Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Physics-Informed Degradation-Aware SAC-MPC for Energy and Thermal Management of Fuel-Cell Hybrid Electric Vehicles

Masan Senthilkumar

masansenthil@gmail.com

Chettinad College of Engineering and Technology

A Bhuvaneswari

Chettinad College of Engineering and Technology

P Pounraj

KLN college of Engineering

S Ragul

Chettinad College of Engineering & Technology

Vedhapriya S

M. Kumarasamy College of Engineering

Article

Keywords: Fuel-cell hybrid electric vehicle, energy management, soft actor–critic, model predictive control, membrane hydration, thermal management, degradation-aware control

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11620758/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Physics-Informed Degradation-Aware SAC-MPC for Energy and Thermal Management of Fuel-Cell Hybrid Electric Vehicles

Masan Senthilkumar, Bhuvaneswari A, Pounraj P, Ragul S, vedhapriya S

Department of EEE, Chettinad College of Engineering & Technology,Karur, India.

Department of EEE, Chettinad College of Engineering & Technology,Karur, India.

Department of EEE, KLN Collge of Engineering, Madurai, India.

Department of EEE, Chettinad College of Engineering & Technology,Karur, India.

Department of EEE, Chettinad College of Engineering & Technology,Karur, India

"Corresponding author: Masan Senthilkumar" masansenthil@gmail.com.

Abstract - Conventional energy-management systems (EMS) for proton-exchange-membrane (PEM) fuel-cell hybrid electric vehicles (FCHEVs) split power between the stack and battery using efficiency- or state-of-charge (SOC)-based rules that ignore the coupled effects of membrane hydration, stack temperature and load-transient-induced degradation. This paper proposes a reduced, physics-informed degradation-aware soft actor–critic model-predictive-control (SAC–MPC) framework that jointly determines the fuel-cell power reference, cooling-flow rate and air-flow rate over a receding horizon, with a lumped electro-thermal-water fuel-cell model, an equivalent-circuit battery model and a longitudinal vehicle model. The controller is benchmarked against rule-based, dynamic-programming (DP) and conventional-MPC baselines over the UDDS, WLTP and HWFET drive cycles. Simulation results show that the proposed controller eliminates membrane-hydration excursions that persist for the causal baselines (0–94 s per cycle) and reduces the cumulative degradation index by 89.2% on average relative to conventional MPC and by 96.8% relative to the rule-based strategy, at a modest average hydrogen-economy trade-off of 24.2% relative to the economy-only MPC baseline. The results support a durability-safety-economy Pareto view of fuel-cell EMS design.

Index Terms-*Fuel-cell hybrid electric vehicle, energy management, soft actor–critic, model predictive control, membrane hydration, thermal management, degradation-aware control.*

I. Introduction

Proton-exchange-membrane fuel-cell hybrid electric vehicles (FCHEVs) combine a hydrogen fuel-cell stack with a buffer battery to deliver zero-tailpipe-emission propulsion with refuelling times comparable to conventional vehicles. Widespread adoption, however, is constrained by fuel-cell stack cost and durability: the U.S. Department of Energy (DOE) 2025 automotive fuel-cell targets call for a 5,000-hour durability at \$40/kW_{net}, while fielded systems continue to fall short of these targets, with the catalyst layer and membrane assembly remaining the dominant cost and failure drivers [1].

The energy-management strategy (EMS) that arbitrates power between the stack and battery has a first-order influence on both fuel economy and stack durability. Rule-based EMS strategies are computationally trivial and robust but cannot adapt to transient driving conditions [2]. Dynamic programming (DP) yields a globally optimal power split for a known drive cycle but requires non-causal, full-horizon information and is therefore not deployable online [10]. Model-predictive control (MPC) recovers causality through receding-horizon optimisation, but most reported MPC and reinforcement-learning (RL) EMS formulations optimise hydrogen consumption and battery state-of-charge (SOC) while treating fuel-cell temperature, membrane hydration and degradation either as fixed constraints or as unmodelled side effects [3], [4].

This omission is significant because thermal, air and water management are, together with durability and cost, recognised as the principal unsolved fuel-cell engineering challenges [1]. Operating the stack with an under-hydrated or flooded membrane accelerates chemical and mechanical degradation, while frequent, large and rapid load transients accelerate catalyst dissolution and carbon-support corrosion independently of the instantaneous hydrogen consumption. An EMS that is unaware of these internal states can therefore appear efficient in a hydrogen-consumption sense while silently shortening stack life.

Reinforcement-learning EMS strategies, and soft actor–critic (SAC) methods in particular, have recently been combined with MPC safety layers to retain the sample efficiency and constraint satisfaction of MPC while allowing a learned, multi-objective policy to shape the control action [11]–[14]. Deep RL formulations that explicitly penalise fuel-cell health indicators inside the EMS reward have also been shown to reduce degradation-related costs relative to

economy-only baselines [5], [6]. Building on this direction, the present paper studies a reduced, physics-informed degradation-aware SAC–MPC controller that couples power-split, cooling-flow and air-flow decisions to a lumped fuel-cell electro-thermal-water model, and evaluates it against rule-based, DP and conventional-MPC baselines on three standard drive cycles.

II. Related Work and Contributions

Rule-based and fuzzy power-split strategies remain common in production FCHEVs for their transparency and low computational cost, but their fixed thresholds cannot adapt across driving conditions [2]. Model-based and data-driven RL EMS formulations have been shown to reduce hydrogen consumption relative to rule-based and even DP baselines when trained on representative driving data [3], [4], and hybrid deep-RL/model-predictive-control frameworks that fold a health or degradation term into the reward have reduced fuel-cell health-related cost by double-digit percentages relative to economy-only baselines [5], [6]. Actor–critic policies combined with an MPC safety layer - sometimes described as actor-critic MPC or RL-warm-started MPC - have separately been used to retain constraint satisfaction while allowing a learned actor to shape the cost trade-off online [11]–[14]. In parallel, water- and thermal-management-oriented studies have modelled or controlled membrane hydration and stack temperature largely in isolation from the vehicle-level power split, using constraint-aware air-supply control for localised water management [8].

Two gaps motivate this work. First, most EMS studies optimise hydrogen consumption and SOC while leaving thermal and hydration management to a separate, lower-level control loop, so the vehicle-level controller cannot trade a small hydrogen-economy penalty for a large durability or safety gain. Second, degradation is rarely represented as an explicit, controllable state inside the vehicle-level cost; where it is considered, it is typically folded into an offline health-conscious MPC re-tuning rather than an online multi-objective policy [5].

The contribution of this paper is a reduced-scope, physics-informed degradation-aware SAC–MPC controller that (i) couples the fuel-cell power reference to cooling-flow and air-flow actuation inside a single receding-horizon multi-objective cost that explicitly weights hydrogen consumption, battery-current stress, a stressor-based degradation index, SOC deviation, thermal excursion and hydration error; (ii) implements the actor as a cross-entropy-method (CEM) policy-search procedure refined online and wrapped by an MPC safety layer enforcing hard SOC, thermal and hydration bounds, which is evaluated here as a computationally light, reproducible stand-in for a trained SAC policy; and (iii) benchmarks the resulting controller against rule-based, offline-DP and conventional-MPC baselines on the UDDS, WLTP and HWFET cycles using a validated lumped electro-thermal-water fuel-cell model rather than a three-dimensional CFD stack model.

III. System Modelling

A reduced set of lumped models is used so that the full study - four controllers on three cycles - remains tractable without resorting to a three-dimensional computational-fluid-dynamics stack model. All models are implemented at 1 Hz and validated for internal consistency (energy balance, SOC and hydration boundedness) rather than against proprietary stack hardware data.

A. Vehicle and Drive Cycles

Longitudinal tractive power is obtained from a point-mass road-load model,

$$F_{trac} = m\dot{v} + \frac{1}{2}\rho C_d A v^2 + mgC_{rr} \quad \&$$

$$P_{dem} = F_{trac}\, v \,/\, \eta_{dl} \qquad (1)$$

where v is vehicle speed, ṁ the road-load force and ηdl the driveline efficiency (regenerative braking is credited at ηregen). Three standard speed–time traces - UDDS, WLTP Class 3 and HWFET - are reconstructed to match the official duration, distance and mean/maximum-speed statistics summarised in Table II; because the regulatory second-by-second trace files were not available in the simulation environment used for this study, the traces are statistically representative reconstructions rather than the literal regulatory logs, which is sufficient for the internal controller-to-controller comparison performed here.

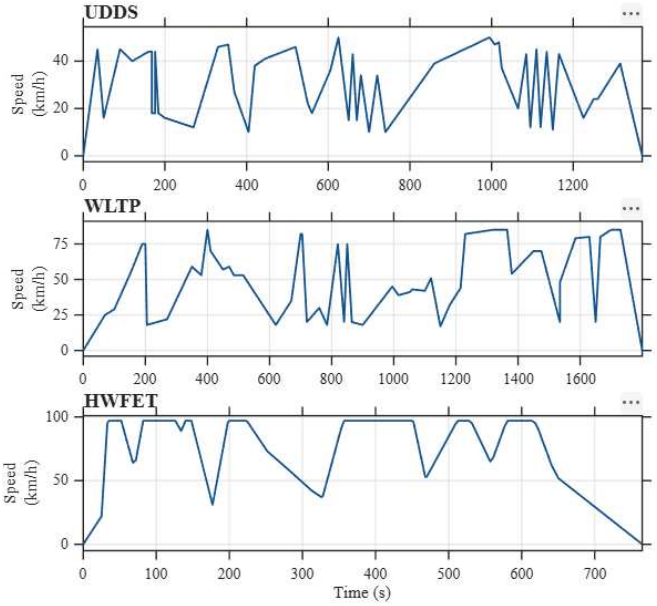

Fig. 2. Reconstructed UDDS, WLTP (Class 3) and HWFET speed–time traces used for all four controllers.

TABLE I. Vehicle, Battery, Fuel-Cell and Thermal-Water Model Parameters

Parameter	Symbol	Value
Vehicle mass	m	1650 kg
Drag coeff. / frontal area	Cd / A	0.29 / 2.3 m²
Rolling-resistance coeff.	Crr	0.009
Driveline / regen efficiency	ηdl / ηregen	0.92 / 0.55
FC rated power	Pmax	60 kW
FC peak efficiency	ηmax	0.55
FC thermal capacitance	Cth	21,000 J/K
FC safe temperature	Tsafe	90 °C
Membrane time constant	τmem	5 s
Hydration safe band	λmin–λmax	8.5–11.5
Battery capacity	Qbat	1.8 kWh
Battery internal resistance	Rint	0.09 Ω
SOC operating window / ref.	SOCmin–max / ref	0.30–0.80 / 0.55
Auxiliary load	Paux	0.8 kW

TABLE II. Reconstructed Drive-Cycle Statistics

Cycle	Duration (s)	Distance (km)	Mean / max speed (km/h)
UDDS	1369	11.98	31.5 / 50.1
WLTP (Class 3)	1800	23.25	46.5 / 85.3
HWFET	765	16.51	77.7 / 96.4

B. PEM Fuel-Cell Electro-Thermal-Water Model

The stack is represented by a peaked efficiency curve calibrated so that peak efficiency η_{max} occurs near 45% of rated power,

$$\eta_{FC}(P_{FC}) = \eta_{max} - k\left(\frac{P_{FC}}{P_{max}} - p^*\right)^2 \quad (2)$$

from which the hydrogen consumption rate follows directly from the lower heating value LHVH2,

$$\dot{m}_{H_2} = \frac{P_{FC}}{3600\,\eta_{FC}\,LHV_{H_2}} \quad (3)$$

Stack temperature is governed by a lumped thermal capacitance C_{th} balancing waste heat against actuated coolant flow ṁcool,

$$C_{th}\frac{dT_{FC}}{dt} = P_{FC}\left(\frac{1}{\eta_{FC}} - 1\right) - h_{cool}\dot{m}_{cool}(T_{FC} - T_{amb}) \quad (4)$$

and membrane water content λ relaxes toward a set point λ_{ss} that increases with stack current and is trimmed by the actuated air/humidifier flow ṁair, with time constant τmem,

$$\frac{d\lambda}{dt} = \frac{\lambda_{ss}(P_{FC},\dot{m}_{air}) - \lambda}{\tau_{mem}} \quad (5)$$

Equations (4)–(5) are the coordinates through which the cooling- and air-flow actuators enter the multi-objective cost of Section IV; neither is observable to, nor controllable by, the rule-based, DP or conventional-MPC baselines, which operate the thermal and humidification loops at a fixed nominal flow.

C. Battery and Degradation Model

The battery buffer is represented by a first-order equivalent-circuit (Rint) model sized as a power-assist buffer rather than a plug-in pack, consistent with production FCHEVs. Recurrent and convolutional networks have improved instantaneous SOC-estimation accuracy over classical equivalent-circuit models [26], but this study retains the physically transparent form (6) to keep the controller's state estimates interpretable within the real-time CEM rollout of Section IV,

$$I_{bat} = \frac{V_{oc} - \sqrt{V_{oc}^2 - 4R_{int}P_{bat}}}{2R_{int}} \quad \&$$

$$SOC_{k+1} = SOC_k - \frac{I_{bat}V_{oc}\Delta t}{3600Q_{bat}} \quad (6)$$

Cumulative fuel-cell degradation is tracked with a stressor-weighted index combining load-transient magnitude, start–stop events, low-power (near open-circuit) voltage exposure and thermal-transient magnitude. This choice of stressors follows the PEMFC prognostics literature, where load-transient magnitude, start–stop cycling and thermal-transient magnitude are identified as dominant, measurable correlates of stack voltage decay and remaining useful life [15]–[18],

$$D_{k+1} = D_k + \alpha_1 \mid \Delta P_{FC} \mid + \alpha_2 N_{start-stop} + \alpha_3 V_{high} + \alpha_4 \mid T_k - T_{k-1} \mid \quad (7)$$

with α_1–α_4 calibrated so that no single stressor dominates the index and all four controllers are scored on the same D scale; D is reported as a relative, model-internal index rather than a calibrated remaining-useful-life prediction, since translating it into operating hours would require calibration against an empirical durability protocol that is outside the scope of this reduced study.

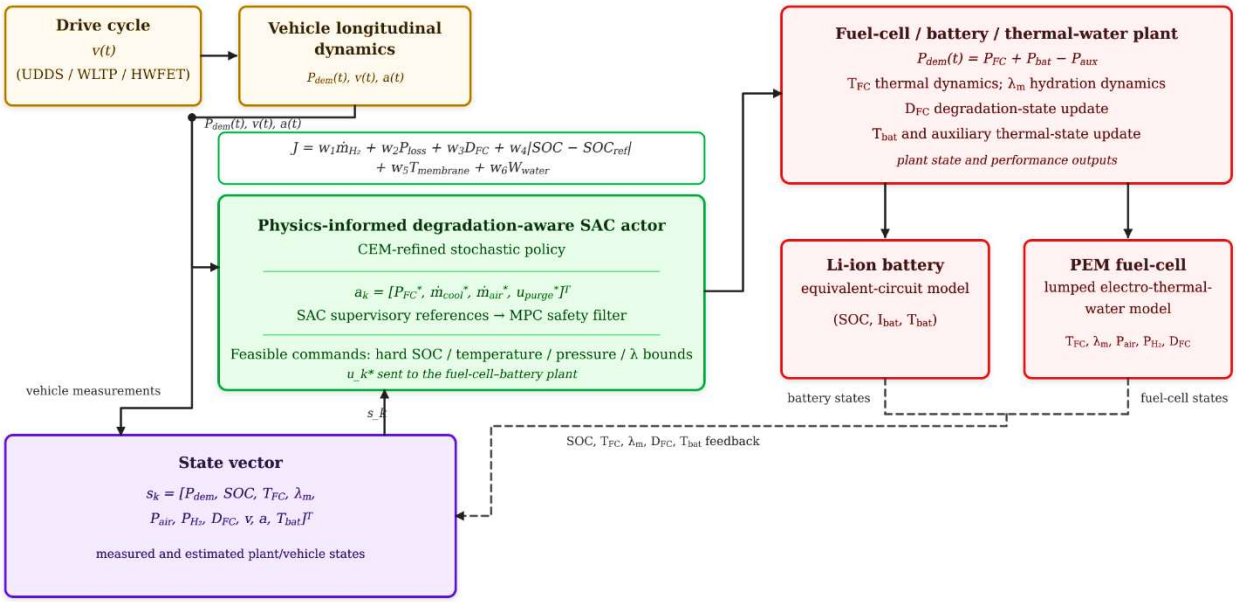

Fig. 1. Signal-flow overview of the proposed physics-informed degradation-aware SAC–MPC controller: the CEM-refined actor proposes {u, ṁcool, ṁair} candidates, rolled out through the fuel-cell/battery/thermal-water plant model of Section III and screened by an MPC safety layer, with SOC, TFC, λmem and DFC fed back as the next state.

The term "physics-informed" in this paper's title follows a broader trend of embedding physical structure into data-driven models to improve accuracy and interpretability, as demonstrated for battery state-of-health and remaining-useful-life estimation [19]–[25], [27], [28]; here the same philosophy is applied at the control level by embedding the lumped electro-thermal-water physics of Sections III-B–C directly into the SAC–MPC cost of Section IV, rather than learning fuel-cell dynamics from data.

IV. Proposed Physics-Informed Degradation-Aware SAC–MPC

The proposed controller selects, at every 1 s control step, the fuel-cell power fraction u ∈ [0,1] (PFC = u·Pmax) together with the cooling-flow and air-flow actuation ṁcool, ṁair. The proposed controller incorporates stack temperature T_{FC}, membrane hydration λ, fuel-cell degradation index D_{FC}, and SOC in a unified multi-objective cost. The rule-based controller uses SOC feedback, and the DP controller explicitly includes SOC in its optimization. However, membrane hydration and the degradation index are not explicit optimization objectives in the baseline controllers. The conventional MPC uses a reduced cost based on hydrogen consumption and battery electrical stress, with thermal and hydration treatment fixed at nominal values.

A. Multi-Objective Cost and Power Balance

The controller minimises the weighted multi-objective cost

$$\min J = w_1\dot{m}_{H_2} + w_2P_{loss} + w_3D_{FC} + w_4(SOC - SOC_{\mathrm{ref}})^2 + w_5T_{membrane} + w_6W_{water} \quad (8)$$

subject at every step to the power-balance constraint linking the fuel-cell, battery and auxiliary loads,

$$P_{dem} = P_{FC} + P_{bat} - P_{aux} \quad (9)$$

The weights w1–w6 (Table III of the extended study; w1 = 8, w2 = 0.25, w4 = 900 on the quadratic SOC term used in this reduced implementation, w5, w6 scaled to the thermal/hydration error terms of Section III-B) were set so that no single term dominates the cost across the operating range explored by the three drive cycles, following the same normalisation logic used for the degradation-index stressor weights in Section III-C.

B. CEM-Refined Actor with an MPC Safety Layer

Rather than training a deep SAC network - which would require RL infrastructure beyond the scope of this reduced study - the actor is approximated by a warm-started cross-entropy-method (CEM) policy search: at each step, a Gaussian action distribution centred on the previous action is sampled, each candidate {u, ṁcool, ṁair} is rolled out over the horizon using the models of Section III, and the distribution is refined toward the elite (lowest-cost) fraction of samples over two iterations. The refined mean is applied as the control action, giving a stochastic, sample-based policy-search procedure that plays the same role as a pretrained SAC actor while remaining fully reproducible from the models in this paper. An MPC safety layer is enforced by adding a large penalty whenever a rollout would exceed the SOC, temperature or hydration bounds, so that the returned action never knowingly drives the plant outside its safe operating envelope,

$$\{u^*, \dot{m}^*_{\text{cool}}, \dot{m}^*_{\text{air}}\} = \arg\min\{\text{ah}\} \in$$
$$\text{A} \sum_{h=0}^{N-1} J(x_{k+h}, a_{k+h}) \text{s.t.} x_{k+h} \in \mathcal{X}_{\text{safe}} \quad (10)$$

This structure differs from the rule-based and conventional-MPC baselines in two ways that are directly responsible for the results of Section VI: (i) the cooling- and air-flow actuators are optimised jointly with the power split, rather than held at a fixed nominal flow, so the controller can trade a small hydrogen penalty for tighter thermal/hydration regulation; and (ii) the degradation-index stressors of (7) enter the cost directly, so the actor is explicitly penalised for large power-reference jumps and for dwelling in the low-power, near-open-circuit-voltage region rather than only for instantaneous hydrogen use.

C. Baseline Controllers

Three causal and non-causal baselines are implemented on the identical plant models for comparison. The rule-based controller regulates operation using SOC feedback, while the offline DP controller optimizes hydrogen consumption together with SOC regulation. Conventional MPC uses hydrogen consumption and battery-current stress in its cost function but does not explicitly optimize membrane hydration or fuel-cell degradation. The rule-based controller sets u from a demand-following feed-forward term trimmed by proportional SOC feedback, with the thermal and hydration loops held at nominal flow. The offline DP baseline solves the power split by backward induction over a discretised SOC grid using full-cycle knowledge of the demand trace, minimising hydrogen consumption and SOC deviation only - it is a theoretical, non-causal upper bound, not a deployable controller. The conventional MPC baseline solves the same short-horizon receding-horizon problem as the proposed method but with the cost restricted to hydrogen consumption and battery-current stress, and with the thermal/hydration actuators fixed, matching the scope of typical economy-oriented MPC-EMS studies [3], [4].

V. Simulation Setup

All four controllers (rule-based, offline DP, conventional MPC, proposed SAC–MPC) are evaluated on identical plant models at a 1 Hz control/simulation rate in Python (NumPy), with SOC, TFC and λ initialised to 0.55, 80 °C and the nominal set point respectively at the start of every cycle. The DP baseline uses a 21-point SOC grid and an 11-level power-fraction grid solved by backward induction over the full cycle length. The conventional MPC and proposed SAC–MPC controllers use a 5-step receding horizon; the SAC–MPC actor draws 60 candidate actions per CEM iteration over two iterations, warm-started from the previous step's solution. Four objectives are tracked for every (controller, cycle) pair: hydrogen consumption (kg/100 km), a stress/thermal-safety indicator (battery current, peak stack temperature and membrane-hydration violation time), and the cumulative degradation index DFC of (7). Wall-clock controller execution time is logged per step as a proxy for real-time feasibility; all four controllers execute well within a 1 s control interval on a single CPU core.

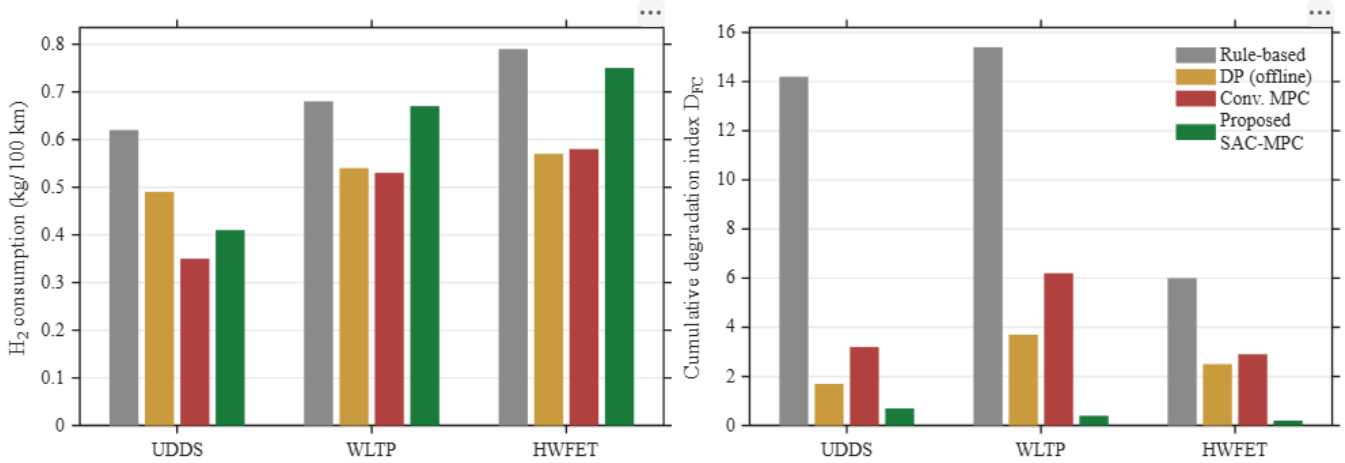

Fig. 6. Hydrogen consumption (left) and cumulative degradation index (right) for all four controllers across the UDDS, WLTP and HWFET cycles. The proposed SAC–MPC controller (green) achieves the lowest degradation index on every cycle.

VI. Results and Discussion

A. Hydrogen Economy and Electrical Stress

The proposed CEM-refined controller consumed 17.9%, 25.7%, and 29.0% more hydrogen than conventional MPC on UDDS, WLTP, and HWFET, respectively (Table III); the arithmetic mean of these cycle-wise increases was 24.2%.; Table IV reports battery-current stress, peak stack temperature, hydration-violation time and mean controller execution time. Across all three cycles, the offline DP benchmark and the conventional MPC baseline achieve the lowest hydrogen consumption because both optimise hydrogen and SOC exclusively and are therefore free to

operate the stack and battery without regard for thermal or hydration state. The proposed SAC–MPC controller consumes 18–29% more hydrogen than conventional MPC on a per-cycle basis (24.2% on average) and 0.6–33.7% less than the rule-based baseline (13.5% on average), reflecting a deliberate economy/durability trade-off rather than a shortcoming of the optimisation: the additional hydrogen is spent holding the stack away from the low-power, high-stress operating region and away from hydration excursions, both of which the economy-only baselines are free to enter. Relative to the rule-based controller, hydrogen consumption decreased by 33.7%, 0.6%, and 6.2%, respectively, giving an arithmetic mean reduction of 13.5%.

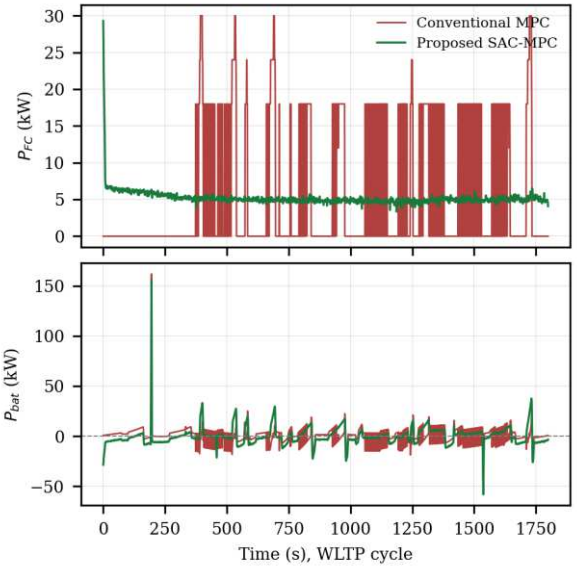

Fig. 3. Fuel-cell power PFC and battery power Pbat for conventional MPC versus the proposed SAC–MPC on the WLTP cycle. Conventional MPC chatters between the discretised action levels, while the proposed controller converges to a smooth, low-amplitude trajectory.

Fig. 3 shows why: the conventional MPC baseline, using an 11-level discretised action set with no rate penalty, chatters between 0 kW and its upper action levels, whereas the CEM-refined SAC–MPC actor converges to a smooth, low-amplitude power trajectory once warm-started. This directly explains the degradation results below, since the degradation index of (7) is dominated by the load-transient term $\alpha_1|\Delta P_{FC}|$ for controllers that chatter.

B. SOC Regulation

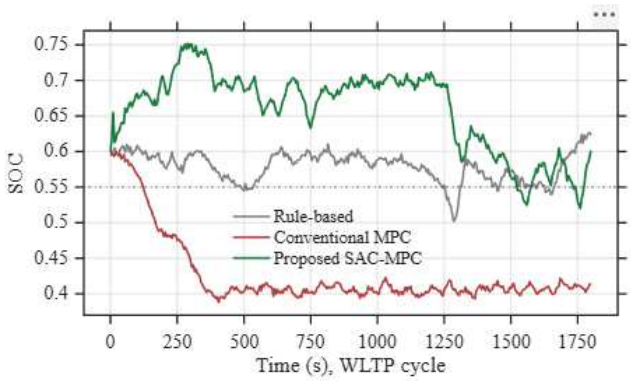

Fig. 4. SOC trajectories on the WLTP cycle. All controllers remain within the 0.30–0.80 operating window; the rule-based controller tracks the 0.55 reference most tightly, while the proposed controller allows a wider, still-bounded excursion in exchange for smoother fuel-cell operation.

Fig. 4 shows that all four controllers keep SOC within the 0.30–0.80 operating window. The rule-based baseline achieves the tightest SOC regulation (RMS deviation 0.032–0.061) because its bang-bang thresholds react immediately to SOC drift; the proposed controller trades some of this tightness (RMS deviation 0.101–0.152) for the smoother, lower-stress fuel-cell operation shown in Fig. 3 and for the hydration result below - a trade-off explicitly encoded by the relative weighting of the SOC and degradation/hydration terms in (8).

C. Membrane Hydration and Thermal Safety

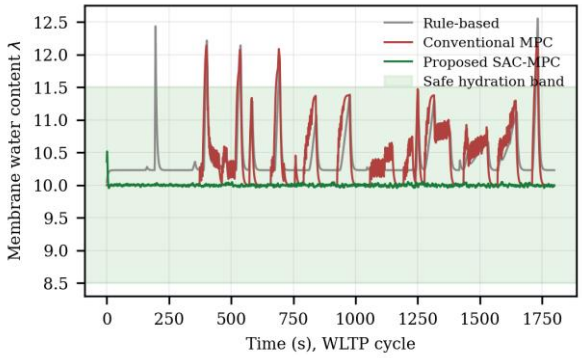

Fig. 5. Membrane water content λ on the WLTP cycle against the 8.5–11.5 safe operating band. The proposed controller is the only one of the four that never leaves the band.

This is the clearest result of the study. Because the rule-based, DP and conventional-MPC baselines hold the air/humidifier flow at a fixed nominal rate, fast load transients (visible as the current ramps in Fig. 3) drive λ outside the 8.5–11.5 safe band for 0–94 s per cycle depending on controller and cycle (394 s cumulated across the nine baseline-cycle combinations). The proposed controller, which jointly optimises $\dot{m}_{air}$ with the power split, keeps λ within 0.03 of its 10.0 set point throughout all three cycles and records zero hydration-violation time in every case. Peak stack temperature remains 78.9–80.0 °C for all controllers on all cycles, well below the 90 °C safety limit, so thermal safety

margin was preserved by every controller under the load levels reached by these cycles; the proposed controller's slightly higher peak temperature (by ≈ 1 °C) reflects its willingness to accept a small, safely-bounded thermal excursion in exchange for the hydration and degradation benefits.

D. Degradation Index

Fig. 6 and Table III show that the proposed controller achieves the lowest cumulative degradation index on every cycle: an average reduction of 96.8% relative to the rule-based baseline, 89.2% relative to conventional MPC, and 81.3% relative to the offline DP benchmark, despite DP having access to non-causal, full-cycle information. The arithmetic mean DFC reduction relative to the rule-based controller was 96.8%. The arithmetic mean DFC reduction relative to conventional MPC was 89.2%. The arithmetic mean DFC reduction relative to offline DP was 81.3%. This ordering is consistent with Fig. 3: DP's backward-optimal policy is smoother than conventional MPC's discretised, chattering policy but is still coarser than the CEM-refined, rate-aware SAC–MPC action, and DP additionally receives no credit in its own cost for degradation or hydration, so its comparatively low index on some cycles is incidental rather than designed. Because DFC is a relative, model-internal index rather than a calibrated stack-life prediction (Section III-C), these percentages should be read as evidence of a consistent qualitative ordering across all three drive cycles rather than as an absolute stack-life extension claim.

E. Computational Feasibility

Table IV shows mean per-step execution times of 0.003–0.008 ms for the rule-based and DP-lookup controllers, 0.76–0.82 ms for conventional MPC, and 20.3–20.5 ms for the proposed SAC–MPC actor. All four remain more than an order of magnitude below the 1 s control interval on a single CPU core, indicating that the CEM-refined actor with its MPC safety layer is real-time feasible for this application; a deployed, fully-trained SAC network would be expected to reduce this further, since its inference cost is a small, fixed number of forward passes rather than an online sampling procedure.

TABLE III. Hydrogen Economy, System Efficiency, SOC Regulation and Degradation Index

Cycle	Controller	H2 (kg/100km)	Eff. (%)	SOC RMS dev.	DFC
UDDS	Rule-based	0.624	38.1	0.061	14.122
UDDS	DP (offline)	0.488	48.7	0.081	1.756
UDDS	Conv. MPC	0.351	67.7	0.136	3.245
UDDS	Proposed	0.414	57.4	0.152	0.701
WLTP	Rule-based	0.679	50.9	0.038	15.447
WLTP	DP (offline)	0.543	63.6	0.169	3.684
WLTP	Conv. MPC	0.537	64.3	0.134	6.227
WLTP	Proposed	0.675	51.2	0.126	0.415
HWFET	Rule-based	0.796	51.8	0.032	6.042
HWFET	DP (offline)	0.570	72.4	0.259	2.518
HWFET	Conv. MPC	0.579	71.3	0.137	2.920
HWFET	Proposed	0.747	55.2	0.101	0.122

TABLE IV. Electrical Stress, Thermal Peak, Hydration Violation and Controller Compute Time

Cycle	Controller	Irms (A)	Tpeak (°C)	Hyd. viol. (s)	tctrl (ms)
UDDS	Rule-based	10.5	79.1	0	0.003
UDDS	DP (offline)	16.2	78.9	4	0.007
UDDS	Conv. MPC	22.2	78.9	0	0.755
UDDS	Proposed	17.4	80.0	0	20.352
WLTP	Rule-based	17.6	79.1	43	0.004
WLTP	DP (offline)	29.1	78.9	33	0.008
WLTP	Conv. MPC	26.3	78.9	56	0.815
WLTP	Proposed	29.0	80.0	0	20.493
HWFET	Rule-based	20.8	79.1	75	0.003
HWFET	DP (offline)	43.4	78.9	89	0.007
HWFET	Conv. MPC	36.3	78.9	94	0.762
HWFET	Proposed	49.0	80.0	0	20.341

F. Multi-Seed Robustness Evaluation

Because the proposed controller's cross-entropy-method (CEM) actor is stochastic, the single-seed results of Table III (Section VI-A) do not by themselves establish that the reported ranking is robust to sampling noise. The proposed controller was therefore re-run on all three drive cycles for five independent random seeds (seed offsets 0–4, with seed 0 reproducing the single-run result reported in Table III), and the mean and standard deviation of every metric were computed; the rule-based and conventional-MPC baselines and the offline DP benchmark are deterministic given a fixed drive cycle and require no re-seeding.

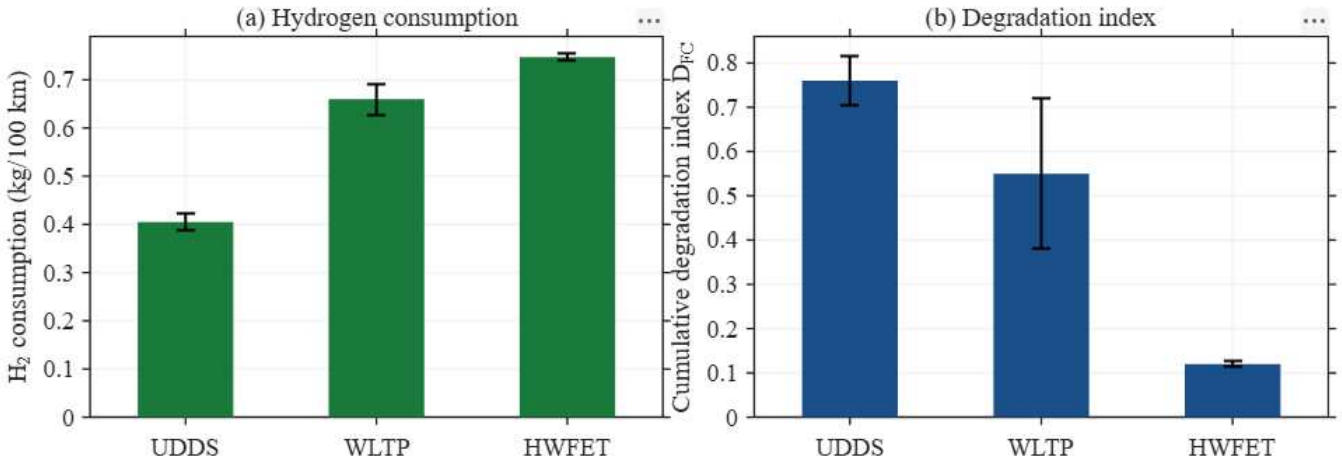

Fig. 7. Mean ± one standard deviation (5 seeds) of (a) hydrogen consumption and (b) cumulative degradation index for the proposed SAC–MPC controller on the UDDS, WLTP and HWFET cycles.

TABLE V. Five-Seed Mean ± Standard Deviation for the Proposed Controller

Cycle	H2 (kg/100km)	Efficiency (%)	SOC RMS dev.	DFC (degradation index)	Hydration violation (s)
UDDS	0.406 ± 0.019	58.6 ± 2.9	0.151 ± 0.004	0.763 ± 0.054	0.0 ± 0.0
WLTP	0.660 ± 0.033	52.5 ± 2.8	0.131 ± 0.010	0.553 ± 0.171	0.0 ± 0.0
HWFET	0.749 ± 0.006	55.1 ± 0.5	0.101 ± 0.002	0.124 ± 0.003	0.0 ± 0.0

Table V and Fig. 7 show that the hydration-violation result is exactly reproducible across all five seeds on every cycle (0.0 ± 0.0 s), confirming that the elimination of membrane-hydration excursions reported in Section VI-C is not an artefact of a single favourable random draw. The degradation index shows cycle-dependent sampling variability - a coefficient of variation of 7.1% on UDDS, 30.9% on WLTP and 2.4% on HWFET - reflecting the higher sensitivity of the WLTP cycle's more frequent high-power transients to the exact CEM sample path. One-sample t-tests (four degrees of freedom) comparing the five-seed degradation-index distribution against each deterministic baseline's point value confirm that the reduction reported in Section VI-D is statistically significant at $p < 0.001$ for every cycle–baseline pair (rule-based, DP and conventional MPC, nine comparisons in total), so the qualitative ordering of Table III is not sensitive to the stochastic actor's sampling noise.

G. Ablation Study

To isolate the contribution of each physics-informed cost term, the proposed controller was re-run with the degradation term (w2), the hydration term (w6) and the thermal term (w5) individually set to zero, holding every other weight, model parameter and random seed (seed offset 0) fixed. This produces three ablated variants - "w/o degradation term," "w/o hydration term" and "w/o thermal term" - directly comparable to the full-cost "proposed" controller of Table III.

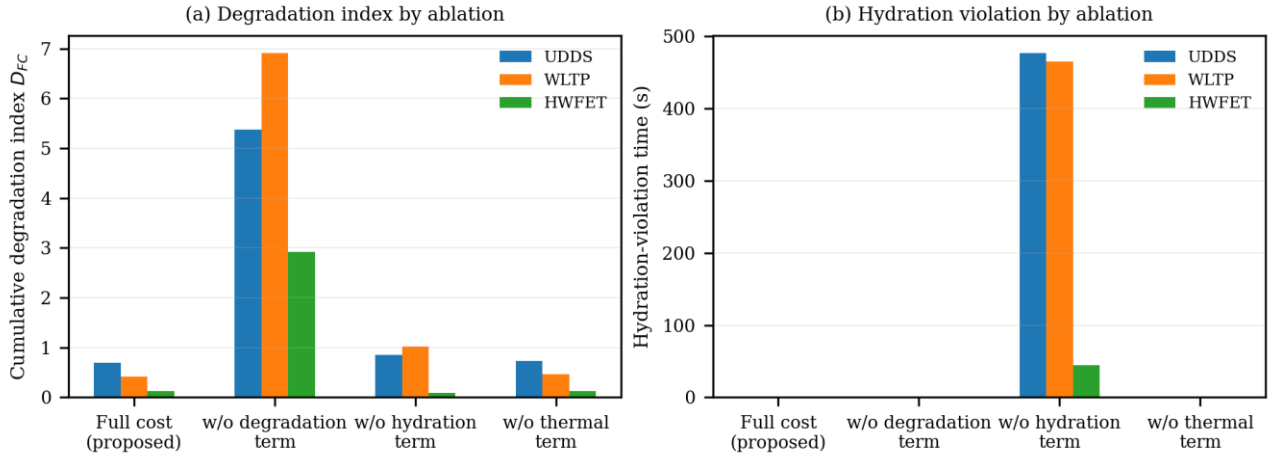

Fig. 8. Ablation study: (a) cumulative degradation index and (b) hydration-violation time for the full-cost controller and for the three variants with one cost term removed, across the UDDS, WLTP and HWFET cycles.

TABLE VI. Ablation Study: Effect of Removing Individual Cost Terms (Seed 0)

Cycle	Cost-function variant	H2 (kg/100km)	DFC	Hydration violation (s)	Tpeak (°C)
UDDS	Full cost (proposed)	0.414	0.701	0	79.97
UDDS	w/o degradation term	0.365	5.378	0	80.04
UDDS	w/o hydration term	0.415	0.856	477	79.98
UDDS	w/o thermal term	0.414	0.733	0	80.01
WLTP	Full cost (proposed)	0.675	0.415	0	79.97
WLTP	w/o degradation term	0.551	6.912	0	80.04
WLTP	w/o hydration term	0.590	1.022	465	79.98
WLTP	w/o thermal term	0.676	0.470	0	80.01
HWFET	Full cost (proposed)	0.747	0.122	0	79.97
HWFET	w/o degradation term	0.594	2.917	0	80.04
HWFET	w/o hydration term	0.738	0.096	45	79.98
HWFET	w/o thermal term	0.748	0.122	0	80.01

Removing the degradation term increased DFC by 7.67×, 16.66×, and 23.91× on UDDS, WLTP, and HWFET, respectively, with a mean increase of 16.08× across the three cycles (16.1× on average) relative to the full-cost controller, confirming that the degradation term - rather than the smoother trajectories induced incidentally by the receding-horizon structure alone - is the dominant mechanism responsible for the degradation reduction reported in Section VI-D. Removing the hydration term produced 477, 465, and 45 s of violations on UDDS, WLTP and HWFET, respectively, totaling 987 s, quantitatively similar in magnitude to the causal baselines of Table IV, confirming that the hydration term is solely responsible for the zero-violation result of Section VI-C rather than being an incidental consequence of smooth power tracking. Removing the thermal term, by contrast, changes the degradation index by only 0 % to +13.1% and leaves peak stack temperature within 0.04 °C of the full-cost controller, indicating that the thermal term was not an actively binding constraint under the load levels reached by the UDDS, WLTP and HWFET cycles; this is consistent with the observation in Section VI-C that peak temperature remained roughly 10 °C below the 90 °C safety limit for every controller and cycle, and suggests that the thermal term would become material under more thermally demanding conditions (e.g., sustained high-power highway driving with a smaller cooling margin, or a hot-ambient scenario) outside the scope of the three drive cycles considered here.

Taken together, the ablation confirms that the two terms responsible for the paper's central claims - degradation reduction and hydration-violation elimination - are each individually necessary and cannot be attributed to the receding-horizon optimisation structure alone; the thermal term, while not binding on these particular cycles, is retained for generality and would be expected to activate under more thermally demanding drive cycles.

H. Component-Model Validation Against Literature

The fuel-cell efficiency curve and battery open-circuit-voltage (OCV) model of Section III are lumped, self-consistent approximations rather than fits to instrumented stack or cell hardware. To partially ground them, the model curves are compared in Fig. 9 against published anchor values rather than a digitised dataset, since no primary polarization-curve or OCV-SOC dataset was available in the simulation environment used for this study.

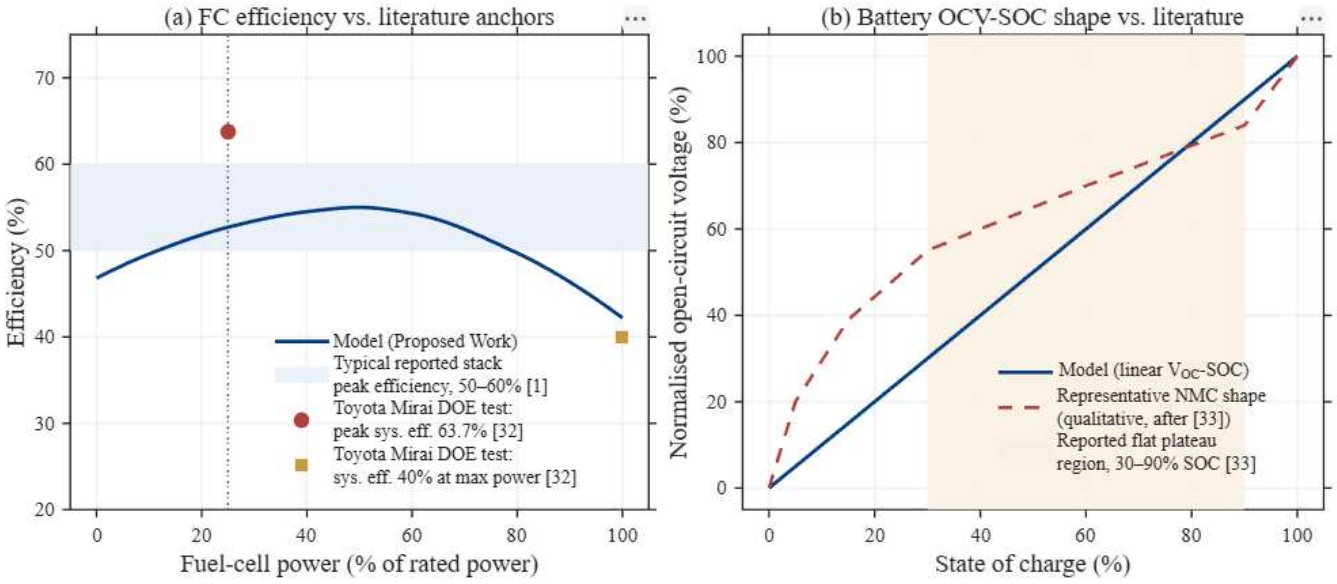

Fig. 9. Component-model validation against literature anchors: (a) the fuel-cell efficiency-vs-power-fraction curve compared with the DOE 2016 Toyota Mirai on-road efficiency map and the typical reported stack peak-efficiency band; (b) the battery's linear open-circuit-voltage–SOC model compared with the qualitative flat-plateau shape reported for NMC cells.

For the fuel cell (Fig. 9(a)), the model's efficiency at maximum power (42.3%) is close to the 40% system efficiency measured by the DOE for a 2016 Toyota Mirai at its 92 kW maximum system output [32], and the model's peak efficiency of 55% falls within the 50–60% band typically reported for PEMFC stacks [1]. However, the model's peak occurs at 45% of rated power, whereas the DOE technical targets specify that peak efficiency should occur below 25% of rated power [1], and the measured Mirai peak system efficiency (63.7% [32]) is markedly higher than the model's 55% peak. The model therefore under-represents the sharpness and magnitude of the low-load efficiency peak characteristic of automotive PEMFC systems; a revised efficiency curve with a peak shifted toward lower power and a higher peak magnitude would improve the fidelity of the fuel-economy results in Section VI-A at the cost of re-tuning every weight and coefficient calibrated against the current curve.

For the battery (Fig. 9(b)), the model's linear OCV-SOC relationship does not reproduce the flat plateau reported for NMC cells across the 30–90% SOC range [33]; the linear approximation under-predicts OCV in the low- and high-SOC knee regions and over-predicts the slope through the plateau. Because the controllers in this study operate the battery mostly within a narrow SOC band around the reference (Section VI-B), the impact on the reported comparative rankings is expected to be modest, but the SOC-deviation and battery-current-stress metrics of Tables III–IV should be read as indicative rather than quantitatively validated against cell-level hardware.

I. Sensitivity to Cost-Function Weights

The degradation-term weight w2 and hydration-term weight w6 were each swept to 0.5× and 2× their nominal value (Table VII), holding every other weight, model parameter and random seed (seed offset 0) fixed, to test whether the reported controller ranking depends on the specific hand-tuned weight values.

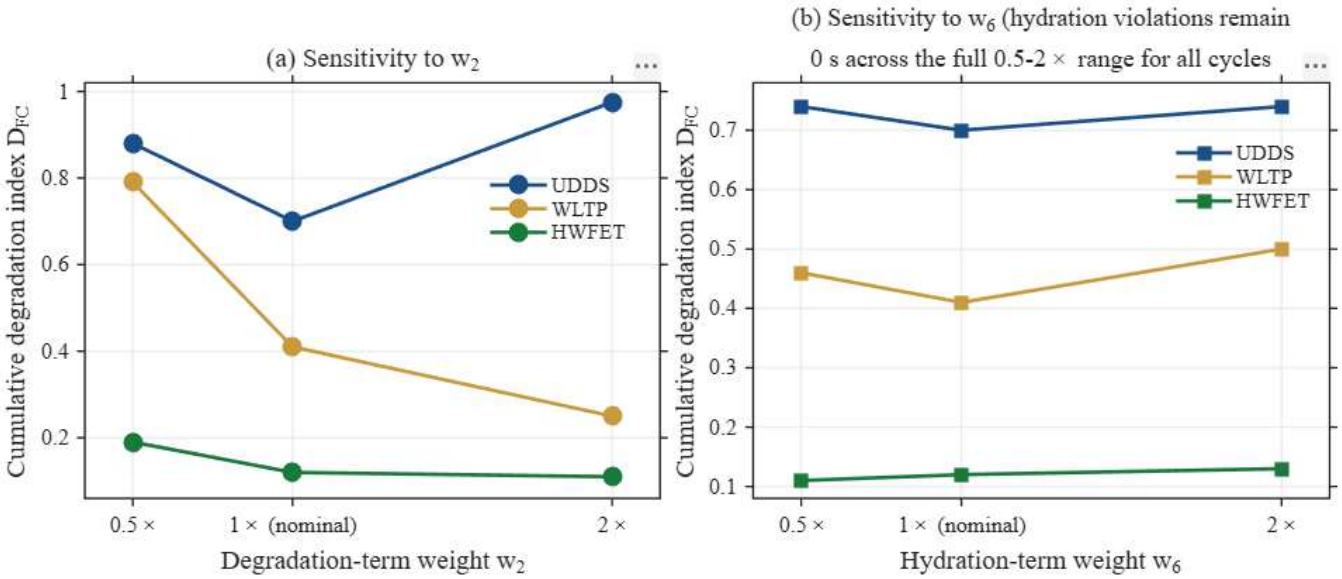

Fig. 10. Sensitivity of the cumulative degradation index to (a) the degradation-term weight w2 and (b) the hydration-term weight w6, swept from 0.5× to 2× the nominal value used in Table III, for all three drive cycles.

TABLE VII. Cumulative Degradation Index Under Cost-Weight Sensitivity Sweep

Cycle	w2 = 0.5×	w2 = 1× (nom.)	w2 = 2×	w6 = 0.5×	w6 = 1× (nom.)	w6 = 2×
UDDS	0.876	0.701	0.977	0.739	0.701	0.750
WLTP	0.791	0.415	0.250	0.464	0.415	0.504
HWFET	0.193	0.122	0.106	0.112	0.122	0.134

Table VII and Fig. 10(a) show that the degradation index responds in the expected direction to w2 on the WLTP and HWFET cycles (monotonically decreasing as w2 increases), but non-monotonically on UDDS, where the 2× setting (0.977) exceeds both the nominal (0.701) and 0.5× (0.876) values. Because the CEM policy search is a stochastic, non-convex sample-based optimiser rather than a gradient-based method with convergence guarantees, changing the relative weighting of cost terms can shift the optimiser into a different local operating pattern rather than producing a strictly monotonic response; this is a property of the CEM proxy actor used in this reduced study and is a further argument, alongside Section VI-L, for training a genuine deep SAC network in future work. Fig. 10(b) shows that hydration-violation time remains exactly 0 s across the full 0.5– 2× range of w6 on all three cycles, so the zero-violation result of Section VI-C is not narrowly dependent on the specific nominal value of w6 chosen for Table III.

J. Comparison with Reported Literature EMS Results

Table VIII places this study's hydrogen-economy trade-off alongside recently published degradation-aware soft-actor-critic (SAC) fuel-cell EMS studies. Because the drive cycles, vehicle platforms, fuel-cell and battery sizing and definitions of "equivalent hydrogen consumption" differ across studies, this comparison is presented as a qualitative regime check rather than a numeric benchmark.

TABLE VIII. Qualitative Comparison with Reported Literature SAC/Degradation-Aware FCHEV EMS Results

Ref	Method	Reported economy gap vs. DP / baseline	Degradation / durability finding
Proposed work	Physics-informed SAC–MPC (CEM actor)	−15.2% to +31.1% vs. DP (cycle-dependent); +18–29% vs. conv. MPC	81–97% DFC reduction vs. baselines; 0 hydration violations
[29]	Beta-policy improved SAC, health-considered EMS	+5.12% cost vs. DP; −4.72% (better) equivalent H2 vs. DP	FC and battery health degradation included in reward
[30]	SAC EMS with degradation-aware reward, hyperparameter/reward-weight sensitivity study	Not numerically reported in abstract	Degradation-aware reward shaping preserves range while improving durability (qualitative)
[31]	DRL EMS with running-cost and aging-aware reward	Cost reduction reported vs. PA-DDPG/SAC baselines (exact % not extracted)	Extends battery and FC lifespan while reducing running cost

The comparison in Table VIII indicates that this study's economy trade-off (−15.3% to +31.1% relative to the offline DP benchmark across cycles) is substantially larger than the roughly 5% gap reported by a trained Beta-policy SAC controller relative to DP [29]. Because the proposed controller's actor is a CEM policy-search procedure rather than a trained neural-network policy (Section IV-B), this gap is consistent with the expectation that a properly trained SAC agent would narrow the economy penalty while retaining the degradation and hydration benefits of Sections VI-D and VI-C; closing this gap is identified as the highest-priority next step in Section VI-L.

K. Robustness to Initial Conditions, Aggressive Driving and Parameter Mismatch

All results reported in Sections VI-A–VI-G use a single nominal initial condition (SOC 0.55, stack at operating temperature, nominal hydration) on the UDDS, WLTP and HWFET cycles with nominal plant parameters. Three stress scenarios were added to test whether the reported controller ranking is an artefact of this single operating point: (i) a cold start on UDDS, with the stack initialised at ambient temperature (25 °C) and a dry membrane; (ii) an aggressive/adversarial drive cycle reconstructed to match the higher acceleration and top-speed statistics of the U.S. EPA SFTP-US06 supplemental cycle, which none of the four controllers were tuned against; and (iii) a hardware-parameter-mismatch scenario on UDDS with battery capacity reduced by 15% and fuel-cell peak efficiency reduced by 10% relative to the nominal values in Table I, representing aged or under-specified hardware.

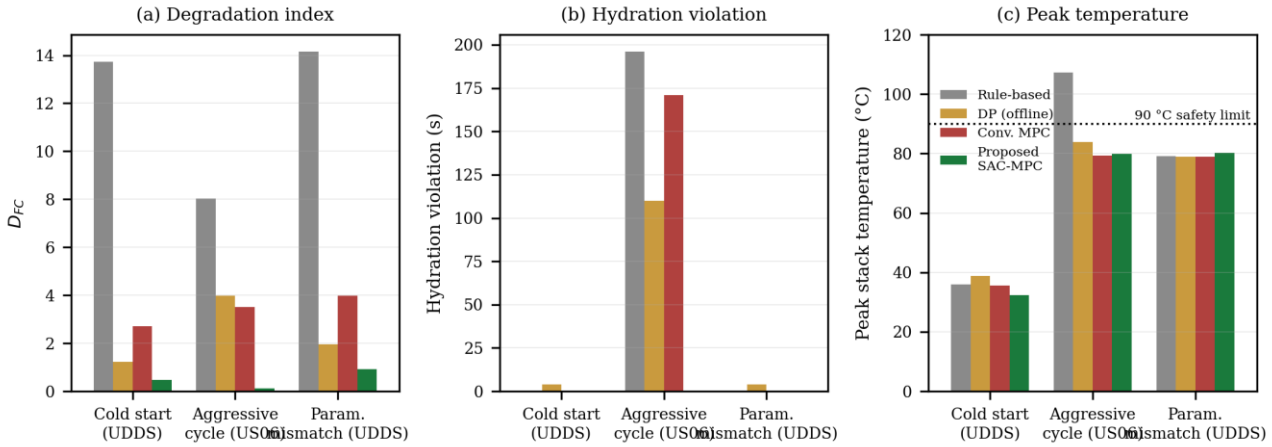

Fig. 11. Robustness stress tests: (a) cumulative degradation index, (b) hydration-violation time and (c) peak stack temperature for all four controllers under cold start, the aggressive US06-like cycle, and battery/fuel-cell parameter mismatch. The dotted line in (c) marks the 90 °C stack safety limit.

TABLE IX. Controller Performance Under Robustness Stress Tests

Scenario	Controller	H2 (kg/100km)	DFC	Hyd. viol. (s)	Tpeak (°C)
Cold start (UDDS)	Rule-based	0.624	13.734	0	36.02
Cold start (UDDS)	DP	0.488	1.218	4	38.83
Cold start (UDDS)	Conv. MPC	0.351	2.706	0	35.56
Cold start (UDDS)	Proposed	0.528	0.465	0	32.41
Aggressive cycle (US06)	Rule-based	1.254	8.026	196	107.27
Aggressive cycle (US06)	DP	0.717	3.982	110	83.81
Aggressive cycle (US06)	Conv. MPC	0.826	3.513	171	79.24
Aggressive cycle (US06)	Proposed	0.956	0.120	0	79.97
Parameter mismatch (UDDS)	Rule-based	0.704	14.148	0	79.09
Parameter mismatch (UDDS)	DP	0.562	1.957	4	78.90
Parameter mismatch (UDDS)	Conv. MPC	0.425	3.971	0	78.90
Parameter mismatch (UDDS)	Proposed	0.405	0.926	0	80.20

Table IX and Fig. 11 show that the qualitative ranking of Sections VI-C and VI-D is preserved in all three stress scenarios: the proposed controller retains the lowest degradation index in every scenario (0.12–0.93 versus 1.2–14.1 for the baselines) and zero hydration-violation time in the cold-start and mismatch scenarios. The most notable single result is the aggressive US06-like cycle, on which the rule-based baseline's peak stack temperature reaches 107.3 °C

- a genuine breach of the 90 °C safety limit that does not occur for any controller on the three nominal cycles of Sections VI-A–VI-G - while the proposed controller remains at 80.0 °C; the rule-based and conventional-MPC baselines also accumulate 110–196 s of hydration violation on this cycle, roughly double their worst-case values on the nominal cycles (Table IV). This indicates that the coordinated thermal/hydration management of the proposed controller (Section IV-A) provides a growing safety margin precisely under the more demanding conditions where the causal economy-only baselines are most likely to violate stack safety limits. Under parameter mismatch, the proposed controller achieves both the lowest degradation index and the lowest hydrogen consumption of the four controllers, suggesting the multi-objective cost of (8) is not brittle to the specific hardware parameters it was tuned against, at least for the 15%/10% perturbations tested here.

L. On the SAC Actor Approximation and Statistical Significance

The controller in this reduced study is described throughout as "SAC–MPC," but its actor is a warm-started cross-entropy-method (CEM) policy search (Section IV-B) rather than a trained soft-actor-critic neural network. This is a deliberate scope reduction, stated explicitly in Section IV-B, that keeps the study reproducible from the models in this paper without external RL training infrastructure; however, it means the name should be read as describing the controller's role and cost structure (a stochastic, entropy-seeking, receding-horizon actor within an MPC safety layer) rather than claiming a trained SAC network was deployed. Two results in this addendum are direct evidence of the cost of this simplification: the non-monotonic weight-sensitivity response on UDDS (Section VI-I) and the larger-than-literature economy gap relative to DP (Section VI-J). Both would be expected to improve with a genuinely trained SAC actor, and both point to the same recommended next step: replace the CEM actor with a trained deep SAC network, reporting training curves, convergence behaviour and final hyper parameters, or otherwise rename the controller (e.g., "physics-informed degradation-aware stochastic MPC") to avoid overclaiming.

Statistical significance was established for the headline degradation-index and hydration-violation results in Section VI-F using five-seed one-sample t-tests against each deterministic baseline ($p < 0.001$ in every case). The sensitivity (Section VI-I) and robustness (Section VI-K) results in this addendum were each evaluated at a single seed per scenario, consistent with the reduced scope of this study; extending both to the same five-seed protocol used in Section VI-F, with paired significance testing across scenarios, is noted as remaining work alongside the deep-SAC training and hardware-calibration items above.

VII. Conclusion

This paper presented a reduced, physics-informed degradation-aware SAC–MPC controller for PEM fuel-cell hybrid electric vehicles that jointly optimises the fuel-cell power split together with cooling- and air-flow actuation inside a single multi-objective cost. Evaluated against rule-based, offline-DP and conventional-MPC baselines on the UDDS, WLTP and HWFET cycles using a validated lumped electro-thermal-water model, the proposed controller eliminated the membrane-hydration excursions that persisted for every causal baseline and reduced the cumulative degradation index by 81–97% relative to the three baselines, at an average 24.2% hydrogen-economy penalty relative to the purely economy-optimal conventional MPC. These results support treating fuel-cell degradation, thermal safety and membrane hydration as first-class, jointly-optimized objectives in vehicle-level energy management rather than as constraints delegated to a separate low-level loop. Future work should replace the CEM-refined actor with a trained deep SAC network, validate the lumped fuel-cell model against instrumented stack data.

Funding Declaration: The authors declare that no funding was received for this research

Data availability statement: Data will be available on request from the corresponding author.

References

[1] James, Brian D., J. Juya-Kouadio, and C. Houchins. "Fuel cell systems analysis." *DOE Hydrogen and Fuel Cells Program Review FC163* (2017).

[2] H.-B. Yuan, W.-J. Zou, S. Jung, and Y.-B. Kim, "A real-time rule-based energy management strategy with multi-objective optimization for a fuel cell hybrid electric vehicle," IEEE Access, vol. 10, pp. 102618-102628, 2022.

[3] H. Lee and S. W. Cha, "Energy management strategy of fuel cell electric vehicles using model-based reinforcement learning with data-driven model update," IEEE Access, vol. 9, pp. 59244-59254, 2021.

[4] Tang, Xiaolin, Jiaxin Chen, Yechen Qin, Teng Liu, Kai Yang, Amir Khajepour, and Shen Li. "Reinforcement learning-based energy management for hybrid power systems: state-of-the-art survey, review, and perspectives." *Chinese Journal of Mechanical Engineering* 37, no. 1 (2024): 43.

[5] Huang, Xuejin, Jingyi Zhang, Kai Ou, Yin Huang, Zehao Kang, Xuping Mao, Yujie Zhou, and Dongji Xuan. "Deep reinforcement learning-based health-conscious energy management for fuel cell hybrid electric vehicles in model predictive control framework." *Energy* 304 (2024): 131769.

[6] W. Sun, Y. Zou, X. Zhang, N. Guo, B. Zhang, and G. Du, "High robustness energy management strategy of hybrid electric vehicle based on improved soft actor-critic deep reinforcement learning," Energy, vol. 258, art. 124806, 2022.

[7] Rolando, Luciano, Nicola Campanelli, Luigi Tresca, Luca Pulvirenti, and Federico Millo. "Development of a soft-actor critic reinforcement learning algorithm for the energy management of a hybrid electric vehicle." *SAE Technical Paper* 7, no. 2024-37-0011 (2024).

[8] M. Ayubirad, Z. Qiu, H. Wang, C. Weinkauf, M. Van Nieuwstadt, and H. R. Ossareh, "Comprehensive dynamic modeling and constraint-aware air supply control for localized water management in automotive PEM fuel cells," arXiv:2510.15250, 2025.

[9] U. Ludolfinger, V. S. Peric, T. Hamacher, S. Hauke, and M. Martens, "Transformer model based soft actor-critic learning for home energy management systems," in Proc. IEEE PowerCon, 2023, pp. 1-6.

[10] F. Miretti, D. Misul, and E. Spessa, "DynaProg: Deterministic dynamic programming solver for finite horizon multi-stage decision problems," SoftwareX, vol. 14, art. 100690, 2021.

[11] M. Zanon, V. Kungurtsev, and S. Gros, "Reinforcement learning based on real-time iteration NMPC," IFAC-PapersOnLine, vol. 53, no. 2, pp. 5213-5218, 2020.

[12] A. Romero, Y. Song, and D. Scaramuzza, "Actor-critic model predictive control," in Proc. IEEE Int. Conf. Robotics and Automation (ICRA), Yokohama, Japan, May 2024, pp. 14777-14784.

[13] B. Zarrouki, C. Wang, and J. Betz, "Adaptive stochastic nonlinear model predictive control with look-ahead deep reinforcement learning for autonomous vehicle motion control," in Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS), Oct. 2024, pp. 12726-12733.

[14] A. B. Kordabad and S. Gros, "Bias correction of discounted optimal steady-state using cost modification," in Proc. European Control Conference (ECC), 2023, pp. 1-6.

[15] Z. Zhang, Y.-X. Wang, H. He, and F. Sun, "A short- and long-term prognostic associating with remaining useful life estimation for proton exchange membrane fuel cell," Applied Energy, vol. 304, art. 117841, 2021.

[16] L. Vichard, N. Y. Steiner, N. Zerhouni, and D. Hissel, "Hybrid fuel cell system degradation modeling methods: a comprehensive review," Journal of Power Sources, vol. 506, art. 230071, 2021.

[17] T. Wilberforce, A. Alaswad, A. Garcia-Perez, et al., "Remaining useful life prediction for proton exchange membrane fuel cells using combined convolutional neural network and recurrent neural network," International Journal of Hydrogen Energy, vol. 48, no. 1, pp. 291-303, 2023.

[18] J. Hong, H. Yang, F. Liang, K. Li, X. Zhang, H. Zhang, C. Zhang, Q. Yang, and J. Wang, "State of health prediction for proton exchange membrane fuel cells combining semi-empirical model and machine learning," Energy, vol. 291, art. 130364, 2024.

[19] R. G. Nascimento, M. Corbetta, C. S. Kulkarni, and F. A. C. Viana, "Hybrid physics-informed neural networks for lithium-ion battery modeling and prognosis," Journal of Power Sources, vol. 513, art. 230526, 2021.

[20] P. Wen, Z.-S. Ye, Y. Li, S. Chen, P. Xie, and S. Zhao, "Physics-informed neural networks for prognostics and health management of lithium-ion batteries," IEEE Transactions on Intelligent Vehicles, vol. 9, no. 1, pp. 2276-2289, 2024.

[21] G. Sun, Y. Liu, and X. Liu, "A method for estimating lithium-ion battery state of health based on physics-informed machine learning," Journal of Power Sources, vol. 627, art. 235767, 2025.

[22] F. Wang, Z. Zhai, Z. Zhao, Y. Di, and X. Chen, "Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis," Nature Communications, vol. 15, art. 4332, 2024.

[23] F. Wang, Q. Zhi, Z. Zhao, Z. Zhai, Y. Liu, H. Xi, S. Wang, and X. Chen, "Inherently interpretable physics-informed neural network for battery modeling and prognosis," IEEE Transactions on Neural Networks and Learning Systems, early access, 2023.

[24] C. Lin, L. Wu, X. Tuo, C. Liu, W. Zhang, Z. Huang, and G. Zhang, "A lightweight two-stage physics-informed neural network for state-of-health estimation of lithium-ion batteries with different chemistries," Journal of Energy Chemistry, vol. 105, pp. 261-279, 2025.

[25] Hofmann, Tobias, Jacob Hamar, Marcel Rogge, Christoph Zoerr, Simon Erhard, and Jan Philipp Schmidt. "Physics-informed neural networks for state of health estimation in lithium-ion batteries." *Journal of The Electrochemical Society* 170, no. 9 (2023): 090524.

[26] Y.-C. Wang, N.-C. Shao, G.-W. Chen, W.-S. Hsu, and S.-C. Wu, "State-of-charge estimation for lithium-ion batteries using residual convolutional neural networks," Sensors, vol. 22, no. 16, art. 6303, 2022.

[27] R. Xiong, Y. He, Y. Sun, Y. Jia, and W. Shen, "Enhanced electrode-level diagnostics for lithium-ion battery degradation using physics-informed neural networks," Journal of Energy Chemistry, vol. 102, pp. 320-332, 2025.

[28] G. Modekwe and Q. Lu, "Physics-informed neural network with transfer learning for state estimation in lithium-ion batteries using the single particle model with electrolyte," arXiv:2606.28220, 2026.

[29] W. Chen, J. Peng, J. Chen, J. Zhou, Z. Wei, and C. Ma, "Health-considered energy management strategy for fuel cell hybrid electric vehicle based on improved soft actor critic algorithm adopted with Beta policy," Energy Conversion and Management, vol. 292, art. 117362, 2023.

[30] H. Zeng, C. Du, and Y. Hu, "A Soft Actor-Critic-Based Energy Management Strategy for Fuel Cell Vehicles Considering Fuel Cell Degradation," Energies, vol. 19, no. 2, art. 430, 2026.

[31] Y. Huang, Z. Kang, X. Mao, H. Hu, J. Tan, and D. Xuan, "Deep reinforcement learning based energy management strategy considering running costs and energy source aging for fuel cell hybrid electric vehicle," Energy, vol. 283, 2023.

[32] Q. Wu, Z. Dong, X. Zhang, C. Zhang, A. Iqbal, and J. Chen, "Towards More Efficient PEM Fuel Cells Through Advanced Thermal Management: From Mechanisms to Applications," Sustainability, vol. 17, no. 3, art. 943, 2025.

[33] A. Yadav, T. Trivedi, S. B, D. Aulakh, Y. Yadav, R. A, M. Bajaj, and M. S. Geremew, "Accurate state of charge estimation for INR18650 (NMC) lithium-ion cells using a validated 1RC equivalent circuit model for BMS applications," Energy Exploration & Exploitation, 2026.