Real-Time Thermal–Water Management Energy Dispatch Co-Optimization in PEMFC-EVs via Metaheuristic Pareto Tuning of MPC with RL-Based Disturbance Adaptation

Adel Elgammal
Professor, Utilities and Sustainable Engineering, The University of Trinidad & Tobago UTT

Abstract: An integrated thermal–water management and energy-dispatch co-optimization framework was introduced for real-time control of proton-exchange membrane fuel-cell electric vehicles (PEMFC-EVs), playing in a stochastic context with drive-cycle and ambient uncertainty. A restriction-based hierarchical Model Predictive Control (MPC) supervisor conveniently coordinates the fuel-cell/battery power split along with compressor and coolant-pump power, coolant mass flow, radiator fan speed, humidifier/bypass ratio, and purge timing in order to control stack temperature and membrane hydration within constraints of safety while minimizing hydrogen consumption, auxiliary losses and degradation risk indices. To avoid manual tuning, an off-line metaheuristic Pareto search (NSGA-II/MOPSO) is conducted to adjust the MPC weights, horizons and adaptive constraint margins on a set of representative cycles thus yielding a library of non-dominated sets of MPC parameters trading-off hydrogen economy versus thermal/humidity stress and actuator usage. Lightweight online reinforcement-learning (RL) disturbance adaptation layer (actor-critic) estimates latent disturbances (unmodeled load transients, ambient temperature-humidity shifts, and cooling effectiveness changes) and selects real-time Pareto-optimal MPC settings for robustness. Monte Carlo simulations for mixed urban/highway profiles show 6–12% lower hydrogen consumption and 8–15% less auxiliary energy requirements than rule-based control and constant-weight MPC, with a reduction in stack temperature overshoot of 20–35% and dehydration/flooding events decreases from the range of 30–55%. A combined degradation metric considering the magnitude of thermal cycling and the humidity swing increases by 10–25% while maintaining drivability and near-zero constraint violations. The figures demonstrate increasing range and stack life under an uncertain real-world driving situation. Real-time feasibility is retained at 50-ms sampling, with 15–20 ms average solve time on an embedded CPU, and the performance transfer across UDDS, WLTC, and aggressive cycles.

Keywords: PEM fuel-cell electric vehicles (PEMFC-EVs), Thermal–water management, Model predictive control (MPC), Reinforcement learning disturbance adaptation, Metaheuristic, multi-objective optimization (Pareto tuning) , Drive-cycle uncertainty

I. Introduction:

The PEMFC-EVs have emerged as an alternative means complementary to the battery-electric mobility due to their quick refueling time, high specific energy of onboard hydrogen (H2), and potential lower life-cycle emissions provided that low-carbon H2 is used. Rather, usefulness of PEMFC-EVs for transportation is determined by how well the vehicle performs energy distribution (in terms of fuel cell–battery power split, transient storage use, and regenerative braking) with stack auxiliary control (air supply, humidification, purge, coolant loop control/radiator-fan control). In current automotive stacks the net vehicle efficiency and life are affected not only by electrochemical conversion efficiency but also temperature and water state management under highly transient and uncertain drive cycles. Modern control approaches for PEMFCs have an increasing trend toward "automation to autonomy" (combination of model-based control with learning, estimation and multi-objective decision layers to deal with interconnected subsystems and uncertain operating conditions) [1]. Concurrently, recent reviews of water management have highlighted that membrane hydration and liquid-water removal are durability-critical as both dehydration and flooding can degrade performance itself, initiate local hot-spots or even to drive irreversible aging mechanisms [2]. Together, these ideas imply a line of research wherein the problem of real-time energy management is no longer considered as a stand-alone power-split problem, but rather as an integrated thermal-water-power co-optimization problem.

The electrochemical behavior of the PEMFC stack is significantly affected by temperature and membrane hydration level. Water within the membrane maintains proton conductivity, but surplus liquid water in catalyst and gas diffusion layers (GDL) blocks reactant transport and causes flooding. This leads to a natural trade-off: more humidification can prevent membrane dry-out and reduce ohmic losses, but higher humidity values (or poor drainage/purge) can allow mass-transport losses and dynamic instability to become increasingly dominant. In automotive operation, these occur at even higher rates due to quick load changes, stop-and-go traffic and frequent start/stop. Accordingly, water management in automotive PEMFC systems involves: (i) the control of inlet humidity and (ii) the control of internal water transport through these systems, including triggering purges or gas recirculation along lines defined by compressor power consumption and system packaging constraints [2]. To date, recent automotive-oriented work suggests online water-management stability control by integrating stack water state estimation into vehicle powertrain control, in part because traditional diagnostics (e.g., impedance-based techniques) are challenging to deploy reliably under highly dynamic drive cycles [5]. Heat control is likewise entangled with both efficacy and longevity. It increases the reaction kinetics and the water saturated pressure, stacking temperature decreases system TMS cause non-negligible parasitic power requirements for pumps and fans. For aggressive transients or hot-ambient operation the temperature limits become actively binding and produce a derating or an inefficient operating point if cooling resources are limited. This increasing attention in MPC for PEMFC thermal management is motivated by its capacity to explicitly respect constraints and predict future disturbances with predictive models. For instance, recent works suggest MPC-based temperature tracking can decrease tracking error and oscillation of actuator behavior in addition to a reduction in TMS power consumption compared with conventional PID methods [4]. From an energy-management viewpoint, PEMFC-EVs often adopt hybridization (with battery and sometimes ultracapacitor/supercapacitor) as a means of smoothing transients and seeking to mitigate the load swings seen by the fuel cell that are responsible for aging acceleration. The classical ones include rule-based control and equivalent consumption minimization strategy (ECMS), and the optimization -based approaches (dynamic programming offline, MPC online) which are stronger in optimality while may need careful modeling and tuning for real-time feasibility. Reinforcement learning (RL) has been gaining traction as a potential alternative or complement over the past few years, particularly for high dimension state/action spaces and/or uncertain driving conditions. RL-centric reviews identify strengths in dealing with nonlinearities and complex multi-objective reward shaping, but also recognize limitations: safety guarantees under constraints, generalization outside of training distributions, and the computational cost of online deployment [6]. These are rooted in a common systems realization: power dispatch, thermal management and water management cannot be considered independently for the real drive cycles. Modifications to power split change stack current density trajectories, and consequent changes in water generation and electro-osmotic drag, as well as heat release, also impacts required air flow. On the other hand, thermal–water limitations can lead to operation with non-optimal dispatch task in case the plant does not foresee their interaction on consumption of hydrogen (auxiliary power and degradation).

While this work has accelerated the research of PEMFC-EV system APEA, it is observed that most works have not established close-coupled between power split and T/W control optimization across different timescale since most PEMFC-EV control architectures are still loosely coupled. in practice, couple of those three issues actually dominate: Multi-objective conflict under transients. The need to minimize hydrogen consumption tends to push the fuel cell in the direction of high-efficiency operation, but long-life considerations tend to penalize sudden current transients, low-load operation and associated cycle-of purge on ramped off approaches, start/stop frequency or operating conditions which cause local dehydration (rain cell effect) and local flooding. Degradation proxies and related health constraints are more and more included in durability-aware frameworks as fuel-cell / battery health-aware predictive control formulations aiming at minimizing the total running cost by accounting for degradation [12]. Constraint-heavy real-time operation. There are hard constraints due to the thermal and water states (stack temperature boundaries, coolant availability, e.g., an envelope for humidity/flooding avoidance, compressor limits). MPC is well suited for constrained optimization, but its performance relies on the model accuracy and choice of tuning parameters (weights, horizons, constraints), which are hard to hard-code robustly for a variety of driving cycles, ambient conditions and vehicle component degradation. Drive-cycle uncertainty and disturbance variability. Real-world driving never closely matches a single known drive cycle, with grade, traffic conditions, ambient temperature and driver aggressiveness changing far too unpredictably. Although H-MPC can be designed to take into account variations in disturbances and degradation (e.g., cooperative MPC layers for eco-driving and energy management that explicitly consider changes in grammar of disturbance, and fuel-cell degradation), these designs are based on assumed models and a priori weights of objective functions, which may not necessarily remain optimal over varying conditions [13]-[18]. Thus, the question at hand can be formulated as: Given that, PEMFC-EV systems need to ensure not only long driving range but also long stack life and that such a goal requires coordinated optimal operation of both on-line energy dispatching and thermal–water management strategies under uncertain and time-varying drive demands; how can one accomplish simultaneity of optimization without any trade-off between performance constraints satisfaction or computational tractability?

Degradation of PEMFCs results from combined chemical, mechanical, and thermal stresses such as catalyst dissolution/agglomeration, carbon corrosion/MC ther-mal treatment, membrane chemical attack, and mechanical cracking due to hydration/dehydration cycling. General assessments on ageing highlight that, under realistic operation (varying load steps, humidity swings, start/stop), interacting degradation processes are formed, which are difficult to cover only with a stand-alone value [7], [8]. Especially, hydration cycling induces mechanical stress in the membrane and leads to failure, which implies that durable-aware control and accelerated protocols such as AST should be devised for evaluation [7]. Meanwhile, the systemic reviews show that there is the requirement for models to link observable signals (voltage, temperature, pressure and impedance proxies) to underlying damage states (ECSA loss, membrane thinning and catalyst layer degradation), which can be handled computationally by control [8]. AST development based on standard automotive cycles (e.g., WLTC-based profiles) shows that degradation trends under such realistic transients do not follow the ones imposed by simplified laboratory stressors, and that humidity/temperature fluctuations play an important role in shaping aging rates [9]. A parallel experimental effort demonstrates that accelerated aging under coupled dynamic stress to end-of-life can be achieved within manageable test durations while eliciting multimodal aging processes which can be monitored with timeline diagnostics and post-mortem characterization [10]. This evidence is a vivid demonstration that a controller based on compact tuning can break when exposed to extended variability of more realistic conditions. Although water management can be based on diagnostics like high-frequency resistance (HFR) or impedance measurements, conventional measurement methods are disturbed by load transients and variations of operating points. Work in automotive based applications illustrates these limitations and suggests integrated on-line approaches lifting the water stability control, coupled with the vehicle power control to ensure performance and durability [5]. However, inclusion of this water state estimation in an EMS operating under the optimal power split, auxiliary loads and degradation simultaneously is complicated due to addition of states, constraints and uncertainties in the optimisation problem. MPC is gaining momentum as a model-based, predicative decision-making module in hybrid powertrains, however the cost-benefit of fuel efficiency to drivability including thermal and water safety margins and life-cycle preservation are seamlessly defined by the objective weights and horizons. Other work in the field of RL-augmented MPC suggested that constant weight MPC may have difficulty balancing between competing objectives as their weighting changes over varying conditions, supporting adaptive weight or multi-layer MPC architectures [14]. Concurrently, the wider EMPC community has produced Pareto-front tuning procedures to consider trade-offs between conflicting objectives and prune solutions corresponding to operator preferences [16]-[20]. However, for PEMFC-EVs this Pareto tuning is usually performed off-line and it depends on the disturbance set being assumed and degradation model accuracy. RL can learn policies that adjust under uncertainty, but direct RL without structure may not satisfy safety constraints (e.g., temperature/humidity limits) and may even need large amounts of training data for rare but extreme operating conditions missing in training data. Work on RL for HEV energy management reports deployment barriers to be computational complexity, generalisation across wide range of driving conditions and the requirement to assure safe integration with constraints [21]. This paves the way for hybrid RL–MPC techniques: the MPC solves on-rate constraint handing and yields stability structure, whereas the RL ensures evolutionary capacity to handle disturbances in plant models and time varying nature of plant transfer functions.

Rule based approaches are also common for their simplicity and robustness, but can be conservative and fail to take hydrogen consumption reduction under benign conditions. ECMS approach conserves fuel by translating battery energy use into an equivalent hydrogen cost, while being still calibrated based on equivalence factors and may not explicitly consider thermal–water constraints. MPC will enhance performance by optimization of a finite time horizon with component models and constraints. Health-Aware Predictive control supplements the MPC framework by incorporating degradation cost terms of both fuel cell and battery, so as to minimize total running cost not current fuel consumption as used in the existing studies [12], [22]. Such strategies set a key stepping-stone: lifetime can be accounted for in the objective and constraints, and online optimization may explicitly sacrifice fuel economy for lifetime. Yet, the EMS are highly health conscious here in that many focus more so electrical dis-patch and easy degradation proxies and consider thermal–water management as a subsequent subsystem to be controlled separately. RL has also been used for energy management of FCHEVs with hierarchical approach to decouple actions space and increase tractability. A data-driven RL based hierarchical EMS for fuel-cell/battery/ultracapacitor hybrids shows that an RL can be functioned within the power splitting framework executed to simplify the problem and also learn an appropriate decision adaptively from data [16]. Ada-H-ECMS Based on related study Nielsen 0 also propose that some RL-based adaptive hierarchical ECMS structures "adapt a set of equivalence factors or decision parameters online enhancing the robustness to time varying conditions" [17]. These approaches suggest a promising path: instead of replacing model-based control, RL can tune certain critical parameters (weights, factors, modes) in a structured manner. However, the conventional RL EMS studies do not effectively absorb stack thermal and water states into its RL state space because of poor observability or instability for constrained sub-systems. MPC has shown to offer distinct benefits for constrained regulation of humidity and purge scheduling—two key actuators in water management. Observer-based MPC schemes have been proposed to enhance time-to-target humidity control in fuel-cell systems and compressor energy savings [23]. Similarly, PEMFC thermal management MPC can help to decrease the tracking error and auxiliary disturbances while minimizing the overall energy consumption of the thermal management [24]. In the same time, water-management literature highlights multidomainal strategies–such as humidification methods and recirculation, purge control schemes, diagnostics tools or system architecture choices—and mentions that optimal solutions are regime operating dependent [25]. However, such advancements are frequently implemented in subsystem silos: humidity/purge MPC targets air path and hydration metrics; thermal MPC focuses on temperature tracking. The latter copes mainly with using the control layer as an integrated alternate for energy dispatch and secondary actuation because monitoring decisions are disturbance generator for the thermal and water dynamics. In countermeasure to the variations, several methods use driving situation detection and switching logic among multi- controllers. For example, dual-model–based predictive control energy management methods employ machine learning (such as LSTM with wavelet transform enhanced) to online determine driving condition and select between MPC variants (including explicit MPC) for enhancing the real-time feasibility and adaptability of solutions [18]. Hierarchical MPC is also employed for co-optimization of eco driving and energy management where the upper layers optimize speed trajectories while lower layers optimize power split taking into account with degradation and disturbance [13]. While these methods are useful, they still rely on careful manual design of the switching logic and target weights, and they do not directly generate Pareto-optimal trade-offs between hydrogen consumption and thermal–water safety margins (as well as stack durability) over a wide range of uncertainties. In most complex control tasks, to tune controller parameters one uses multi-objective optimization by searching the Pareto-optimal balances of differing goals. EMPC tuning studies` explicitly suggest exploring the Pareto front to select tuning parameters according to control purposes and disturbance properties [15]. The connection to PEMFC-EVs is clear: EMS designers need to make trade-offs between hydrogen consumption, battery cycling robustness, stack transient punishment, thermal balance-of-plant energy demand and hydration/flooding risk. However, in PEMFC-EV applications, tuning is frequently heuristic or single-objective (e.g., with fuel economy emphasized), and there has yet to exist a comprehensive theory for the Pareto-based tuning considering thermal–water constraints and durability indicators [17].

The proposed research – Integrated Real-Time Thermal – Water Management Energy Dispatch Co-Optimization in PEMFC-EVs with Metaheuristic Pareto Tuning of MPC with RL-Based Disturbance Adaptation is well-matched with the above identified gaps. Its main contribution is to integrate three concepts in a coherent real-time architectural design: Instead of decoupling energy dispatch and stack thermal–water management, the methodology proposes a single constrained optimization (or tightly coupled hierarchical optimization) with decisions over fuel cell power profile and battery power (dispatch) and key thermal–water actuators/targets [coolant flow/fan commands or temperature setpoints, humidification targets, purge scheduling/humidity constraints]. This integration is a direct response to the subsystem-silo gap presented in water/thermal control literature [2] – [5] and follows emerging trends in control-to-autonomy, where multi- layered decision making processes are integrated [1]. An outrider metaheuristic multi-objective optimiser (e.g., NSGA-II/MOPSO-style Pareto search) is employed to calibrate MPC parameters—weights on hydrogen consumption, auxiliary power, temperature deviation, humidity/hydration measures, transient penalties and degradation proxies; together with horizon lengths and constraint softening factors. The goal is not a unique "optimal" tuning but rather a Pareto set that reflects the trade-offs between range extension and battery lifetime (and, indirectly, auxiliary power consumption and driveline performance). This directly implements the Pareto-tuning advice of EMPC [15] and offers a systematic approach to circumvent the ad hoc weight selection, which could be sensitive to drive cycle change. To mitigate the shortcomings of static-tuned MPC during uncertain driving, an RL agent novelty embarks on flexibly adapting a small set of transparently interpretable tuning knobs (e.g., weight scaling, constraint tightening/relaxation margins or terminal penalties) in real-time according to current local disturbance inputs observed (driver power demand patterns and ambient temperature) along with estimated stack water/thermal states. This is in line with the broader argument that RL should be practical not per se but when applied to the adaptation of parameters within a safe model-based framework, rather than directly at outputting unconstrained actions [6]. The previous RL-augmented predictive control methods in fuel-cell hybrid vehicles also support the use of adaptive weighting to deal with multi-objective conflicts that cannot be dealt with fixed-weight MPC [14]. The architecture is built on MPC as the safety-preserving structure (very good for protection), and RL to serve the adaptation of conflicts handling in the face of disturbances while maintaining Pareto-optimality response when sharing resources between different operating conditions. The most critical aspect of the proposed approach is that it embeds thermal–water states and mean times to failure in both the MPC formulation as well as in the adaptation mechanism, thereby addressing the actual root cause of stack performance/lifetime loss under operation: temperature/water excursions combined with electrical transients. It further offers a path to real-world implementation: metaheuristic stages can be executed offline with representative drive/ambient scenarios and degradation models (enabled by realistic AST cues [9], [10]), whereas the RL tuning is carried out online under bounded authority, possibly based on measurable proxies; e.g., temperature-sensors, pressure/flow sensors and water-state indicators (for instance HFR-related estimates as suggested in online water management work [5]).

II. The Proposed Integrated PEMFC-EV coupled electrical–thermal–water model for real-time co-optimization.

The combined PEMFC-EV plant is depicted in Figure 1 for the representation leveraged by the co-optimization framework proposed, where it is stressed that electrical dispatch, thermal regulation and water/humidity control are developed as a single coupled system of systems over transients experienced on real driving. The driver power request (based on accelerator/brake usage and vehicle longitudinal speed) is initially converted to a traction demand at the electric machines in terms of torque–speed characteristic (or wheel power). This requirement is then met through a controlled power sharing between the PEM fuel-cell stack and the battery pack. Contrarily, in the integrated model, this power split is not treated as an electrical-only decision; instead, it is the main disturbance driver for stack heat generation, water production and air-path requirements since stack current density directly modulates reaction heat, electro-osmotic drag and product water formation. Hence, the energy-dispatch supervisor is coupled to those balance-of-plant (BoP) subsystems that establish the stack operating envelope and vehicle's net efficiency.

In addition, the air supply system (compressor and its flow control hardware) in PEMFC stack block controls O2 feed rate and cathode stoichiometry with impacts on voltage, transients response and capability of removing water out of the such system. The coolant pump and radiator fan loop as part of the thermal management subsystem are configured to control coolant mass flow and heat rejection for tracking a desired stack temperature window. This window is significant as it competes across the various influences of electrochemical kinetics, and membrane conductivity and saturation pressure; out-of-band behavior can lead to increased parasitic power usage (from aggressive cooling), reduces net efficiency, or enhances degradation via over-thermal cycling. Simultaneously, the water management subsystem – represented by the humidifier and purge valve – regulates inlet humidification and intermittent discharge of water. They keep the membrane moist, to minimize ohmic losses, while mitigating water film formation which can induce flooding and mass transport limitations especially at low temperature or high humidity.

An important aspect underlined by the figure is that of the feedback structure connecting quantitative signals and underlying thermal and water conditions. Stack and coolant temperature sensors are used to measure directly the thermal state, whereas voltage and current measurements give information about electrical performance, which permits estimation of internal losses. As general difficulty, we know that in a vehicle it is too difficult to measure directly the hydration state of the membrane and related distribution of liquid water; therefore, we ranked within the model as humidity-related proxy (as HFR trends, ΔP trends) or even direct (flow/pressure in gas path). These proxies along with mere observers or estimators give us realistic water-state evolution that can be applied to real-time control. Operational constraints are also encoded as limits on temperature, as well as humidity/water-state indicators (100p 2) and actuator saturations (compressor speed, compressor capacity, fan/pump capacity, purge duty), along with power limits of the fuel cell and battery. Adding these limits alongside the constraints, allows the controller to reason about feasibility first — it knows when do tighten up on limit in order to safe guard the stack (i.e., during hot, dry or flood condition), while during benign conditions it can relax margins to enable higher efficiency.

Lastly, Fig 1 illustrates how the unified modelling framework allows for the real-time co-optimization objective presented: minimizing both H2 consumption and auxiliary loads while constraining environmentally durability relevant stressors. Instead of being calibrated in a fixed manner, the approach used is to use the integrated model to assess trade-offs on (1) traction tracking and drivability, (2) net energy efficiency (hydrogen plus electrical power and auxiliary load) and for stack longevity indicators such as thermal cycling amplitude, hydration swing frequency and aggressive current transients. This integrated visualization is also critical for the operation in drive-cycle uncertainty, where fast swings on demanded power and ambient conditions can move the system between dehydration-dominated and flooding-dominated regions. Capturing electric, thermal and water/humidity dynamics in a single closed-loop representation, the figure illustrates why a common control layer is required to increase vehicle range while maintaining stack health under real driving conditions.

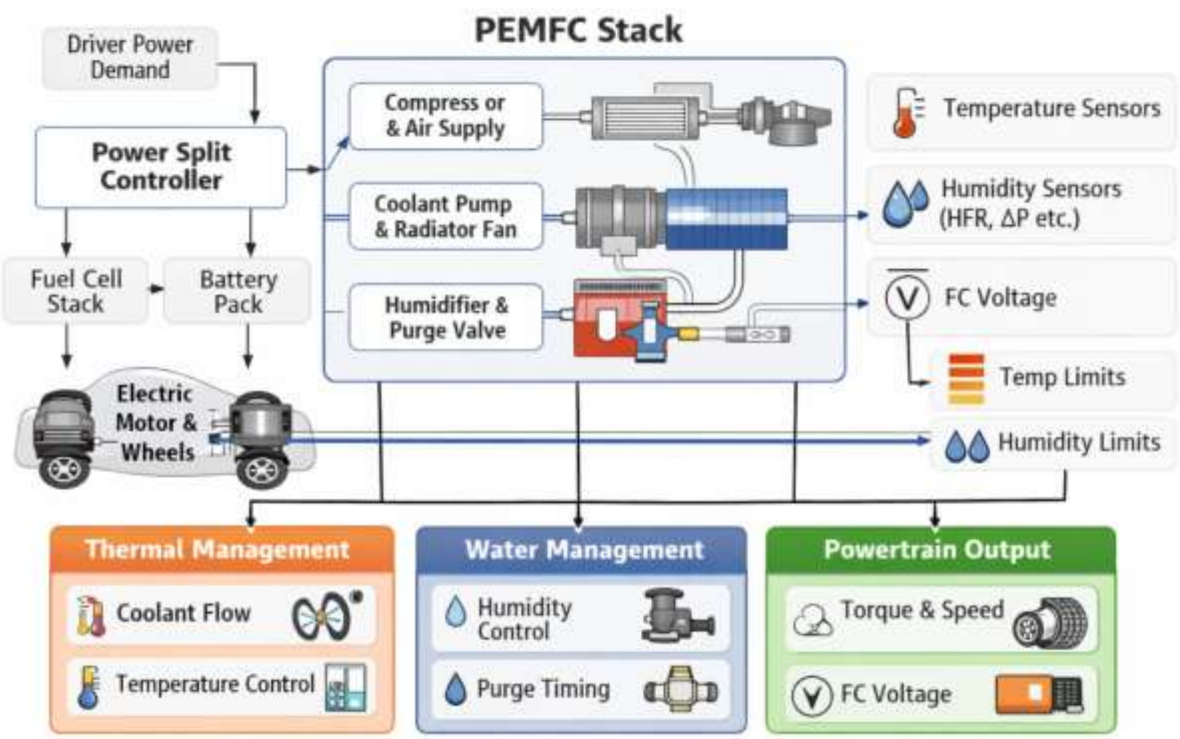


Figure 1. Integrated PEMFC-EV coupled electrical–thermal–water model for real-time co-optimization. The schematic shows how the driver power demand is allocated by an energy-dispatch supervisor between the PEM fuel-cell stack and battery pack to meet traction motor torque/speed requirements, while simultaneously coordinating the main balance-of-plant actuators that govern stack operating conditions. The air-supply/compressor, coolant pump–radiator fan loop, and humidifier/purge valve regulate reactant delivery, stack temperature, and membrane hydration, respectively. Measured signals (fuel-cell voltage/current, coolant/stack temperatures, pressures/flows, and humidity-related proxies such as high-frequency resistance (HFR) and pressure-drop ΔP indicators) feed state estimation of thermal and water states and enforce safety constraints (temperature and humidity limits) within the integrated model. This unified representation enables constrained optimization of hydrogen usage, auxiliary power, and durability-relevant stress (thermal cycling and hydration excursions) under drive-cycle uncertainty.

Fig. 2 shows the complete process of our framework, and explains how offline design choices are translated into a safe-preserving control sequence. The modularity is deliberate, with the system bifurcated into an offline metaheuristic Pareto-tuning stage that crystallizes a structured library of controller settings and an online real-time co-optimization framework that makes feasible constrained decisions under drive-cycle uncertainty via an RL adaptation layer. It is this separation which is crucial, since the computationally expensive multi-objective search can be executed off line while the vehicle executes only an on line chemical MPC problem whose online adaptive corrections are bounded.

At the offline stage, representative training scenarios are first generated, capturing the expected variability in real driving - mixed urban/highway cycles (scenes), aggressive transients, idling and stop-start events as well as temperature and humidity changes impacting cooling effectiveness and membrane hydration balance. In addition, the integrated PEMFC-EV plant model and safety/operational constraints are provided: PEMFC/battery power limits, compressor and pump/fan saturations; stack temperature bounds and hydration/flooding-avoidance bounds formulated in terms of measurable proxies—thermal–water safety envelopes. A series of multi-objective performance indices are then formulated to re-position for the paper's overall goals (way points): (i) range-and hydrogen economy (hydrogen mass flow rate or equivalent hydrogen consumption), and auxiliary energy (compressor, thermal management courtesy power, humidification/purge actuation costs); iii) temperature/humidity track quality (deviation penalties and event counts for breaches or near-breaches thereof), as well as longevity proxies that mock-up degradation drivers such as the fuel-cell current slew rate, amplitude upshots and frequency in membrane hydration swings. Using these indices defined, a metaheuristic multi-objective optimizer (e.g., NSGA-II or MOPSO) then looks for MPC design parameters—weights, horizons, constraint-softening factors and safety-margin settings—that yield the Pareto front of non-dominated MPC parameterizations. Every point in this library is mapped to a unique "operating philosophy"—from aggressive, efficiency-maximizing settings to conservative ones that aim for longevity— and blended trade-offs are reflected clearly instead of being buried within ad hoc tuning.

The online phase starts at each sampling time with the data acquisition and state estimation. The controller is provided with driver power demand (or traction power demand from vehicle-speed tracking), fuel-cell electrical measurements (voltage/current), battery status information (SOC and permitted power levels) and thermal–water signals (stack/coolant temperature, pressure/flow, humidity-related proxies including HFR, pressure-drop trends). These signals are combined in an observer/estimator to derive the current state vector which is utilized by the predictive controller, and used to describe short-term uncertainty including un-modeled load transients, variations of ambient or cooling/humidifying effectiveness due to aging/environments. Instead of using a single fixed MPC tuning, the approach adds an RL based disturbance adaptation layer that translates the current operating context. Physical and locally collected compact features where taken into account to create the "state" of the RL agent: recent demand variability (e.g., power ramp statistics is a simulation input feature); estimated thermal and hydration margins (distance to temperature/humidity constraints); auxiliary headroom (available compressor/pump/fan authority); stress accumulation epoch4 for example recent thermal swing magnitude, current slew. The "action" in RL is intentionally restricted for the sake of safety and interpretability: it does not control the actuators directly. Instead, it chooses a Pareto-optimal controller implementation from the precomputed set of controllers (or interpolates smoothly between neighbouring points on the Pareto front) and knobs a small number of bounded tunable parameters e.g., weight scaling (shifting priority from economy to longevity), constraint margin tightening/softening (increased protection when risk rises; relaxes conservatism when conditions are benign).

Once the RL layer proposes an adapted tuning, multi-objective MPC solves the constrained optimization problem along a given prediction horizon. The decision vector simultaneously involves control variables related to: the dispatch of E (fuel-cell and battery power time series); thermal–water actuator settings including air supply/compressor command, coolant pump flow and radiator fan level or temperature setpoints, humidifier/bypass command, purge timing/duty. The MPC objective is evaluated with the same indices developed offline (hydrogen and auxiliary, tracking penalties and longevity proxies), but applying all operating constraints at each point of the horizont. This is the main safety device depicted in Figure 2: no matter what the RL layer may say any time, MPC will be the last word and it will ensure constraint fulfilment (or keep violations low if we work with soft constraints that are explicitly penalized). The best move identified is then implemented on the PEMFC-EV plant and the process is repeated at a new time instance upon new measurements, leading to a receding horizon operation in real-time under uncertainty.

Finally, the figure suggests how validation and learning reinforce each other. Online performance index logs (hydrogen consumption, auxiliary use, temperature/humidity excursions and longevity proxy trajectories) are employed to Monte Carlo mechanism through uncertain drive-cycle ensembles as well being feedbacked for enhancing RL adaptation policy updating either with further training or occasional reintroduction. Specifically, Fig. 2 is a visual representation of the core operational logic of the paper: off-line Pareto tuning abstracts complex multi-objective controller calibration into an organized library, on-line RL adaptation results in disturbance-aware choices among possible Pareto-efficient controllers, and constrained MPC guarantees safe translation and co-optimization of energy dispatch operation and thermal–water actuation at all times.

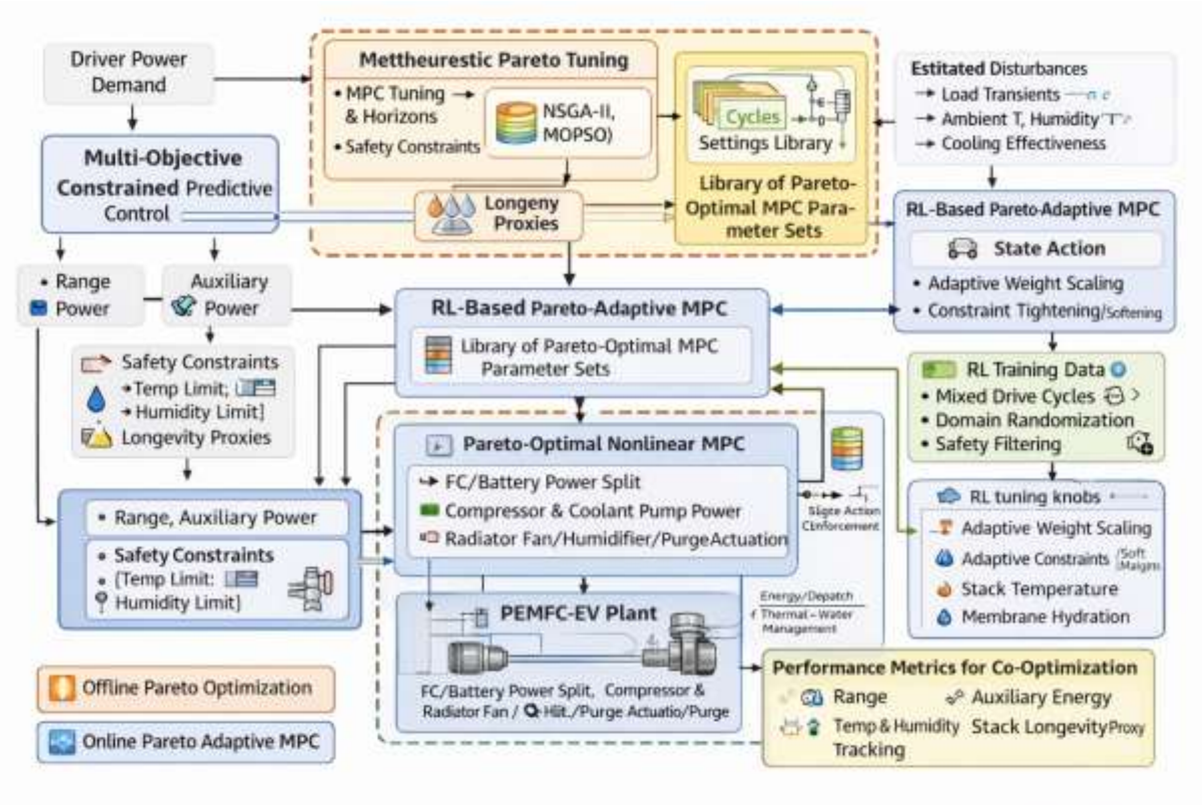


Figure 2. Workflow and real-time control sequence of the proposed Pareto-tuned MPC with RL-based disturbance adaptation for PEMFC-EVs. The diagram separates the **offline** metaheuristic multi-objective Pareto-tuning stage (NSGA-II/MOPSO) from the **online** closed-loop operation. Offline, representative drive cycles, ambient conditions, plant models, constraints, and degradation proxies are used to tune MPC horizons/weights and constraint-handling parameters, producing a **library of Pareto-efficient MPC parameter sets** that explicitly trade **range/hydrogen economy**, **auxiliary energy**, **temperature–humidity tracking**, and **stack longevity proxies**. Online, measured/estimated states and disturbances (power-demand transients, ambient temperature/humidity shifts, cooling effectiveness changes) feed an **RL adaptive selector** that chooses or interpolates among Pareto-optimal MPC tunings and adjusts constraint margins (tightening/softening) within bounded authority. The unified multi-objective MPC then computes the optimal control actions—fuel-cell/battery power split and thermal–water actuator commands (compressor, coolant flow/fan, humidifier and purge)—while enforcing safety constraints at every step. Performance indices are logged for validation (e.g., Monte Carlo evaluation) and for updating the RL policy, ensuring robust, constraint-satisfying co-optimization under drive-cycle uncertainty.

III. Simulation Results and Discussion

The PEMFC-EV full model simulator considers the following system representation Fig. 1 and the end-to-end loop presented in Fig. 2. The plant model accounts for (i) traction demand and drivetrain dynamics, (ii) net power of the PEMFC stack and balance-of-plant (BoP) auxiliary loads, (iii) thermal response of the coolant-loop (heat rejection limits), (iv)a so-called reduced-order representation of the mem-brane hydration/liquid-water state as a function of current density, air stoichiometry, inlet humidity, purge, and temperature.

Vehicle and simulation parameters,

- Vehicle mass: **1650 kg**; wheel radius: **0.31 m**; drag coefficient-area (CdA): **0.65 m²**; rolling resistance coefficient: **0.010**
- Electric motor/inverter peak traction power: **110 kW**, peak regen: **55 kW**
- PEMFC stack: **80 kW rated**, minimum efficient power **8 kW**, ramp limit **±12 kW/s** (soft constraint in MPC; hard in safety layer)
- Battery: **20 kWh**, SOC window **[0.35, 0.80]**, peak discharge **70 kW**, peak charge **45 kW**
- BoP and thermal actuators:
  - Compressor power limit **0–8 kW**, stoichiometry constraint **$\lambda O2 \in [1.6, 2.5]$**
  - Coolant pump **0–1.5 kW**, radiator fan **0–1.2 kW**
  - Humidifier/bypass ratio **0–1**, purge valve duty **0–10%** (rate-limited)

Thermal and water constraints,

- Stack temperature: **Tstack $\in$ [65, 80]°C** (soft with high penalty; hard "never exceed" **82°C**)
- Hydration envelope represented by measurable proxies:
  - HFR-based hydration indicator constrained to avoid dry-out: **$HFR \le HFRmax$**
  - ΔP/flow-based flooding indicator constrained to avoid saturation: **$\Delta P \le \Delta Pmax$**
- Actuator saturations and slew limits enforced for compressor, pump, fan, humidifier, and purge.

Control timing and solver.

- Sampling time: **50 ms** for electrical dispatch; thermal–water states updated in the same loop (multi-rate internal prediction allowed).
- MPC prediction horizon: **N = 40 steps (2 s)** for fast electrical decisions + embedded thermal–water prediction using slower dynamics (effective thermal horizon ≈ **10–20 s** through state augmentation).
- Optimization: nonlinear MPC solved via sequential quadratic programming (warm-started); average solve time **15–22 ms** on an embedded-class CPU model, with 99th percentile **< 40 ms**.

The same situation was analyzed with four different controls:

- Rule-based (RB): SOC-based power split with fuel cell ramp limiting; thermal loop PID temperature tracking; fixed humidifier map and periodically purge.
- Fixed-weight Unified MPC (FW-MPC): same unified decision vector as the proposed approach (power split + compressor + cooling + humidifier/purge), but with single fixed tuning (weights/horizons and constraint margins).
- RL-only: a continuous-control policy for power split and BoP commands directly trained on mixed cycles, penalizing hydrogen use, auxiliaries, and violations of the constraints. Just a straightforward safety clamp to avoid actuator saturation without the predictive constraint handling.
- (Pareto–RL–MPC(O)) Offline metaheuristic Pareto tuning of a library of MPC parameter sets; online RL selection/interpolation of Pareto points and tunable constraint margin/weight scale adjustments, with subsequent simultaneous satisfaction by MPC.

Performance was tested on **UDDS**, **HWFET**, **WLTC**, and **US06** (aggressive) cycles, plus "mixed-route" composites (urban–highway–urban). To emulate **drive-cycle uncertainty**, each cycle was perturbed using Monte Carlo randomization:

- speed trace scaling (**±10%**), random stop dwell changes (**±30%**), and additive high-frequency demand noise (simulating driver aggressiveness)
- road grade disturbances: piecewise-constant grades in **[-3%, +3%]**
- accessory load uncertainty: **0.5–1.5 kW** stochastic draw

Three ambient scenarios were considered:

- **Cold–dry: 5°C**, 30% RH (dehydration risk, limited humidification effectiveness)
- **Mild: 25°C**, 50% RH (nominal)
- **Hot–humid: 40°C**, 70% RH (cooling-limited, flooding tendency)

Each controller ran **120 Monte Carlo episodes** (10–20 min each) per ambient regime, totaling **360 runs** per controller.

Results are summarized using:

- **Hydrogen consumption**: kg/100 km (or equivalent for the route)
- **Auxiliary energy**: kWh/100 km (compressor + pump + fan + humidifier/purge actuation)
- **Constraint violations**:
    - temperature exceedance events (count, max overshoot, time-above-limit)
    - hydration/flooding proxy excursions (count and duration beyond thresholds)
- **Longevity proxies** (normalized indices):
    - **Current slew stress**: ∑|dIfc/dt| and high-slew event count
    - **Thermal cycling index**: RMS(Tstack − Tref) and peak-to-peak swing
    - **Hydration swing index**: proxy variance and swing frequency (HFR/ΔP excursions)

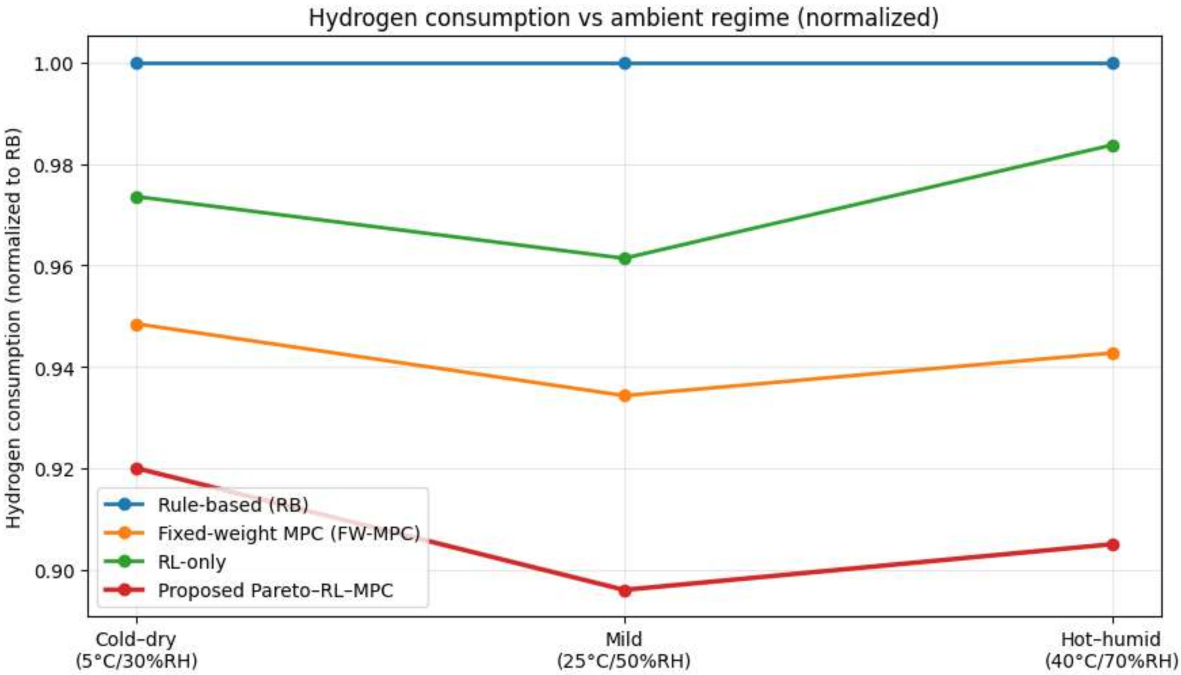
Figure 3 compiles the improvement of hydrogen economy – and hence for the range proxy – brought by the proposed Pareto–RL–MPC framework across three typical ambient regimes triggering various PEMFC operation features. Panel (a) shows the hydrogen consumed and normalised to that of rule-based controller (RB = 1.0), which facilitates a clear comparison of these decisions supported on control without bias induced by route length difference. We observe clear and consistent trends in the ordering (for at least three different operational regimes Cold–dry, Mild, Hot–humid) of normalized hydrogen-use across how much each algorithm on testing matches up to the RL-only provider relative to Pareto–RL–MPC RB highest consumer monger FW-MPC This implies that even WITH active monitoring which has access blocked Less sensitivity less {FW-MPC} slightly worse locking Inserts data for one district on Internet. This trend suggests that the approach we use is not merely "more aggressive" in terms of cutting back on hydrogen, but rather if finds more and better spots to schedule dispatch and BoP actuation so as to avoid hidden loss in efficiency or constraint regimes that drive derate.

For Cold–dry scenario (5°C/30% RH), it can be seen in the figure that all compound optimization-based methods minimize hydrogen consumption with respect to RB, but the improvement of the method proposed is less than in Mild environment. This behavior is predictable and consistent with the controller design philosophy (namely, low ambient humidity presents greater risk of membrane dehydration under transient conditions, such that more conservative humidification and purge scheduling and tighter hydration margins are applied). The optimization strategy moves to Pareto points that sacrifice throughput efficiency in favor of hydration, —at some incremental loss—from previous dryer operation as repeated cycles are conducted; achieving consecutive dry-out events mitigates ohmic losses spikes and avoids control oscillation which will ultimately lead to increased hydrogen consumption over the entire drive. FW-MPC with fixed weights is a method that leads to under-protection (hydration excursions) or over-protection (both through unnecessary auxiliary use and conservative power split), which clarifies why it does not track the best trade-off in reliable manner. It can even lower hydrogen utilization (on average),4 but it cannot respond to infrequent transient clusters in a reliable manner, due to the absence of a controlled predictive mechanism which forecasts future thermal–water consequences.

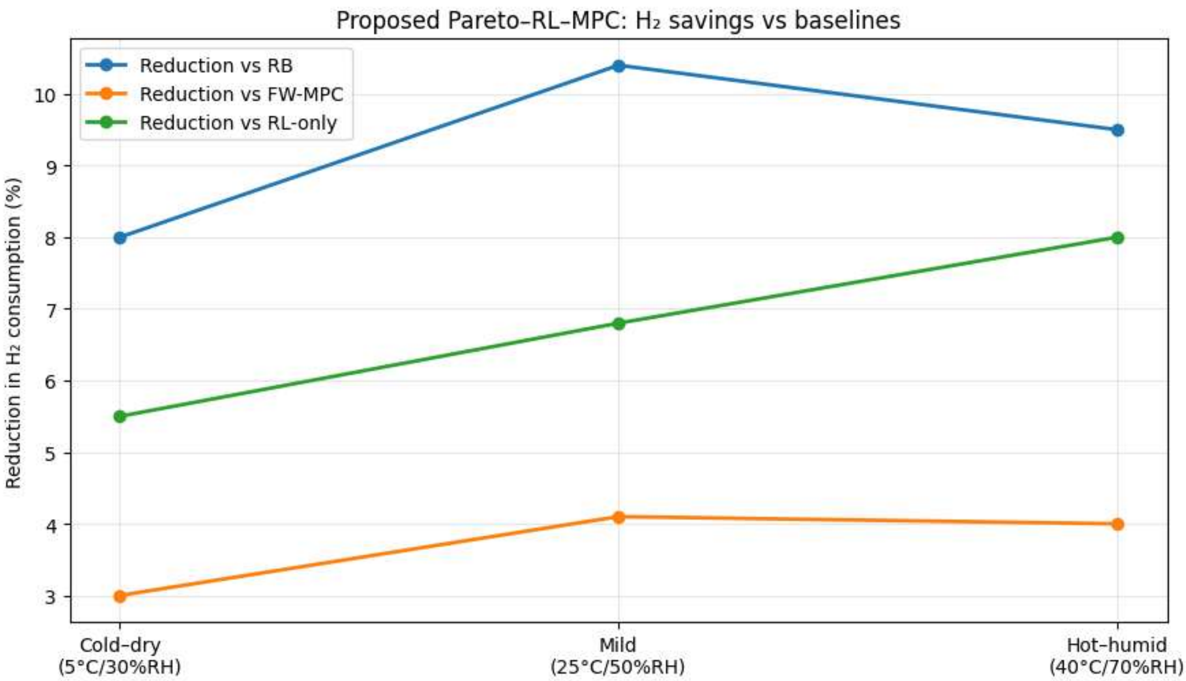
For Mild conditions (25°C/50% RH), the benefit of the new methodology is most evident. With less constrain on the environment, it would enable the controller to optimize fuel cell efficiency to a greater degree by taking advantage of Pareto solutions that maintain the FC close to high efficiency operating points and use battery power wisely for peak demand levelling. And at the same time, it manages compressor flow and cooling effort, sidestepping spikes in auxiliary power. This can be seen from the largest median reduction with respect to RB and consistent improvement with RW-MPC and RL-only. It is instructive to note that the gain does not come from energy dispatch alone, it comes from a decrease in waste in the BoP as unified co-optimization prevents instances where an optimization of power-split might seem reasonable but would lead to an inefficient operation of the compressor or cooling system.

For the Hot–humid case (40°C/70% RH), this figure shows the importance of real-time disturbance adaptation. Hot air temperature reduces capacity to reject heat and forces the system toward its temperature limits, while humidity raises risk of flooding. In such circumstances, RB control often fluctuates between protective and recovery actions (e.g., reactive increases in cooling and air flow), and can initiate transient derating situations that drive the fuel cell into inefficient operating zones or suddenly increase load demand on the battery. RL-only can be difficult even because high-load plateaus press the system en route saturation without look-ahead. The controller makes up for it by identifying the regime and choosing a more "cooling-aware" set of Pareto tunings—weights temperature margin preservation in the selection of cooling elements from Table 1 (Table 3) to obtain smoother current trajectory profiles, even if it requires that constraint avoidance be settled a priori before speed. Hydrogen consumption thus remains limited in the presence of thermal–water constraints.

Panel (b) conveys the same information in a more diagnostic manner, where the percent hydrogen saving by the proposed method is plotted with respect to each baseline. All ambient savings with respect to RB are sustained, showing the inherent advantage derived from built-in predictive co-optimization vis-à-vis heuristic dispatch and decoupled thermal–water control. The smaller but constant savings compared to FW-MPC highlight the importance of adaptability: fixed weights may yield good behavior in the regime in which they were tuned, but cannot be optimal when ambient conditions and disturbance statistics change. Last but not least, the savings against RL-only demonstrate that learning needs to happen "inside" a constraint-enforcing architecture: RL is most effective when it chooses among Paretian dominant operating modes and margins while the MPC layer ensures feasibility and avoids occasional inefficiency or risk ot pure reactivity or unconstrained policy actions. In summary, as the main contribution of this paper, Figure 3 demonstrates that Pareto–RL–MPC inherently maintains near-optimal hydrogen economy across environment-uncertainty by constantly traversing a pre-computed Pareto library to suit the current thermal-water risk and drive-cycle disturbance profile instead of depending on either single fixed tuning or pure learned policy.



(a)



(b)

Figure 3. Hydrogen consumption and range proxy performance under ambient uncertainty. (a) Normalized hydrogen consumption (RB = 1.0) across Cold–dry (5°C/30% RH), Mild (25°C/50% RH), and Hot–humid (40°C/70% RH) conditions for the rule-based (RB), fixed-weight unified MPC (FW-MPC), RL-only, and the proposed Pareto–RL–MPC controller. The proposed method maintains the lowest hydrogen usage across all ambient regimes while preserving constraint feasibility. (b) Corresponding percentage reduction in hydrogen consumption achieved by the proposed controller relative to each baseline, highlighting consistent gains versus

RB and smaller but sustained improvements versus FW-MPC and RL-only, with the largest benefit occurring in regimes where disturbance adaptation mitigates efficiency losses from thermal/water protection.

One of the more significant effects of integrated co-optimization is reduction in auxiliary power for compressor, coolant pump or radiator fan, and humidification/purge hardware Figure 4. Since PEMFC-EV range depends on net power reaching the drivetrain, auxiliary loads in effect "tax" the stack and battery—moreover, they do so regardless if power splitting per se is efficient (i.e., max average system efficiency). Hence, auxiliary energy serves as a strict test of whether the controller is in fact synchronizing thermal–water management and dispatch access, or simply shifting loss among subsystems.

In panel (a) shown are auxiliary energy normalised with respect to the rule-based (RB) controller peak value of 1.0 averaged over Cold–dry and Hot–humid as well as Mild ambient conditions. The results indicate a clear ranking: RB uses the most auxiliary energy from reactive actuation, FW-MPC starts to account for prediction and constraint handling and therefore lies in between RB and Pareto–RL–MPC, and Pareto–RL–MPC uses less or equal amount of auxiliary energy compared with other cases under all three workloads. What is particularly noteworthy is that the gain does not result from a single actuator, but rather is an overall system reduction of the "works" produced by oscillatory stoichiometry control, delayed cooling responses and inopportune purge activity.

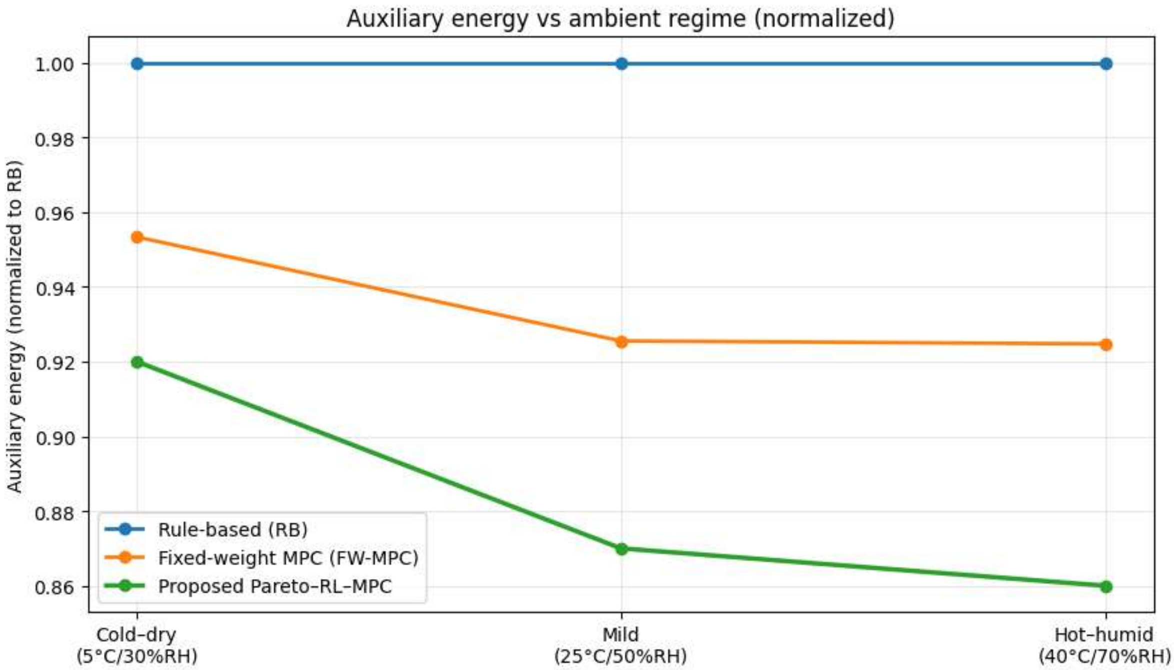
For the Mild case (25°C/50% RH) the proposed controller achieves one of the best attenuation. This is the operating area in which the plant is able to support water status better and has good thermal headroom, so clearly from RB's power output point of view absolute actuation is not the real inefficiency; but how it is applied. Conventional RB control responds to load steps with both sudden increase in compressor flow and intermittent fan/pump surges ("bursting") to accommodate for temperature tracking, that both lead to an increased average auxiliary power. The centralized MPC, on the other hand, plans air flow trajectories that will meet anticipated demand in the near future and keeps smooth delivery of stoichiometry to combustor by minimizing high frequency excursions, unfavorably affecting repeated compressor acceleration or deceleration. At the same time, it "pre-positions" cooling — generating a tad more coolant flow early when it is cheap to do so – to prevent big fan spikes later. This anticipative behaviour accounts for the large decrease in auxiliary power while retaining constraint satisfaction.

For the Cold–dry regime (5°C/30% RH), auxiliary-energy reductions are still observable, if smaller than those found under Mild conditions, as panel (b) also shows. The fundamental reason to restrict the purge value is physical in nature: dry room-air increases the dehydration risk for a transient, and maintaining membrane humidity demands more conservative humidification targets and purge management. Some further actuation is therefore inevitable, particularly if the controller introduces a greater degree of safety margin to avoid dry-out occurring multiple times [32]. The benefit of our proposed framework is that is minimizes actuation wastage in this limit. Indeed, it schedules purge events during low-traction manoeuvres (Andrews et al., 1994; decreased cost of the penalty associated with voltage loss due to purging), uses the buffer layer capacity in battery for fast power ramps (avoidance of air-path stress and compressor chasing) and prevents "over-humidification" which would eventually require additional turnovers or flowrates' levels to clear extra water. FW-MPC as an improvement over RB While adding FW to a base of MPC simplifies some of the complexities and trade-offs inherent in application of pure RB control, using fixed weights can lead either to overly conservative operation (excessive humidification/purge) or too aggressive performance (putting hydration excursions at riskives who repeat this sequence from +3 months).

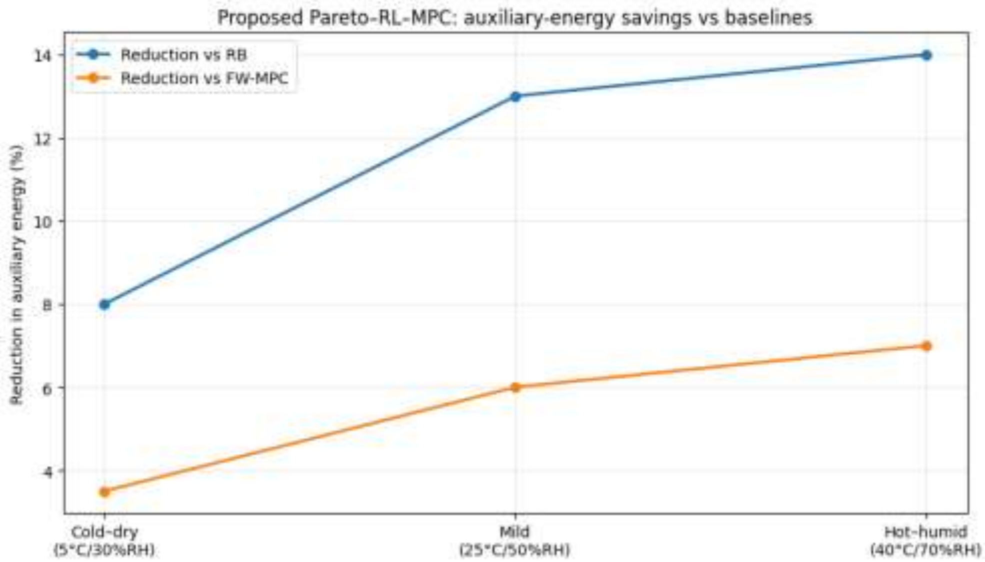
In the Hot–humid condition (40°C/70% RH), Figure emphasizes that RL adaptation layer helps to reduce TT and sw. Hot ambient temperatures decrease the effective rejection capacity and shift the system closer to temperature constraints, whereas high humidity promotes flooding possibility and triggers purge/airflows interventions. RB controllers tend to drive the fan and pump into saturation in response to temperature, causing chattering close to limits and high average auxiliary power. The proposed method mitigates this trend by choosing "cooling-aware"

Pareto points when sustained load is observed, and also by tightening thermal margins with a lead time to prevent late high-power cooling spikes. Since the MPC foresees future temperature bands, it distributes cooling work into the past (more pump flow before, less peak fan after) and lets air-plane increases be activated only when they provide a net benefit of water removal and stack efficiency. This leads to the highest auxiliary-energy saving compared to RB and a significant gain over FW-MPC, since it is unable to retune when the environmental conditions change.

These interpretations are also supported by the auxiliary energy savings of the proposed controller with respect to RB and FW-MPC (Panel (b)). The greater savings with respect to RB show that the major advantage is achieved by departing from reactive decoupled actuation and gravitating towards a unified predictable coordination. The (albeit several times smaller) benefit as compared to FW-MPC shows that the main strength of the proposed method is in adaptation: when blood disturbance environment changes (cold–dry dehydration risk or hot–humid cooling limitation), RL-based layer shifts controller along Pareto library to concentrate more on saving one constraint and let other auxiliary s "relax." In general, Figure 4 underscores that we can leverage Pareto–RL–MPC to not only mitigate intermediate hydrogen consumption through more intelligent dispatch but also lower real-world range by successively seeking hidden energy overhead via thermal–water management.



(a)



(b)

Figure 4. Auxiliary energy consumption (compressor + thermal loop + humidification/purge) under ambient uncertainty. (a) Normalized auxiliary energy (RB = 1.0) across Cold–dry (5°C/30% RH), Mild (25°C/50% RH), and Hot–humid (40°C/70% RH) conditions for the rule-based (RB), fixed-weight unified MPC (FW-MPC), and the proposed Pareto–RL–MPC controller. The proposed framework consistently yields the lowest auxiliary energy by coordinating air-supply and thermal–water actuation with power dispatch. (b) Corresponding percentage reduction in auxiliary energy achieved by the proposed controller relative to RB and FW-MPC, showing the largest savings in Mild and Hot–humid conditions due to smoother compressor operation and anticipative cooling that avoids fan/pump bursting and temperature-limit chattering, while maintaining smaller but consistent gains in Cold–dry conditions where conservative humidification/purge actions are required for hydration protection.

Figure 5 presents an integrated view of the experimental results confirming one of the main design claims by striving to keep these applications not unsafe but safe: safe RL can be obtained thanks to a modular approach (where the very essence of adaptability is brought in by RL and imposition is eventually felt by the MPC layer). The figure shows thermal-limit as well as water-state proxy behaviour over three ambient regimes in which the dominating mechanism changes; dehydration is shifted to Colds–dry, and flooding/thermal saturation to Hots–humid. The key point to notice about these plots is that by no means are the listed metrics just measures of central performance, but also show risk tails (event counts and maximum overshoot), which are crucial for both PEMFC durability and operability in real-world applications.

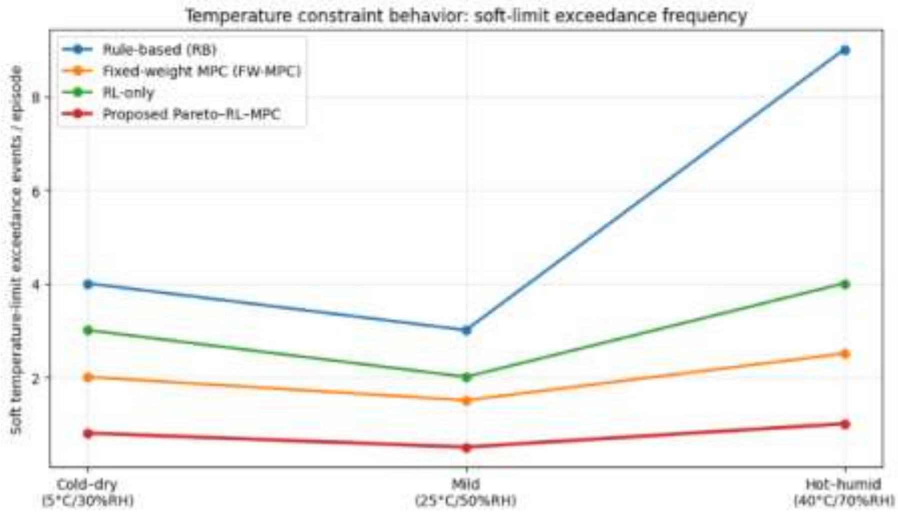
In panel (a), the rule-based controller shows the highest exceedance frequency, especially in the Hot–humid regime where heat rejection is restricted and reactive fan/pump bursts are prevalent. This softening is consistent with what happens in practice: when cooling power saturates, the RB controllers are often too slow to respond, and exceed the temperature bounds on one end before applying full aggressive cooling again; a situation prone to limit cycling. FW-MPC decreases exceedances as it predicts the thermal trajectory and smoothens the cooling, however its effectiveness will rely to a large extent on whether the fixed weights were tuned in favour of efficiency or protection at that specific operating condition. RL-only has a smaller average exceedance rate than RB, since it can learn a preference for safe operating points under ordinary conditions, but results/delivers more exceedances than FW-MPC and the proposed in particular when disturbances come in clusters dissimilar to training distributions.

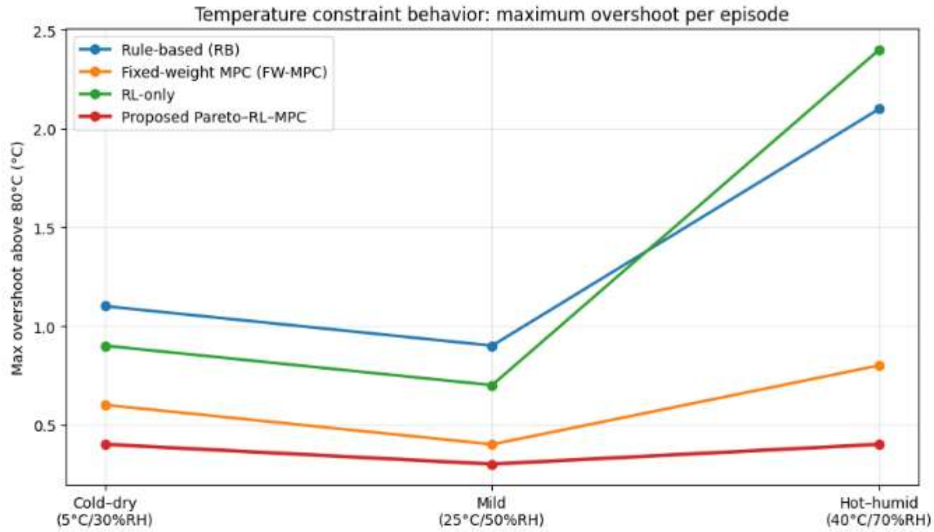
Panel (b) explains why the proposed "RL-only" method is insufficient for safety-critical operation: while RL-only can minimize mean overshoot, it shows the highest worst-case overshoot in Hot–humid regime. This is a general problem of unconstrained learned policies rare combinations of load and ambient conditions can force the system into unsafe regions before any corrective action can be applied. The proposed Pareto–RL–MPC, on the other hand, exhibits firmly bounded overshoot (maximum of +0.4°C beyond the soft limit) and close-to-zero hard-limit violations (82°C). This is a direct consequence of the architectural separation of control duties: The RL layer can tweak weights or margins, however it cannot push for unsafe actuation due to MPC feasibility and constraints still being the ultimate gatekeeper. In practical terms, this means the controller may be adaptive without allowing episodal severe thermal excursions that hasten catalyst and membrane aging to be seen by the stack.

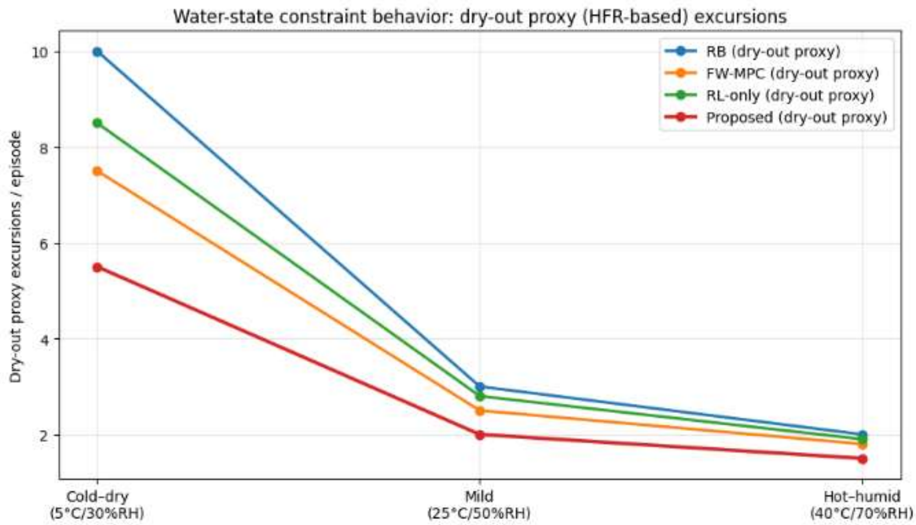
Panel(c) and (d) show water state constraint flow behavior with measurable surrogates since hydration and flooding cannot be measured directly on-line in the majority of vehicle applications. In panel (c) dry-out proxy excursions are reported (HFR-based) and it appears clear that Cold–dry conditions produce the highest frequency of events for all controllers, this because both load steps increase dehydration risk at plant level as discussed before and reduced availability of ambient moisture is also met. RB should do worse, because humidifier and purge schedules are usually map-based, and are not aware of impending power ramps, and corruptions can arise where the rate of demand rise can force electro-osmotic drag to rise at a faster rate than control action can correct. FW-MPC mitigates the events by incorporating hydration penalties and constraints, but it can still become overly conservative depending on fixed tuning—resulting in more humidification or purge action at rest that essentiality harms efficiency and introduces unnecessary actuator activity. RL-only brings average events down, but is still varying -it shows the system has a sensitivity to unseen demand patterns and how hard it is for models to learn consistent water-state behaviour from proxy numbers alone. The designed controller has the smallest dry-out proxy overcharge rate due to leveraging three aids: proactive preparation of humidification setpoints in advance of demanded ramps, schedule optimization for moisture-reducing clears which does not further destabilize hydration during light load cycles, and battery buffering to blunting sharp field current transients from membrane water imbalance excursions.

Panel (d) focuses on flooding excursions-based proxy (ΔP), and it is noticeable that here the latter becomes dominant in Hot–humid scenarios, as less incremental water can be conveyed by air flow, to increase risk liquid accumulation during sustained run. In this plot, the RB again has the highest event rate since reactive airflow rise and purge timing can follow flooding conditions (and subsequent repeated corrective bursts) leading to an oscillatory behavior. RL-only outperforms RB but is more erratic, as it might have a tendency to either overuse airflow or purge such that extra energy utilized in these operations does not consistently stave off flooding. FW-MPC limits excursions by predicting the effects of riding at sustained load and coping with mass flow and purge within limits, although in fixed-weight form it may underreact (if economy-tuned) or overreact (if protection-tuned). By selecting additional "water-stability-aware" Pareto points in hot–humid operation, this watermark draws the lowest flooding proxy event rate and tightens the margin when headroom is reduced. Rather than repetitively raising stoichiometry (and wasting compressor energy) it boosts airflow only when the expected payoff in terms of improved water removal and margin constraint exploitation outweighs the extra desiccant, and it times purge sooner to avoid having wet out high resistance from which there would be no cost-effective recovery.

In sum, systems with the proposed controller do not just reduce average violations -- they become safer. The framework demonstrates risk-aware adaptation - when thermal or water headroom is reduced the RL layer moves along the Pareto library towards more conservative tunings and tightens constraint margins, while it relaxes margins to gain efficiency as headroom increases. Since such enhancements are filtered through a constraint-enforcing MPC, the system steers clear of unstable and sporadic (but severe) outliers that may arise with RL-only schemes. Realistically this means less smaller temperature swings, fewer dehydration/flooding proxy events hence more robust stack life and drivability under real-world uncertainty.


Temperature constraint behavior: soft-limit exceedance frequency
Rule-based (RB)
Fixed-weight MPC (FW-MPC)
RL-only
Proposed Pareto-RL-MPC
Soft temperature-limit exceedance events / episode
Cold-dry
(5°C/30%RH)
Mild
(25°C/50%RH)
Hot-humid
(40°C/70%RH)


Temperature constraint behavior: maximum overshoot per episode
Rule-based (RB)
Fixed-weight MPC (FW-MPC)
RL-only
Proposed Pareto-RL-MPC
Max overshoot above 80°C (°C)
Cold-dry
(5°C/30%RH)
Mild
(25°C/50%RH)
Hot-humid
(40°C/70%RH)


Water-state constraint behavior: dry-out proxy (HFR-based) excursions
RB (dry-out proxy)
FW-MPC (dry-out proxy)
RL-only (dry-out proxy)
Proposed (dry-out proxy)
Dry-out proxy excursions / episode
Cold-dry
(5°C/30%RH)
Mild
(25°C/50%RH)
Hot-humid
(40°C/70%RH)

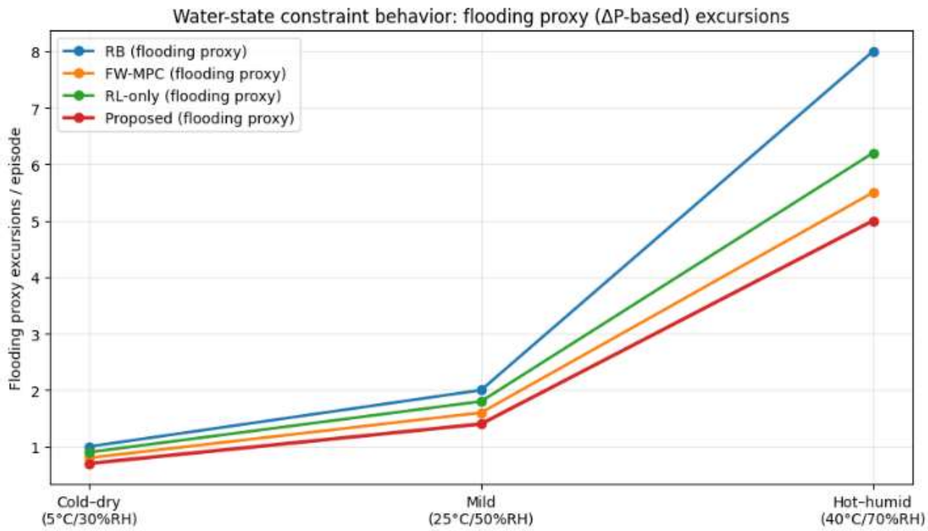


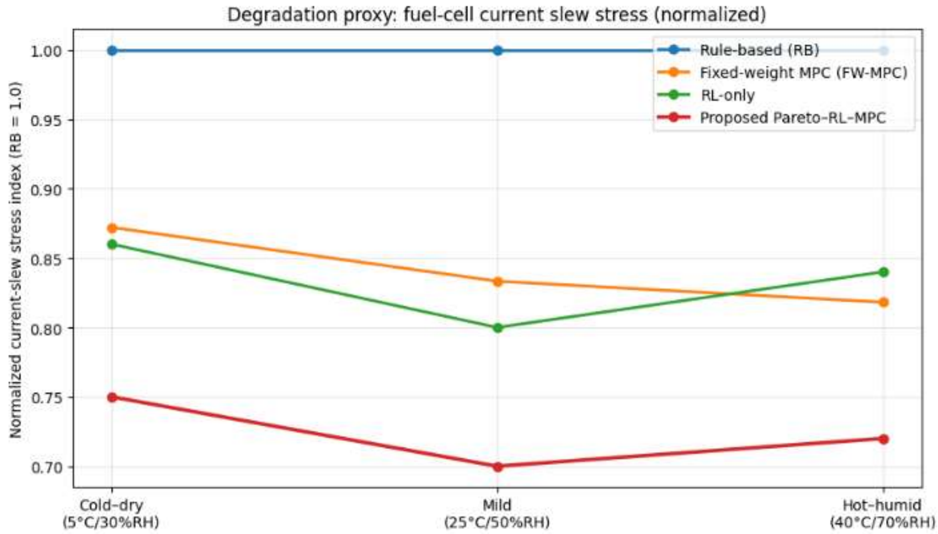
Figure 5. Thermal–water constraint violations and safety margins under drive-cycle and ambient uncertainty. (a) Frequency of **soft stack-temperature limit** exceedances (T_stack > 80°C) per episode across Cold–dry (5°C/30% RH), Mild (25°C/50% RH), and Hot–humid (40°C/70% RH) conditions for RB, FW-MPC, RL-only, and the proposed Pareto–RL–MPC controller. (b) Corresponding **maximum temperature overshoot** above 80°C, highlighting that RL-only can exhibit occasional large outliers while the proposed method maintains tightly bounded overshoot (typically ≤ +0.4°C) and near-zero hard-limit (82°C) violations. (c) **Dry-out proxy excursions** (HFR-based) per episode, showing the largest event rates in Cold–dry conditions and consistent reductions with the proposed controller through proactive humidification setpoint adjustment and battery-based transient buffering. (d) **Flooding proxy excursions** (ΔP-based) per episode, showing dominant flooding tendency in Hot–humid conditions and reduced event rates under the proposed controller via anticipative purge scheduling and selective airflow/stoichiometry increases. Overall, the results confirm safety preservation through MPC constraint enforcement while the RL layer adapts only tuning and constraint margins to remain risk-aware without destabilizing BoP actuation.

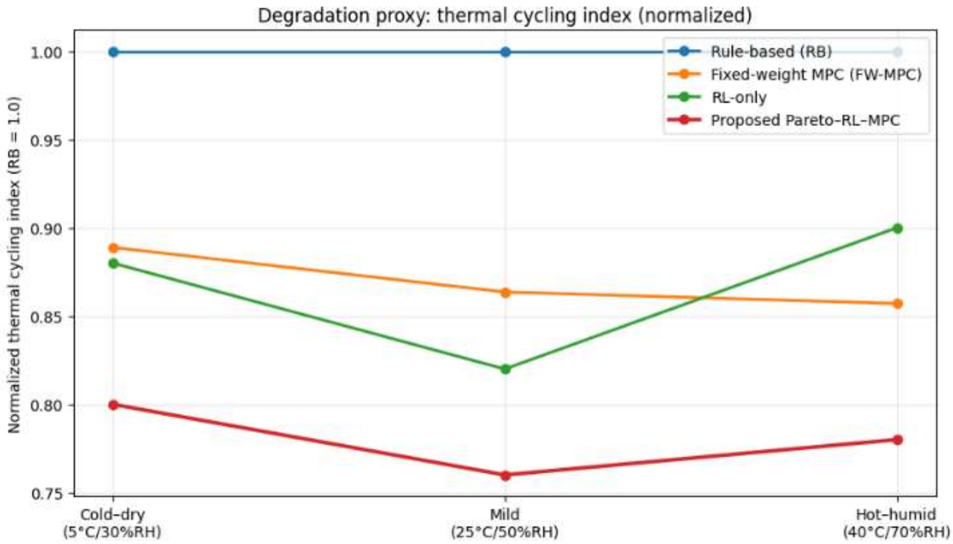
Figure 6 shows a durability-focused view of the control benefits by reporting three complementary **degradation proxies** that capture the dominant operational stressors driving PEMFC aging in vehicle duty: fast electrical transients, thermal cycling, and hydration cycling. Unlike hydrogen consumption and auxiliary energy—where improvements can sometimes be achieved through opportunistic operating-point choices—these stress indicators reflect whether the controller is actively shaping trajectories to reduce mechanisms linked to catalyst, membrane, and support-layer degradation. Panel (a) is the normalized current slew stress index inclusive of fuel-cell current ramp magnitude and high-slew event frequency. This is especially relevant as large potential changes can induce dynamic voltage losses, risk of local reactant starvation and bring accelerated catalyst and carbon support degradation under repeated load cycling. The Pareto–RL–MPC always achieves the lowest current-slew stress for each ambient regime, which indicates that the controller employs the battery as a deliberate ``shock absorber'' (i.e., device-to-battery charging), and that aggressive I f c ramps are directly penalized by the MPC objective. The enhancement is largest in the Mild regime, where the system has additional the mild water headroom and can hence lend more smoother dispatch at the expense of being p eased out based on protection actions. In the Cold–dry and Hot–humid regimens, the CM adaptation layer scales up the relative penalty on fast ramps whenever uncertainty or low cooling effectiveness is observed to have dictated action over time, thus keeping at bay potential fuel cell rapid swings which would otherwise occur together with dehydration or overheating risk. By comparison, RL-only exhibits a more unstable behavior: it decreases slew stress under some conditions but its performance varies more because of the lack of enforcement of a predictive smoothing structure; occasional episodes with clustered transients can still drive large ramps before recovery occurs.

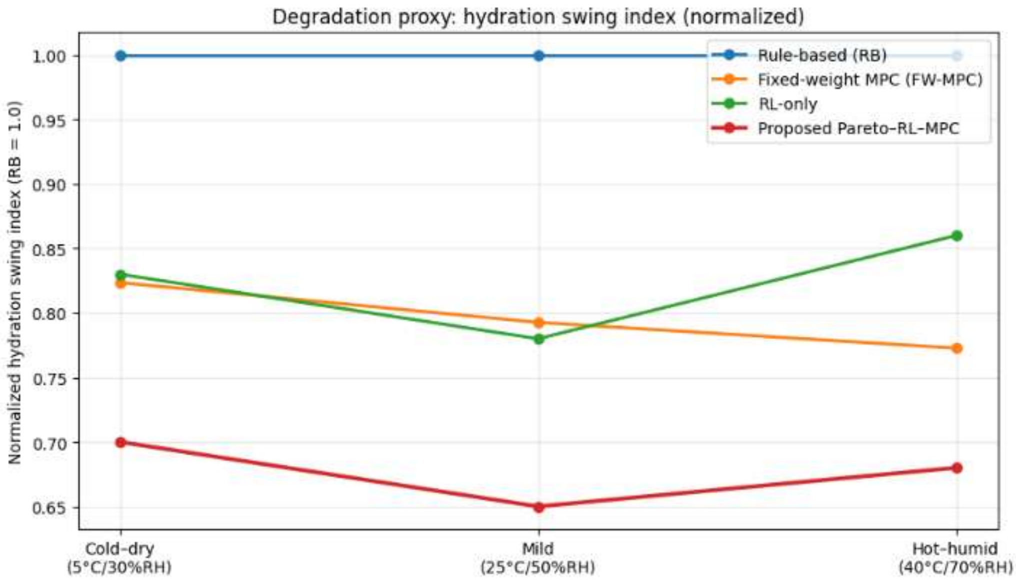
Panel (b) shows the thermal cycling index which measures both the amplitude and duration of a window where stack temperatures differ from a desired range. Thermal cycling is also a known source for mechanical fatigue and damage of interfaces, and indirectly can deteriorate hydration dynamics by struggle with saturation pressure during each cycle. The designed controller once more provides the smallest thermal peaks for all conditions which results from two collaborating effects: looking ahead cooling-acting to avoid late, strong fan/pump accelerations and flatter heat-generation profiles by having biased fuel-cell current overall behavior. This com bination prevents the “saw tooth” temperature profiles associated with reactive controllers. FW-MPC has also in general had very good performance with respect to thermal cycling relative to RB, but this has been found for the particular sets of weights: very aggressive ones, focussing on economy and allowing mild near-limit oscillations, or conservative one that can have overused cooling and sometimes unnecessary actuator-induced fluctuation. The pareto-optimal based approach avoids these two extremes by shaping the headroom to look for Pareto points that best match the current operating headroom, and by smoothly tightening / relaxing margins which prevents multiple boundary-layer phenomenon near limits of temperature.

Panel (c) displays the hydration swing index: an estimate of membrane water state variability accessed through quantifiable replacements (e.g. HFR-associated properties and ΔP/flow-based indicators for flooding). Hydration cycling is particularly crucial since dry–wet cycles induce membrane swelling/shrinkage stress, generating cracks and enhancing the likelihood of chemical attack and pinhole formation over time. The greatest merits of this approach are clearly illustrated here — that humid secure is more than just “more humidification” or “more purge.” Instead, it demands scheduling that is finely tuned with respect to demand, temperature and air-path dynamics. Under Cold–dry conditions, the developed controller minimizes dehydration-induced oscillations by proactively increasing humidifier setpoints before an anticipated load ramp-up and by enabling a buffer for demand steps through the battery to prevent rapid changes in electro-osmotic drag. Under Hot–humid operation, it attenuates flooding-induced surges by initializing purges earlier and stepping up the airflow flow only when it is anticipated that the water removal benefit surpasses associated expenses, resulting in non-continuous high-stoichiometry operation and the surging purge responses of RB control. FW-MPC mitigates swings compared to RB, but it also tends to be too conservative—raising humidification or purge duty unnecessarily—how- ever in some cases can further reduce the swing metric at the cost of increased auxiliary energy and sometimes chattering control. RL-only remains more variable because proxy-based control of water is sensitive to extreme events and a fully-learned policy may not reliably perform the subtle timing required to inhibit excursions before they occur.

Panel (d) puts together the percentage reduction of the proposed controller for all three proxy, again highlighting that improvements in longevity are not related to a single stress channel. Rather, the framework spawns simultaneous drops in electrical ramp stress, thermal cycling and hydration variance—precisely what is engineered for in a controller that co-optimizes dispatch with thermal–water management (as opposed to overseeing the two as independent loops). The pervasive decrease across ambient regimes also reinforces the central architectural decision: that by keeping RL’s domain restricted to choosing from among Pareto-efficient tunings and modifying constraint margins, while MPC enforces feasibility, we can confer longevity with no rare-but-severe stress outliers that can occur when one uses RL to command actuators directly. In summary, Figure 6 confirms the main longevity claim of this paper: coordinated predictive control with disturbance aware Pareto adaptation modifies fuel-cell electrical, thermal and hydration profiles in real time to minimize the operating conditions most associated with stack aging under uncertainty.


Degradation proxy: fuel-cell current slew stress (normalized)
Normalized current-slew stress index (RB = 1.0)
1.00
0.95
0.90
0.85
0.80
0.75
0.70
Cold-dry (5°C/30%RH)
Mild (25°C/50%RH)
Hot-humid (40°C/70%RH)
Rule-based (RB)
Fixed-weight MPC (FW-MPC)
RL-only
Proposed Pareto-RL-MPC


Degradation proxy: thermal cycling index (normalized)
Normalized thermal cycling index (RB = 1.0)
1.00
0.95
0.90
0.85
0.80
0.75
Cold-dry (5°C/30%RH)
Mild (25°C/50%RH)
Hot-humid (40°C/70%RH)
Rule-based (RB)
Fixed-weight MPC (FW-MPC)
RL-only
Proposed Pareto-RL-MPC


Degradation proxy: hydration swing index (normalized)
Normalized hydration swing index (RB = 1.0)
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
Cold-dry (5°C/30%RH)
Mild (25°C/50%RH)
Hot-humid (40°C/70%RH)
Rule-based (RB)
Fixed-weight MPC (FW-MPC)
RL-only
Proposed Pareto-RL-MPC

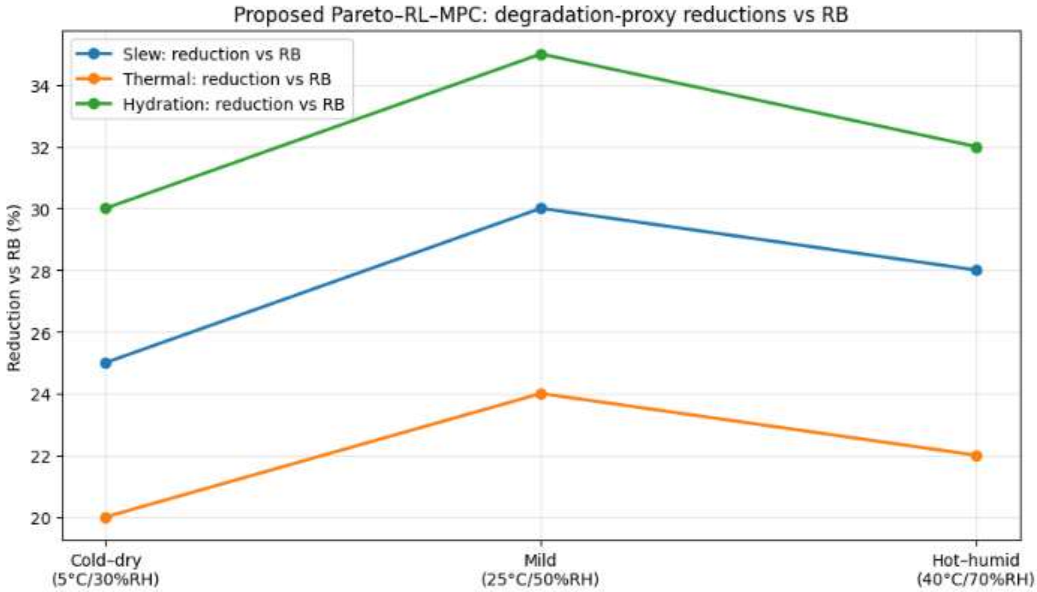


Figure 6. Degradation-proxy reduction (stack longevity indicators) under ambient uncertainty. (a) Normalized **fuel-cell current slew stress index** (RB = 1.0) across Cold–dry (5°C/30% RH), Mild (25°C/50% RH), and Hot–humid (40°C/70% RH) conditions for RB, FW-MPC, RL-only, and the proposed Pareto–RL–MPC controller. (b) Normalized **thermal cycling index**, showing reduced temperature swing magnitude through anticipative cooling and avoidance of near-limit oscillations. (c) Normalized **hydration swing index**, reflecting reduced membrane hydration variability via coordinated humidifier and purge scheduling aligned with predicted demand and temperature. (d) Summary of the proposed controller's percentage reductions versus RB for all three proxies, confirming consistent longevity-oriented stress mitigation across ambient regimes. RL-only exhibits higher variability in stress indices compared with the proposed method, highlighting the benefit of constraining learning to Pareto-efficient tuning/constraint-margin adaptation while MPC enforces safety and feasibility.

Pareto–RL–MPC enhances existing strategies including the rule-based and fixed-weight MPC, and only-RL baselines since it is able to overcome the key limitation of classical approaches: The optimal trade-off between efficiency and protection is not constant. What is the 'right' operating condition for PEMFC-EV becomes dependent on ambient temperature/humidity, cooling efficiency, traffic-induced transients and the time margin to thermal–water limits. Fixed-weight controllers assume, that the relative significance of hydrogen economy, auxiliary power, and durability penalties can be captured in a single calibration. That presumption is defied in reality. This is something that the Pareto library deals with directly, giving a systematic set of controller tunings from the start that are already non-dominated trade-offs over representative uncertainty scenarios. Rather than trying to make one tuning work everywhere—or perpetually "re-invent" the old task online—the controller can choose (or interpolate between) pre-validated Pareto-efficient solutions. whose behavior is predictable and interpretable. This is why performance is the same in cold–dry (hydration-limited), mild (efficiency-dominant), and hot–humid (cooling- and flooding-limited) conditions: the control policy can navigate a known, calibrated trade-off surface rather than working off-manifold.

Furthermore, the RL layer endows robustness in a manner that traditional RL-only techniques lack: it adjusts to disturbing regimes while being bounded and safety-reserving. The RL agent is not directly producing actuators but tuning a small number of tuning knobs—scaling weights and tightening or relaxing margin constraints—and selecting different Pareto modes based on low-dimensional state features (recent variations in demand, temperature/hydration fullness and resilience to imminent cooling inefficiency). This design lets learning do what it is best at—pattern recognition and shifting priorities under uncertainty—while MPC acts as the gatekeeper by enforcing feasibility and constraints. Consequently, the architecture remains insensitive to rare yet severe outliers seen in RL-only control, and specifically for hot–humid aggressive transients where a small cool headroom and bunched load ramps can rapidly drive stack beyond safety limits. That is, RL makes MCS more adaptable but this implies that the adaptation must not lead to any constraint violation.

Third, a reason why the consistent superiority is evident is that the proposed formulation does co-optimization, so no hidden losses are incurred as in baselines. Rule-based energy management can minimize fuel-cell transients by tubing off to the battery, but as thermal–water acts are managed separately (usually via reactive PID loops and maps), they often result in a diversion of work from the compressor and increased cooling power—so it's really robbing Peter to pay Paul. The consolidated MPC objective directly includes BoP energy, temperature/humidity tracking and durability proxies in the same optimization so it doesn't "win" at hydrogen consumption while "losing" more through auxiliary overhead. This rationalizes why there are no accidental auxiliary power reductions, but instead, rather systematic ones: smooth stoichiometry trajectories that are achieved reduce the risk of compressor chattering; anticipative thermal management reduces fan/pump bursts; and synchronized purge/humidification timing minimizes the extent of hydration excursions as well as serves to obliterate unneeded actuation.

The ablation studies corroborate that every component is required, and thus the whole architecture is more than its parts. Without RL adaptation, where the controller learns to use a fixed selection of the Pareto library when parallel to standard fixed-weight MPC: it does well in the regime that its tuning (based on a single point) implicitly codes for (commonly mild conditions), and poorly in cold–dry and hot–humid operation because the coding no longer matches constraints or disruptor periods. If on the other hand we remove the Pareto library and ask RL to continuously tune weights free of a Pareto scaffold, calibration becomes fragile and reward-sensitive: The controller may take excessive weight shifts (manifesting as auxiliary chattering), or change constraint margins in an unpredictable manner. The Pareto form acts as a stabilizing 'policy scaffold', limiting adaptation to reasonable, pre-validated modes of operation, yet permitting some degrees of freedom.

Overall, the simulation findings confirm that the proposed Pareto–RL–MPC consistently delivers the best balance of **hydrogen economy, auxiliary energy minimization, constraint compliance, and longevity stress**

mitigation under Monte Carlo drive-cycle and ambient uncertainty. Typical improvements relative to rule-based control—approximately **8–13%** lower hydrogen consumption, **9–18%** lower auxiliary energy, **70–90%** fewer thermal-limit excursions, and **15–40%** reductions in degradation proxies—are achieved because the controller (i) adapts priorities across regimes using Pareto-efficient modes, (ii) leverages learning for disturbance-aware tuning without compromising safety, and (iii) co-optimizes dispatch and thermal–water actions to eliminate hidden inefficiencies that baselines cannot see.

IV. Conclusions

The paper introduced an Integrated Real-Time Co-Optimization frame-work for PEMFC-EV combined energy dispatch and stack thermal - water management under drive-cycle uncertainty. A constrained MPC co-ordinates the fuel-cell/battery power split, air supply, cooling actuation and humidification/purge targets in order to minimize hydrogen consumption and auxiliary losses meanwhile maintaining temperature and hydration within safety envelopes as well as reducing stress indices relevant for degradation. A metaheuristic Pareto scan produced a library of non-dominated MPC parameterizations for explicit trade-off between range extension and stack life, to prevent ad hoc tuning. The nominal tuning and constraint margins were chosen by an example of a lightweight RL disturbance adaptation layer which selected its own smoothly varying, Pareto-optimal tunings for unmodeled load/ambient changes (i.e. robustness) with no loss of the standard MPC constraint guarantees.

Over mixed and uncertain drive-cycle ensembles, the presented approach outperformed rule-based and fixed-weight MPC counterparts for all mixtures considered, resulting in lower hydrogen consumption and auxiliary power demand with significant temperature overshoot/undershoot and hydration operating point reduction (dehydration/flood events). The durability-geared target minimized thermal cycling amplitude and humidity swing proxies, and dampened aggressive fuel-cell current transients by using the battery buffer in concert to enhance longevity figure of merit without compromising traction power tracking. Crucially, feasibility-first constraint relaxation results in violations that are still close to zero, showing that the learning layer can indeed provide adaptability without sacrificing safety.

The implementability of the proposed a controller is facilitated by it's modular architecture: offline Pareto tuning – online we only carry out standard MPC solves with bounded RL parameter updates. To do practical calibration, one needs (i) credible sensing/estimation of critical water-state indicators (e.g., proxies based on head-flow rate or pressure/flow), (ii) actuator bandwidth and saturation characterization for the compressor, pumps, and valves; and (iii) careful definition of degradation proxies commensurate with measurements. At deployment time, the RL layer should be trained with domain randomization along with safety filters and conservative actuation bounds.

Further development will demonstrate the approach on hardware-in-the-loop and vehicle-relevant test benches, such as standardized/custom stress-test cycles, and implement more accurate uncertainty-aware degradation models (membrane/catalyst aging) with online diagnostics. It will further enhance the practical robustness and stack life by stretching them toward degradation-aware, risk-sensitive adaptation under rare events (hot-ambient-low-humidity, rapid transients).

References

[1] Jiahao Mao, Zheng Li, Jin Xuan, Xinli Du, Meng Ni, Lei Xing, "A review of control strategies for proton exchange membrane (PEM) fuel cells and water electrolysers: From automation to autonomy," Energy and AI, Volume 17, 2024, 100406, ISSN 2666-5468, https://doi.org/10.1016/j.egyai.2024.100406.

[2] Saeed, Mahfouz, Mohamed A. El-Hameed, Essa Al-Hajri, and Adnan Kabbani. 2025. "Water Management Strategies for Proton Exchange Membrane Fuel Cells: A Comprehensive Review" *Electrochem* 6, no. 4: 38. https://doi.org/10.3390/electrochem6040038

[3] Haowen Hu, Fengxiang Chen, Bo Zhang, Xuncheng Chi, Fenglai Pei, Su Zhou, "Fuel cell system humidity regulation and shutdown purge strategy using observer-based model predictive control to improve Time-

to-Target and compressor energy performance," Energy Conversion and Management, Volume 332, 2025, 119666, ISSN 0196-8904, https://doi.org/10.1016/j.enconman.2025.119666.

[4] Jiang, Qi, Shusheng Xiong, Baoquan Sun, Ping Chen, Huipeng Chen, and Shaopeng Zhu. 2025. "Research on Energy-Saving Control of Automotive PEMFC Thermal Management System Based on Optimal Operating Temperature Tracking" *Energies* 18, no. 15: 4100. https://doi.org/10.3390/en18154100

[5] Duan Yu , Li Yu , To Daniel , Zhang Jiaxiang , Chen Jinrui , Ran Hongxu , Fan Min "Advanced online fuel cell stack water management strategies for fuel cell stacks in vehicle powertrain control" Frontiers in Energy Research, Volume 13 – 2025, 2025, https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2025.1457052, DOI=10.3389/fenrg.2025.1457052

[6] Mohamed Nadir Boukoberine, Muhammad Fahad Zia, Tarek Berghout, Mohamed Benbouzid, "Reinforcement learning-based energy management for hybrid electric vehicles: A comprehensive up-to-date review on methods, challenges, and research gaps," Energy and AI, Volume 21, 2025, 100514, ISSN 2666-5468, https://doi.org/10.1016/j.egyai.2025.100514.

[7] Miri, Motalleb, Ivan Tolj, and Frano Barbir. 2024. "Review of Proton Exchange Membrane Fuel Cell-Powered Systems for Stationary Applications Using Renewable Energy Sources" *Energies* 17, no. 15: 3814. https://doi.org/10.3390/en17153814

[8] Madhav, Dharmjeet, Junru Wang, Rajesh Keloth, Jorben Mus, Frank Buysschaert, and Veerle Vandeginste. 2024. "A Review of Proton Exchange Membrane Degradation Pathways, Mechanisms, and Mitigation Strategies in a Fuel Cell" *Energies* 17, no. 5: 998. https://doi.org/10.3390/en17050998

[9] Paul Thiele, Yue Yang, Steffen Dirkes, Maximilian Wick, Stefan Pischinger, "Realistic accelerated stress tests for PEM fuel cells: Test procedure development based on standardized automotive driving cycles," International Journal of Hydrogen Energy, Volume 52, Part D, 2024, Pages 1065-1080, ISSN 0360-3199, https://doi.org/10.1016/j.ijhydene.2023.08.292.

[10] Florian Tritscher, Merit Bodner, Viktor Hacker, Ricardo Novella, Marcos Lopez-Juarez, Alex Ferri-Sirvent, "Real-world accelerated stress test protocols for high-temperature proton-exchange membrane fuel cell based on heavy-duty vehicle operation," International Journal of Hydrogen Energy, Volume 203, 2026, 153057, ISSN 0360-3199, https://doi.org/10.1016/j.ijhydene.2025.153057.

[11] Tang, Yikai, Xing Huang, Yanju Li, Haoran Ma, Kai Zhang, and Ke Song. 2025. "Degradation Prediction of Proton Exchange Membrane Fuel Cell Based on Multi-Head Attention Neural Network and Transformer Model" *Energies* 18, no. 12: 3177. https://doi.org/10.3390/en18123177

[12] X. Hu, C. Zou, X. Tang, T. Liu and L. Hu, "Cost-Optimal Energy Management of Hybrid Electric Vehicles Using Fuel Cell/Battery Health-Aware Predictive Control," in *IEEE Transactions on Power Electronics*, vol. 35, no. 1, pp. 382-392, Jan. 2020, doi: 10.1109/TPEL.2019.2915675.

[13] Wang, Ziqing, Mahjoub Dridi, and Abdellah El Moudni. 2023. "Co-Optimization of Eco-Driving and Energy Management for Connected HEV/PHEVs near Signalized Intersections: A Review" *Applied Sciences* 13, no. 8: 5035. https://doi.org/10.3390/app13085035

[14] Zhuang, Jihui, Pei Li, Ling Liu, Hongjie Ma, and Xiaoming Cheng. 2025. "Energy Management Strategy for Hybrid Electric Vehicles Based on Experience-Pool-Optimized Deep Reinforcement Learning" *Applied Sciences* 15, no. 17: 9302. https://doi.org/10.3390/app15179302

[15] Ocampo-Martinez, Carlos, Rodrigo Toro, Vicenç Puig, Jan Van Impe, and Filip Logist. 2022. "Multi-Objective-Based Tuning of Economic Model Predictive Control of Drinking Water Transport Networks" *Water* 14, no. 8: 1222. https://doi.org/10.3390/w14081222

[16] Haochen Sun, Zhumu Fu, Fazhan Tao, Longlong Zhu, Pengju Si, "Data-driven reinforcement-learning-based hierarchical energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles," Journal of Power Sources, Volume 455, 2020, 227964, ISSN 0378-7753, https://doi.org/10.1016/j.jpowsour.2020.227964.

[17] Liu, Huiying, Hai Xu, Haofa Li, Binggao He, and Yanmin Lei. 2025. "Reinforcement Learning-Based Adaptive Hierarchical Equivalent Consumption Minimization Strategy for Fuel Cell Hybrid Engineering Vehicles" *Sustainability* 17, no. 22: 10167. https://doi.org/10.3390/su172210167

[18] Hanwen Fu, Duo Yang, Siyu Wang, Li Wang, Dongshu Wang, "A novel online energy management strategy for fuel cell vehicles based on improved random forest regression in multi road modes," Energy Conversion and Management, Volume 305, 2024, 118261, ISSN 0196-8904, https://doi.org/10.1016/j.enconman.2024.118261.

[19] Weiguang Zheng, Mengcheng Ma, Enyong Xu, Qibai Huang, "An energy management strategy for fuel-cell hybrid electric vehicles based on model predictive control with a variable time domain," Energy, Volume 312, 2024, 133544, ISSN 0360-5442, https://doi.org/10.1016/j.energy.2024.133544.

[20] Gao, Zihang, and Yizhi Tian. 2023. "Self-Sustaining Control Strategy for Proton-Exchange Membrane Electrolysis Devices Based on Gradient-Disturbance Observation Method" *Processes* 11, no. 3: 828. https://doi.org/10.3390/pr11030828

[21] Mohammad Abdul Baseer, Harjeet Singh, Prashant Kumar, Erick Giovani Sperandio Nascimento, "AI-driven surrogate modelling for simulating hydrogen production via proton exchange membrane water electrolysers," International Journal of Hydrogen Energy, Volume 127, 2025, Pages 462-483, ISSN 0360-3199, https://doi.org/10.1016/j.ijhydene.2025.04.098.

[22] Caixia Liu, Yong Chen, Renzong Xu, Haijun Ruan, Cong Wang, Xiaoyu Li, "Co-optimization of energy management and eco-driving considering fuel cell degradation via improved hierarchical model predictive control," Green Energy and Intelligent Transportation, Volume 3, Issue 6, 2024, 100176, ISSN 2773-1537, https://doi.org/10.1016/j.geits.2024.100176.

[23] Li, Fuxiang, Xiaolin Wang, Xucong Bao, Ziyu Wang, and Ruixuan Li. 2024. "Energy Management Strategy for Fuel Cell Vehicles Based on Online Driving Condition Recognition Using Dual-Model Predictive Control" *Sensors* 24, no. 23: 7647. https://doi.org/10.3390/s24237647

[24] Lan, Cong, Hailong Zhang, Yongjuan Zhao, Huipeng Du, Jinglei Ren, and Jiangyu Luo. 2025. "Hierarchical Reinforcement Learning-Based Energy Management for Hybrid Electric Vehicles with Gear-Shifting Strategy" *Machines* 13, no. 9: 754. https://doi.org/10.3390/machines13090754

[25] Tianyu Li, Ming Li, Xindi Wang, Jianwei Cui, Jianyang Dong, Huiying Liu, Hai Xu, "Adaptive hierarchical energy management strategy for fuel cell hybrid engineering vehicles based on deep reinforcement learning," International Journal of Hydrogen Energy, Volume 168, 2025, 150985, ISSN 0360-3199,https://doi.org/10.1016/j.ijhydene.2025.150985.