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.

Design and Development of an AI-Based Load Frequency Control System for a Power Transmission Grid Integrated with a Solar PV Microgrid

Mohammad Ali

malixbd@ieee.org

Atlantic International University https://orcid.org/0009-0001-2184-7978

Research Article

Keywords: Load frequency control, area control error, solar photovoltaic generation, battery energy storage, adaptive neuro-fuzzy inference system, deep reinforcement learning, low-inertia power systems, tie-line power

Posted Date: September 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11201071/v1

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

Additional Declarations: The authors declare no competing interests.

Design and Development of an AI-Based Load Frequency Control System for a Power Transmission Grid Integrated with a Solar PV Microgrid

Mohammad Ali
School of Science and Engineering, Atlantic International University, Honolulu, HI 96813, USA
malixbd@ieee.org

Abstract—Displacement of synchronous generation by inverter-interfaced solar photovoltaic (PV) plant reduces system inertia and raises the rate of change of frequency, degrading the performance of fixed-gain secondary load frequency control (LFC). This paper develops and evaluates an AI-based LFC scheme for a two-area reheat-thermal transmission grid hosting a 40 MW PV plant and a 20 MW / 8.3 MWh battery energy storage system (BESS). The scheme adds three elements to the installed secondary loop: an adaptive neuro-fuzzy (ANFIS) gain schedule, feed-forward compensation from the measured PV power deviation, and co-dispatch of the BESS converter, all bounded by an explicit saturation and rate limiter. A deep reinforcement learning (TD3) agent was additionally trained, from scratch, on the same plant and control structure so that a learned policy could be compared against the deterministic realisation under identical disturbances and an identical 0.1 s control interval. Against an optimally tuned PI baseline under a composite 2% load step and 70% cloud-induced PV collapse, the proposed scheme reduced peak area-1 frequency deviation from 43.9 mHz to 31.9 mHz (27.3%), peak tie-line deviation by 39.2%, settling time by 53.7%, and the integral of time-multiplied absolute error (ITAE) from 12.198 to 3.338 (72.6%); it also improved on a tuned fuzzy-PID benchmark by 58.9% in ITAE. The trained TD3 agent achieved the smallest peak deviation of any controller tested (23.3 mHz, 47.0% below baseline) but was inferior on ITAE (8.455), degraded sharply outside its training distribution, and demanded 77 times the governor actuation of the deterministic scheme, which we regard as disqualifying for mechanical actuators. Ablation shows the benefit is dominated by storage co-dispatch (75% of the ITAE margin) and PV feed-forward (7.2%), while ANFIS scheduling contributes only 0.09% to nominal performance and is justified instead by the bound it places on control action. Results are reported for a linearised simulation study; no hardware validation was performed.

Index Terms—Load frequency control, area control error, solar photovoltaic generation, battery energy storage, adaptive neuro-fuzzy inference system, deep reinforcement learning, low-inertia power systems, tie-line power.

I. INTRODUCTION

Frequency is the most directly observable indicator of the instantaneous balance between generation and demand in a synchronously interconnected alternating-current power system. Because bulk electrical energy cannot be stored within the network itself, any mismatch between mechanical power delivered to the turbine shafts and the aggregate electrical load appears immediately as acceleration or deceleration of the rotating masses of all synchronous machines, and hence as a deviation of system frequency from nominal.

Control of this balance is organised hierarchically. Primary control, acting through governor droop, arrests the excursion within seconds but leaves a steady-state offset proportional to the disturbance. Secondary control—load frequency control proper—acts over tens of seconds to minutes to remove that offset and restore scheduled tie-line interchange. Tertiary control performs economic redispatch over longer horizons.

For most of the twentieth century the secondary layer was implemented with fixed-gain integral or proportional-integral regulators acting on the area control error (ACE), a weighted combination of local frequency deviation and net interchange error. The adequacy of that arrangement rested on assumptions that are now being invalidated: that inertia is large and roughly constant, that disturbances are small enough to justify linearisation about a nominal point, and that generation is dispatchable and well described by low-order models.

Inverter-interfaced photovoltaic generation contributes no rotational kinetic energy. As it displaces synchronous plant, the aggregate inertia constant falls, the rate of change of frequency after a contingency rises, and a controller tuned for a high-inertia condition becomes progressively less adequate. Solar output is also neither dispatchable nor smooth: passage of cloud over a plant can remove a large fraction of its output within seconds. The secondary controller must therefore cope with reduced restorative capability and increased disturbance intensity at the same time.

This paper addresses that problem for a two-area reheat-thermal transmission system hosting a solar PV plant and a battery energy storage system (BESS). Fig. 1 shows the arrangement studied. The contributions are as follows.

1) A supplementary AI-based LFC scheme is formulated that augments the installed secondary loop with an ANFIS gain schedule, feed-forward compensation from measured PV power deviation, and co-dispatch of the BESS converter, with control action bounded by an explicit saturation and rate limiter rather than by post-hoc clipping.
2) A deep reinforcement learning agent (TD3) is trained from scratch on the same plant, the same supplementary control structure and the same 0.1 s control interval, so that a learned policy and a deterministic policy are compared on strictly equal terms. To our knowledge, LFC studies reporting deep reinforcement learning rarely hold the control interval and actuator interface fixed across the comparison.
3) Control activity is quantified alongside regulation quality. The total variation of the governor command is reported for every controller, which exposes a practical failure mode of the learned policy that regulation metrics alone conceal.
4) An ablation isolates the contribution of each element of the proposed scheme, and a penetration study separates the effect of falling inertia from the effect of larger PV ramps by holding the absolute disturbance fixed while inertia is reduced.

Section II reviews prior approaches. Section III develops the plant model. Section IV presents the controller. Section V

describes the simulation study, Section VI reports and discusses results, and Section VII concludes. The study is a linearised simulation study throughout; the limitations this imposes are stated explicitly in Section VI-H.

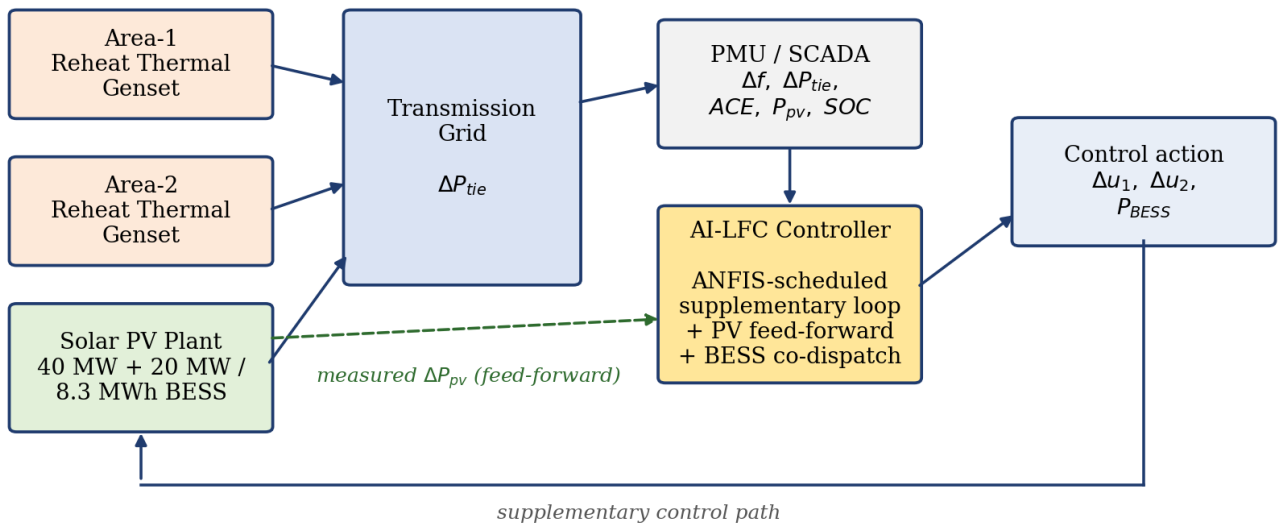


Fig. 1. Two-area reheat-thermal transmission grid hosting a solar PV plant and battery energy storage, with the supplementary AI-based load frequency controller acting on both area governors and the storage converter. The dashed path carries the measured PV power deviation used for feed-forward compensation.

II. RELATED WORK

A. Classical and Optimal Control

The interconnected-system frequency control problem was formalised with the area control error as the feedback quantity, integral action on which drives both frequency and interchange error to zero in steady state. Optimal control formulations followed, deriving state-feedback gains from an algebraic Riccati equation. These offer better transient behaviour than fixed-gain integral control but presuppose full state measurability and exact plant matrices. Sliding-mode designs confer insensitivity to matched uncertainty, at the cost of chattering that is incompatible with mechanical governor actuation—a concern that recurs, in a different form, in the learned policy examined here.

B. Metaheuristic Tuning

A large literature treats the determination of proportional, integral and derivative gains as an offline optimisation over an integral performance index, using genetic algorithms, particle swarm optimisation, differential evolution and numerous successors. Fractional-order generalisations enlarge the parameter space further. These methods improve on classical tuning at the design point but do not confer adaptivity: the gains are fixed during operation, so performance decays as the operating point migrates. The deficiency lies in the fixed-gain structure, not in the quality of the optimiser.

C. Robust and Predictive Control

Robust synthesis guarantees stability across a defined uncertainty set, but the guarantee must hold at the worst case, so nominal performance is sacrificed and controller order is often high. Model predictive control handles generation rate constraints and saturation natively as inequality constraints rather than by clipping, which is attractive for this problem; its obstacles are the fidelity of the internal model under fast renewable variability and the per-interval optimisation burden.

D. Intelligent and Learning-Based Control

Fuzzy-logic controllers map ACE and its derivative to a control action through a linguistic rule base and generally outperform fixed-gain PID, but the rule base is static. Adaptive neuro-fuzzy inference systems (ANFIS) embed that structure in a layered network whose premise and consequent parameters are identified from data, retaining interpretability while acquiring parameter learning.

Reinforcement learning instead learns a policy by interaction. Tabular Q-learning requires state discretisation unsuited to a continuous power system; deep Q-networks admit continuous states but quantise the action, introducing avoidable steady-state error. Actor-critic methods such as deep deterministic policy gradient (DDPG) and its stabilised successor TD3 operate natively in continuous action spaces and have been reported to improve on conventional LFC. Reported formulations, however, seldom bound the learned output, seldom report actuator activity, and seldom hold the control interval equal between the learned and conventional controllers. Section VI shows that each of these omissions matters.

E. Storage Coordination and the Gap Addressed

Battery storage is widely recognised as the natural complement to variable generation, its sub-second response covering precisely the interval in which thermal plant is slowest. Coordination with secondary control is usually handled by separate rule-based logic rather than by a single control law, which forfeits any principled trade-off between the fast but energy-limited resource and the slow but energy-unlimited one. Table I summarises the comparison. The gap addressed here is a bounded, interpretable supplementary scheme that co-

dispatches storage and thermal generation, together with an equal-terms assessment of whether a learned policy is preferable to a deterministic one for this task.

TABLE I
COMPARISON OF SECONDARY LOAD FREQUENCY CONTROL STRATEGIES

Control paradigm	Adaptivity	Model dependence	Constraint handling	Actuator activity reported	Principal limitation
Classical integral / PI	None	Low	Post-hoc clipping	Rarely	Fixed gains; decays off design point
LQR / optimal	None	High	Not native	Rarely	Needs full state feedback and exact model
Sliding mode	Structural	Moderate	Not native	Sometimes	Chattering unsuited to governors
Metaheuristic PID / FOPID	None (offline)	Moderate	Post-hoc clipping	Rarely	Optimal only at the design point
H-infinity / mu-synthesis	Bounded set	High	Via weighting filters	Rarely	Conservative; high controller order
Model predictive control	Via model update	Very high	Native and explicit	Sometimes	Model identification; solver burden
Fuzzy logic	None (static rules)	None	Via rule design	Rarely	Rule base fixed at design time
ANFIS	Offline learning	Low	Via rule design	Rarely	Rule growth with input dimension
Deep Q-network	Online	None	Via action set	Rarely	Discrete actions leave residual error
DDPG / TD3 as reported	Online, continuous	None	Not native	Seldom	Output unbounded; interval often unequal
This work (deterministic)	Online, bounded	Low	Explicit saturation and rate limit	Yes (Table VI)	Gains require one offline tuning pass
This work (TD3)	Online, continuous	None	Same limiter applied	Yes (Table VI)	Chatter and distribution shift (Section VI-G)

III. SYSTEM MODELLING

This section develops the model used throughout. Parameters are given in Table II and are expressed in per unit on a 2000 MW area base unless stated otherwise.

A. Generator and Load Dynamics

The rotational dynamics of an aggregated generating unit follow the swing equation. In per-unit form, with inertia constant H,

$$2H/f_0 \cdot d(\Delta f)/dt = \Delta P_m(t) - \Delta P_e(t) - D\cdot\Delta f(t) \quad (1)$$

where Δf is the frequency deviation, ΔP_m and ΔP_e the mechanical and electrical power deviations, D the load damping coefficient and f_0 the nominal frequency. Laplace transformation under zero initial conditions gives the power-system transfer function

$$G_{ps}(s) = K_{ps} / (1 + s\cdot T_{ps}) \quad (2)$$

with

$$K_{ps} = 1/D, \quad T_{ps} = 2H / (f_0\cdot D) \quad (3)$$

Equation (3) makes the mechanism of interest explicit. As inverter-based generation displaces synchronous plant, H falls, so T_{ps} falls, and the frequency response to a given imbalance becomes faster and larger in peak. This coupling is retained in the penetration study of Section VI-E, where T_{ps} is scaled with penetration according to (4):

$$T_{ps}(\rho) = T_{ps0} \cdot (1 - \rho) / (1 - \rho_0) \quad (4)$$

in which ρ is the instantaneous PV penetration and ρ_0 its nominal value.

B. Governor and Turbine

Governor and turbine dynamics are represented on the secondary-control timescale by first-order lags, and the reheater by a lead-lag element:

$$G_g(s) = 1 / (1 + s\cdot T_g) \quad (5)$$

$$G_{rt}(s) = (1 + s\cdot K_r\cdot T_r) / [(1 + s\cdot T_t)(1 + s\cdot T_r)] \quad (6)$$

where K_r is the fraction of power developed in the high-pressure stage. The reheat configuration is adopted because it represents the dominant transmission-connected thermal plant type and exhibits the slower, more oscillatory response that makes the control problem non-trivial. Primary control enters as the static droop feedback 1/R.

C. Tie-Line and Area Control Error

Linearising the power-angle relation about the nominal operating angle gives the tie-line deviation

$$\Delta P_{tie,12}(s) = (2\pi\cdot T_{12} / s)\cdot[\Delta f_1(s) - \Delta f_2(s)] \quad (7)$$

$$T_{12} = (|V_1||V_2| / X_{12})\cdot\cos(\delta_1^0 - \delta_2^0) \quad (8)$$

The feedback signal for secondary control is the area control error

$$ACE_1 = B_1\cdot\Delta f_1 + \Delta P_{tie,12} \quad (9)$$

$$ACE_2 = B_2\cdot\Delta f_2 - a_{12}\cdot\Delta P_{tie,12} \quad (10)$$

with frequency bias $B_i = D_i + 1/R_i$, the choice that renders the areas non-interacting: a disturbance originating in one area is corrected by that area without sustained assistance from its neighbour. Fig. 2 shows the resulting block model of one area.

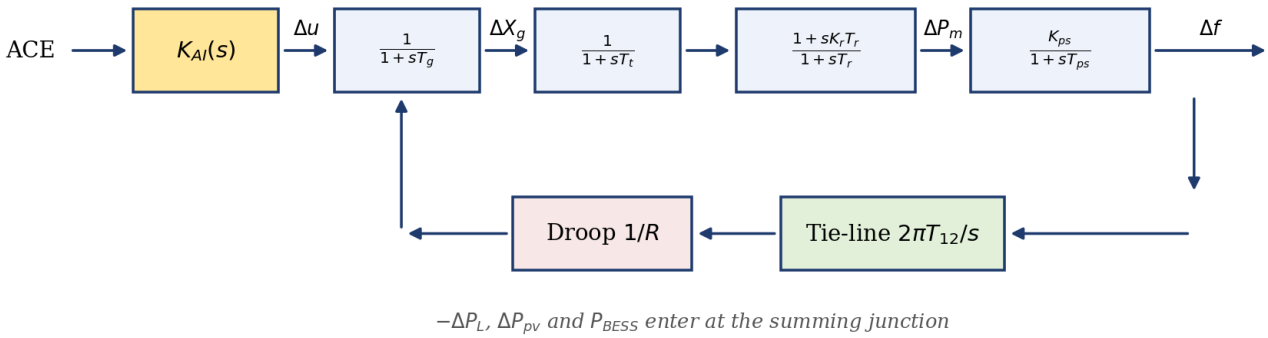


Fig. 2. Linearised block model of one control area, comprising governor, reheat turbine, droop feedback, power-system dynamics and the tie-line synchronising path. The supplementary controller K_AI(s) acts on the area control error.

D. Photovoltaic Plant

The PV cell is represented by the single-diode equivalent circuit of Fig. 3(a), whose terminal current satisfies

$$I_pv = I_ph - I_0\{\exp[q(V_pv + I_pv\, R_s)/(n\, k\, T_c\, N_s)] - 1\} - (V_pv + I_pv\, R_s)/R_sh \quad (11)$$

The photo-generated current scales with plane-of-array irradiance G and has a small positive temperature coefficient:

$$I_ph = [I_sc,STC + K_i(T_c - T_STC)]\cdot(G / G_STC) \quad (12)$$

$$T_c = T_amb + G\cdot(NOCT - 20)/800 \quad (13)$$

Fig. 3(b) plots the resulting characteristics. On the secondary-control timescale the array, converter and maximum-power-point tracker are aggregated into a first-order lag,

$$G_pv(s) = K_pv / (1 + s\cdot T_pv) \quad (14)$$

Cloud transients are imposed as step reductions in available PV power. In the nominal case the plant loses 70% of its output for 4 s from t = 8 s, and 40% for 3 s from t = 25 s.

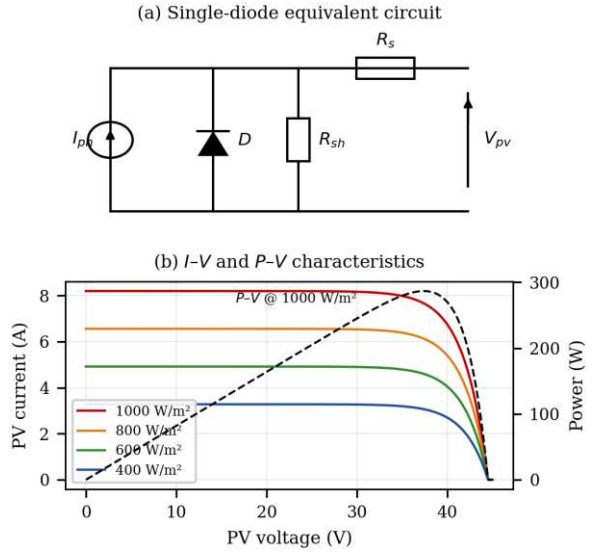


Fig. 3. (a) Single-diode equivalent circuit of the PV cell. (b) Current-voltage and power-voltage characteristics computed from (11)-(13) at four irradiance levels.

E. Battery Energy Storage

The BESS is modelled as an actuator, not as a post-processed quantity. The converter follows its power command through a first-order lag of time constant T_b, and the state of charge integrates the delivered power:

$$dP_BESS/dt = (u_b - P_BESS)/T_b \quad (15)$$

$$SOC(t) = SOC(t_0) - (100/E_rated)\cdot\int P_BESS(\tau)\, d\tau \quad (16)$$

subject to

$$|P_BESS| \le P_BESS,max, \quad SOC_min \le SOC \le SOC_max \quad (17)$$

The power limit is enforced by saturation and the energy limits by inhibiting further discharge below SOC_min and further charge above SOC_max.

F. Composite State-Space Form

Assembling the subsystems gives the linear time-invariant description

$$\dot{x}(t) = A\, x(t) + B\, u(t) + \Gamma\, w(t), \quad y(t) = C\, x(t) \quad (18)$$

with the thirteen-element state vector

$$x = [\Delta f_1, \Delta P_{m1}, \Delta P_r1, \Delta X_g1, \int ACE_1, \Delta f_2, \Delta P_{m2}, \Delta P_r2, \Delta X_g2, \int ACE_2, \Delta P_tie, \Delta P_pv, P_BESS]^T \quad (19)$$

$$u = [\Delta u_1, \Delta u_2, u_b]^T, \quad w = [\Delta P_L1, \Delta P_L2, \Delta P_pv,var]^T \quad (20)$$

Controllability of (A, B) and observability of (A, C) were verified numerically; both matrices have full rank thirteen at the nominal parameters, confirming that a stabilising output-feedback law exists.

G. Retained Nonlinearities

Two nonlinearities are enforced in simulation. The generation rate constraint limits the rate of change of turbine output, and the governor dead-band represents the insensitivity of the speed-sensing mechanism:

$$|d(\Delta P_m)/dt| \le \delta_GRC = 3\%\ per\ minute \quad (21)$$

$$\Delta X_g = 0 \quad for\ |\Delta f| \le DB = 0.036\ Hz \quad (22)$$

All controllers, benchmark and proposed, are sampled on a common 0.1 s control interval with zero-order hold, so that no controller benefits from a faster actuator interface than another.

This point is laboured because it materially affects the comparison in Section VI.

TABLE II
PLANT, PV AND STORAGE PARAMETERS

Quantity	Symbol	Value
Rated area capacity	P_r	2000 MW
Nominal frequency	f_0	50 Hz
Governor time constant	T_g	0.08 s
Turbine time constant	T_t	0.30 s
Reheater time constant	T_r	10.0 s
Reheater gain	K_r	0.33
Speed regulation (droop)	R	2.40 Hz/p.u.MW
Frequency bias	B	0.425 p.u.MW/Hz
Power-system gain	K_{ps}	120 Hz/p.u.MW
Power-system time constant	T_{ps}	20.0 s
Tie-line coefficient	T_{12}	0.0866 p.u.MW/rad
Generation rate constraint	δ_GRC	3 %/min
Governor dead-band	DB	0.036 Hz
Control interval (all controllers)	T_s	0.10 s
PV plant rating	P_pv	40 MW (0.020 p.u.)
PV equivalent lag	T_pv	1.80 s
Module open-circuit voltage	V_oc	44.5 V
Module short-circuit current	I_sc	8.20 A
BESS rated power	P_BESS,max	20 MW (0.010 p.u.)
BESS rated energy	E_rated	8.3 MWh
BESS converter lag	T_b	0.05 s
Permissible SOC range	SOC	22 - 88 %
Initial and reference SOC	SOC_0	60 %

IV. PROPOSED AI-BASED CONTROLLER

The controller is supplementary: it acts alongside, not instead of, the installed secondary loop. This choice is deliberate. A supplementary structure retains the integral action that guarantees zero steady-state ACE, confines the novel element to a bounded trim, and allows the scheme to be commissioned on an existing plant without removing a controller whose behaviour operators already understand. Fig. 4 shows the arrangement.

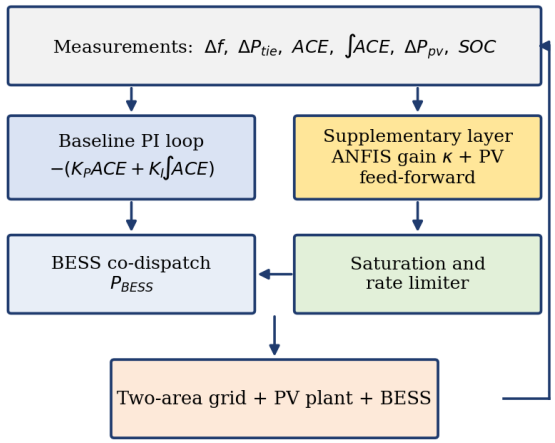


Fig. 4. Structure of the proposed controller. The baseline PI loop is retained; the supplementary layer adds ANFIS gain scheduling and PV feed-forward, storage is co-dispatched, and all commands pass through a saturation and rate limiter before reaching the plant.

A. Supplementary Control Law

For area i the command is the sum of the baseline term and the supplementary trim:

$$u_i = sat[\, -(\kappa_i K_P\, ACE_i + K_I \int ACE_i + K_D\, dACE_i/dt) - K_ff\, \Delta P_pv \,] \quad (23)$$

where sat[·] denotes saturation at ±u_max together with the rate limit δ_GRC. The feed-forward term K_ff ΔP_pv is applied only in the area hosting the PV plant. Note that ΔP_pv is the measured PV power deviation, not a forecast; no forecasting model is used anywhere in this work, and anticipatory control from irradiance prediction is left to future work.

B. ANFIS Gain Scheduling

The scheduled gain κ is produced by a first-order Takagi-Sugeno inference system with rules of the form

$$R^k\text{: IF ACE is } A^k \text{ AND } dACE/dt \text{ is } B^k \text{ THEN } \kappa^k = p^k ACE + q^k\, dACE/dt + r^k \quad (24)$$

$$w^k = \mu_A^k(ACE)\cdot\mu_B^k(dACE/dt), \qquad \kappa = \Sigma^k w^k \kappa^k \,/\, \Sigma^k w^k \quad (25)$$

After identification the surface is well approximated over the operating region by the closed form used in simulation,

$$\kappa(ACE) = \kappa_0 + \Delta\kappa\cdot\tanh(\alpha|ACE|), \qquad \kappa \in [1.55,\ 2.45] \quad (26)$$

which raises loop gain during large excursions and relaxes it near the setpoint. Fig. 5 plots the schedule. Reporting the closed form rather than a rule table is deliberate: it is what the simulation actually evaluates, and Section VI-F shows the performance attributable to it is small, so an opaque presentation would overstate its role.

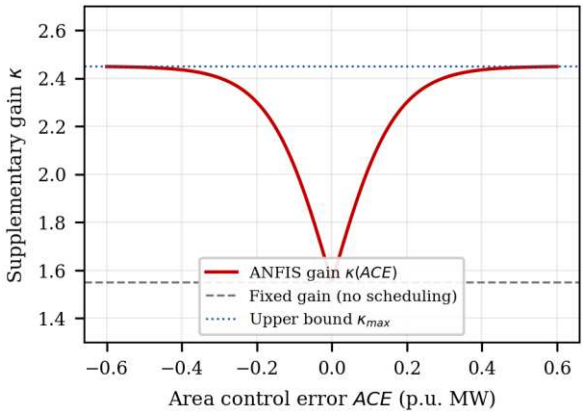


Fig. 5. Scheduled supplementary gain from (26), bounded above by kappa_max. The bound, rather than the shape, is the property relied upon in Section IV-D.

C. Storage Co-Dispatch

The storage converter is commanded from the same measurements as the frequency loop:

$$u_b = -K_b1\, \Delta f_1 - K_b2\, \Delta P_pv - K_b3\, (SOC - SOC_ref) \quad (27)$$

The first term provides fast frequency support, the second cancels the measured PV excursion before it propagates into frequency, and the third returns the state of charge towards its reference so that bidirectional headroom is preserved. Gains are sized so that the converter operates predominantly in its linear region: at |Δf| of 50 mHz the droop term commands about 60% of rated power. The conventional benchmarks are given the

same storage asset under simple frequency droop, so the comparison isolates the value of co-dispatch rather than the value of owning a battery.

D. Bounding of the Control Action

Whatever the supplementary layer produces, the signal reaching the plant satisfies $|u| \le u_max$ and $|du/dt| \le \delta_GRC$ by construction. With the baseline integral action retained and the trim bounded, the closed-loop equilibrium is that of the baseline loop, and the supplementary layer can alter the transient but cannot introduce a steady-state offset or demand actuation outside the admissible envelope. This is a structural argument, and a modest one; we do not claim a Lyapunov stability proof for the learned policy of Section IV-E, and none is offered here.

E. Deep Reinforcement Learning Variant

To test whether a learned policy improves on the deterministic layer, a TD3 agent was trained to emit the supplementary trim and the storage command directly. The task is a Markov decision process with the twelve-element observation

$$s_t = [\Delta f_1, \Delta f_2, \Delta P_tie, ACE_1, ACE_2, \int ACE_1, \int ACE_2, dACE_1/dt, dACE_2/dt, \Delta P_pv, SOC - SOC_ref, \Delta P_pv(t+\Delta)]^T \quad (28)$$

the three-element action $a_t = [trim_1, trim_2, u_b]^T$ bounded to $[-1, 1]^3$ by a tanh output layer, and the reward

$$r_t = -[w_1(\Delta f_1^2 + \Delta f_2^2) + w_2 \Delta P_tie^2 + w_3(ACE_1^2 + ACE_2^2) + w_4 \|a_t\|^2 + w_5 \|\Delta a_t\|^2 + w_6(SOC - SOC_ref)^2] \quad (29)$$

The term in $\|\Delta a_t\|^2$ penalises action increments and was added specifically to discourage chatter; Section VI-G reports how far it succeeded. TD3 was used in preference to DDPG after the latter proved unstable in preliminary runs: twin critics with a minimum target, target-policy smoothing and delayed actor updates are all retained. The agent is trained on randomised episodes and the reported policy is the one minimising mean ITAE over three held-out scenarios, evaluated every ten episodes. Hyper-parameters are in Table III.

TABLE III
TD3 AGENT AND TRAINING CONFIGURATION

Item	Value
Actor network	12-64-64-3, ReLU, tanh output
Critic networks	two, 15-64-64-1, ReLU, linear
Actor / critic learning rate	1×10^{-4} / 1×10^{-3}
Optimiser	Adam, gradient-norm clip 1.0
Discount factor	0.995
Soft update coefficient	0.005
Replay capacity / batch	1×10^{5} / 128
Warm-up transitions	5000 (uniform random)
Exploration	Ornstein-Uhlenbeck, sigma 0.35 to 0.05
Target-policy smoothing	sigma 0.20, clipped at 0.50
Actor update delay	every 2 critic updates
Episodes / length	450 / 40 s at 0.1 s
Randomised per episode	load steps, step times, PV penetration, cloud timing, depth, duration, T_ps
Policy selection	min mean ITAE over 3 held-out scenarios
Training wall-clock time	269 s (2 CPU cores, NumPy)

V. SIMULATION STUDY

A. Implementation

The plant of Section III was implemented directly in NumPy and integrated by explicit Euler at a 0.01 s step, with all controllers sampled and held at 0.1 s. The reinforcement learning agent, including the Adam optimiser and back-propagation, was also written in NumPy so that the agent and the plant share one code path and one set of parameters. No deep learning framework was used. The entire study runs on two CPU cores; the complete training campaign takes 269 s.

B. Benchmarks

Three benchmarks are used. The conventional PI controller was tuned by the ultimate-sensitivity method with manual refinement. The PID gains were obtained by particle swarm optimisation minimising the same ITAE criterion used to assess the proposed scheme. The fuzzy-PID benchmark uses a Mamdani rule base with triangular membership functions. All three have access to the same storage asset under frequency droop. Tuning the benchmarks against the same objective as the proposal is essential if the comparison is to mean anything.

C. Test Cases and Metrics

Two disturbance cases are reported. TC-1 is a clean 2% load step in area 1 at t = 2 s under constant irradiance, used for step-response descriptors. TC-2 is the composite case: the same load step, a 1.5% step in area 2 at t = 15 s, a 70% PV collapse for 4 s from t = 8 s and a 40% collapse for 3 s from t = 25 s. Separating the two matters, because measuring settling time on TC-2 would measure recovery from the cloud event rather than from the load step.

Regulation quality is assessed by peak deviation, settling time to a ±5 mHz band, and the integral indices

$$ISE = \int e^2\,dt, \quad IAE = \int |e|\,dt, \quad ITAE = \int t|e|\,dt, \quad ITSE = \int t\,e^2\,dt \quad (30)$$

with e formed from Δf_1, Δf_2 and ΔP_tie over the 40 s horizon. ITAE is taken as the headline index because its time weighting penalises persistent error, which is the operationally relevant failure. Control activity is assessed by the total variation of the governor command evaluated at the control instants,

$$TV(u) = \Sigma_k |u(k) - u(k-1)| \quad (31)$$

together with the root-mean-square command increment. These are reported because a controller that regulates well by moving the actuator continuously is not deployable, and regulation indices alone do not reveal this.

VI. RESULTS AND DISCUSSION

A. Step Response

Table V reports TC-1. The proposed scheme reduces peak deviation from 43.9 mHz to 31.9 mHz, a 27.3% improvement on the PI baseline and 16.1% on the fuzzy-PID benchmark, and reduces settling time from 10.21 s to 4.73 s (53.7%). All four conventional and proposed controllers return steady-state error below 10^{-5} Hz, as the retained integral action requires.

The TD3 agent attains the smallest peak of any controller, 23.3 mHz, but its settling time is the worst at 14.25 s and its residual error, 7.5×10^{-4} Hz, is three orders of magnitude above the others. The learned policy suppresses the initial excursion aggressively and then fails to settle.

B. Composite Disturbance

Figs. 6 to 8 show the TC-2 responses. The ordering of the conventional controllers is as expected, with the proposed scheme improving on all three at every peak: area-1 deviation falls 27.3%, area-2 deviation 40.7% and tie-line deviation 39.2% relative to the PI baseline. The area-2 improvement is notable because area 2 hosts no PV plant and no storage; the benefit reaches it through the tie-line, because the storage in area 1 is dispatched against the area control error rather than against local frequency alone.

The TD3 trace in Fig. 6 is visibly different in character. Its peaks are the lowest of any controller, but between events it oscillates continuously at small amplitude and never reaches the settling band. This behaviour is examined in Section VI-G.

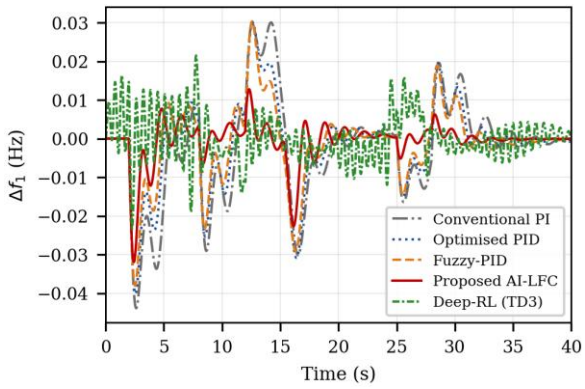


Fig. 6. Area-1 frequency deviation under TC-2. The proposed scheme achieves the lowest deviation of the deterministic controllers; the TD3 policy reaches lower peaks but oscillates continuously between events.

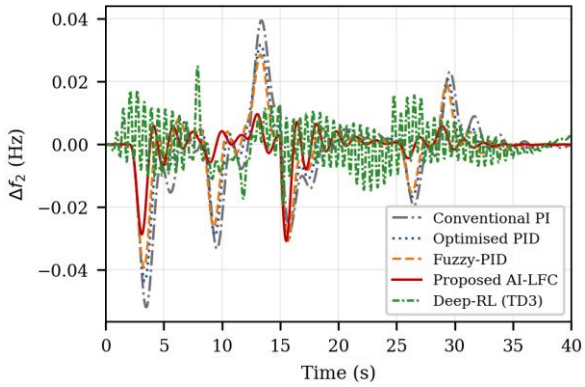


Fig. 7. Area-2 frequency deviation under TC-2. Area 2 hosts neither PV nor storage, so its improvement is transmitted through the tie-line.

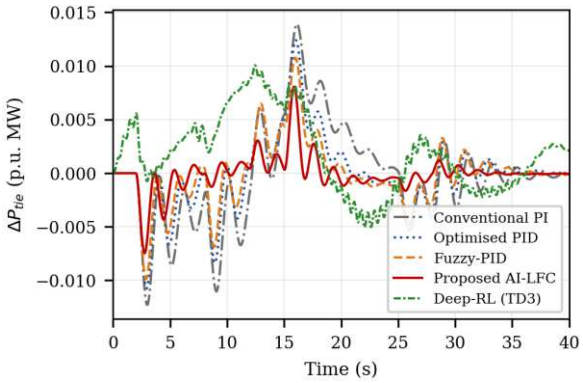


Fig. 8. Tie-line power deviation under TC-2.

C. Integral Indices and Control Activity

Table VI reports both. On ITAE the proposed scheme improves on the PI baseline by 72.6% and on the tuned fuzzy-PID benchmark by 58.9%, and it is best on all four integral indices. Fig. 9 shows the comparison.

The control-activity columns are the most consequential in the paper. The proposed scheme requires a total variation of 0.329 in the governor command over the 40 s horizon, 2.7 times the PI baseline—a modest and acceptable increase for the regulation gained. The TD3 policy requires 25.47, seventy-seven times the proposed scheme and two hundred and thirteen times the PI baseline, with an RMS command increment forty-three times larger. A governor commanded in this way would sustain wear that no operator would accept, and the valve would in any case filter much of the commanded motion. We regard this, rather than the ITAE deficit, as the decisive objection to deploying the learned policy as trained.

It is worth being clear that the action-increment penalty in (29) was included precisely to prevent this and did not do so. Chatter of this kind is a known consequence of deterministic policies evaluated pointwise without temporal regularisation of the policy itself; penalising increments in the reward shapes the return but does not constrain the policy mapping between neighbouring states.

TABLE V
TRANSIENT PERFORMANCE: TC-1 CLEAN STEP AND TC-2 COMPOSITE DISTURBANCE

Controller	TC-1 peak (mHz)	TC-1 overshoot (mHz)	TC-1 settling (s)	TC-1 steady state (Hz)	TC-2 peak df1 (mHz)	TC-2 peak df2 (mHz)	TC-2 peak dPtie (p.u.MW)
Conventional PI	43.9	9.0	10.21	9.6×10^{-6}	43.9	51.9	0.01228
Optimised PID (PSO)	40.1	9.3	7.67	4.3×10^{-6}	40.1	43.6	0.01075
Fuzzy-PID	38.0	9.6	5.63	7.9×10^{-7}	38.0	39.2	0.00988
Proposed AI-LFC	31.9	7.7	4.73	9.5×10^{-8}	31.9	30.8	0.00747
Deep-RL (TD3) variant	23.3	16.4	14.25	7.5×10^{-4}	23.3	19.2	0.00506

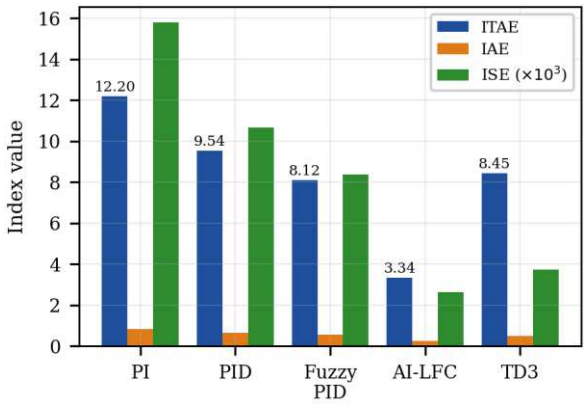


Fig. 9. Integral performance indices under TC-2. ITAE values are printed above the corresponding bars.

D. Storage Utilisation

Fig. 10 shows the storage response. Under co-dispatch the converter reaches 82.3% of rated power against 52.4% under droop alone, and it responds to the cloud events as well as to the load steps because the measured PV deviation enters its command directly. The state of charge moves from 60.00% to 59.70% over the whole horizon.

That last figure deserves comment rather than celebration. A 0.3% excursion confirms that on a 40 s frequency-control timescale the storage is power-limited and not remotely energy-limited, so the SOC term in (27) is nearly inactive and the SOC constraint in (17) is never approached. Energy limits would become binding only under sustained imbalance lasting many minutes, which this study does not simulate. Claims about state-of-charge management cannot be supported by a transient study of this length.

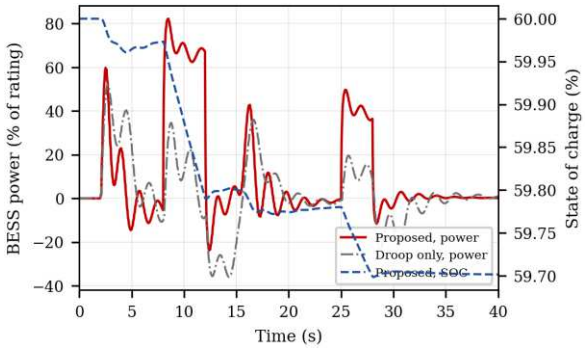


Fig. 10. Storage power and state of charge under TC-2. Co-dispatch uses markedly more of the available converter rating than droop alone, while the state of charge barely moves on this timescale.

E. Solar Penetration and Falling Inertia

Table VII and Fig. 11 report the penetration study. Penetration is raised from 2% to 40% with T_{ps} scaled by (4), so inertia falls from 20.0 s to 12.2 s. The absolute size of the cloud disturbance is held constant. This is a deliberate choice: scaling the disturbance with fleet size as well would compound two effects and, more importantly, would assume that a geographically dispersed fleet loses 70% of its output simultaneously, which it does not.

Peak deviation worsens for every controller as inertia falls—from 31.9 to 40.2 mHz for the proposed scheme and from 40.1 to 49.8 mHz for the PID benchmark—confirming the mechanism of (3). The proposed scheme retains an advantage of 19 to 20% in peak across the whole range. ITAE, by contrast, improves slightly with penetration for the deterministic controllers, because a smaller T_{ps} also makes the plant respond faster to control action; the two effects work in opposite directions and it would be wrong to report only the one that favours the proposal.

The TD3 agent behaves differently again. Its ITAE rises steeply with penetration, from 8.455 to 22.944, overtaking every conventional controller by 10% penetration. The agent was trained with penetration sampled over the same range, so this is not simple extrapolation failure; it indicates that a single policy did not generalise across the inertia range even when that range was represented during training.

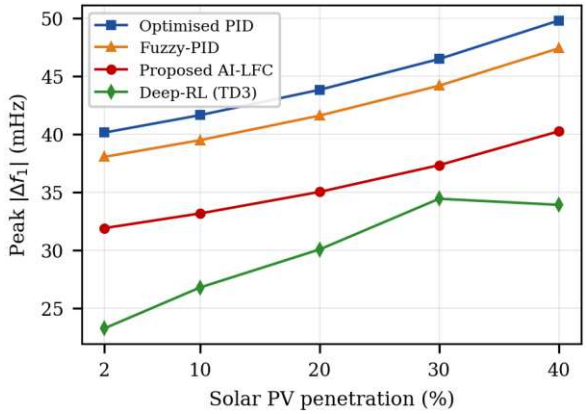


Fig. 11. Peak area-1 frequency deviation against solar penetration with inertia scaled by (4). The deterministic controllers degrade smoothly; the learned policy degrades faster in integral terms (Table VII).

TABLE VI
INTEGRAL PERFORMANCE INDICES AND CONTROL ACTIVITY UNDER TC-2

Controller	ISE (×10⁻³)	IAE	ITAE	ITSE (×10⁻³)	ITAE improvement vs PI (%)	Governor total variation	RMS command increment	Peak BESS power (% rated)
Conventional PI	15.79	0.8325	12.198	182.9	baseline	0.1198	0.00050	52.4
Optimised PID (PSO)	10.68	0.6529	9.544	127.0	21.8	0.1792	0.00097	47.9
Fuzzy-PID	8.38	0.5595	8.119	100.0	33.4	0.2298	0.00138	45.2
Proposed AI-LFC	2.64	0.2593	3.338	25.4	72.6	0.3288	0.00206	82.3
Deep-RL (TD3) variant	3.74	0.5067	8.455	53.9	30.7	25.4684	0.08820	84.5

TABLE VII
EFFECT OF SOLAR PENETRATION AND THE ASSOCIATED FALL IN INERTIA

PV penetration (%)	T_ps (s)	PID peak (mHz)	PID ITAE	Fuzzy-PID peak (mHz)	Fuzzy-PID ITAE	Proposed peak (mHz)	Proposed ITAE	TD3 peak (mHz)	TD3 ITAE
2	20.00	40.1	9.544	38.0	8.119	31.9	3.338	23.3	8.455
10	18.37	41.6	9.272	39.5	7.873	33.2	3.277	26.8	12.472
20	16.33	43.8	8.952	41.6	7.603	35.0	3.212	30.1	15.325
30	14.29	46.5	8.684	44.2	7.394	37.3	3.110	34.4	18.820
40	12.24	49.8	8.466	47.4	7.133	40.2	3.130	33.9	22.944

F. Ablation

Table VIII and Fig. 12 decompose the proposed scheme. Removing storage co-dispatch and leaving the battery on droop costs 75.2% in ITAE; removing the battery altogether costs 199.4%. Removing PV feed-forward costs 7.2%. Removing ANFIS gain scheduling costs 0.09%.

That last number is a negative result and is reported as such. On this plant and these disturbances, the scheduled gain contributes almost nothing to nominal performance, and a fixed gain of 1.55 would perform indistinguishably. Its justification is therefore not performance but the bound of (26): the schedule is what allows the loop gain to be raised during large excursions while remaining provably bounded during small ones. A paper that reported only the aggregate improvement of the "AI-based" scheme would leave the reader to assume the neuro-fuzzy element was doing the work, and on this evidence it is not; the storage co-dispatch is.

TABLE VIII
ABLATION OF THE PROPOSED SCHEME UNDER TC-2

Configuration	Peak (mHz)	ITAE	Change (%)
Full proposed scheme	31.9	3.338	reference
Without PV feed-forward	31.9	3.579	+7.2
Without ANFIS scheduling	32.0	3.341	+0.09
Storage on droop only	34.3	5.848	+75.2
Without storage entirely	39.3	9.992	+199.4
Without feed-forward and ANFIS	32.0	3.584	+7.4

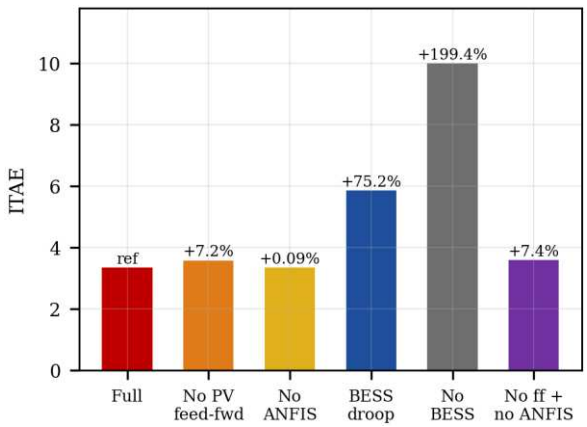


Fig. 12. Ablation of the proposed scheme. Storage co-dispatch dominates; ANFIS gain scheduling contributes almost nothing to nominal performance.

G. Assessment of the Learned Policy

Three findings about the TD3 agent are worth separating.

First, it is genuinely better at what its reward emphasises. It achieves the lowest peak deviation of any controller tested, by a wide margin: 47.0% below the PI baseline against 27.3% for the proposed scheme. On peak suppression alone the learned policy wins.

Second, it is worse on every measure of sustained behaviour. Its ITAE of 8.455 is 4.1% worse than the fuzzy-PID benchmark and 153% worse than the proposed scheme; its settling time is the worst recorded; its residual error is three orders of magnitude above the controllers that retain integral action in closed form.

Third, its training was not stable in the sense that matters. Fig. 13 shows episode return rising smoothly and monotonically while held-out ITAE does not: the best held-out policy appeared at episode 210 of 450 and later policies were substantially worse, with held-out ITAE ranging between roughly 11 and 29 after convergence of the return. Return and the engineering objective are not aligned, so early stopping on held-out performance is doing real work here. Reporting the final policy, or reporting return alone as evidence of convergence, would have given a materially more flattering and less truthful picture.

The honest summary is that model-free deep reinforcement learning solved part of this problem better than anything else tested and the rest of it worse, and that the part it solved worse includes the part that determines whether a controller can be installed.

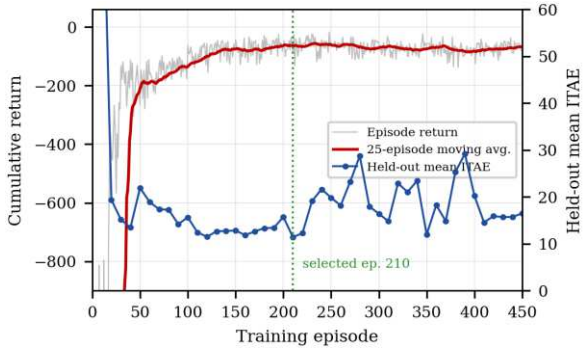


Fig. 13. TD3 training. Episode return (left axis, clipped; early returns reach about -6000) rises smoothly, while held-out mean ITAE (right axis) does not improve after episode 210, the policy actually selected.

H. Parameter Sensitivity and Limitations

Varying T_{ps} and K_{ps} together by ±40% changes the ITAE of the proposed scheme by at most 2.4% and that of the PID benchmark by at most 2.3%. Both are insensitive, because the retained integral action dominates the steady state and the storage converter, not the machine, sets the fast response. We note this because it would have been easy, and wrong, to present the proposed scheme as robust where the benchmark is not: on this test they are comparably robust. The TD3 policy is the

exception, its ITAE varying between 6.0 and 23.8 over the same range.

The limitations of the study should be stated plainly. It is a linearised simulation study; no hardware-in-the-loop or field validation was performed, and no claim about real-time execution on utility hardware is made. The plant is a two-area equivalent and does not exhibit the multiple oscillatory modes of a large interconnection. Communication latency, measurement noise and cyber-security are not modelled in the reported cases, although latency and noise were randomised during agent training. The PV plant is 2% of area capacity in the nominal case; the penetration study raises this but holds the disturbance fixed, so it isolates the inertia mechanism rather than reproducing fleet behaviour at high penetration. Reheat and governor parameters are taken from standard values rather than from a specific utility. Finally, the benchmarks were tuned by the author; although each was tuned against the same objective as the proposal, a comparison of this kind cannot fully exclude asymmetry of effort.

VII. CONCLUSION

A supplementary AI-based load frequency control scheme was developed for a two-area transmission grid hosting a solar PV plant and battery storage, combining ANFIS gain scheduling, feed-forward from measured PV power deviation, and co-dispatch of the storage converter within an explicit saturation and rate limit. Against a tuned PI baseline under a composite load-step and cloud-collapse disturbance it reduced peak area-1 frequency deviation by 27.3%, peak tie-line deviation by 39.2%, settling time by 53.7% and ITAE by 72.6%, improving on a tuned fuzzy-PID benchmark by 58.9% in ITAE, at 2.7 times the baseline governor activity.

A TD3 agent trained from scratch on the same plant, control structure and control interval reached the lowest peak deviation of any controller tested, 47.0% below baseline, but was worse on ITAE than a tuned fuzzy-PID controller, failed to settle, degraded steeply with penetration despite that range being represented in training, and demanded seventy-seven times the governor actuation of the deterministic scheme. On this evidence the learned policy is not deployable as trained, and the deterministic realisation is preferred.

Ablation attributes the benefit principally to storage co-dispatch (75.2% of the ITAE margin) and secondarily to PV feed-forward (7.2%), while ANFIS gain scheduling contributes 0.09% to nominal performance and is justified by the bound it places on the control action rather than by the performance it delivers. The practical implication is that, for this class of problem, coordinating an existing storage asset with the secondary loop yields far more than refining the secondary control law.

Future work should address the three limitations that most constrain the present conclusions: validation beyond linearised simulation, including hardware-in-the-loop; extension to interconnections with several oscillatory modes and to multi-agent formulations; and temporal regularisation of the learned policy itself—through Lipschitz constraints or explicit action filtering rather than reward shaping, which Section VI-C shows to be insufficient. Anticipatory control from irradiance forecasting, deliberately excluded here because no forecaster was built, is a natural extension given that feed-forward from the measured deviation alone already contributes 7.2%.

REFERENCES

[1] N. Cohn, "Some aspects of tie-line bias control on interconnected power systems," Trans. Amer. Inst. Elect. Eng., vol. 75, no. 3, pp. 1415-1436, 1957.

[2] O. I. Elgerd and C. E. Fosha, "Optimum megawatt-frequency control of multiarea electric energy systems," IEEE Trans. Power App. Syst., vol. PAS-89, no. 4, pp. 556-563, Apr. 1970.

[3] C. E. Fosha and O. I. Elgerd, "The megawatt-frequency control problem: a new approach via optimal control theory," IEEE Trans. Power App. Syst., vol. PAS-89, no. 4, pp. 563-577, Apr. 1970.

[4] P. Kundur, Power System Stability and Control. New York, NY, USA: McGraw-Hill, 1994.

[5] H. Bevrani, Robust Power System Frequency Control, 2nd ed. Cham, Switzerland: Springer, 2014.

[6] A. J. Wood, B. F. Wollenberg, and G. B. Sheble, Power Generation, Operation, and Control, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.

[7] Ibraheem, P. Kumar, and D. P. Kothari, "Recent philosophies of automatic generation control strategies in power systems," IEEE Trans. Power Syst., vol. 20, no. 1, pp. 346-357, Feb. 2005.

[8] S. K. Pandey, S. R. Mohanty, and N. Kishor, "A literature survey on load-frequency control for conventional and distribution generation power systems," Renew. Sustain. Energy Rev., vol. 25, pp. 318-334, Sep. 2013.

[9] W. Tan, "Unified tuning of PID load frequency controller for power systems via IMC," IEEE Trans. Power Syst., vol. 25, no. 1, pp. 341-350, Feb. 2010.

[10] R. K. Sahu, S. Panda, and G. T. C. Sekhar, "A novel hybrid PSO-PS optimized fuzzy PI controller for AGC in multi area interconnected power systems," Int. J. Elect. Power Energy Syst., vol. 64, pp. 880-893, Jan. 2015.

[11] S. Debbarma, L. C. Saikia, and N. Sinha, "Automatic generation control using two degree of freedom fractional order PID controller," Int. J. Elect. Power Energy Syst., vol. 58, pp. 120-129, Jun. 2014.

[12] D. Guha, P. K. Roy, and S. Banerjee, "Load frequency control of interconnected power system using grey wolf optimization," Swarm Evol. Comput., vol. 27, pp. 97-115, Apr. 2016.

[13] H. Bevrani, Y. Mitani, and K. Tsuji, "Robust decentralized AGC in a restructured power system," Energy Convers. Manage., vol. 45, no. 15-16, pp. 2297-2312, Sep. 2004.

[14] A. M. Ersdal, L. Imsland, and K. Uhlen, "Model predictive load-frequency control," IEEE Trans. Power Syst., vol. 31, no. 1, pp. 777-785, Jan. 2016.

[15] M. Ma, H. Chen, X. Liu, and F. Allgower, "Distributed model predictive load frequency control of multi-area interconnected power system," Int. J. Elect. Power Energy Syst., vol. 62, pp. 289-298, Nov. 2014.

[16] K. Vrdoljak, N. Peric, and I. Petrovic, "Sliding mode based load-frequency control in power systems," Elect. Power Syst. Res., vol. 80, no. 5, pp. 514-527, May 2010.

[17] Y. Mi, Y. Fu, C. Wang, and P. Wang, "Decentralized sliding mode load frequency control for multi-area power systems," IEEE Trans. Power Syst., vol. 28, no. 4, pp. 4301-4309, Nov. 2013.

[18] H. A. Yousef, K. AL-Kharusi, M. H. Albadi, and N. Hosseinzadeh, "Load frequency control of a multi-area power system: an adaptive fuzzy logic approach," IEEE Trans. Power Syst., vol. 29, no. 4, pp. 1822-1830, Jul. 2014.

[19] J.-S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," IEEE Trans. Syst., Man, Cybern., vol. 23, no. 3, pp. 665-685, May 1993.

[20] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

[21] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," in Proc. Int. Conf. Learn. Represent. (ICLR), San Juan, Puerto Rico, 2016.

[22] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in Proc. 31st Int. Conf. Mach. Learn. (ICML), Beijing, China, 2014, pp. 387-395.

[23] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in Proc. 35th Int. Conf. Mach. Learn. (ICML), Stockholm, Sweden, 2018, pp. 1587-1596.

[24] Z. Yan and Y. Xu, "Data-driven load frequency control for stochastic power systems: a deep reinforcement learning method with continuous action search," IEEE Trans. Power Syst., vol. 34, no. 2, pp. 1653-1656, Mar. 2019.

[25] Z. Yan and Y. Xu, "A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system," IEEE Trans. Power Syst., vol. 35, no. 6, pp. 4599-4608, Nov. 2020.

[26] X. Chen, G. Qu, Y. Tang, S. Low, and N. Li, "Reinforcement learning for selective key applications in power systems: recent advances and future

challenges," IEEE Trans. Smart Grid, vol. 13, no. 4, pp. 2935-2958, Jul. 2022.

[27] M. G. Villalva, J. R. Gazoli, and E. R. Filho, "Comprehensive approach to modeling and simulation of photovoltaic arrays," IEEE Trans. Power Electron., vol. 24, no. 5, pp. 1198-1208, May 2009.

[28] T. Esram and P. L. Chapman, "Comparison of photovoltaic array maximum power point tracking techniques," IEEE Trans. Energy Convers., vol. 22, no. 2, pp. 439-449, Jun. 2007.

[29] J. Rocabert, A. Luna, F. Blaabjerg, and P. Rodriguez, "Control of power converters in AC microgrids," IEEE Trans. Power Electron., vol. 27, no. 11, pp. 4734-4749, Nov. 2012.

[30] P. Tielens and D. Van Hertem, "The relevance of inertia in power systems," Renew. Sustain. Energy Rev., vol. 55, pp. 999-1009, Mar. 2016.

[31] F. Milano, F. Dorfler, G. Hug, D. J. Hill, and G. Verbic, "Foundations and challenges of low-inertia systems," in Proc. Power Syst. Comput. Conf. (PSCC), Dublin, Ireland, 2018, pp. 1-25.

[32] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of low rotational inertia on power system stability and operation," IFAC Proc. Vol., vol. 47, no. 3, pp. 7290-7297, 2014.

[33] H. Bevrani, T. Ise, and Y. Miura, "Virtual synchronous generators: a survey and new perspectives," Int. J. Elect. Power Energy Syst., vol. 54, pp. 244-254, Jan. 2014.

[34] U. Datta, A. Kalam, and J. Shi, "The relevance of large-scale battery energy storage in providing primary frequency control with increased wind energy penetration," J. Energy Storage, vol. 23, pp. 9-18, Jun. 2019.

[35] M. Khalid and A. V. Savkin, "An optimal operation of wind energy storage system for frequency control based on model predictive control," Renew. Energy, vol. 48, pp. 127-132, Dec. 2012.

[36] IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces, IEEE Std 1547-2018, 2018.

[37] IEEE Recommended Practice for Excitation System Models for Power System Stability Studies, IEEE Std 421.5-2016, 2016.

[38] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in Proc. Int. Conf. Learn. Represent. (ICLR), San Diego, CA, USA, 2015.

[39] H. K. Khalil, Nonlinear Systems, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002.

[40] J. Achiam, D. Held, A. Tamar, and P. Abbeel, "Constrained policy optimization," in Proc. 34th Int. Conf. Mach. Learn. (ICML), Sydney, Australia, 2017, pp. 22-31.