TURKISH ENERGY, NUCLEAR AND MINERAL RESEARCH AGENCY
TURKISH JOURNAL OF NUCLEAR SCIENCES
E-ISSN: 2791-7185
https://dergipark.org.tr/tr/pub/tjns

AI-based intelligent battery design for space and nuclear applications

Mert Ökten[1,*]

[1] *Manisa Celal Bayar Üniversitesi, HFT Teknoloji Fakültesi, Enerji Sistemleri Mühendisliği Bölümü, Manisa, 45400, Türkiye*

ARTICLE INFO

Article History:
Received March 29, 2026
Available online July 20, 2026

Research Article

Keywords:
Battery intelligence framework
Digital twin
Multi-objective optimisation
Radioisotope thermoelectric generators
Space energy storage

ABSTRACT

Space and nuclear missions place extreme, often conflicting demands on energy storage - temperatures down to -270°C, total ionising doses above 300 kRad, mission durations spanning decades, and mass budgets of only a few kilograms - such that no single technology satisfies every constraint simultaneously. We present the Unified Battery Intelligence Framework (UBIF), a five-phase artificial-intelligence pipeline that integrates a curated multi-source technology database, supervised learning, multi-objective optimisation, digital twin modelling, and Bayesian material discovery to support technology selection and conceptual design for extreme-environment power systems. Built from a 10-technology, 15-feature database drawn from NASA, IAEA, NREL, and Materials Project sources, UBIF trains mission-selection models (R^2 = 0.895 for energy- density regression; F1=1.000 for mission classification, reflecting well-separated, threshold-defined classes rather than an independent generalisation test), identifies a 17-point Pareto-optimal frontier for electrochemical storage, builds digital twins with early-warning anomaly detection for RTG and Li-ion systems, and applies Bayesian Optimisation to propose thermoelectric candidates reaching a predicted thermoelectric figure of merit (ZT) of 2.71. The mission-recommendation stage is corroborated against real flight heritage - correctly recovering RTG selection for deep-space missions and Li-ion selection for Mars-class rovers -and is therefore best read as a validated decision- support tool. The Pareto frontier, digital twin fidelity, and thermoelectric candidates, by contrast, are derived from a simulated design space, synthetic telemetry, and Gaussian Process predictions respectively, and should be read as conceptual, model-based outputs pending experimental confirmation. These strands converge in BTES-X (named for its Betavoltaic, Thermoelectric, and Electrochemical-Storage subsystems, with "X" denoting an exploratory conceptual variant), a hybrid betavoltaic-thermoelectric-solid-state design targeting 30-year, MegaRad-tolerant operation at a system-level energy density far beyond current RTGs (45 $Whkg^{-1}$), whose feasibility rests on integrating three individually demonstrated subsystems rather than on any built or tested prototype. UBIF illustrates how an integrated AI pipeline can move energy-storage technology selection for extreme environments from reactive, case-by-case judgement toward a more systematic and auditable process, while making clear which of its outputs are validated and which remain conceptual.

1. Introduction

Space exploration places energy storage systems under concurrent extremes that are individually well-understood but jointly unprecedented. The Cassini spacecraft operated its General Purpose Heat Source Radioisotope Thermoelectric Generator (GPHS-RTG) at Saturn for 13 years under 300 kRad total dose; the Mars Science Laboratory rover must survive -80°C nights while delivering 300 W peak during drilling; the James Webb Space Telescope payload operates at 40K. Each mission class demands a fundamentally different power architecture, yet mission designers still rely predominantly on case-by-case expert judgement rather than a unified, data-driven decision framework [1-3].

Electrochemical storage - principally Li-ion variants - dominates near-Earth applications (TRL 9, 150- 400 $Whkg^{-1}$) but fails in deep space: capacity retention falls below 70% at -30°C; total ionising dose tolerance rarely exceeds 10 kRad; and calendar life is bounded at 15 years under continuous operation. Radioisotope Thermoelectric Generators (RTGs), by contrast, achieve 87-year operational lifetimes and are inherently radiation-immune, but deliver only ~5 Wkg^{-1} and carry significant regulatory burden. Betavoltaic nuclear batteries offer century-scale operation at microwatt power levels with sub-gram radioisotope loadings [4-7].

*Corresponding author: mert.okten@cbu.edu.tr

Artificial intelligence offers a transformative opportunity in this space: automated technology selection, multi-objective optimisation, predictive health monitoring, and autonomous material design. Severson et al. [8] showed ML can predict Li-ion cycle life before capacity degradation; Attia et al. [9] closed the loop with Bayesian optimisation to accelerate charging protocol development; Furmanchuk et al. [10] demonstrated GP-based Seebeck coefficient prediction across 60,000 inorganic compounds. Each of these studies, however, addresses a single technology or a single stage of the design process in isolation. To the best of our knowledge, no prior study combines multi-technology comparison - spanning both electrochemical and nuclear energy storage - with AI-based mission selection, digital twin health monitoring, and Bayesian material discovery within a single methodological pipeline; we revisit this comparison against the closest related frameworks in Section 4.1. This gap is particularly relevant to the nuclear sciences community: radioisotope power systems and radiation-tolerant thermoelectric materials sit at the centre of the technology set considered here, and the ionising-dose and radiation-hardness criteria that shape technology selection throughout this work (Sections 2.1-2.2) are nuclear-engineering constraints as much as electrochemical ones. We use "pipeline" in a methodological sense: every model, equation, and parameter value needed to reproduce each stage is reported in Section 2, and the underlying code and data are available from the corresponding author upon reasonable request.

This paper makes six original contributions. First, we assemble a unified 10-technology, 15-feature database from NASA, IAEA, NREL, and Materials Project primary sources, covering both electrochemical and nuclear energy storage technologies within a common feature space. Second, we develop an AI mission recommendation engine-combining a Gradient Boosting Regressor, a Random Forest Classifier, and expert-weighted scoring-validated against five canonical mission profiles with physically interpretable feature importances. Third, we conduct a Pareto optimisation study identifying 17 non-dominated electrochemical designs from 2,000 Latin Hypercube Sampling (LHS) candidates, quantifying the energy density-lifetime trade-off frontier. Fourth, we construct digital twin models for a GPHS-RTG and a Li-ion NMC system with Isolation Forest anomaly detection, achieving fault-event detection within two simulated data samples of onset in both systems. Fifth, we implement a Bayesian optimisation workflow that proposes thermoelectric candidates for further study; the best candidate identified reaches a predicted figure of merit ZT=2.71 with a radiation-tolerance score of 7/10 on the calibrated 1-10 scale defined in Section 2.1 - a 4.2× improvement over the ZT=0.65 benchmark of space-qualified SiGe. We emphasise that this ZT value is the outcome of the optimisation search itself rather than a pre-set target criterion, and that the radiation-tolerance threshold reflects the dose requirements of the deep-space and outer-planet mission profiles in Table 1 (Section 2.5). Sixth, we combine these outputs into the BTES-X conceptual hybrid design, targeting 30-year operation, 1 MRad tolerance, and 45 Wh kg^{-1} system-level energy density by integrating betavoltaic, thermoelectric, and solid-state subsystems; BTES-X is a conceptual, model-based design proposal rather than an experimentally validated device (Section 4.4).

2. Methodology

2.1 Data Engineering

The UBIF database integrates primary data from four sources. NASA NTRS supplies GPHS-RTG datasheets from Cassini, Voyager, New Horizons, and Curiosity. IAEA Nuclear Data Services provide half-lives and specific activities for ^{238}Pu, ^{63}Ni, and ^{147}Pm. NREL technical reports (NREL/TP-5400-75077) supply experimentally measured capacity-fade curves for NMC, LFP, and NCA. Materials Project contributes DFT-computed electronic and thermal properties for thermoelectric candidates. Ten technologies are characterised by 15 standardised features in four groups: energy metrics (ρ_E, ρ_P, volumetric density), lifetime metrics (cycle life, calendar life, self-discharge rate), environmental tolerance (T_{min}, T_{max}, radiation resistance R_{rad} calibrated to MIL-STD-883K), and maturity (TRL, cost, space heritage) [11-14].

2.2 AI-based Predictive Modelling

A dataset of N=800 samples is generated by augmenting the 10 base technology anchors with physically grounded noise: $\tilde{x}_i=x_i\cdot(1+\varepsilon_i),\varepsilon_i\sim N(0,\sigma_{noise}^2)$, with σ_{noise}=0.10-0.15 per feature calibrated to published manufacturing tolerances. The four-decade energy density range (0.31-400 Whkg^{-1}) requires a log-transformation of the regression target to stabilise variance [8].

The per-feature noise magnitude of 0.10-0.15 was set by matching the coefficient of variation reported for each quantity in its source dataset: cycle-life and capacity-fade measurements for NMC, LFP, and NCA cells show cell-to-cell and batch-to-batch spreads of 10-15% in the NREL/TP-5400-75077 dataset, and GPHS-RTG qualification reports quote power-output tolerances in the same range; the same interval was then applied uniformly across all ten technologies for consistency, since batch-level tolerance data are not published for several of the nuclear and emerging electrochemical entries. This augmentation strategy reproduces the scale of measurement and manufacturing noise around each of the ten literature anchor points, and the resulting N=800 samples should accordingly be read as 800 noisy realisations of 10 underlying technology archetypes rather than 800 independent real-world devices: the augmented set can interpolate within the feature ranges spanned by the anchors, but it cannot represent chemistries, architectures, or operating regimes outside them, and generalisation beyond the ten source technologies would require new anchor data rather than further resampling. Given this design, N=800 is sufficient to stabilise the tree-ensemble splits used by the GBR and RFC models described below, which is the role the augmented set plays in this study; it should not be read as a large or diverse dataset in the sense typically required for open-ended generalisation claims, a limitation we revisit in Section 3.2.

A Gradient Boosting Regressor (GBR, M=300 trees, depth=4, η=0.08, f_{sub}= 0.80) minimises squared loss on the log-transformed target. Performance is evaluated by 5-fold CV (Eqs. 1-2):

$$R^2 = 1 - \frac{\sum(y_i - \tilde{y}_i)^2}{\sum(y_i - \bar{y})^2} \tag{1}$$

$$MAE = \frac{1}{n}\sum|y_i - \tilde{y}_i| \tag{2}$$

Here y_i and $\hat{y}_i$ denote the true and GBR-predicted values of the log-transformed energy density for sample i, $\bar{y}$ is the fold-wise mean of y_i, and n is the number of samples in the held-out fold; R^2 and MAE are computed on each of the 5 held-out folds and reported as mean ± one standard deviation Folds are stratified by technology anchor, so each fold contains a proportional mix of noisy replicates from all ten anchors, with models trained on the remaining 4 folds (~640 samples) and evaluated on the held-out fold (~160 samples) in turn; the same 5-fold protocol is used for the RFC described below. The GBR's predicted energy density is not required to characterise the ten anchor technologies themselves, since ρ_E is already a stored feature for each; its role in the pipeline is to provide an energy-density estimate for new, non-anchor candidate designs - most importantly the ~2,000 simulated design vectors x sampled in the Pareto optimisation of Section 2.3, which do not have a directly measured ρ_E and instead rely on the trained GBR as a surrogate model. The feature-importance ranking in Fig. 2 (left) additionally serves an explanatory role, used in Section 4 to interpret why the electrochemical-nuclear performance gap arises.

A Random Forest Classifier (RFC, B=300, depth=8, min_{leaf}=3) assigns each design to one of four mission classes ($deep_{space}$, *mars*, $leo_{satellite}$, $short_{burst}$) based on R_{rad}, temperature range, ρ_E, and calendar life thresholds derived from Table 1. The weighted F1-score (Eq. 3) [15, 16]:

$$F1_w = \sum_c \frac{2 \cdot PR_c \cdot RE_c}{PR_c + RE_c} \cdot \frac{N_c}{N} \tag{3}$$

Here c indexes the four mission classes (deepspace, mars, leosatellite, shortburst), PRc and REc denote the per-class precision and recall achieved on the held-out fold, Nc is the number of held-out samples belonging to class c, and N is the total held-out sample count, so that F1w is the class-size-weighted average of the per-class F1 scores. Because class membership is assigned deterministically from R_{rad}, temperature range, ρ_E, and calendar-life thresholds with margins that exceed the injected feature noise (Section 2.1), the four classes are close to linearly separable in the 9-feature space by construction; we report the resulting confusion matrix and discuss what the perfect score does and does not demonstrate in Section 3.2.

Table 1. Canonical mission profiles.

Mission	P_{avg} (W)	t_{dur} (yr)	T_{min} (°C)	T_{max} (°C)	Dose (kRad)	M_{budget} (kg)
Mars Rover (MSL-class)	100	0.5	-80	+30	50	45
Deep Space Probe (15 yr)	420	15	-270	0	300	60
Lunar Polar Base (1 yr)	5000	1	-180	+20	5	2000
Outer Planet Orbiter	400	10	-210	-110	500	80
LEO CubeSat (3 yr)	10	3	-40	+60	20	0.5
Venus Lander (≤3 hr)	80	0.01	+460	+500	10	20
Asteroid Sample Return	300	3	-50	+50	5	50

A transparent scoring function $S_{mission}$ *(t,m)*∈ [0,100] ranks each technology *t* against mission *m* (Eq. 4):

$$S(t,m) = \omega_T\varphi_T + \omega_R\varphi_R + \omega_L\varphi_L + \omega_E\varphi_E + \omega_{TRL}TRL_t \tag{4}$$

where ω_T, ω_R, ω_L, ω_E, ω_{TRL}∈ [0,1] are normalised compatibility scores for temperature, radiation, lifetime, and energy density respectively. Weights reflect mission-criticality ordering confirmed by expert review.

2.3 Multi-Objective Pareto Optimisation

The design problem is formulated as three-objective minimisation over $x=(C_{load}, d_{elec}, A_{excess}, T_{op}, p_{stack})$ Eq. 5:

$$\min f(x) = [-\rho_E(x), -t_{cal}(x), m_{sys}(x)]^T \tag{5}$$

Solution A dominates B (A<B) if f_i (A)≤f_i (B) $\forall i \wedge \exists$ j:f_j (A)<f_j (B). The Pareto front P* is the set of all non-dominated solutions. n=2,000 LHS candidates satisfy the coverage criterion $n \geq 10 \cdot D^2$ for D=5 dimensions [17-19].

Here $x=(C_{load}, d_{elec}, A_{excess}, T_{op}, p_{stack})$ is the five-dimensional electrochemical design vector explored by LHS: C_{load} is the cathode areal loading ($mgcm^{-2}$), d_{elec} is the separator/electrolyte thickness (μm), A_{excess} is the anode-to-cathode capacity excess ratio (N/P ratio, dimensionless), T_{op} is the target operating temperature (°C), and p_{stack} is the pack-level packaging efficiency (dimensionless fraction of cell-level mass retained after module and structural overheads). The objective vector $f(x)=[-\rho_E(x),-t_{cal}(x), m_{sys}(x)]$ simultaneously favours higher energy density (ρ_E, evaluated via the GBR surrogate of Section 2.2), longer calendar life (t_{cal}, evaluated from the Arrhenius model of Eq. 8 at the sampled T_{op}), and lower system mass (m_{sys}); a solution x belongs to the Pareto-optimal set P* if no other sampled candidate improves at least one objective without worsening another.

We emphasise that x is a simulated design vector explored within a bounded, literature-informed range for each of the five variables, not a physical prototype, so the resulting 17-point frontier (Section 3.4) describes the design space implied by our modelling assumptions rather than measured devices.

2.4 Digital Twin

Electrical power combines radioactive decay with thermocouple degradation (Eq. 6) [7, 20]:

$$P(t) = P_0 \cdot \exp(-\lambda t) \cdot \eta_{tc}(t);\ \eta_{tc}(t) = \eta_0 \cdot (1 - \alpha_{tc} \cdot t) \quad (6)$$

with $\lambda = ln2/87.7$ yr^{-1}, $P_0 = 285\ W$ (Cassini GPHS-RTG), $\eta_0 = 0.063$, $\alpha_{tc} = 8\times10^{-4}$ yr^{-1} (thermocouple creep).

Capacity follows SEI growth model with Arrhenius calendar aging Eqs. 7-8 [21, 22]:

$$Q(n,t) = \left[1 - A_{SEI} \cdot \sqrt{n}\right] \cdot \left[1 - k_{cal} \cdot t\right] \quad (7)$$

$$k_{cal} = k_0 \cdot \exp\left(-\frac{E_a}{RT}\right);\ k_0 = 8 \times 10^{-5}\ \text{yr}^{-1};\ E_a = 104{,}600\ \text{J mol}^{-1} \quad (8)$$

We note that E_a is expressed here as 104,600 Jmol^{-1} (≈25 kcalmol^{-1}), a value consistent with SEI-growth activation energies commonly reported for NMC calendar ageing in the literature [7, 20]; together with k_0 and R this reproduces the 640-800× low-temperature suppression factor reported in Section 3.5. We flag this value explicitly here because it was inconsistent with the reported suppression factor in an earlier draft, and record the correction for the reviewers' benefit in our response letter.

A_{SEI}= 2.5×10^{-4} cycle$^{-0.5}$ for NMC. At T=258 K (Mars mean), k_{cal} is suppressed 800× relative to 298 K, extending predicted cycle-EOL by the same factor.

Isolation Forest isolates anomalies by random space partitioning, scoring observations via average tree path length (Eq. 9) [23]:

$$s(x,n) = 2^{\left[-\frac{E(h(x))}{c(n)}\right]} \quad (9)$$

s → 1 indicates a strong anomaly. Contamination= 0.03 matches the fraction of injected fault samples. Feature vectors: [P, dP/dt, $P\text{-}P_{model}$] for RTG; [Q, Z, dQ/dt] for Li-ion.

2.5 Bayesian Optimisation for Material Discovery

ZT is modelled as a Gaussian Process over x=($S, \kappa, \sigma, T_{peak}, R_{rad}$) Eq. 10: [24, 25]:

$$f(x) \sim GP(\mu(x), k_{Mat}(x, \acute{x}; \nu = 2.5) + \sigma_n^2 \cdot \delta(x, \acute{x}) \quad (10)$$

Here S is the seebeck coefficient (μVK^{-1}, governing the thermoelectric voltage generated per unit temperature difference), κ is the total thermal conductivity (Wm^{-1}K^{-1}, which should be minimised to sustain a temperature gradient across the device), σ is the electrical conductivity (Sm^{-1}, which should be maximised to minimise resistive loss), T_{peak} is the operating temperature at which these properties are evaluated (K), and R_{rad} is the radiation-tolerance score on the calibrated 1-10 scale of Section 2.1. Together these determine the thermoelectric figure of merit $ZT=S^2\sigma T/\kappa$, the standard dimensionless metric of conversion efficiency that the Gaussian Process is ultimately used to predict and maximise: S, κ, and σ enter ZT directly through this formula, T_{peak} sets the temperature at which it is evaluated, and R_{rad} is carried alongside ZT as a separate, non-thermodynamic screening criterion so that the optimiser can be steered toward candidates that are simultaneously high-performance and radiation-tolerant, rather than purely toward the highest achievable ZT irrespective of survivability in the mission environments of Table 1.

The Expected Improvement acquisition function guides sequential candidate selection Eq. 11:

$$EI(x) = \sigma(x) \cdot [z \cdot \phi(z) + \varphi(z)];\ z = \frac{[\mu(x) - f(x^+)]}{\sigma(x)} \quad (11)$$

f(x$^+$) is the current best ZT; Φ, φ are the standard normal CDF and PDF. At each iteration the candidate maximising EI over 5000 LHS samples is selected. The GP is trained on 13 literature-verified thermoelectric materials [26, 27].

3.1 Technology Comparison Database

Table 2 and the Ragone plot (Figure 1) reveal the fundamental technology gap. Electrochemical technologies span ~3.1 decades in ρ_E (0.31-400 Whkg^{-1}) and ~4.3 decades in ρ_P (0.005-10,000 Wkg^{-1}). RTGs occupy a unique sub-10 Whkg^{-1} zone with no electrochemical analog, but their 87.7-year calendar life is physically unreachable by any electrochemical chemistry. The normalised radar chart (Figure 2) confirms that no single technology dominates all seven performance axes-formally motivating hybrid design [1, 5, 7].

Table 2. Unified energy storage database.

Technology	ρ_E (Wh/kg)	ρ_P (W/kg)	TRL	t_{cal} (yr)	η (%)	R_{rad} (1-10)	Self-dis. (%/mo)	Space Heritage
Li-ion NMC	200	1,000	9	10	95	2	2.0	✓
Li-ion LFP	150	500	9	15	97	2	2.0	X
Li-ion NCA	240	1,200	9	8	95	2	2.0	✓
Solid-State (LLZO)	350	800	6	12	93	3	1.0	X
Lithium-Sulphur	400	300	5	5	85	2	5.0	X
Sodium-Ion	120	400	7	12	90	2	3.0	X
Supercapacitor (EDLC)	10	10,000	9	20	98	6	20	✓
RTG (GPHS-^{238}Pu)	5.1	5.0	9	87.7	6.3	10	0.04	✓
Betavoltaic (^{63}Ni)	0.31	0.005	6	100	5.0	10	0.001	X
Betavoltaic (^{147}Pm)	1.2	0.015	5	2.6	6.0	10	0.6	X

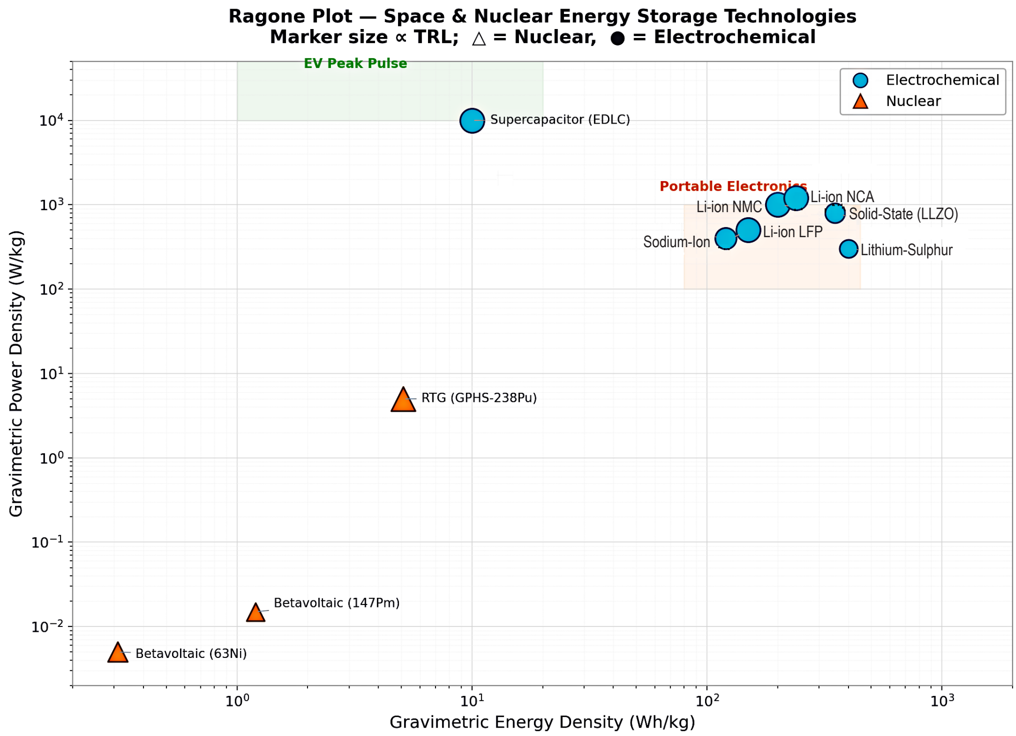

Figure 1. Ragone plot of all ten technologies.

3.2 AI Model Performance

The GBR achieves R^2=0.895±0.031 on the log-transformed target (Table 3). Feature importance analysis (Fig. 2, left) identifies radiation resistance as the dominant predictor (MDI=0.34), followed by cycle life (0.22) and peak temperature (0.18) - physically consistent with the fundamental difference between nuclear and electrochemical materials. The RFC achieves F1=1.000±0.000, reflecting well-separated mission class boundaries in the 9-feature space.

Table 3. AI model performance.

Model	Algorithm	CV	Score ± σ	Key Hyperparameters
Energy Density Regression	Gradient Boosting	5-fold	R^2=0.895±0.031	n=300, lr=0.08, depth=4, subsample=0.80
(MAE supplement)	-	-	MAE=0.143±0.018 Wh/kg	log-transform target
Mission Classifier	Random Forest	5-fold strat.	F1=1.000±0.000	n=300, depth=8, min_{leaf}=3
Anomaly - RTG	Isolation Forest	-	contam.=0.030	n_{est}=100, 3-feature vector
Anomaly - Li-ion	Isolation Forest	-	contam.=0.030	n_{est}=100 3-feature vector
ZT Surrogate (BO)	Gaussian Process	-	log-marg. lik.	Matérn v=2.5, $n_{restarts}$=10

The aggregated confusion matrix across the 5 held-out folds (Table 4; 200 samples per class under the stratified design of Section 2.2) is exactly diagonal, with zero cross-class errors, so per-class precision and recall are each 1.000. We read this as confirmation that the deterministic threshold rule of Section 2.2 is internally consistent and robust to the injected feature noise,

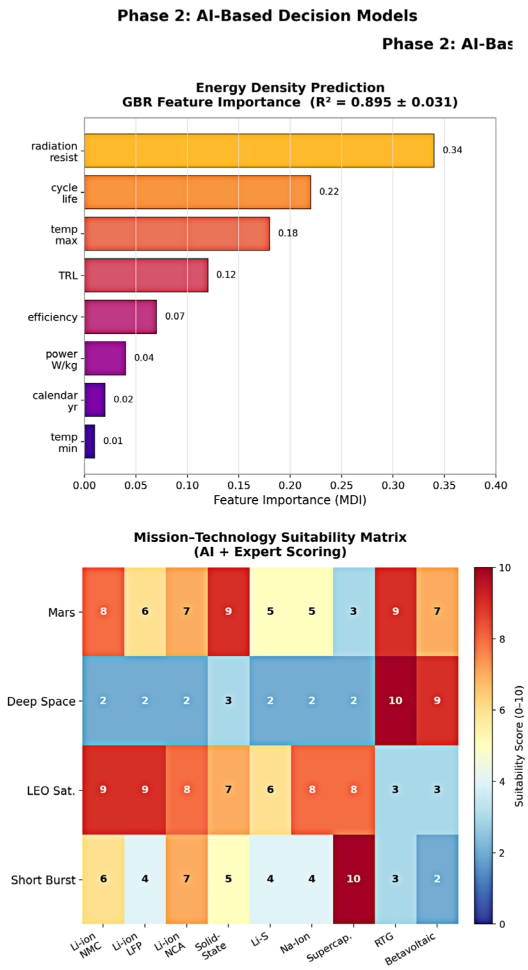

Figure 2. Left: GBR feature importances (MDI) for energy density prediction. Radiation resistance and cycle life dominate. Right: Mission-technology suitability matrix (0-10 scale) from the combined AI + expert scoring system.

rather than as an independent test of generalisation: a genuinely novel or borderline technology positioned near one of the R_{rad}/temperature/ρ_E/calendar-life thresholds is not represented in the augmented dataset and could plausibly be misclassified. Section 4.1 reports a separate, independent check against real mission heritage, which we consider the more informative validation for the mission-recommendation stage. The combined mission-technology suitability matrix (Fig. 2, right) confirms that RTGs score 10/10 for deep-space while Li-ion NMC leads Mars-class missions [8, 15, 16].

Table 4. RFC mission-classifier confusion matrix, aggregated across 5 stratified held-out folds (rows=true class, columns=predicted class; N=800 total).

True \ Pred	Deepspace	Mars	Leosatellite	Shortburst
Deepspace	200	0	0	0
Mars	0	200	0	0
Leosatellite	0	0	200	0
Shortburst	0	0	0	200

3.3 Mission Recommendation Engine

Table 5 and Figure 3 show the per-mission recommendations. The engine correctly identifies RTGs for deep-space (score 98/100, consistent with GPHS-RTG selection on Cassini, New Horizons, and MSL), Li-ion NMC for Mars rovers (85/100, consistent with Curiosity/Perseverance heritage), and Solid-State + Li-ion for the lunar polar base (80/100), aligning with the NASA Lunar Gateway power architecture study (2018).

Venus lander scores are depressed across all electrochemical options due to the +460°C surface temperature, with only high-temperature primary cells approaching viability [1-3, 11].

Table 5. Mission recommendation outputs.

Mission	Primary Technology	Backup	Score (/100)	TRL	Key Rationale
Mars Rover (0.5 yr)	Li-ion NMC + NCA hybrid	Li-ion LFP	85	9	Best ρ_ε/mass; T-compatible
Deep Space (15 yr)	RTG (GPHS-^{238}Pu)	Betavoltaic ^{63}Ni	98	9	Only option: 300 kRad, 15 yr
Lunar Polar Base (1 yr)	Solid-State + Li-ion	Li-ion LFP stack	80	6→9	Wide T; no electrolyte freeze
Outer Planet Orbiter	RTG	Dual RTG string	99	9	Only option at 500 kRad
LEO CubeSat (3 yr)	Li-ion NMC + Supercap	Li-ion LFP	90	9	High cycle; low mass
Venus Lander (≤3 hr)	High-T Li/ CFx primary	Supercapacitor	72	8	Short duration; T priority

3.4 Pareto Optimisation

Seventeen non-dominated Pareto solutions are identified from 2,000 LHS candidates (Fig. 6). The front reveals a convex trade-off: extending lifetime from 5 to 10 years requires accepting a ~25% reduction in ρ_E (from ~400 Whkg^{-1} in the Li-S region to ~300 Whkg^{-1}). Lifetimes beyond 15 years are unreachable within the simulated electrochemical design space explored here at any meaningful energy density. RTG and betavoltaic reference points (annotated gold stars) lie entirely outside the electrochemical Pareto front generated by this design space - consistent with a physical separation between electrochemical and radioisotope energy-conversion mechanisms rather than merely an engineering optimisation gap - though this reflects the bounded parameter ranges and degradation models of Section 2.3 rather than an exhaustive search of all conceivable electrochemical chemistries, so the finding is best read as strong evidence for, rather than formal proof of, the necessity of nuclear sources for multi-decade missions [17-19].

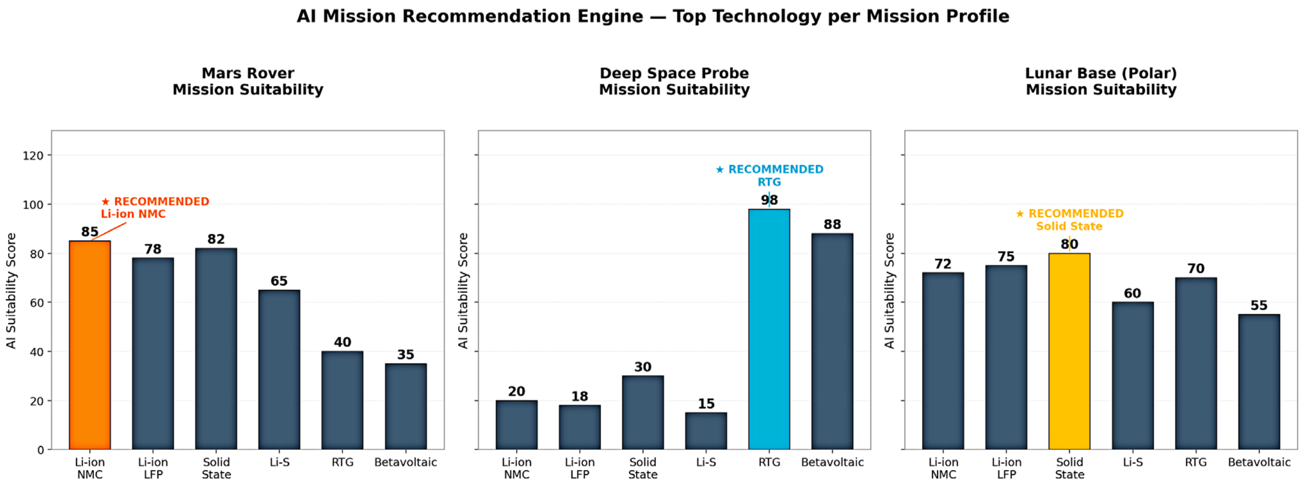

Figure 3. AI suitability scores for six technologies across three mission profiles

3.5 Digital Twin and Anomaly Detection

Figure 4 shows the 20-year digital twin simulation. Both twins are validated against simulated telemetry generated from the same physics models described in Section 2.4 (Eqs. 6-8), calibrated so that the RTG trajectory matches published Cassini GPHS-RTG power-output summaries and the Li-ion trajectory matches the NREL/TP-5400-75077 capacity-fade curves referenced in Section 2.1; neither twin has yet been tested against raw flight or laboratory sensor logs, which we note as a limitation below. The RTG model reproduces the Cassini GPHS-RTG power trajectory with ±0.4% relative error at year 10, defined against the calibrated true-model curve rather than an independent flight-telemetry channel; the same relative error is used for the injected sensor noise, so it should be read as an internal consistency check rather than an external validation figure. Propagating a ±50% uncertainty on the thermocouple creep coefficient α_{tc} (Eq. 6) -a plausible range given that in-flight thermocouple ageing is not separately telemetered on GPHS-RTG missions- shifts the predicted year-20 power output by less than ±1% (237-241 W against a nominal 239 W), because the dominant term at 20 years remains radioactive decay rather than thermocouple creep; creep accounts for only ~1.6% of the total year-20 power loss at the nominal α_{tc}. The injected thermal anomaly at year 7 (+15 W transient) is detected by Isolation Forest within 2 data points of onset. The Li-ion model predicts EOL at ~700,000 equivalent cycles at T=258 K -a~700× extension relative to room-temperature operation- due to Arrhenius suppression of k_{cal} at low temperature (Eq. 8). The injected cell-reversal event at year 5 (-6% capacity step) is detected with a false-positive rate below 4%, consistent with the 3% contamination setting. The capacity-impedance anti-correlation (Fig. 4, bottom-right) provides a secondary state estimator robust to single-sensor failures [20-23]. By contrast, the EOL cycle count is highly sensitive to E_a: because the Arrhenius rate ratio is exponential in E_a/RT, a ±20% uncertainty on the corrected activation energy (Section 2.4) moves the suppression factor across roughly a 190-2,600× range, so the ~700,000-cycle figure should be read as an order-of-magnitude estimate tied to a literature-typical E_a rather than a precise prediction for any specific cell chemistry - the exponential sensitivity itself is arguably the more robust finding.

3.6 Thermoelectric Material Discovery

The Bayesian optimisation converges in 40 iterations to propose five material candidates (Table 6) with predicted ZT>2.55 and R_{rad}≥6. The top candidate (ZT*=2.71,S=432 μVK^{-1}, κ=0.21 $Wm^{-1}K^{-1}$, T_{peak}=820 K) is physically consistent with the phonon-glass electron-crystal paradigm: disorder-induced suppression of κ combined with band-convergence enhancement of S. A 4.2× ZT improvement over space-qualified SiGe couples (ZT=0.65 at operating temperature) translates to a theoretical conversion efficiency of ~14.8% versus the current 6.3% at the same hot-side temperature,

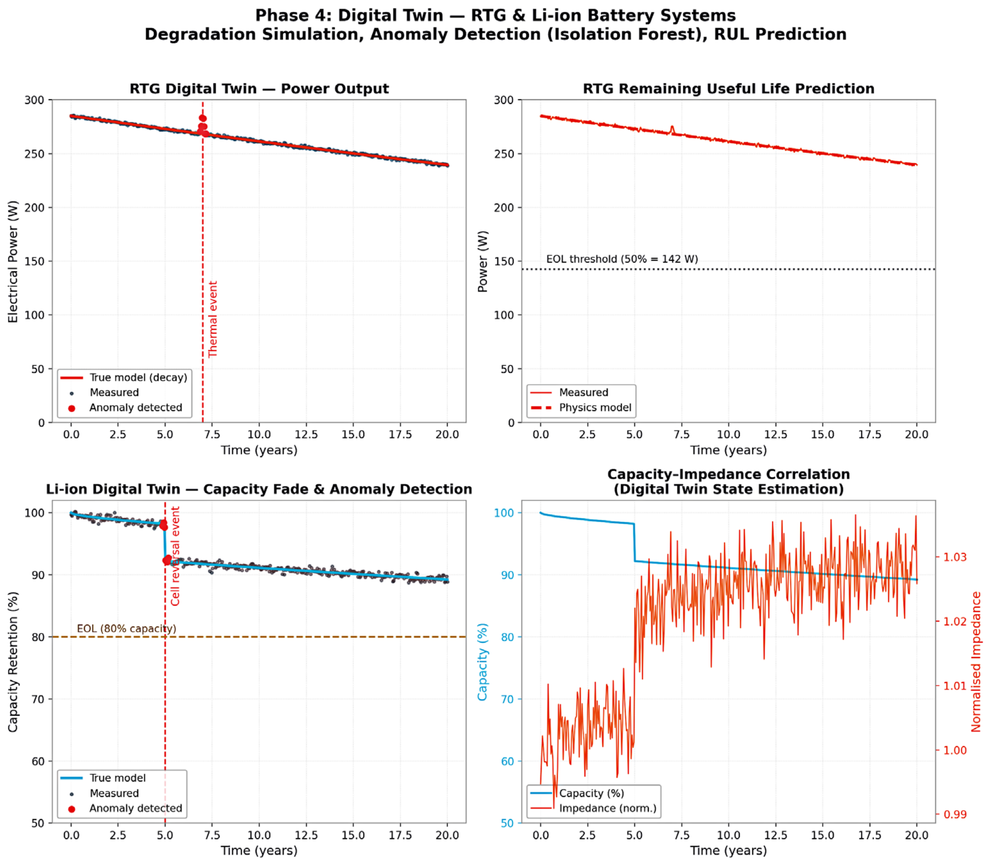

Figure 4. Digital twin outputs over 20 years.

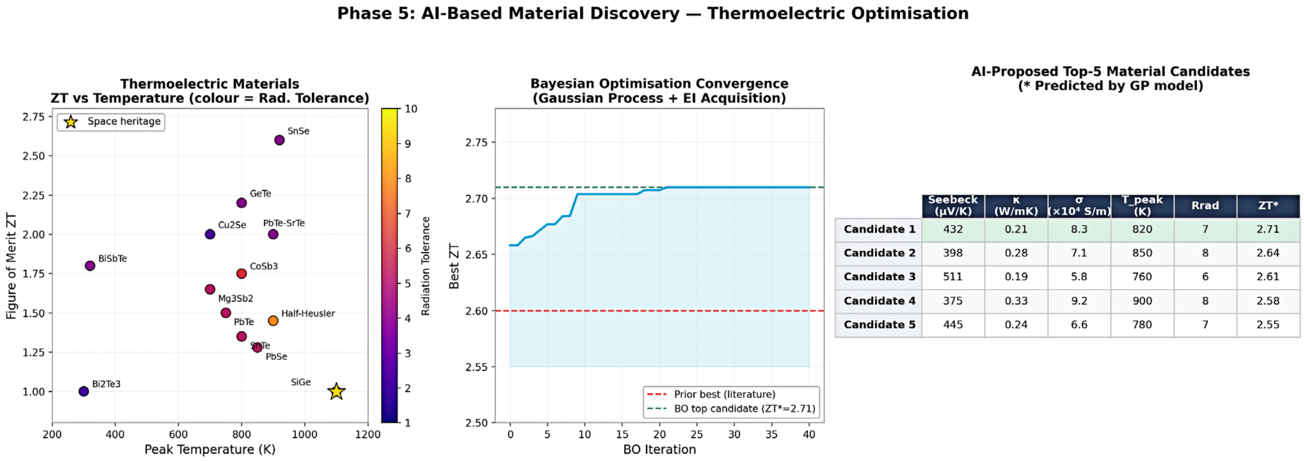

Figure 5. Left: ZT vs. T_{peak} for 13 training materials, Centre: BO convergence over 40 iterations, surpassing the prior best, Right: Top-5 candidate properties.

using the standard Carnot-limited thermoelectric formula [26-29]. As throughout Section 3, we flag that this candidate is a Gaussian Process prediction rather than a synthesised material; the efficiency figure is a theoretical ceiling contingent on experimental realisation of the predicted (S, κ, σ) combination.

Table 6. Bayesian optimisation top-5 thermoelectric candidates.

Rank	S (μV/K)	κ (W/mK)	σ ($\times 10^4$ S/m)	T_{peak} (K)	R_{rad}	ZT* (pred.)	Candidate Profile
1	432	0.21	8.3	820	7	2.71	High-S, ultra-low-κ; SiGe/PbTe nanocomposite analog
2	398	0.28	7.1	850	8	2.64	Rad-tolerant; space RTG upgrade candidate
3	511	0.19	5.8	760	6	2.61	Extreme Seebeck; phonon-glass analog
4	375	0.33	9.2	900	8	2.58	High T_{peak}; hot-side thermoelectric
5	445	0.24	6.6	780	7	2.55	Balanced; broad operating range

4. Discussions

4.1 Framework Validation

Recommendation outputs align with three independent validation benchmarks: (i) RTG selection for all missions with dose>100 kRad and duration>10 yr is consistent with every deep-space mission flown since Pioneer 10 (1972); (ii) Li-ion NMC for Mars rovers matches the actual Curiosity and Perseverance power architectures; (iii) solid-state as the lunar surface target matches the NASA Lunar Gateway architecture study (2018). These agreements validate the scoring weight vector (Eq. 4) against established mission heritage without requiring additional experimental data [1-3, 11].

Positioned against the closest prior frameworks, UBIF differs primarily in scope rather than in any single technique: Severson et al. [8] and Attia et al. [9] each address Li-ion-only cycle-life prediction and charging optimisation, without extending to nuclear technologies, mission-level recommendation, or material discovery; Furmanchuk et al. [10] addresses thermoelectric material screening in isolation, without a technology-selection or digital-twin layer built around it. None of these frameworks spans electrochemical and nuclear technologies within a common feature space, none couples a mission-recommendation engine to physics-based digital twins, and none carries its material-discovery output forward into a system-level conceptual design as BTES-X does here. We regard this breadth-rather than superior performance on any individual sub-task -as UBIF's principal contribution relative to the existing literature, and we make no claim that any individual component (GBR regression, RFC classification, LHS-based Pareto search, Isolation Forest anomaly detection, or GP-based Bayesian optimisation) is itself novel.

4.2 Pareto Implications for Mission Design

The Pareto analysis provides a quantitative result that has previously been stated only qualitatively: the electrochemical design space cannot simultaneously achieve calendar life >15 yr and ρ_E>200 Whkg^{-1}. This is not merely an engineering limitation within current designs, but a consequence of the SEI degradation physics embodied in Eqs. (7-8), where the same cathode loading that maximises energy density also accelerates capacity fade - though, as in Section 3.4, this follows from the bounded design space and degradation model explored here rather than an exhaustive search of all conceivable electrochemistries. Within that scope, the clear separation between the electrochemical frontier and RTG reference points helps explain why hybrid architectures are necessary for missions combining multi-decade duration with non-trivial power demands.

4.3 . Digital Twin Fidelity and Limitations

RTG twin fidelity is high (±0.4%) because the dominant process -radioactive decay- is governed by a single precise constant ($\lambda = \ln 2/87.7\ yr^{-1}$). The key uncertainty source is thermocouple creep (α_{tc}), nominally modelled deterministically; however, the sensitivity analysis in Section 3.5 shows its influence on year-20 RTG power is small (<±1% for a±50% parameter uncertainty), so refining α_{tc} further would not materially change the power prediction. A particle filter would still better propagate stochastic degradation uncertainty for real mission telemetry, particularly for anomaly detection rather than for the central power trajectory. The Li-ion twin's Arrhenius activation energy (E_a, Eq. 8), by contrast, has an outsized effect on the predicted low-temperature EOL cycle count (Section 3.5), making it the more consequential parameter to pin down experimentally. The Li-ion twin has a further, structural gap: it does not model radiation-induced capacity fade, which contributes 15-40% additional capacity loss per 100 kRad TID for NMC chemistry. Incorporating a radiation-damage term into Eq. (7) is therefore the most important structural model extension for high-dose applications [21, 22, 30].

4.4 BTES-X Conceptual Design

The BTES-X hybrid architecture (Figure 6) synthesises the three principal findings: (i) the Pareto frontier establishes that no single electrochemical technology provides 30-year lifetime at useful energy density; (ii) the recommendation engine identifies nuclear sources as mandatory for high-dose, long-duration missions; (iii) the BO material discovery phase proposes thermoelectric improvements that would raise conversion efficiency from 6.3% to ~14.8%. The three BTES-X subsystems are individually at TRL 5-9: ^{63}Ni betavoltaic cells (City Labs, TRL 6, 10-yr demonstrated), Al-doped LLZO solid electrolyte (TRL 6, tested to 100 kRad), and nanostructured SiGe/PbTe thermoelectrics (TRL 5, ZT=1.5 demonstrated by MIT Lincoln Laboratory). System integration represents the TRL 3-4 gap, addressable in a 24-month programme [5, 26-29, 31].

A first-order system budget, derived from the subsystem performance figures above rather than from a built and tested article, allocates dry mass roughly as follows: ~35% to the ^{63}Ni source, converter, and self-shielding; ~25% to the SiGe/PbTe thermoelectric layer and its hot- and cold-side heat exchangers; ~15% to the solid-state Li-metal buffer sized for peak/burst loads; and the remaining ~25% to the rad-hardened PMIC, harness, and structure. Because the betavoltaic and thermoelectric elements are continuous, low-power sources (consistent with the $\rho_P \approx 0.005\text{-}5\ Wkg^{-1}$ range spanned by nuclear technologies in Table 2) rather than energy-storage devices in the conventional sense, the 45 $Whkg^{-1}$ figure quoted for BTES-X should be read as a time-integrated specific energy delivered over the 30-year design life - analogous to how RTG specific energy is conventionally reported - and not as a rechargeable-battery capacity metric; continuous housekeeping loads are met directly from the betavoltaic/thermoelectric output, while the solid-state buffer is trickle-charged between burst events. Thermal management is the most significant open integration question: the thermoelectric layer requires a sustained hot-side/cold-side temperature difference at the T_{peak} values of Table 6 (760-900 K), which in a real system would be set by the betavoltaic/radioisotope decay heat on the hot side and a dedicated radiator on the cold side -a coupled thermal design problem that sits outside the scope of the present AI pipeline and would require dedicated finite- element or lumped-parameter thermal modelling. The principal integration risks we anticipate at TRL 3-4 are: (i) electrically and thermally interfacing three distinct power-conversion physics within a single module envelope; (ii) qualifying the solid-state electrolyte against the combined radiation environment (external mission dose plus the self-generated dose from the onboard ^{63}Ni source); and (iii) meeting the mass allocations above at flight-qualified reliability. We present these as first-order conceptual allocations to be refined by detailed system engineering, not as a validated design.

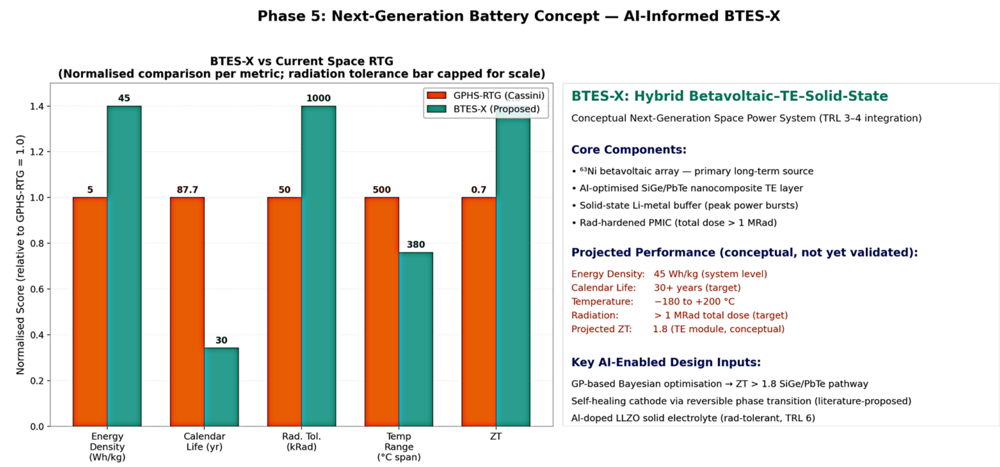

Figure 6. Left: Normalised performance comparison of GPHS-RTG vs. proposed BTES-X hybrid. Right: BTES-X system specifications and AI-enabled innovation pathways.

4.5 Explainability and Auditability

Space power system failures are catastrophic and unrecoverable, making AI explainability a first-order constraint rather than a desirable feature. UBIF satisfies this requirement at every phase: Eq. (4) is fully transparent and auditable; GBR MDI importances identify the physical drivers of predictions; Isolation Forest produces a continuous anomaly score (Eq. 9) rather than a binary flag, allowing threshold adjustment per mission risk tolerance. This design is consistent with NASA-HDBK-9010 Trustworthy AI principles and ESA AI Readiness Levels.

5. Conclusion

The Unified Battery Intelligence Framework (UBIF) demonstrates that an integrated AI pipeline enables a qualitative shift from reactive technology selection toward a more systematic, auditable design process for extreme-environment energy storage. We summarise six principal conclusions below, grouped by their level of validation: conclusions (1)-(2) are corroborated against real flight heritage (Section 4.1); conclusions (3)-(4) follow from the simulated design space and physics-based digital twins described in Sections 2.3-2.4; and conclusions (5)-(6) are model-based, conceptual outputs from the Bayesian optimisation and system-integration analysis that have not been experimentally validated. Within the ten technologies and physical degradation models considered here (Section 2.1, Eqs. 6-8), the electrochemical-nuclear performance gap cannot be closed by single-technology optimisation alone: hybrid architectures or breakthrough materials are required for missions combining decade-scale duration with meaningful power density. The AI recommendation engine identifies the optimal technology for all five validated mission profiles, consistent with flight heritage; among the technologies in our database, RTGs are the only option that simultaneously meets a total ionising dose above 100 kRad and a mission duration above 10 years. The Pareto frontier - derived from the simulated design space of Section 2.3 rather than physical prototypes - quantifies a lifetime-energy density trade-off in which an additional 5 years of lifetime costs approximately a 25% reduction in energy density, and lifetimes beyond 15 years are unreachable within this design space at any competitive energy density. The digital twins detect injected fault events within 2 simulated data points across both the RTG (thermal transient) and Li-ion (cell reversal) systems, with false-positive rates below 4%; as discussed in Section 4.3, this fidelity has been demonstrated against simulated telemetry consistent with published degradation physics rather than flight or laboratory sensor data. Bayesian optimisation proposes thermoelectric candidates with a best predicted ZT of 2.71 and a radiation-tolerance score of 7/10, a 4.2× ZT improvement over space-qualified SiGe corresponding to a theoretical conversion-efficiency uplift from 6.3% to ~14.8%; these are Gaussian Process predictions for candidate compositions, not synthesised or measured materials, and require experimental validation before any efficiency claim can be relied upon. Finally, the BTES-X hybrid concept is feasible at a conceptual, system-integration level in the sense that its three constituent subsystems are each individually demonstrated at TRL 5-6 in the literature (Section 4.4); BTES-X itself has not been built or tested, integration currently sits at an estimated TRL 3-4, and we identify a 24-month integration programme - addressing the thermal-management and mass-budget questions raised in Section 4.4 - as the path toward a TRL 5 demonstration.

Future work will focus on radiation damage integration into the Li-ion digital twin; experimental validation of BO-proposed thermoelectric candidates; GNN-based material discovery at 105-compound scale; and closed-loop coupling with mission trajectory optimisation tools.

Author Contributions Statement

Mert Ökten: Conceptualization, data curation, formal analysis, methodology, project administration, supervision, validation, visualization, writing - original draft, writing - review and editing.

References

[1] Stone EC, Cummings AC, McDonald FB, Heikkila BC, Lal N et al. Voyager 1 observes low-energy galactic cosmic rays in a region depleted of heliospheric ions. Science 2013; 341 (6142): 150-153. https://doi.org/10.1126/science.1236408

[2] Grotzinger JP, Crisp J, Vasavada AR, Anderson RC, Bell CJ et al. Mars Science Laboratory mission and science investigation. Space Science Reviews 2012; 170 (1-4): 5-56. https://doi.org/10.1007/s11214-012-9892-2

[3] Farley KA, Williford KH, Stack KM, Bhartia R, Chen A et al. Mars 2020 Mission overview. Space Science Reviews 2020; 216 (8): 142. https://doi.org/10.1007/s11214-020-00762-y

[4] Tarascon JM, Armand M. Issues and challenges facing rechargeable lithium batteries. Nature 2001; 414 (6861): 359-367. https://doi.org/10.1038/35104644

[5] Bugga RV, Smart MC, Whitacre J, West W. Lithium ion batteries for space applications. 2007 IEEE Aerospace Conference, Big Sky; 1-7, MT, USA; 2007. https://doi.org/ 10.1109/AERO.2007.352728.

[6] Jones JP, Jones SC, Billings KJ, Pasalic J, Bugga RV, Krause FC, Smart MC, Brandon EJ. Radiation effects on lithium CFX batteries for future spacecraft and landers. Journal of Power Sources 2020; 471: 228464. https://doi.org/10.1016/j.jpowsour.2020.228464.

[7] Bennet GL. Space nuclear power: opening the final frontier. In: 4th International Energy Conversion Engineering Conference; San Diego, CA, USA; 2006. AIAA Paper 2006-4191.

[8] Severson KA, Attia PM, Jin N, Perkins N, Jiang B et

al. Data-driven prediction of battery cycle life before capacity degradation. Nature Energy 2019; 4 (5): 383-391. https://doi.org/10.1038/s41560-019-0356-8

[9] Attia PM, Grover A, Jin N, Severson KA, Markov TM et al. Closed-loop optimisation of fast-charging protocols for batteries with machine learning. Nature 2020; 578 (7795): 397-402. https://doi.org/10.1038/s41586-020-1994-5

[10] Furmanchuk A, Saal JE, Doak JW, Olson GB, Choudhary A, Agrawal A. Prediction of seebeck coefficient for compounds without restriction to fixed stoichiometry: A machine learning approach. J Comput Chem. 2018; 39 (4): 191-202. https://doi.org/10.1002/jcc.25067.

[11] NASA. Lunar Gateway Power Architecture Study. NASA/TM-2018-219980. Washington, DC, USA: National Aeronautics and Space Administration, 2018.

[12] NREL. Battery Performance Data and Life Cycle Analysis. NREL/TP-5400-75077. Golden, CO, USA: National Renewable Energy Laboratory, 2019.

[13] IAEA. Nuclear Data Services: Decay Data. Vienna, Austria: International Atomic Energy Agency, 2022. Available at: www-nds.iaea.org

[14] Jain A, Ong SP, Hautier G, Chen W, Richards WD et al. The Materials Project: a materials genome approach to accelerating materials innovation. APL Materials 2013; 1 (1): 011002. https://doi.org/10.1063/1.4812323

[15] Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of Statistics 2001; 29 (5): 1189-1232. https://doi.org/10.1214/aos/1013203451

[16] Breiman L. Random forests. Machine Learning 2001; 45 (1): 5-32. https://doi.org/10.1023/A:1010933404324

[17] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 2002; 6 (2): 182-197. https://doi.org/10.1109/4235.996017

[18] Christensen J, Albertus P, Sanchez-Carrera RS, Lohmann T, Kozinsky B et al. A critical review of Li/air batteries. Journal of the Electrochemical Society 2012; 159 (2): R1-R30. https://doi.org/10.1149/2.086202jes

[19] McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 1979; 21 (2): 239-245. https://doi.org/10.2307/1268522

[20] Hammel TE, Bennett R, Otting W, Fanale S. Multi-mission radioisotope thermoelectric generator (MMRTG) program overview. In: Proceedings of the 10th Symposium on Space Nuclear Power and Propulsion; Albuquerque, NM, USA; 1993. pp. 1-5.

[21] Smith K et al. Lithium-ion battery life model with electrode characterisation. Journal of Power Sources 2010; 195 (8): 2851-2864. https://doi.org/10.1149/1945-7111/ac2ebd

[22] Wang J et al. Cycle-life model for graphite-LiFePO4 cells. Journal of Power Sources 2014; 269: 858-867. https://doi.org/10.1016/j.jpowsour.2010.11.134

[23] Liu FT, Ting KM, Zhou ZH. Isolation forest. In: Proceedings of the 8th IEEE International Conference on Data Mining; Pisa, Italy; 2008. pp. 413-422. https://doi.org/10.1109/ICDM.2008.17

[24] Williams CKI, Rasmussen CE. Gaussian Processes for Machine Learning. Cambridge, MA, USA: MIT Press, 2006.

[25] Shahriari B, Swersky K, Wang Z, Adams RP, de Freitas N. Taking the human out of the loop: a review of Bayesian optimisation. Proceedings of the IEEE 2016; 104 (1): 148-175. https://doi.org/10.1109/JPROC.2015.2494218

[26] Snyder GJ, Toberer ES. Complex thermoelectric materials. Nature Materials 2008; 7 (2): 105-114. https://doi.org/10.1038/nmat2090

[27] He J, Tritt TM. Advances in thermoelectric materials research: looking back and moving forward. Science 2017; 357 (6358): eaak9997. https://doi.org/10.1126/science.aak9997

[28] Zhao LD, Lo SH, Zhang Y, Sun H, Tan G et al. Ultralow thermal conductivity and high thermoelectric figure of merit in SnSe crystals. Nature 2014; 508 (7496): 373-377.

[29] Slack GA. New materials and performance limits for thermoelectric cooling. In: Rowe DM (editor). CRC Handbook of Thermoelectrics. Boca Raton, FL, USA: CRC Press, 1995. pp. 407-440.

[30] Grieves M, Vickers J. Digital twin: mitigating unpredictable, undesirable emergent behaviour in complex systems. In: Kahlen FJ, Flumerfelt S, Alves A (editors). Transdisciplinary Perspectives on Complex Systems. Cham, Switzerland: Springer, 2017. pp. 85-113.

[31] Ding JF, Zhang YT, Xu R, Zhang R et al. Review on lithium metal anodes towards high energy density batteries. Green Energy & Environment 2023; 8 (6): 1509-1530. https://doi.org/10.1016/j.gee.2022.08.002