached at any p
Firmware Integrity Checks	Cryptographic verification firmware matches vendor-signed binary	Attack delivered through vendor's own update infrastructure using vendor's own signing c
Statistical Anomaly Detection	Z-score, IQR-based outlier detection on individual channels	Each individual value remains within one standard deviation of its channel distribution. Dri
Bayesian Virtual Lab (M-W)	Physics-constrained VAE on Riemannian manifold; [illegible]	DETECTS. Demonstrated on real fleet data 8 March 2026. [illegible] VERDICT [illegible] CLEAN

S2: Cloud Platform Poisoning — Urban Grid Reliability

Reference Document	BVL-Global-Infra-2026-01-CascadiaAdvisory
Target	Urban BESS fleet (Cascadia / Mumbai metro); cloud analytics platform of dominant BMS vendor

Reference Document	BVL-Global-Infra-2026-01-CascadiaAdvisory
Attacker Profile	Sophisticated state or criminal actor; target is cloud processing layer, not hardware
Attack Vector	Compromise of vendor cloud platform that processes raw BMS telemetry before forwarding to national grid control centre. Phys
Dormancy Period	24 months; attacker has real-time visibility of true pack state throughout

Attack Phases / Consequences

Stakeholder	During Dormancy	At Activation
Grid Operator	Data stream shows flawless fleet performance	Multiple units trip within hour 1 of peak dispatch
Asset Owner	Quarterly compliance reports clean	Thermal runaway — fire service response in dense urban areas
Public	City clean energy transition narrative positive	Major commercial corridors lose grid support
Grid (systemic)	Resource adequacy forecasts rely on BESS capacity	Emergency voltage reduction; peaker emergency-start

Scale of Exposure

Metric	Estimate	Source
Mumbai island grid constraint	Tata Power / Adani / BEST — constrained load pocket	MERC tariff orders 2025
Mumbai peak ambient temp	38-43 deg C summer	IMD Mumbai climatology
Indian BESS operational	500 MW+ operational, multi-GW pipeline	SECI, CERC 2025
Cloud platform audit rights	Absent — no contractual right to algorithm audit	Standard BMS vendor contracts
Detection by passive CAN tap	Possible — raw telemetry bypasses cloud layer	M-W Framework architecture

Detection Gap — Why Deployed Tools Fail

The attack occurs upstream in the vendor's cloud, before the data reaches any monitoring layer the operator controls. The M-W framework, deployed as an independent monitor reading raw telemetry directly from the site BMS via a passive tap, sits entirely outside the compromised vendor stack. It compares what the cloud platform reports against what the physics of the raw data actually shows. The divergence between the vendor's reported SOH and the M-W physics projection is the detection signal — visible weeks before activation, not hours after thermal runaway. There is no compromised hardware to find. The attack lives entirely in software the grid operator does not own.

Tool / Method	Detection Approach	Why It Fails Against This Attack
Claroty / Dragos / Nozomi	Network packet inspection, OT protocol co[illegible]	[illegible] protocol violation occurs. The attack is entirel
SCADA / EMS Dashboards	SOH, SOC, temperature, voltage — all rea[illegible]	[illegible] reports. The BMS is the compromised component. No indep
BMS Threshold Alerts	Overvoltage (>3.65V), undervoltage (<2.8V[illegible]	[illegible] — 311 mV below ceiling. No threshold breached at any p
Firmware Integrity Checks	Cryptographic verification firmware matche[illegible]	[illegible]gh vendor's own update infrastructure using vendor's own signing c
Statistical Anomaly Detection	Z-score, IQR-based outlier detection on in[illegible]	[illegible]value remains within one standard deviation of its channel distribution. Dri
Bayesian Virtual Lab (M-W)	Physics-constrained VAE on Riemannian m[illegible]	[illegible]

S3: Capacity Market Financial Attack

Reference Document	BVL-MarketAttack-EU-India-8Mar2026
Target	Grid-scale BESS enrolled in EU FCR-D / UK BM / India CERC ancillary services; capacity payments contingent on verified SOH
Attacker Profile	Nation-state actor, financially motivated criminal group, or market manipulator; financial dimension requires no grid failure — on
Attack Vector	Same firmware compromise as S1, timed to peak demand season with thin grid reserves
Dormancy Period	12-18 months; assets appear exceptionally healthy, clearing all capacity market tests

Attack Phases / Consequences

Consequence	Mechanism	Scale
1 — Capacity Failure	Assets cannot sustain rated output; multiple units trip undervoltage	Hundreds of MW disappear at peak need

Consequence	Mechanism	Scale
2 — Thermal Runaway	High discharge + high ambient + elevated real IR = thermal limit breach	EUR 40-80M asset destruction; casualty risk
3 — Market Collapse	Fast-response capacity loss during shortage drives LMP to EUR 2000+/MWh	Multi-hundred-million-euro settlement in one day
4 — Regulatory Destruction	BESS earning capacity payments for SOH they no longer possess	Multi-year regulatory freeze; project finance withdrawal

Scale of Exposure

Market	Asset Type	Capacity Revenue	SOH Requirement
EU FCR-D (Nordic)	BESS 1-100 MW	EUR 60-120k/MW/yr	Typically >80% verified quarterly
UK Balancing Mechanism	BESS 10-500 MW	GBP 50-200k/MW/yr	Performance-based, tested via dispatch
Germany FCR/aFRR	BESS 1-200 MW	EUR 40-100k/MW/yr	>80% SOH, monthly self-reported
India CERC DSM	BESS 10-250 MW	INR 15-40L/MW/yr	SOH verification framework under development

Detection Gap — Why Deployed Tools Fail

The entire revenue stream is contingent on reported SOH staying above the contractual threshold. An attacker who inflates reported SOH keeps a degraded asset earning full capacity payments for performance capacity that physically no longer exists. A 200 MW BESS portfolio enrolled in Nordic FCR-D earns approximately EUR 16-24 million per year in capacity payments. The financial dimension is the most insidious consequence — a grid failure recovers in hours, but a market manipulation event that destroys confidence in BESS SOH reporting could set back grid-scale storage deployment by a decade.

Tool / Method	Detection Approach	Why It Fails Against This Attack
Claroty / Dragos / Nozomi	Network packet inspection, OT protocol conformance, CAN/Modbus anomaly detection	All CAN frames are correctly formatted. No protocol violation occurs. The attack is entirel
SCADA / EMS Dashboards	SOH, SOC, temperature, voltage — all read from BMS-reported values	SCADA shows what the BMS reports. The BMS is the compromised component. No indep
BMS Threshold Alerts	Overvoltage (>3.65V), undervoltage (<2.8V), overcurrent, overtemperature	Max attacked voltage 3.3384 V — 311 mV below ceiling. No threshold breached at any p
Firmware Integrity Checks	Cryptographic verification firmware matches vendor-signed binary	Attack delivered through vendor's own update infrastructure using vendor's own signing c
Statistical Anomaly Detection	Z-score, IQR-based outlier detection on individual channels	Each modified value remains within one standard deviation of its channel distribution. Dri
Bayesian Virtual Lab (M-W)	Physics-constrained VAE on Riemannian manifold; trajectory consistency score, latent physics layers (IC, drift, DCIR)	DETECTS. Demonstrated on real fleet data 8 March 2026. VERDICT: CLEAN

THE DETECTION PRINCIPLE

Why an Attacker Who Controls the BMS Cannot Control the Physics

A battery pack cannot get healthier as it ages. This is not a heuristic or a threshold — it is a consequence of irreversible electrochemical processes: SEI layer growth, lithium plating, active material loss, electrolyte decomposition. No operating condition, calibration update, or firmware version produces an increase in State of Health over time in a real LFP cell.

When the Bayesian Virtual Lab projects a suspect pack's State of Health forward on the Riemannian degradation manifold learned from real fleet data, and observes either (a) a reported Full Capacity above the rated pack capacity, or (b) a projected SOH trajectory that diverges upward from the clean baseline trajectory, the system flags the telemetry as geometrically inconsistent with any valid LFP degradation geodesic.

This detection is independent of the BMS vendor, firmware version, communication protocol, CAN frame format, or network architecture. It operates on the physics of the battery, not the behaviour of the software reporting it. An attacker who can compromise BMS firmware cannot compromise the laws of electrochemistry.

Proof of Detection — 8 March 2026

Metric	Clean Telemetry	Attack-Manipulated Telemetry	Verdict
Status at baseline	16 Normal, 0 Critical	16 Normal, 0 Critical	Attack invisible at snapshot level
Mean FC (reported)	100.00 Ah (= rated)	100.00 Ah (within range)	Marginal
Projected SOH at 1000d	94.44%	94.44%	Trajectory diverges
Threshold alert triggered	N/A	NOT DETECTED — max voltage 3.3384 V, no threshold breached	All thresholds blind
M-W physics detection	N/A — this is ground truth	DETECTED — CLEAN (HIGH)	Physics-layer catches what network tools cann