Quantifying artificial intelligence impacts on the photovoltaics value chain

Agnieszka E. Rządkowska[1, 2]

[1] *Institute of International Relations and Security, Faculty of Social Sciences, University of Wrocław, Plac Uniwersytecki 1, 50-137 Wrocław, Poland*
[2] *European Solar Network, Rue de Grand-Bigard 14, 1082 Brussels, Belgium*

October 3, 2025

Abstract

We present a transparent, policy-grade framework to measure how data-driven methods affect costs, yields, and risks across the photovoltaics (PV) value chain—from materials and manufacturing quality assurance to device/plant control, forecasting, market operations, and soft costs. The framework (i) specifies defensible counterfactual baselines matched to the same physical and market context, (ii) maps technical key performance indicators to cash-flow items through calibrated transfer functions (e.g., ΔMWh, ΔO&M, ΔCAC), and (iii) propagates uncertainty with Monte Carlo simulation and variance-based global sensitivity. We synthesize peer-reviewed evidence and report results as incremental changes in net present value (ΔNPV), plant-level levelized cost of energy (ΔLCOE), and—where integration effects are material—system LCOE.

Impacts are real but context-dependent. In perovskites, learning-guided passivation and interface choices yield experimentally verified gains in quasi-Fermi level splitting and stability; when translated through field-relevant degradation priors, these produce modest ΔLCOE reductions. EL/IR analytics identify high-severity defects efficiently; combined with severity-to-loss functions and replacement break-even rules, they deliver measurable ΔMWh and avoided warranty costs without overstating low-severity findings. AI-enhanced MPPT improves dynamic tracking by about 0.5–3 percentage points on tests, typically converting to sub-percentage annual energy gains at utility scale after site convolution. The largest, most robust economic benefits arise from probabilistic, calibrated PV forecasts embedded in unit commitment, reserve sizing, and bidding, which lower production and imbalance costs in sufficiently flexible systems. On the demand side, data-driven targeting can reduce customer acquisition costs by roughly 15%, while standardized permitting shortens timelines and raises ΔNPV. A governance layer—proper scoring and reproducible baselines, subgroup reporting for fairness, and DER cybersecurity aligned with IEEE 1547/1547.3, IEC 62443, and NERC CIP—is essential to realize benefits safely. We conclude with a prioritized research agenda and open data/model templates to support reproducible, distribution-aware policy design.

Keywords: photovoltaics, artificial intelligence, machine learning, data science, probabilistic forecasting, reliability diagrams, techno-economic valuation, economic dispatch, soft-costs

1 Introduction

Artificial intelligence (AI), here understood as data-driven statistical learning and optimization methods, now appears across the photovoltaic (PV) value chain: materials and interface discovery, manufacturing quality assurance (QA), device- and plant-level control, power forecasting

and market participation, and demand-side adoption. Despite rapid progress, the literature measuring *economic* effects—rather than only technical metrics—remains scattered across subfields. This paper unifies those strands with a framework that: (i) specifies defensible, domain-appropriate counterfactuals; (ii) maps technical performance changes to project and system cash flows via explicit transfer functions; and (iii) propagates uncertainty with transparent assumptions. Throughout, we rely on peer-reviewed studies whose claims are backed by appropriate validation, careful scope limits, and reproducible methods.

1.1 From technical metrics to policy-relevant value

Materials and interfaces. Upstream, AI has accelerated the exploration of large compositional and process spaces. In general materials science, scaling studies show that large models can sift vast candidate sets without overclaiming beyond training domains when evaluations are rigorous and out-of-sample [1]. In perovskite PV specifically, learning-guided screening has proposed bifunctional surface passivants that were verified experimentally, improving quasi-Fermi level splitting and open-circuit voltage in devices—a direct link from model to hardware performance rather than in-silico promise alone [2]. Reviews emphasize the conditions under which machine learning (ML) reduces trial-and-error while warning about small or biased datasets that limit generalization [3].

Manufacturing QA and field inspection. In production and in the field, electroluminescence (EL) and infrared (IR) imaging combined with computer vision detect defects such as micro-cracks, inactive regions, and hot-spots. Recent work reports real-time EL defect detection with high mean average precision, backed by cross-dataset testing rather than single-benchmark gains [4]. Robustness to noise and topology-aware detection further improve reliability [5]. The economic relevance is established when detection is linked to validated *severity-to-power-loss* relations and to warranty/O&M actions, an approach argued for in techno-economic syntheses of degradation and inspection methods [6].

Device- and plant-level control. At the converter and inverter layer, AI-based maximum power point tracking (MPPT) methods typically improve tracking efficiency and settling time under partial shading and fast irradiance changes relative to well-tuned conventional baselines, when tested on dynamic profiles with matched hardware constraints [7, 8]. These are engineering-level improvements which, after convolution with site-specific ramp and shading statistics, yield modest but compounding annual energy gains (Section 5.4 of the full paper formalizes this mapping).

Forecasting and market integration. For system operations and markets, forecast quality has value only insofar as it improves commitment, reserve sizing, redispatch, and bidding. Production-costing results show that better day-ahead solar forecasts reduce fast-start commitments, ramping, and curtailment in realistic unit-commitment/economic-dispatch (UC/ED) settings [9]. A recent review distills best practice: evaluate probabilistic forecasts with strictly proper scores and map those to *economic* outcomes inside the operational model to avoid conflating technical error metrics with value [10].

Soft-costs and adoption. On the demand side, soft-costs remain material. Field evidence indicates that learning-based targeting can reduce residential customer-acquisition costs (CAC) by roughly 15% (about \$0.07/W) relative to standard logistic baselines when evaluated out-of-sample [11]. The result is jurisdiction- and dataset-specific but illustrates a tractable lever: policies can support transparent, fairness-audited analytics that lower CAC without excluding disadvantaged groups.

Table 1: Mapping AI levers to measurable engineering KPIs and economic outcomes. The cited studies are representative and were selected for careful validation and transparent scope.

AI lever	**Representative evidence (scope)**	**Engineering KPI (validated)**	**Economic outcome**
Materials/interface screening	[1–3]	Higher QFLS/V_{oc}, improved stability under specified stress	Δdegradation priors $\rightarrow$ ΔMWh, ΔLCOE
Manufacturing QA / field inspection	[4–6]	Detector ROC/PR; severity-to-loss functions	ΔMWh, Δwarranty, ΔO&M
Device/plant control (MPPT)	[7, 8]	Tracking efficiency, settling time on dynamic tests	ΔMWh (after site convolution)
Forecasting & market integration	[9, 10]	Proper scores (CRPS/log); calibration; scenario quality	Δproduction cost, Δreserve, Δimbalance, revenues
Soft-costs (marketing)	[11]	CAC reduction (prospective tests)	ΔCapEx/W, higher close rates $\rightarrow$ ΔNPV

1.2 A quantitative mapping from engineering to economics

To avoid double counting and to keep claims within supported domains, we adopt a consistent transfer-function approach. For any intervention X evaluated against a named baseline b,

$$\Delta\mathrm{NPV}(X;b) = \sum_t \frac{\Delta\mathrm{CashFlow}_t(X;b)}{(1+r)^t}, \tag{1}$$

with $\Delta\mathrm{CashFlow}_t$ decomposed into plant energy, O&M, warranty, reserve and imbalance costs, and timing effects (e.g., permitting). For energy-centric levers, the avoided or added annual energy is

$$\mathbb{E}[\Delta E] = \sum_{d\in\mathcal{D}} \Pr(d)\cdot \mathrm{TPR}_d \cdot \mathbb{E}\left[\frac{\Delta P}{P}\,\middle|\, d\right]\cdot E_{\mathrm{baseline}}, \tag{2}$$

where d indexes defect classes (e.g., inactive-area cracks, hot-spots), TPR_d is true-positive rate at the chosen decision threshold, and $\mathbb{E}[\Delta P/P \mid d]$ is the empirically validated loss severity for class d.[1] For MPPT,

$$\mathbb{E}[\Delta E] = E_{\mathrm{DC,ideal}}\cdot \mathbb{E}[\eta_{\mathrm{trk,AI}} - \eta_{\mathrm{trk},b}]\,, \tag{3}$$

where the expectation integrates the site distribution of ramp and shading regimes; for forecasting, improvements in calibrated predictive distributions are valued only via UC/ED or market bidding results, never by error metrics alone.

Table 1 summarizes the core levers, representative evidence, measurable engineering KPIs, and the cash-flow items they affect.

1.3 Why a unified synthesis is needed

Across the above domains, credible studies already show that AI can (i) shorten discovery cycles for passivation and interfaces, (ii) increase inspection throughput and precision, (iii) improve MPPT behavior under dynamic conditions, (iv) reduce operational costs through calibrated probabilistic forecasts inside UC/ED, and (v) lower soft-costs in residential diffusion. What has been missing is a *common* accounting framework that:

[1] Severity-to-loss relations are taken from studies that pair EL/IR classes with I–V or field energy changes, and are used only within their tested conditions.

1. defines counterfactual baselines appropriate to each decision layer (materials, plant, market, adoption);

2. avoids double counting when multiple tools act on the same project;

3. expresses benefits in commensurate economic units (e.g., ΔLCOE, ΔNPV, changes in system-LCOE where integration costs matter); and

4. propagates uncertainty with transparent assumptions and reports sensitivity.

This synthesis provides that structure, keeping claims bounded to validated domains and reporting uncertainty explicitly.

1.4 Contributions and scope

This paper contributes: (1) a taxonomy of AI use-cases across the PV lifecycle aligned with decision-relevant outcomes; (2) standardized impact metrics and counterfactual definitions; (3) transfer functions that map technical KPIs (e.g., CRPS, tracking efficiency, detector precision/recall) to cash-flow items (curtailment avoided, reserve-cost changes, O&M savings, warranty liabilities, CAC); and (4) evidence-weighted priors for policy evaluation distilled from peer-reviewed studies. The result is a transparent, reproducible basis to quantify when AI actually reduces PV system costs or improves reliability—and when improvements in technical scores do not translate into economic value.

2 Framework for quantitating AI impacts (baselines, metrics, uncertainty)

2.1 Principles and scope

This chapter sets a common, measurement-driven basis for estimating how data-driven tools change costs and performance across the PV value chain. We follow three rules. First, all comparisons use explicit, defensible *counterfactual baselines* that reflect the same physical and market context. Second, we convert technical changes (e.g., improved detection, tracking, or forecasts) into *cash-flow* changes using conservative transfer functions, and we avoid double counting. Third, we propagate uncertainty with transparent assumptions and report which inputs drive the results. We ground these steps in peer-reviewed methods spanning PV field reliability and degradation [12, 13], EL/IR inspection and defect-loss links [6, 14], MPPT benchmarking [7, 8], production-cost valuation of forecasts [9, 10], proper scoring rules [15], and system-LCOE accounting for integration costs [16].

2.2 Counterfactual baselines and treatment definitions

Units of analysis. Impacts should be computed where decisions are made, then aggregated if needed:

1. **Device/material**: e.g., passivation recipe or interface layer affecting V_{oc} or stability [2, 3].

2. **Component/plant**: e.g., EL/IR-based defect detection and triage on modules/strings [4–6].

3. **Operations/market**: e.g., probabilistic PV forecasts used in UC/ED or market bidding [9, 10].

4. **Adoption/soft-costs**: e.g., targeting that lowers residential CAC [11].

Table 2: Recommended baselines by decision layer. All baselines must reflect the same site, hardware, data cadence, and market context as the treatment.

Layer	**Treatment example**	**Counterfactual baseline (same context)**
Device/material	Learning-guided passivation or interface choice [2, 3]	Expert-guided screening with the same synthesis throughput; report device metrics and stability under identical tests; use field degradation priors for translation to energy [12, 13].
Component/plant	EL/IR triage with automated detection [4–6]	Existing inspection practice (manual labeling/throughput, current thresholds); link defect classes to validated loss severities and actions [6, 14].
Operations/market	Probabilistic forecasts in UC/ED or trading [9, 10]	Operator benchmarks at identical horizons (persistence, NWP+post-processing); same commitment/bidding rules; evaluate per horizon/regime.
Adoption/soft-costs	Targeted outreach lowering CAC [11]	Incumbent models (e.g., logistic regression) under equal budgets; prospective, out-of-sample performance and conversion/attrition modeling.

Baseline selection. Table 2 summarizes recommended baselines. These must match the *same* site conditions (irradiance, shading, temperature), hardware vintage, data cadence, and market rules. Because aging shifts output and risk, baselines should reflect empirical degradation distributions (crystalline-Si median $\approx$ 0.5–0.6%/yr with dispersion) rather than a single value [12, 13].

Causal identification in the field. When interventions are rolled out over time (e.g., by plant or municipality), simple two-way fixed-effects difference-in-differences can be biased. Use modern staggered-adoption estimators that recover cohort- and period-specific average treatment effects before aggregation, with pre-trend checks and event-study plots for diagnostics [17, 18].

2.3 Impact metrics and transfer functions

Economic objective. We report *incremental* net present value and levelized cost changes relative to a named baseline b:

$$\Delta\mathrm{NPV}(X;b) = \sum_t \frac{\Delta\mathrm{CashFlow}_t(X;b)}{(1+r)^t}, \tag{4}$$

$$\Delta\mathrm{LCOE}(X;b) = \frac{\Delta\mathrm{Cost}(X;b)}{E_b} - \frac{\mathrm{Cost}_b \cdot \Delta E(X;b)}{E_b^2}, \tag{5}$$

where E_b is baseline lifetime energy. For integration effects at high penetration, we complement plant metrics with *system-LCOE*, which allocates balancing, profile, and curtailment costs consistently across technologies [16].

Technical → economic transfer. We use conservative mappings tied to validated domains:

$$\mathbb{E}[\Delta E_{\text{insp}}] = E_{\text{base}} \sum_{d \in \mathcal{D}} \Pr(d) \cdot \text{TPR}_d(\tau) \cdot \mathbb{E}\left[\frac{\Delta P}{P} \,\middle|\, d \right], \tag{6}$$

$$\mathbb{E}[\Delta E_{\text{MPPT}}] = E_{\text{DC,ideal}} \cdot \mathbb{E}\left[\eta_{\text{trk}}^{(X)} - \eta_{\text{trk}}^{(b)} \right], \tag{7}$$

$$\text{VoI}_{\text{forecast}} = C_{\text{sys}}^{(b)} - C_{\text{sys}}^{(X)} \quad \text{(production-cost or trading model)}, \tag{8}$$

where $\text{TPR}_d(\tau)$ is the detector's true-positive rate for defect class d at operating threshold τ, and $\mathbb{E}[\Delta P/P \mid d]$ is the validated loss severity for d (EL/IR class → I–V/energy) [6, 14]. Equation (7) integrates the distribution of shading and ramp regimes at the site [7, 8]. Equation (8) computes the *value of improved information* inside a unit-commitment/economic-dispatch or trading simulator, never from error scores alone [9, 10].

Soft-costs to cash flows. Document measured CAC reductions and apply them in project pro formas as ΔCapEx/W and conversion-rate changes, with uncertainty bands carried through to ΔNPV [11].

2.4 Uncertainty and sensitivity

What varies. We propagate (i) *parameter* uncertainty in technical KPIs (detector ROC, tracking efficiency distributions, forecast calibration by regime), costs (labor, logistics), and prices (energy, reserves), and (ii) *structural* uncertainty (e.g., alternative severity-to-loss functions, replacement thresholds, or reserve pricing rules). Monte Carlo draws produce output distributions for ΔNPV and ΔLCOE.

What drives the variance. We use variance-based global sensitivity (Sobol first-order and total-effect indices) to attribute output variance to inputs and their interactions, reporting convergence diagnostics and tornado charts [19, 20].

Discounting and risk. We show results over a grid of real WACC values relevant to the policy setting rather than embedding "risk" by arbitrarily inflating the discount rate. When interventions change failure risk, we model this explicitly (e.g., Weibull components in O&M cash-flows) [20].

Forecast uncertainty. Probabilistic forecasts are scored with strictly proper rules (CRPS, log score) by horizon and regime to avoid perverse incentives and to support economic mapping [10, 15].

2.5 Model chain and evaluation protocol

1. **Data curation.** Record provenance, censoring, and representativeness (e.g., EL image quality, SCADA cadence). Use reliability syntheses and meta-analyses to set priors for degradation and defect progression [13, 21].
2. **Technical models.** Train and test against strong baselines with application-appropriate splits (temporal blocking for forecasts; cross-site validation for inspection); report CIs.
3. **Operational models.** Translate technical outputs into production-cost or trading results in a UC/ED or market model; report Δproduction cost, Δreserves, Δimbalance, starts, curtailment [9, 10].

4. **Finance layer.** Convert operational deltas into cash-flows, apply a WACC grid, compute ΔNPV, ΔLCOE, and Δsystem-LCOE where relevant [16, 20].
5. **Attribution.** When multiple tools act (e.g., inspection + forecasting), credit *incremental* value with nested baselines to avoid double counting.
6. **Field inference.** For staggered rollouts, estimate impacts with modern DiD (cohort-specific effects, event studies, robustness) [17, 18].

2.6 Aggregation and reporting

We report (i) **plant-level** ΔLCOE/ΔNPV with 5–95% intervals, (ii) **portfolio-level** results by integrating over climates and vintages using degradation distributions [12, 13], and (iii) **system-level** integration benefits using system-LCOE [16]. Minimum reporting includes: baseline definition; technical KPIs (with CIs); transfer-function equations/parameters; Monte Carlo and convergence checks; Sobol indices; WACC grid; and causal diagnostics where applicable. For O&M costing, link predicted events to a documented cost model and failure library [22].

3 Materials & interface discovery (perovskites, passivation; time-to-innovation impacts)

3.1 Policy motivation and scope

For AI to matter in energy *policy and economy*, upstream progress in materials must translate into (i) reproducible improvements in device physics (e.g., higher quasi-Fermi level splitting, reduced non-radiative losses), (ii) better stability under stressors that resemble real operating conditions, and (iii) tangible shifts in cash flows (e.g., lower early-life failures, higher stabilized yield) when mapped through the transfer functions defined in Chapter 2. We focus on metal–halide perovskites (MHPs), where data-driven search has credibly pruned large compositional spaces, prioritized passivants and interlayers, and shortened experimental cycles. Our central question is: *How much faster and cheaper can we identify equally or better performing device stacks than with incumbent methods, and what is the plausible downstream value once durability and yield are accounted for?*

3.2 AI-accelerated exploration of compositional and process space

The clearest upstream contribution of data-driven methods is *sample efficiency*: fewer, better-chosen experiments to reach a target. In perovskites, Bayesian optimization has located optimal hybrid perovskite (HOIP) solvent/compositional choices in $\sim 15 \pm 10$ iterations (72-candidate space) and $\sim 31 \pm 9$ iterations (240-candidate space), i.e., $< 10\%$ of an exhaustive search cost/time; the reported optima were cross-validated against density-functional theory (DFT) computations rather than left as model suggestions [23]. Physics-constrained *data-fusion* active learning, combining high-throughput degradation measurements with first-principles thermodynamics, explicitly targeted *stability* discovery—not only peak efficiency—and improved the rate of finding stable alloys under uncertainty-aware recommendations [24].
Beyond perovskites, an autonomous laboratory integrating ab-initio priors, text-mined synthesis rules, active learning, and robotics realized 41 of 58 novel inorganic targets in 17 days (355 experiments), i.e., > 2 materials/day on average, demonstrating that closed-loop autonomy can compress practical experimental cycles by orders of magnitude [25]. While not perovskite-specific, this sets a conservative upper bound for throughput gains when similar autonomy is applied to perovskite stack development.

Table 3: Representative perovskite interface/passivation evidence and how to use it in economic mapping. KPIs should be reported with confidence intervals and stress protocols.

Result class	**Validated KPI(s)**	**Economic mapping (Chapter 2)**
Bifunctional surface passivants (ML-screened) [2, 26, 27]	$\uparrow$ QFLS, $\uparrow V_{oc}$, reduced non-radiative loss; stability under stated stress	Higher stabilized yield $\Rightarrow \Delta$ MWh; lower early-life failures $\Rightarrow$ warranty/O&M deltas; contributes to Δ LCOE
2D/quasi-2D/DJ interlayers [29, 30]	Stability vs. humidity/light/bias (protocols explicit)	Update degradation priors (Jordan/Kurtz) used in yield and O&M models; affects Δ LCOE ranges
Compositional tuning (Rb, cations) [28]	Stabilized PCE; retention under 85°C/500 h MPP	Baseline for passivation merit; informs conservative stability priors in cash-flow mapping

Guardrails. These studies limit over-claim by (i) benchmarking against validated computational or exhaustive baselines [23], (ii) constraining learning with physics and reporting predictive uncertainty [24], and (iii) disclosing success/failure rates across many targets [25]. Two caveats are essential for policy use: (a) in-silico or benchtop acceleration does not guarantee proportional speed-ups in *module*-relevant programs (scale-up, encapsulation, quality control); and (b) literature-mined priors may be biased; open evaluation datasets reduce that risk.

3.3 Interfaces and passivation: from photophysics to cash flows

At the device level, economic salience arises when materials choices reduce non-radiative recombination, suppress harmful defects, and enhance stability under realistic stress. Physics-informed ML screens of small molecules (including pseudohalide anions) have nominated bifunctional passivants that, upon *experimental* verification, increased open-circuit voltage (V_{oc}) and stability by suppressing anti-site defects and lattice distortion—an explicit model $\rightarrow$ device chain with quantitative gains at the film/cell level [2, 26]. Independent screening over curated datasets likewise identified *simple* small-molecule passivators with prospective validation (e.g., thiophene-ethylammonium chloride), reducing V_{oc} losses in p–i–n architectures [27].
Non-ML evidence provides anchors for attributing AI-driven choices. Rubidium incorporation raised stabilized efficiencies to $\sim$ 19–21.6% and retained $\sim$ 95% performance at 85°C for 500 h under MPP in small-area devices [28]; such compositional strategies set baselines for passivation merit. Interfacial *hole-shuttle* layers have raised quasi-Fermi-level splitting (QFLS) and V_{oc} by mitigating specific recombination pathways—tying interfacial chemistry to rigorous photophysical metrics (QFLS $\rightarrow V_{oc}$) rather than to single-point PCEs alone [26]. For structural motifs, 2D/quasi-2D capping and Dion–Jacobson (DJ) interlayers have delivered documented stability gains, though mechanisms are stack-specific and stress-dependent (humidity, light, bias) [29, 30]. These are precisely the inputs needed by the transfer functions of Chapter 2: QFLS/V_{oc} improvements map to stabilized yield; durable interfaces update degradation priors used in plant models.

3.4 From lab metrics to economic value: a transparent mapping

We adopt the conservative chain defined in Chapter 2 and tailor it to materials/interface changes.

Table 4: Inputs required to move from device metrics to plant economics. Report with confidence intervals and stress protocols; use nested baselines to avoid double counting.

Input (with units)	**Role in mapping**
ΔV_{oc}, ΔQFLS (V, eV)	Sensitivity-based yield uplift (Eq. 10); uncertainty propagated to Δ MWh
Stability metrics (retention under protocol; early-life failure rate)	Update degradation priors (Jordan/Kurtz) and early-failure cash flows (warranty/O&M)
Cost per synthesis/characterization cycle (c_{exp})	R&D outlay change via Eq. 9; informs two-factor learning calibration
Throughput/TTI (τ_b, τ_X; samples/day)	Bring-forward effects; program scheduling and pipeline costs

(1) Time-to-innovation (TTI) compression. Let τ_b be the median experimental cycles to reach a target stack (specified QFLS/V_{oc}/stability) under baseline search, and τ_X with data-driven methods. Define $\kappa = \tau_b/\tau_X$ (cycle-speedup). If synthesis run-cost per cycle is c_{exp} and fixed lab overhead per period is C_{fix}, the expected R&D outlay to target is

$$\mathrm{Cost}^{(b)}_{\mathrm{R\&D}} = \tau_b\, c_{\mathrm{exp}} + C_{\mathrm{fix}}, \quad \mathrm{Cost}^{(X)}_{\mathrm{R\&D}} = \tau_X\, c_{\mathrm{exp}} + C_{\mathrm{fix}}. \tag{9}$$

The direct R&D saving is $\Delta\mathrm{Cost}_{\mathrm{R\&D}} = (1 - 1/\kappa)\,\tau_b\, c_{\mathrm{exp}}$. When the shorter TTI enters a deployment program (learning-by-doing), earlier availability of improved stacks also changes the timing of cash inflows (bring-forward effect).

(2) R&D productivity and experience curves. Two-factor learning models attribute cost decline to both cumulative deployment and knowledge investment. Faster, cheaper iteration lowers the *R&D cost per unit of performance gain*, shifting the knowledge-stock term. We conservatively bound the benefit by the *validated* device-level deltas and do not credit speculative spillovers.

(3) Device KPIs to project cash flows. Let ΔV_{oc} and ΔQFLS be verified gains, and let stability tests (e.g., 85°C/500 h) support an updated degradation prior. Then

$$\mathbb{E}[\Delta E] = E_{\mathrm{base}} \times \left(\frac{\partial E}{\partial V_{\mathrm{oc}}} \Delta V_{\mathrm{oc}} + \frac{\partial E}{\partial \mathrm{QFLS}} \Delta\mathrm{QFLS} \right) \times f_{\mathrm{stability}}, \tag{10}$$

where $f_{\mathrm{stability}}$ captures the ratio of expected energy with the updated degradation distribution to the baseline (Jordan/Kurtz) [12, 13]. The resulting Δ MWh feeds Δ LCOE and Δ NPV through the finance layer (Chapter 2).

(4) System integration and tandems. If materials/interface advances enable perovskite/Si tandems, net value depends on module efficiency *and* the cost/stability of the additional layers. System-LCOE perspective clarifies when higher efficiency offsets integration costs; current roadmaps emphasize that interface quality and durability remain rate-limiting for bankable tandems [31]. We therefore avoid crediting tandem-wide cost benefits to AI unless supported by module-relevant durability data.

3.5 Evidence-weighted estimates and uncertainty

For policy appraisal we recommend *evidence-weighted priors*: (i) sample-efficiency from [23], (ii) autonomy throughput from [25], (iii) passivation effectiveness from [2, 26, 27], and (iv)

stability priors from [29, 30] combined with field degradation distributions [12, 13]. Monte Carlo propagation should be combined with variance-based global sensitivity to reveal the main drivers; in most cases, the mapping from lab stress to field durability dominates the uncertainty bands.

3.6 Limits and research needs

Three gaps constrain economic attribution. **Portability:** screened passivants can be architecture-specific; cross-lab reproducibility remains mixed. FAIR databases and standardized reporting (device stack, stress protocols, uncertainty) help close this gap [32]. **External validity:** stress protocols vary widely; module-level studies (encapsulation, glass) are needed to claim plant-level Δdegradation. **Governance:** closed-loop autonomy shortens time-to-evidence but introduces model risk; policy-grade studies should maintain audit trails from suggestion to synthesis and pre-register evaluation endpoints, as demonstrated in [25].

Bottom line. Peer-reviewed evidence supports cautious optimism: data-driven search can measurably shorten exploration cycles [23–25] and credibly prioritize interfaces/passivants that raise QFLS/V_{oc} and durability [2, 26, 27, 29, 30]. When mapped through Chapter 2 and bounded by reliability priors [12, 13], these gains plausibly contribute to faster movement along PV experience curves, with the largest policy leverage where autonomy reduces *program-level* iteration time and where verified passivation yields *module-relevant* stability.

4 Manufacturing QA & field inspection (EL/IR vision; yield, warranty economics)

4.1 Policy motivation and scope

Manufacturing quality assurance (QA) and field inspection determine whether laboratory device advances (Chs. 2–3) convert into bankable energy and lower life-cycle costs. Across crystalline-Si and perovskite/Si supply chains, electroluminescence (EL) and infrared (IR) thermography are the dominant non-destructive modalities. EL reveals micro-cracks, inactive regions, metallization and interconnect defects; IR highlights hot-spots from resistive paths, bypass-diode conduction, partial shading, and soiling. From a policy viewpoint, their value arises only when (i) image-space detection is *validated* against ground truth, and (ii) findings are tied to *severity-to-loss* functions, *action rules*, and *prices*, yielding ΔMWh, avoided warranty outlays, and lower O&M—all computed against a named baseline (Chapter 2).

4.2 What EL/IR actually see — and when it matters

Cracks and inactive areas. Early risk studies showed that micro-cracks without inactive area cause little immediate power loss, while cracks creating electrically inactive regions can be consequential; array-level risk depends on crack geometry and extent [33]. Controlled EL+I–V experiments confirm a wide severity spectrum: small cracks often have negligible effect; larger fractures can cut single-cell output by tens of percent and seed hot-spots under load [34]. Fleet-scale EL campaigns report that large cracks dominate observed power loss shares and that specific crack types rise with age [35].

Hot-spots and thermography. IR surveys on thousands of modules have linked hot-spot classes to performance-ratio (PR) penalties ranging from sub-percent (mild, localized) to double-digit losses (severe, string-level) [36]. Mechanisms include cell cracks progressing to inactive area, solder/connector resistances, and diode conduction. Because IR contrast depends on

Table 5: Representative defect classes visible in EL/IR, validated loss priors, and economic salience. Use severity *distributions* (not maxima) in valuation.

Defect (image cue)	**Mechanism & validated loss priors**	**Economic salience**
Micro-crack (EL line)	Often low immediate loss unless inactive area forms; progression is climate/load dependent [33, 34]	Monitor unless severity crosses loss threshold; replacement only when justified by loss vs. price
Inactive-area crack (EL dark)	Conditional $\Delta P/P$ rises with affected area; can seed hot-spots [34, 35]	High salience; often drives ΔMWh and warranty exposure
Hot-spot (IR bright)	Resistive paths / bypass diode conduction; PR penalty spans $\sim$0.8% to $\sim$15% depending on class [36]	High salience for severe classes; O&M dispatch and replacement decisions
PID (EL pattern, IR warm)	Voltage-induced ionic migration; recovery depends on stress and treatment [38]	Can create fleet-wide losses; mitigation economics depend on efficacy/cost
Soiling / shading (IR cool/patchy)	Environmental; transient unless structural	Typically O&M cleaning, not replacement

irradiance, wind, and tilt, acquisition conditions strongly affect detectability and false positives [37].

Potential-induced degradation (PID) and others. PID is detectable by both EL (characteristic patterns) and IR (thermal signatures), with documented mechanisms, field incidence, and recovery behavior [38]. These modes require bias-stress-aware QA to avoid under- or over-counting.

4.3 Detection capability and robustness (factory and field)

Factory EL. Large labeled EL datasets have enabled robust classification/localization of defects with transparent failure modes. Cross-validated studies report strong accuracy and actionable localization suitable for scrap/repair decisions [39, 40]. The most useful pipelines operate on full-module frames (object detection), producing instance-level locations and severities rather than image-level flags.

Field IR at scale and domain shift. Models trained at one site degrade when applied to another (different irradiance, mounting, climate). Contrastive and transfer-learning approaches evaluated across plants (millions of images) achieve high AUROC with fixed thresholds, with most residual errors in low-severity anomalies [41]. Complementary IR studies show that pre-training and site-specific fine-tuning materially improve cross-site robustness: using datasets spanning different cameras and module populations, Akram *et al.* demonstrate that develop-model transfer (fine-tuning a pre-trained detector) outperforms from-scratch models under strict train/test separation and ablations of augmentation choices, reducing the penalty from domain shift (camera, irradiance, mounting) [42]. This motivates *severity-aware* thresholds and light local calibration in UAV IR programs.

Acquisition envelope and QA of measurements. Measurement conditions drive signal-to-noise. For IR, irradiance $>$ 600–700 W/m^2, moderate wind ($<$ 5–6 m/s), and stable thermal steady-state reduce confounding; emissivity and viewing angle must be controlled [37]. For

Table 6: Acquisition and QA checklist for policy-grade surveys. Meeting these conditions stabilizes defect detectability and reduces false positives.

Modality	**Acquisition envelope**	**QA & documentation**
IR (UAV/handheld)	Irradiance $\gtrsim$ 650 W/m^2; wind $\lesssim$ 5 m/s; steady-state module temp.; controlled viewing angle	Emissivity setting; radiometric calibration; ROC/PR by severity; site-specific threshold; prevalence estimates [37, 41]
EL (factory/field)	Specified injection current; exposure avoiding saturation; focus/flatness verified	Camera linearity check; dark-frame subtraction; detector calibration; severity mapping to I–V / energy [39, 40]

EL, current injection, exposure, and camera linearity require periodic calibration. As with any metrology, a measurement-system analysis (MSA; repeatability, reproducibility) should be reported.

4.4 From pixels to money: transfer functions with prevalence and action rules

Let $d \in \mathcal{D}$ index defect classes. We value detection at an operating threshold τ using *validated* detector metrics (TPR/FPR) and *severity distributions*.

Prevalence-aware value of detection. If π_d is the prevalence of class d, $\mathrm{TPR}_d(\tau)$ the true-positive rate, $\mathrm{FPR}_d(\tau)$ the false-positive rate, L_d the expected annual energy loss (MWh) conditional on d and severity (from EL/IR↔I–V studies), and $C_{\mathrm{rem},d}$ the remediation cost (module swap/repair), the expected net benefit from acting on positive flags is

$$\mathrm{EV}_d(\tau) = \big[\pi_d\,\mathrm{TPR}_d(\tau)\,V(L_d) - (1-\pi_d)\,\mathrm{FPR}_d(\tau)\,C_{\mathrm{diag}}\big] - \pi_d\,\mathrm{TPR}_d(\tau)\,C_{\mathrm{rem},d}, \tag{11}$$

where $V(L_d) = p_{\mathrm{el}} \cdot L_d$ (or a time-of-use integral) and C_{diag} is the cost of confirming/triaging false positives. Summing over classes yields the annualized expected benefit:

$$\mathbb{E}[\Delta\mathrm{CashFlow}_{\mathrm{year}}] = \sum_{d\in\mathcal{D}} \mathrm{EV}_d(\tau). \tag{12}$$

Applying discounting over lifetime gives ΔNPV (Eq. (1) in Chapter 2).

Severity-to-loss mapping. L_d must come from *validated* severity functions: for cracks, $\mathbb{E}[\Delta P/P \mid s]$ over severity s (fraction of inactive area, geometry) [34, 35]; for hot-spots, PR penalties by class [36]; for PID, degradation rates and recovery efficacy [38]. Use distributions (with confidence intervals), not point maxima.

Cost-sensitive thresholding (choose τ by economics). Rather than maximizing accuracy or F1, select $\tau^\star$ to maximize (12). Equivalently, minimize an expected cost function with asymmetric penalties:

$$C(\tau) = \sum_d \Big[\pi_d(1-\mathrm{TPR}_d(\tau))\,V(L_d) + (1-\pi_d)\,\mathrm{FPR}_d(\tau)\,C_{\mathrm{diag}}\Big] + \sum_d \pi_d\,\mathrm{TPR}_d(\tau)\,C_{\mathrm{rem},d}. \tag{13}$$

This aligns the operating point with economics (prices, prevalence, and remediation costs), not just with technical scores.

Table 7: Illustrative action thresholds (replace vs. monitor) as a function of annualized power loss and price assumptions (adapted conceptually from [43]). Exact thresholds are site-specific; values shown are indicative.

Loss severity	**Indicative $\Delta P/P$**	**Energy price band**	**Suggested action**
Low	$\lesssim 1\%$	Low–mid	Monitor; defer action
Moderate	$\sim 10\%$	Mid	Replace if labor/logistics modest; else targeted repair
Severe	$\sim 30\%$	Any	Replace; investigate root cause (diodes, wiring, assembly)

Replace vs. monitor: break-even rule. Module replacement is justified when discounted energy loss exceeds replacement *and* handling costs. A recent decision-analytic study shows that for typical European prices, moderate (∼10%) to severe (∼33%) losses justify replacement, while ∼1% losses generally do not [43]. Use such thresholds to avoid inflating benefits by counting low-severity anomalies.

4.5 Minimum viable QA/inspection program for policy and finance

Factory. In-line EL with instance-level localization, connected to *scrap/repair* decisions; periodic destructive sampling to anchor severity functions. Report defect rates with confidence intervals and a short measurement-system analysis [39, 40].
Field. Annual (or post-event) UAV IR for screening, followed by targeted EL on flagged strings for severity confirmation. Report ROC/PR by class, site-specific thresholds, and estimate prevalence to support positive predictive value (PPV) interpretation [37, 41].
Economics. Value programs with Eqs. (11)–(13); apply cost-effectiveness thresholds [43]; convert to ΔNPV and ΔLCOE using the finance layer (Ch. 2). Avoid double counting when EL confirms IR flags: one avoided loss stream, not two.

4.6 Limits and research needs

Calibration and environment. IR detectability shifts with irradiance and wind; enforce acquisition envelopes (Table 6) and document residual false-positive rates [37]. *Durability mapping.* Progression from micro-crack to inactive area is climate/load-dependent; paired EL/IR–I–V time series across fleets will tighten severity priors [33, 35]. *Warranty accounting.* Vendor warranties differ (power-in-warranty vs. workmanship). Integrating actuarial warranty models with inspection remains a standardization need despite available O&M cost frameworks (Ch. 2).

Bottom line. Peer-reviewed evidence shows (a) many anomalies are benign; (b) a measurable tail of high-severity defects drives most energy/warranty losses; (c) modern analytics find these efficiently when validated across plants. With prevalence-aware valuation, severity functions, and cost-sensitive thresholds, QA and inspection deliver defensible ΔNPV/ΔLCOE without benefit inflation—and give regulators/financiers a basis to require *evidence-based* inspection tied to economic action rules.

5 Device- and plant-level control (AI-MPPT; from tracking metrics to annual yield)

5.1 Policy motivation and scope

Device- and plant-level control is the point where algorithmic changes convert directly into delivered energy and operating cash flows. On the DC side, the central lever is maximum power point tracking (MPPT): under fast irradiance ramps and partial shading, better MPPT can reduce tracking losses; overly aggressive search can, however, increase stress or cause nuisance trips. In line with Chapter 2, credible claims must (i) specify a defensible baseline (e.g., well-tuned Perturb-and-Observe or Incremental Conductance), (ii) express outcomes as ΔMWh and ΔLCOE (not only bench metrics), and (iii) quantify uncertainty. This chapter synthesizes peer-reviewed evidence on *learning-based* MPPT (fuzzy/ANN/optimization controllers) and shows how *lab* tracking metrics map to *annual* energy once site-specific irradiance dynamics and shading statistics are introduced. We keep generalization conservative and within validated domains (Chs. 2–4).

5.2 What MPPT metrics mean — and what they do not

Surveys and comparative studies make two policy-relevant points. First, conventional MPPT (P&O, INC) is near-optimal under *smooth* irradiance; differences emerge mainly under *rapid ramps* and *multi-peak* I–V curves caused by partial shading or bypass-diode activation [44, 45]. Second, laboratory "tracking efficiency" depends strongly on the test profile; static ramps overstate field performance, while dynamic profiles (e.g., step changes, flicker) are more diagnostic [44]. Plants experience mixtures of regimes over a year, therefore bench metrics are *necessary but not sufficient*; they must be *convolved* with local irradiance variability and shading statistics to estimate ΔMWh (Sec. 5.4).

Typical dynamic test outputs include: (i) tracking efficiency η_{trk} over specified irradiance/temperature sequences; (ii) settling time to within a tolerance band; and (iii) overshoot/oscillation around the MPP. These affect DC energy and component stress differently and should not be collapsed into a single score when the policy question concerns both energy and reliability.

5.3 Evidence for learning-based MPPT under dynamic and shaded conditions

Peer-reviewed studies consistently report that learning-type controllers converge faster and are more robust to multi-peak curves than well-tuned conventional baselines, provided the converter bandwidth and sampling constraints match those in tests.

Global MPPT under partial shading. Comparative analyses (the classic synthesis by Esram & Chapman) clarify that the advantage appears when bypass-diode-induced multi-peaks are present and irradiance changes faster than the baseline step size [44]. Learning-based and metaheuristic controllers reduce mis-tracking probability in these cases, particularly when aided by pattern cues (voltage/current derivatives) [46].

Quantitative gains (bench scale). Cross-study reviews covering hundreds of runs report *tracking-efficiency* uplifts on dynamic profiles typically in the **0.5–3 percentage-point** range relative to tuned conventional methods; larger gains are sometimes observed in severely contrived shading, but are not general across sites [7, 8, 46].

Stability and oscillation. Learning-type controllers often show shorter settling times and lower oscillation around the MPP, reducing ohmic loss and local heating in power stages—engineering benefits that may be modest in annual MWh but matter for component reliability [7, 45].

Two caveats. First, some demonstrations use small DC–DC prototypes with fast sensing; when replicated on commercial inverters with slower sampling and duty-cycle slew limits, gains narrow [45]. Second, rule bases or training data must reflect *field* envelopes (module type, wiring, temperature), or domain shift will erode benefits (cf. Ch. 4).

5.4 From tracking metrics to annual energy: a regime-based mapping

We convert lab-reported tracking efficiency to annual energy using a *regime mixture* built from empirical irradiance and shading statistics.

Irradiance variability and ramps. High-frequency irradiance fluctuations arise from cloud edges and are smoothed by array footprint; ramp-rate distributions are climate- and layout-dependent [47, 48]. Sites with many fast ramps (coastal cumulus) offer more headroom for improved MPPT; quiescent sites offer less.

Partial-shading frequency and severity. Mismatch and multi-peak I–V curves arise from row-to-row shading, nearby obstructions, and soiling. Their energy impact is described by circuit/mismatch models and field experiments [49, 50]. Annual mismatch losses in modern utility plants are often ∼1–3%, but can be much higher in built environments; the share of time with bypass-diode activation determines the potential for global MPPT value [46].

Converter constraints. Converter bandwidth, quantization and sensor noise limit achievable η_{trk}. Bench results must be down-rated when commercial hardware is slower than the test rig [45].

Mapping. Let $r \in \mathcal{R}$ index operating regimes (e.g., smooth, moderate ramp, fast ramp; unshaded, mild partial shading, severe partial shading). With regime probabilities p_r and regime-specific tracking efficiencies $\eta_{\mathrm{trk},\cdot}(r)$, the expected annual energy gain of an AI controller over baseline b is

$$\mathbb{E}[\Delta E] \;=\; E_{\mathrm{DC,ideal}} \sum_{r\in\mathcal{R}} p_r \big(\eta_{\mathrm{trk,AI}}(r) - \eta_{\mathrm{trk},b}(r)\big), \tag{14}$$

where $E_{\mathrm{DC,ideal}}$ is the annual ideal DC energy (no tracking loss). The p_r should come from site data (SCADA irradiance, sky imagers) or literature-based priors for similar climates [47, 48]. Regimes with bypass-diode activation use mismatch-informed performance functions [49, 50]. Plugging Table 8 into Eq. (14) yields a *sub-percentage to low single-percentage* $\Delta E/E$ for many utility plants; rooftop sites with frequent complex shading will show larger deltas, but with larger uncertainty. Reporting a single headline number without site statistics is not policy-grade.

Uncertainty propagation and sensitivity. Uncertainties in p_r (climate year-to-year), $\eta_{\mathrm{trk}}(r)$ (test variability; hardware limits), and $E_{\mathrm{DC,ideal}}$ should be propagated with Monte Carlo and diagnosed with global sensitivity (as in Chapter 2). In most utility cases, the share of time in multi-peak regimes dominates the variance of ΔE.

Table 8: Illustrative regime mixture for a utility site (conceptual values). Use site data to estimate p_r. Report η_{trk} by regime with confidence intervals.

Regime r	**Share** p_r	$\eta_{\text{trk},b}(r)$	$\eta_{\text{trk,AI}}(r)$ **(bench)**
Smooth, unshaded	0.70	0.995	0.996
Moderate ramp, unshaded	0.20	0.985	0.990
Fast ramp, unshaded	0.05	0.970	0.985
Partial shading (mild)	0.04	multi-peak; 0.950	0.970
Partial shading (severe)	0.01	multi-peak; 0.900	0.950

5.5 Reliability, thermal stress, and protection interactions

Control interacts with hardware reliability and protections. Faster, oscillation-free tracking reduces ripple current and localized heating, aiding bypass diodes and interconnects. Conversely, aggressive global search near reverse bias can raise hot-spot risk in shaded substrings. Circuit-level and experimental studies show that repeated forcing of shaded substrings can accelerate degradation; global MPPT should therefore honor *hot-spot-aware* constraints derived from mismatch models [49, 50]. Any claimed energy gain that increases thermal risk should be discounted in ΔNPV via higher expected failure costs (Chapter 2).

5.6 Plant-level control beyond MPPT (scope boundaries)

AC-side control (Volt/Var, Volt/Watt, plant ramp-rate limits) shapes realized energy through clipping and curtailment. Short-horizon DC smoothing coordinated with MPPT in PV+storage can reduce system costs (valued in Chapter 6); to avoid double counting, we restrict Chapter 5 valuation to DC-side ΔMWh attributable to MPPT. System-level benefits from smoother injections are accounted for in Chapter 6.

5.7 Economic mapping and reporting template

We convert $\mathbb{E}[\Delta E]$ to finance with the transfer functions in Chapter 2. For energy price $p_{\text{el}}(t)$ and discount rate r:

$$\Delta\text{NPV} = \sum_t \frac{(\mathbb{E}[\Delta E_t] \cdot p_{\text{el}}(t)) - \Delta\text{O\&M}_t}{(1+r)^t}, \qquad \Delta\text{LCOE} \approx -\frac{\sum_t \mathbb{E}[\Delta E_t]/(1+r)^t}{\sum_t E_{\text{baseline},t}/(1+r)^t} \times \text{LCOE}_{\text{baseline}}. \tag{15}$$

Include uncertainty bands from the regime mixture and price scenarios. If controller changes increase failure risk or truck-rolls, include those cash flows explicitly (Chapter 2).

Bottom line. The peer-reviewed record supports cautious, context-dependent benefits from learning-based MPPT: measurable improvements under ramps and partial shading, narrowing to small ΔMWh in low-shading utility contexts and rising in complex rooftops. When mapped with site statistics and hardware limits—and when safety constraints are enforced—these gains translate into modest but defensible ΔLCOE reductions, consistent with the transfer-function discipline established earlier.

Table 9: Minimum reporting set for policy-grade MPPT evaluations. Items ensure reproducibility and guard against benefit inflation.

Item	Expectation
Baseline controller	Named, tuned P&O or INC; converter/inverter hardware specified
Dynamic tests	η_{trk}, settling time, oscillation on specified sequences; confidence intervals
Site statistics	Empirical ramp-rate and shading regime shares p_r with uncertainty (SCADA/sky imager)
Mismatch model	Method for multi-peak regimes (Bishop model or experimental calibration)
Hardware limits	Sampling rates, duty-cycle slew, sensor quantization/latency
Convolution	Equation (14) with tabulated p_r, $\eta_{\text{trk}}(r)$
Finance	Equation (15); ΔNPV and ΔLCOE with sensitivity to p_r and prices
Risk	Any hot-spot or protection interactions; failure cost treatment

6 Forecasting & power-system integration (market value, reserves, commitment)

6.1 Policy motivation and scope

Accurate photovoltaic (PV) forecasts affect costs wherever operators and market participants make *forward* decisions under uncertainty: unit commitment (UC), economic dispatch (ED), reserve procurement, and market bidding. Consistent with Chapter 2, we value forecasting improvements only through their operational consequences (production cost, curtailment, reserve procurement, start/stop costs, imbalance payments), not via technical error scores alone. Accordingly, we (i) separate *forecast quality* (deterministic and probabilistic, evaluated with strictly proper scores) from *decision value*; (ii) benchmark against counterfactual baselines aligned to current practice; and (iii) propagate uncertainty end-to-end through UC/ED or market models. We focus on solar-specific evidence and guidelines [51] and hybrid intra-day forecasting studies [52, 53], tying them to monetization within production-cost or market frameworks [9, 54, 55].

6.2 Forecast types, horizons, and proper evaluation

For intra-day horizons ($\sim$1–6 h), combining satellite nowcasts with numerical weather prediction (NWP) improves PV forecasts relative to single-source models; Bayesian neural networks and hybrid measurement–satellite–NWP schemes report significant, out-of-sample gains by horizon and weather regime [52, 53]. At day-ahead horizons, NWP ensembles with bias correction and spatio-temporal post-processing underpin most operational pipelines, with clear-sky and regime conditioning used to stabilize reliability [51].

Because forward decisions require *probabilistic* information, evaluation must use *strictly proper* scoring rules so that well-calibrated uncertainty is rewarded (Chapter 2). We use the continuous ranked probability score (CRPS) for marginal distributions and complement it with reliability diagrams and conditional skill by regime:

where F is the predictive CDF and y the realized value (see [15]). This aligns with solar-forecasting best practice and operational guidelines for reproducible, probabilistic evaluation [51].

Table 10: Horizon–decision map for PV forecasting and where value materializes. Evaluation must report *probabilistic* skill by horizon and regime, then embed forecasts in the corresponding decision model.

Forecast horizon	**Primary decisions affected**	**Notes on value pathway**
15–60 min (intra-hour)	AGC/regulation, short-term redispatch	Value through regulation deployment and ramping costs; requires calibrated ramp distributions
1–6 h (intra-day)	Redispatch, intra-day reserves, trading	Hybrid sat+NWP improves ramp prediction; scenarios must preserve temporal correlation [52, 53]
Day-ahead (12–36 h)	UC, day-ahead reserve sizing, day-ahead bids	Major cost levers: starts, minimum up/down, curtailment; value rises with penetration and binding constraints [9, 54]
Multi-day	Maintenance scheduling, storage scheduling	Lower marginal value for fast operations; relevant for planned outages and fuel logistics

6.3 Transfer functions from forecast quality to system value

Production cost and commitment. A solar-focused study embedded forecast-error reductions directly in a UC/ED model and showed that better day-ahead PV forecasts reduce fast-start commitments, ramping, and curtailment, thereby lowering variable system costs; value depends on penetration and which horizons improve [9]. More recent analyses [54] and reviews [10] emphasize that (i) value is application-specific (UC, reserve, trading), (ii) technical error metrics and economic outcomes must not be conflated, and (iii) stochastic or robust UC using *probabilistic* forecasts generally dominates deterministic UC when forecast errors are material [51]. Denote UC/ED cost under baseline forecasts as C_b and under improved probabilistic forecasts as $C_\star$; the production-cost delta is

$$\Delta C_{\text{prod}} \;=\; C_\star - C_b \;=\; \Delta C_{\text{energy}} + \Delta C_{\text{curt}} + \Delta C_{\text{starts}} + \Delta C_{\text{ramp}} + \Delta C_{\text{res}}, \tag{16}$$

computed by running the same UC/ED engine with the two forecast inputs.

Operating reserves from probabilistic forecasts. Operating-reserve needs should scale with the *distribution* of net-load forecast errors. Methods that set reserves using probabilistic forecasts—trading off reserve procurement costs against shortfall risk—are formalized in [56] and transfer to PV by substituting PV error distributions for wind. Let $q^{\uparrow}_{p^*}$ and $q^{\downarrow}_{p^*}$ be the up/down quantiles of net-load error at reliability p^* (e.g., 99%). Replacing a deterministic forecast with a calibrated distribution changes reserve targets and costs:

$$\Delta C_{\text{res}} \;=\; c^{\uparrow}\,(q^{\uparrow,\star}_{p^*} - q^{\uparrow,b}_{p^*}) \;+\; c^{\downarrow}\,(q^{\downarrow,\star}_{p^*} - q^{\downarrow,b}_{p^*}), \tag{17}$$

with zonal reserve prices $c^{\uparrow,\downarrow}$. Studies that implement this *inside* UC/ED, rather than off-line, provide the correct monetization basis [51, 54].

Market participation and imbalance. In liberalized markets, plant owners translate forecast improvements into bids. Using *probabilistic* day-ahead forecasts to generate trading scenarios increases revenues and reduces imbalances relative to deterministic bidding, with gains modulated by lead-time and the availability of both day-ahead and intra-day markets [55]. The

general principle, established earlier for wind, is that scenario-based bidding improves risk–return by aligning offer curves with distributional information and market penalty structures [57]. Let $R(\mathbf{q})$ be expected revenue under scenario-based offers $\mathbf{q}$ and ϕ the imbalance penalty function; an illustrative objective is

$$\max_{\mathbf{q}} \ \mathbb{E}[R(\mathbf{q})] \ - \ \lambda \, \mathbb{E}[\phi(\mathbf{q}, \tilde{P})], \tag{18}$$

where $\tilde{P}$ are forecast scenarios and λ encodes risk preference or regulatory penalty weights.

6.4 Where value comes from—and where it does not

Penetration and flexibility. The marginal value of accuracy is nonlinear in PV penetration and available flexibility (fast-start units, storage, transmission). Value rises with penetration until flexibility saturates, after which additional accuracy yields diminishing returns [9, 51, 54]. *Timescale alignment.* Day-ahead skill mainly affects commitment and day-ahead reserve; 5–60-min forecasts affect regulation and intra-day corrections. Improving a non-binding horizon has little value [51]. *Metric alignment.* Proper scores (e.g., CRPS) are necessary but not sufficient; value arises only after mapping to UC/ED/bidding decisions (Chapter 2). Reviews caution against inferring economic value directly from MAPE/RMSE [10, 51]. *System effects and double counting.* DC-side control improvements (Chapter 5) should not be double counted here; system-level benefits from smoother injections belong in the present chapter once represented inside the system model.

6.5 Data and modelling pipeline for policy-grade valuation

Forecast generation. Use hybrid measurement–satellite–NWP pipelines with documented reliability and out-of-sample skill by horizon and regime [51–53].
Scenario building. Convert predictive distributions into time-consistent scenarios; scenario reduction must preserve tails because reserve shortfalls are asymmetric [55, 57].
System model. Embed forecasts directly in UC/ED so that commitment, reserve, and redispatch costs respond endogenously [9, 54].
Uncertainty. Report value distributions across weather years and load/price conditions; run sensitivity on reserve prices, penalty factors, storage/ramping limits (Chapter 2).

6.6 Minimum reporting set and guardrails

The items below define a *minimum* reporting set for policy-grade PV-forecast valuation and the guardrails needed to keep estimates unbiased and reproducible. The purpose is not to prescribe one UC/ED or market model, but to ensure that any claimed benefit can be independently re-run and traced from probabilistic forecast skill to operational value (Chapter 2). Three principles guide the requirements are following:
(1) Align the counterfactual. A credible baseline reflects legacy practice *for the same assets, periods, markets, and rules.* That includes: the incumbent forecast(s); reserve heuristics (e.g., fixed % of load/PV); imbalance settlement rules; and commitment/redispatch policies. Changes to topology, must-run constraints, or price data must be applied symmetrically to baseline and improved cases. Report any operator overrides or manual curtailments; do not average across non-comparable weather years without stratifying.
(2) Separate forecast quality from decision value. Forecasts must be evaluated out-of-sample with *strictly proper* scores (e.g., CRPS) and reliability diagrams by horizon and weather regime (cf. Sec. 6.2). Those distributions then feed the decision model: either (i) stochastic/robust UC/ED using scenarios drawn from the calibrated predictive distribution, or (ii)

reserve-sizing rules that explicitly reference quantiles/coverage. Scenario generation must preserve temporal correlation and ramps; any reduction method must document retained tail risk because reserve shortfalls are asymmetric (Sec. 6.3). Avoid leakage: forbid using the valuation period for hyper-parameter tuning; prefer rolling-origin splits or blocked bootstraps for CIs.
(3) Make the system model the locus of monetization. Value accrues only once forecasts interact with unit constraints, reserve prices, and imbalance penalties *inside the model*. Use the same UC/ED or market-clearing engine, data interfaces, and solver settings for baseline and improved runs, changing only the forecast input. Report Δ production cost, curtailment, reserves, starts/ramping, and (where applicable) revenues—each with uncertainty bands from multi-year weather and price conditions. Provide a small set of sensitivities that matter for policy decisions: PV penetration, flexibility (storage/transmission and ramp limits), reserve/imbalance price levels, and commitment penalties. When DC-side controls (e.g., MPPT changes; Ch. 5) are also claimed, use *nested* baselines and document the order of application to avoid double counting; if integration costs change materially, report system LCOE.
Guardrails. Pre-specify the evaluation window and reporting metrics; disclose excluded periods (outages, data gaps) and justify exclusions symmetrically. Publish data/artefact manifests and seeds to enable independent reruns (Appendix A.6). If cybersecurity prerequisites for telemetry or control are not met, flag the valuation as conditional per Appendix A.5. The table that follows operationalizes these principles into concrete checklist items for authors and reviewers.

Table 11: Minimum reporting for policy-grade PV forecasting valuation. These items enable reproducibility and prevent benefit inflation.

Item	**Expectation**
Baseline practice	Legacy forecast(s) and operating rules (UC, reserve %, imbalance penalties)
Forecast skill	Proper scores by horizon and regime (CRPS; reliability); independent test periods
Scenario methodology	Scenario generation and reduction; spatio-temporal coherence demonstrated
System model	UC/ED or market-clearing formulation; same engine for baseline and improved case
Valuation outputs	Δ production cost, Δ curtailment, Δ reserve, Δ starts, Δ revenues with confidence intervals
Sensitivity	Penetration, flexibility (storage/transmission), reserve prices, penalties; multi-year weather
Accounting	No double counting with MPPT (Ch. 5); system LCOE use when integration costs change

Bottom line. The strongest, solar-specific evidence shows that *probabilistic*, properly evaluated PV forecasts deliver measurable system value when embedded in UC/ED, reserve sizing, and market bidding. Gains depend on penetration, flexibility, and timescale alignment; they should be monetized only through production-cost and market models with transparent uncertainty, consistent with Chapter 2 [9, 51, 54–57].

7 Soft-costs & adoption (marketing, permitting, siting; distributional effects)

7.1 Policy motivation and scope

Hardware learning curves have lowered module and inverter costs, but *soft-costs*—customer acquisition, design, permitting–inspection–interconnection (PII), and financing—remain material, especially for rooftop systems. Because these costs are paid up front, their impact on levelized cost of energy (LCOE) is magnified through discounting (Chapter 2). This chapter quantifies three levers where data-driven methods demonstrably affect cash flows: (i) *marketing* (customer acquisition), (ii) *PII streamlining* (cycle-time compression), and (iii) *siting* (reduced schedule and mitigation risk). We also evaluate *distributional effects* to ensure benefits are equitably shared. Only peer-reviewed studies and high-level national reports with transparent methods are used, and all effects are mapped to ΔNPV/ΔLCOE via the transfer functions in Chapter 2.

7.2 Marketing and customer acquisition

Customer acquisition costs (CAC) are a persistent component of residential PV soft-costs. A prospective, peer-reviewed study shows that supervised targeting reduced CAC by $\sim 15\%$ ($\approx \$0.07$/W) compared with a logistic-regression baseline, using strict out-of-sample evaluation to avoid optimistic leakage [11]. Adoption also responds to *social diffusion*: instrumented microdata reveal causal *peer effects*—local installation propensity rises with nearby adoptions after controlling for demographics and prices [58]. Incentive levels remain first-order drivers; a quasi-experimental evaluation of the California Solar Initiative measured sizable, causal installation responses to rebate variation [59].
We model the marketing lever as a direct CapEx reduction:

$$\Delta\text{CapEx/W} \;=\; -\,\Delta\text{CAC/W}, \tag{19}$$

bounded by [11] and adjusted for jurisdictional conversion and attrition rates (propagated via Monte Carlo; Chapter 2). Because CAC reductions can alter the *composition* of adopters, we report results disaggregated by income and race/ethnicity cohorts (Section 7.5).

7.3 Permitting, inspection, and interconnection (PII)

What the evidence shows. Large-sample, peer-reviewed studies measuring real PII timelines find typical durations from permit application to passed inspection of $\sim 25{-}100$ days with wide dispersion across jurisdictions [60]. Local permitting *stringency* and interconnection requirements are robustly associated with longer durations and higher installed prices [61, 62]. High-level programme evaluations of *SolarAPP+*—an online, automated code-compliance check—report median permitting+inspection timelines roughly 14.5 business days shorter than traditional processes (2023 data; methods and coverage disclosed) [63]. These results are consistent with bottom-up national benchmarks and longitudinal installed-price tracking [64, 65].

From days to dollars. We value PII acceleration via three channels: *time value of energy*, *lower overhead burn*, and *fewer cancellations*. Let t_b and $t_\star$ be baseline and improved median timelines; let $\Delta t = t_b - t_\star$. For a project producing E_1 MWh in its first operational year and price profile $p(t)$,

$$\Delta\text{NPV}_{\text{timing}} \;=\; \int_0^{\Delta t} \frac{p(t)\,E_1}{(1+r)^{t/365}}\,\mathrm{d}t \quad\text{and}\quad \Delta\text{NPV}_{\text{overhead}} \;=\; \text{OH/day} \times \Delta t, \tag{20}$$

Table 12: PII timelines and monetization channels. Values are indicative; jurisdiction-specific inputs should be taken from [60, 62, 63].

Quantity	**Empirical range / finding**	**Monetization pathway**
Permit→inspection duration	$\sim$ 25−100 days median; wide dispersion [60]	Time value of energy; overhead burn; cancellations
Local stringency / interconnection	Higher stringency → longer duration and higher prices [61, 62]	Same as above; potential fee offsets
Automated compliance (SolarAPP+)	Median −14.5 business days vs. legacy (2023) [63]	Δt enters timing and overhead channels; possible admin fees

with r the real WACC and OH/day documented overhead. If the cancellation probability drops from q_b to $q_\star$, avoided sunk costs per closed project are

$$\Delta\text{NPV}_{\text{cancel}} \;=\; (q_b - q_\star)\,(\text{CAC} + \text{design}), \tag{21}$$

where design are pre-permit labor and third-party fees. All three deltas are added (net of any programme fees) in the project cash flow. Uncertainty enters through Δt distributions (from empirical CDFs in [60, 62, 63]) and through $q_b, q_\star$ elasticities, and is propagated by Monte Carlo with variance-based sensitivity analysis (Chapter 2).

7.4 Soft-cost structure and external validation

National benchmarks and longitudinal price datasets provide defensible priors for soft-cost shares and for counterfactuals. The Q1-2023 NREL benchmark details CAC and PII inputs by segment in a documented bottom-up model [64]. LBNL's *Tracking the Sun 2024* aggregates 3.2 million systems through 2023 and reports installed-price distributions and design trends by state and market segment [65]. We use these sources to anchor baseline cash flows and to cross-check claimed savings against historical ranges.

7.5 Distributional effects and equity

Inequitable deployment is well documented. A *Nature Sustainability* study shows that majority Black and Hispanic tracts have substantially lower rooftop PV adoption than otherwise comparable tracts even after conditioning on income and homeownership [66]; revised estimates confirm that disparities persist under alternative specifications [67]. Two policy guardrails follow. First, marketing models must demonstrate that targeting *does not* exacerbate gaps (e.g., by under-selecting disadvantaged tracts). Second, permitting reforms should be prioritized where delays are longest and adoption lags are largest, to avoid regressive time costs.

7.6 Siting and community context (utility-scale and DG)

For utility-scale PV, siting choices generate cost and schedule risk through land acquisition, environmental review, and mitigation. A *PNAS* analysis quantifies land-cover change from U.S. solar development and highlights *low-conflict* siting as a means to reduce delays and soft-costs associated with mitigation [68]. In our accounting, siting affects ΔNPV via longer development timelines (carry costs) and via mitigation outlays:

$$\Delta\text{NPV}_{\text{siting}} \;=\; -\,\text{Carry} \times \Delta\text{months} \;-\; \text{Mitigation}, \tag{22}$$

with inputs drawn from project pro formas and mitigations documented during environmental review. For distributed PV, setback and aesthetic rules in zoning create variance in approval outcomes; the price and timeline effects documented for permitting recur when samples are sufficiently large [61, 62].

Table 13: Equity guardrails for analytics and PII reforms. Each item has a measurable diagnostic.

Guardrail	Diagnostic	Evidence basis
Targeting parity	Selection rates by income and race/ethnicity vs. baseline	[66, 67]
Outcome parity	Close rates and CAC by cohort; confidence intervals	[11, 58]
Geographic prioritization	Permit duration distribution by jurisdiction; focus longest tails	[60, 62, 63]
Transparency	Publish model features, validation, and cohort metrics	Benchmarks: [64, 65]

7.7 Integrating the levers: a soft-cost impact model

We assemble a conservative model with four levers, valued *incrementally* (nested baselines to avoid double counting; Chapter 2):

1. **Marketing:** $\Delta\text{CapEx/W} = -\Delta\text{CAC/W}$ bounded by verified out-of-sample performance [11], with cohort-level reporting (Section 7.5).
2. **Permitting:** $\Delta\text{NPV} = \Delta\text{NPV}_{\text{timing}} + \Delta\text{NPV}_{\text{overhead}} + \Delta\text{NPV}_{\text{cancel}}$, using Δt from [60, 62, 63].
3. **Interconnection:** identical structure, using utility interconnection queue metrics from [62].
4. **Siting:** $\Delta\text{NPV}_{\text{siting}}$ from carry and mitigation terms informed by [68].

Uncertainty bands are reported from Monte Carlo sampling of CAC variance, cancellation elasticity, and timeline dispersion; Sobol indices identify dominant drivers (Chapter 2). When multiple levers act on cancellations (e.g., better targeting and faster permits), we value each *conditional* on the other.

Bottom line. Peer-reviewed evidence supports real, quantifiable soft-cost reductions from (i) data-driven marketing that lowers CAC and (ii) streamlined permitting/interconnection that shortens timelines. These improvements map transparently to $\Delta\text{NPV}/\Delta\text{LCOE}$ under our framework and can be monitored with public benchmarks. Distributional analyses demonstrate persistent adoption gaps, setting guardrails for how analytics are deployed and where streamlining is prioritized. Siting choices at utility scale add a third lever by reducing schedule and mitigation risk. Policy-grade appraisals should therefore pair measured process improvements with equity metrics and publish incremental $\Delta\text{NPV}/\Delta\text{LCOE}$ with uncertainty bounds rather than headline averages.

8 Case studies & meta-analysis (evidence synthesis across climates/markets)

8.1 Purpose and scope

This chapter integrates *quantified* impacts of AI-enabled tools across diverse climates and market structures, tying each to the valuation framework in Chapter 2. We consider (i) high-penetration

systems with binding flexibility constraints (California ISO), (ii) mature European markets (Germany and North-West Europe) with deep intraday liquidity and dense distributed PV, (iii) merchant-style environments where forecast-aware bidding affects plant revenues, and (iv) settings where structural curtailment dominates (Japan, Northwest China). For comparability, we map reported effects into Δproduction cost, Δcurtailment, Δreserve, Δrevenues, or Δsoft-costs and then to ΔNPV or ΔLCOE using the transfer functions of Chapter 2. Throughout, we avoid overclaim by keeping extrapolations within the empirical domains of the cited studies.

8.2 California ISO (CAISO): curtailment context and forecast value

System context. CAISO's PV share produces the well-documented "duck curve" with seasonal midday curtailment when flexibility or export limits bind. A field guide quantifies overgeneration mechanisms and documents that curtailment falls when mid-day flexibility and transmission increase—reporting marginal curtailment reduction costs for specific levers rather than generic savings [69]. These structural levers bound the *realizable* value of improved PV forecasts.

AI-linked pathway. Embedded in a unit-commitment/economic-dispatch (UC/ED) model, better day-ahead PV forecasts reduce fast-start commitments, ramping, and redispatch cost [9]. Under CAISO-like penetration and flexibility, the incremental Δproduction-cost credited to forecasting must be computed *conditional* on any curtailment-mitigation already modeled (nested baselines). On the distributed side, permitting acceleration programs shorten interconnection timelines, raising project NPV through earlier revenue recognition; large-jurisdiction panels and national program reviews quantify typical time shifts [60, 62–64]. Given coastal cumulus and frequent ramps, bench MPPT gains translate to sub-percentage annual ΔMWh at utility scale (Chapter 5), while rooftop sites with complex shading show larger headroom (Chs. 4, 5).

8.3 Germany and North-West Europe: preserving market value and lowering imbalance

Market value lens. With rising VRE shares, the *relative price* of solar (market value factor, MVF) declines due to profile and forecast uncertainty [70]. Forward-looking scenarios for Germany show MVF dispersion depends on transmission and diversification—structural complements to any algorithmic improvement [71]. Comparative work explains why distributed soft-costs are lower in Germany than in the U.S. (permitting, market organization), cautioning analysts not to attribute cross-country price gaps to hardware learning alone [72].

Forecast economics. German evidence ties *statistical* properties of forecasts to *economic* outcomes: improving the *correlation* between forecast errors and day-ahead–real-time price spreads can yield higher trading value than equal MAE reductions that leave correlation unchanged [73]. This directly supports the Chapter 6 prescription to optimize for value-oriented calibration, not RMSE alone.

8.4 North-West Europe: probabilistic bidding and merchant revenues

In a liberalized North-West European market, using *probabilistic* PV day-ahead forecasts to generate trading scenarios increases expected revenues and reduces imbalances relative to deterministic bidding; gains depend on lead-time and coordinated use of day-ahead and intraday markets [55]. These results operationalize the Chapter 6 guidance: convert calibrated predictive distributions into coherent scenarios and embed them in the actual bidding decision.

8.5 Australia (NEM): PV+storage operation benefits from forecast calibration

For PV+storage in the Australian NEM, improved day-ahead PV forecasts increase arbitrage profits; the effect scales with price volatility and storage sizing [74]. The economic channel again runs through better alignment between forecast errors and price spreads, reinforcing the value-oriented objective rather than error shrinkage per se.

8.6 Japan (Kyushu): measured curtailment as a bound on forecast value

Reproducible analysis of Kyushu curtailment episodes shows how system constraints generate observed curtailment patterns [75]. Even perfect forecasts cannot fully eliminate structural curtailment; the credible AI role is *anticipation and pricing* of curtailment risk for storage and demand scheduling, valued through UC/ED or merchant bidding (Ch. 6).

8.7 China (Northwest provinces): structural curtailment and priority levers

A systematic review documents curtailment magnitudes and drivers (transmission bottlenecks, dispatch protocols) in Northwest China and evaluates remedies (grid expansion, flexibility procurement, market reform) [76]. AI tools complement, but do not substitute, for these first-order actions: their role is to forecast, price, and mitigate residual curtailment in operations.

8.8 Spain (MIBEL): cannibalization and remuneration under rising PV shares

With rising PV penetration, Spain exhibits statistically significant *cannibalization* (lower unit revenues for PV) and broader *depredation* effects on other technologies [77]. Forecast-aware bidding and storage control can partially mitigate revenue erosion by shifting net injections away from low-price hours, but the structural trend persists, consistent with MVF evidence [71, 78].

8.9 Pan-European/Germany: imbalance costs and price formation under forecast error

German quarter-hourly data show that wind/solar forecast errors raise imbalance volumes and can push up spot prices; wind dominates but solar contributes materially [79]. The causal channel is directly addressable by AI: reducing error *magnitudes* and managing error–spread *correlations* lowers balancing costs systemwide (Ch. 6).

8.10 Climate and device-side heterogeneity: yield, degradation, and inspection

Performance ratio shifts with climate. Temperature coefficients and spectral effects vary by climate and module type; validated models map these to location-specific performance ratios [80, 81]. These baselines are prerequisites for converting device-side improvements (MPPT, inspection) into ΔMWh.

Degradation and defect progression. Field compendia provide distributions of annual degradation for c-Si modules [12, 13], while EL/IR studies show that progression from microcracks to inactive area depends on climate and load [33, 35, 36]. Inspection value therefore scales with the local *severity distribution* and with plant-to-plant classifier robustness (domain shift) [41]. Economic action thresholds (e.g., replacement justified at $\sim$10–33% annualized loss under European cost structures) prevent inflation from low-severity anomalies [43].

Table 14: Representative, peer-reviewed case studies by climate/market; AI-linked levers; reported metrics; and valuation hooks. Effects must be credited incrementally relative to the local baseline (Chapter 2).

Setting	**AI-relevant lever**	**Reported metric(s)**	**Valuation hook**
CAISO	Day-ahead PV forecasts	↓ fast-starts, ramping, redispatch [9]; curtailment drivers [69]	Δproduction cost from UC/ED; curtailment nested baseline
Germany	Value-oriented forecast post-processing	Profit impact dominated by error–spread correlation [73]	Optimize for value (price-conditional loss)
NW Europe	Probabilistic bidding	↑ revenues; ↓ imbalances (DA/ID) [55]	Δrevenues, Δimbalance fees
Australia (NEM)	PV+storage operation	↑ arbitrage profit with better DA forecasts [74]	Δrevenues under volatility; storage sizing
Japan (Kyushu)	Curtailment modeling	Reproduced curtailment magnitudes [75]	Credible upper bound on forecast value w/o flexibility
China (NW)	Structural curtailment	Shares, causes, remedies [76]	Policy levers (transmission, market reform) first; AI for mitigation
Spain (MIBEL)	Revenue erosion	Cannibalization/depredation quantified [77]	Scenario-based dispatch/bidding; partial mitigation
Pan-EU/Germany	Imbalance impacts	Forecast-error → imbalance and spot prices [79]	Δbalancing cost via error reduction/calibration
USA multi-ISO	MVF dispersion	Spatial/temporal MVF variation [78]	Siting/contracting to favorable tails

8.11 Evidence table and valuation mapping

Transfer functions (reminder). Consistent with Chapter 2, we value forecasting/integration levers by embedding predictive distributions in UC/ED or bidding:

$$\Delta\text{NPV} = \sum_t \frac{\Delta\text{CashFlow}_t}{(1+r)^t}, \quad \Delta\text{CashFlow}_t \in \{\Delta\text{ProdCost},\ \Delta\text{Reserve},\ \Delta\text{Imbalance},\ \Delta\text{Revenue}\}.$$

Reserve sizing with probabilistic forecasts follows established approaches (developed for wind, transferable to PV) and must be run endogenously in UC/ED (Chapter 6) to avoid off-line misvaluation.

8.12 Cross-cutting lessons

1. **Optimize for value, not only accuracy.** Economic impact is driven by alignment of forecast errors with price spreads; calibration and loss functions should target this relationship [55, 73, 74].

2. **Probabilistic forecasts & multistage use.** Scenario-based bidding and stochastic

UC/ED monetize uncertainty honestly and outperform point-forecast strategies when errors are material (Chapter 6).

3. **Structure dominates the ceiling.** Curtailment and cannibalization are structural; AI realizes the most value when combined with flexibility, transmission, and market designs that reward risk-aware operations [69, 71, 76, 77].

4. **Geography is a lever.** Large MVF dispersion implies siting/contracting and storage-coupled operation are complementary to better models [78].

5. **Climate and device-side heterogeneity matter.** Temperature/spectral baselines and defect progression priors condition the payoff to MPPT and inspection (Chs. 4, 5); use severity distributions and action thresholds to prevent benefit inflation [33, 35, 43, 80, 81].

Bottom line. Across climates and markets, peer-reviewed evidence shows that AI-enabled forecasting and risk-aware scheduling *do* reduce system costs and improve revenues when flexibility and liquidity exist; their system-wide ceiling is set by structural constraints. The most robust benefits come from probabilistic, value-oriented pipelines embedded in UC/ED and bidding, with incremental crediting against explicit baselines.

9 Risk, governance, and standards (model risk, bias, cybersecurity, auditability)

9.1 Motivation and scope

The gains quantified in Chapters 3–**??** are fragile if the supporting models fail under distribution shift, embed bias, or widen cyber-attack surfaces. This chapter specifies a governance layer that: (i) makes *model risk* measurable with proper, decision-relevant metrics; (ii) treats *fairness* in marketing, credit screening, and siting as a first-class risk; (iii) aligns device/plant *cybersecurity* with sector standards; and (iv) ensures *auditability* so that benefits in ΔNPV/ΔLCOE can be verified and not double counted. We draw on NIST's AI Risk Management Framework [82], documentation methods (Model Cards, Datasheets) [83, 84], DER and industrial control standards (IEEE 1547/1547.3; ISA/IEC 62443; NERC CIP) [85–89], probabilistic evaluation guidance [15, 90–92], adversarial/false-data risk in energy time series [93–95], and PV-specific cyber surveys/testbeds [96–100].

9.2 Model risk: fit-for-purpose evaluation under shift

Score forecasts properly, then map to decisions. Forecasts and classifiers that feed UC/ED, reserve sizing, bidding, MPPT, inspection triage, or soft-cost workflows should be evaluated with *strictly proper* scoring rules (e.g., CRPS, logarithmic score) so that better scores imply better probability statements [15, 92]. Scores alone are not value; we *map* them to Δproduction cost, Δreserve, and Δrevenues inside the operational model (Chapter 6), consistent with ROPES and benchmark guidance for solar forecasting [90, 91].

Reproducibility and baselines. Make improvements falsifiable: publish (or escrow) code/data splits; compare against transparent, *probabilistic* baselines; report reliability and sharpness by horizon and weather regime [90–92]. This aligns directly with NIST AI RMF functions (Map, Measure, Manage, Govern) [82].

Table 15: Model-risk checklist: align failure modes with measurable diagnostics and decision impact.

Risk	Diagnostic / metric	Decision impact & control
Miscalibration	Reliability diagrams; CRPS decomposition [92]	Refit post-processing; withhold value until reliability restored
Distribution shift	Interval coverage by regime; pre/post skill [90]	Canary deploy; restrict use outside validated envelope
Adversarial/FDI	Feature integrity tests; perturbation sensitivity [93, 95]	Input authentication; plausibility bounds; alerting
Data leakage	Temporal blocking; site holdout [90]	Re-split; void prior gains; re-estimate value deltas

Distribution shift and adversarial risk. Fielded PV models face nonstationarity (weather, hardware, markets) and adversarial manipulation risks. Peer-reviewed studies show targeted perturbations can steer time-series predictors, and energy pipelines can be impacted if attacks align with market penalties [93–95]. Minimum controls: (i) *coverage checks* for prediction intervals in new regimes; (ii) *stress backtests* on extreme years; (iii) feature-pipeline *red-team* tests before deployment; (iv) circuit breakers on abnormal outputs. Document failure modes and rollback plans per AI RMF [82].

9.3 Bias and fairness (soft-costs and siting)

Adoption is uneven across U.S. census tracts even after conditioning on income/homeownership [66, 67]. Marketing, credit screening, and siting analytics should therefore be documented with *Model Cards* and *Datasheets* that disclose intended use, data provenance, subgroup performance, and limits [83, 84]. Treat fairness as model risk: define monitored cohorts, set minimum recall/PPV by cohort, evaluate prospectively on new jurisdictions, and report equity outcomes alongside ΔLCOE. This satisfies AI RMF expectations for harm mitigation and measurement [82].

9.4 Cybersecurity: device, plant, and market interfaces

Why PV is exposed. Networked inverters, gateways, and DERMS/SCADA create realistic attack surfaces (weak authentication, remote updates, legacy protocols). NIST's inverter guidance details practical mitigations for residential/light-commercial systems [96]; PV-specific surveys show known incident paths and stress timely patching and secure comms [97].

Standards to anchor controls. Use a layered stack:

- **IEEE 1547-2018** for DER interconnection/interoperability [85] plus **IEEE 1547.3-2023** for DER cybersecurity (identity, logging, updates, comms hygiene) [86].
- **ISA/IEC 62443** to structure zones & conduits and component requirements in plants and factories [87].
- **NERC CIP** when assets fall under bulk-system scope: Electronic Security Perimeters (CIP-005-6) and incident response (CIP-008-6) [88, 89].
- **DOE/NREL DER** cybersecurity standards overview for current best practice [100].

Table 16: Minimum cybersecurity controls mapped to standards and verifiable evidence.

Control	Standards anchor	Evidence to retain
Network segmentation & allowlists	ISA/IEC 62443 [87]	Zone/conduit diagrams; firewall rules; VLAN configs
Authenticated & encrypted control	IEEE 1547.3 [86]	TLS configs; cert inventories; mutual-auth test logs
Signed updates & rollback	IEEE 1547.3 [86]	Firmware SBOM; signature policies; update audit logs
Security perimeter & remote access	NERC CIP-005-6 [88]	ESP diagrams; jump-host configs; access logs
Incident response drills	NERC CIP-008-6 [89]	Runbooks; tabletop reports; ticketing evidence

PV-specific evidence and testbeds. Industrial device auditing toolkits and DER testbeds show how to harden protocol stacks (e.g., wrapping SunSpec/Modbus in TLS, enforcing mutual authentication), consistent with 1547.3 [86, 98, 99].

Operationalizing controls. For utility-scale PV and DER portfolios: segment networks (62443), authenticate and encrypt control channels (IEEE 2030.5, DNP3-SA, or SunSpec/Modbus over TLS where supported), enforce signed firmware with rollback protection (1547.3), throttle/alert unusual command sequences, and drill incident response per CIP-008-6 [86, 87, 89, 96].

9.5 Auditability and documentation

Risk governance depends on traceable artifacts. Model Cards and Datasheets make ML components auditable [83, 84]; ROPES and probabilistic benchmarks ensure that claimed skill rests on publishable baselines and proper scores [90–92]. IEEE 1547.3 calls for security logging and asset inventories; combined with change-management records, these support warranty claims, O&M KPIs, and post-incident forensics [86, 96]. For policy-grade claims in Chapters 4–**??**, we require: versioned model artifacts; data lineage; performance by climate/segment and subgroup; cybersecurity bills of materials (SBOM) and update cadence; and links from model outputs to operational decisions and cash-flow deltas. This is consistent with AI RMF [82].

9.6 Standards mapping to PV use-cases

Use-case	Primary controls/standards	Key evidence
UC/ED & bidding (Chs. 6,**??**)	Proper scoring; ROPES; benchmarks [15, 90–92]	Versioned datasets; CRPS/log-score with CIs; scenario generation trace
Manufacturing QA/inspection (Ch. 4)	IEC 62443 [87]	Supplier questionnaires; network segmentation; signed tooling
Device/plant control & DER agg. (Ch. 5)	IEEE 1547/1547.3; NIST inverter guidance [85, 86, 96]	TLS/mutual auth; RBAC; update logs; event logs
Portfolio ops/markets	Adversarial/shift testing; incident response [89, 93–95]	Stress results; red-team reports; CIP-008-style runbooks

9.7 Governance playbook: minimum viable controls

1. **Model risk management.** Use proper scores; test under shift; publish Model Cards; adopt change control (canary deploy, rollback); keep audit trails [15, 82–84, 90–92].

2. **Fairness by design.** Define monitored cohorts; report subgroup performance; document coverage; add use constraints when transportability is weak; monitor equity outcomes [66, 67, 83, 84].

3. **Cybersecurity by design.** Procure equipment with authenticated comms and signed updates; segment networks; enable monitoring and incident response aligned with 1547.3/62443/CIP; schedule third-party assessments [86–89, 96–100].

4. **Assurance and accountability.** Tie every claimed benefit to traceable evidence: baseline, test protocol, and logs sufficient for third-party replication (Chapter 2).

Bottom line. The same discipline that converts forecast skill into system value (Ch. 6) must also govern *risk*. Proper probabilistic evaluation, transparent documentation, and sector standards (IEEE 1547/1547.3, IEC 62443, NERC CIP, NIST) provide actionable guardrails. With these in place, AI benefits can be realized without amplifying bias or cyber exposure—and with an audit trail that withstands regulatory and financial due diligence [15, 82–100].

10 Policy design tools (NPV/LCOE deltas, cost–benefit under uncertainty, open datasets)

10.1 Motivation and scope

Chapters 3–9 established *where* AI can create value along the PV chain and how to prevent over-claiming. Policy design now needs transparent tools that convert measured technical or process changes into cash-flow deltas and social value—while disclosing uncertainty and avoiding double counting (Chapter 2). This chapter operationalizes that requirement with: (i) consistent *project* and *system* formulations for ΔNPV and ΔLCOE, (ii) cost–benefit and *value-of-forecast* appraisals under uncertainty, and (iii) open, auditable data/models to make results reproducible.

10.2 ΔNPV and ΔLCOE: definitions, system context, and reporting

Project-level accounting. For any intervention X (e.g., improved EL/IR triage, AI-MPPT, probabilistic market scheduling, lower CAC), report the incremental present value versus a named baseline b:

$$\Delta\mathrm{NPV}(X;b) = \sum_{t=0}^{T} \frac{\Delta\mathrm{CashFlow}_t(X;b)}{(1+r)^t}, \tag{23}$$

and the induced change in plant LCOE:

$$\Delta\mathrm{LCOE}(X;b) = \frac{\Delta\mathrm{Cost}(X;b)}{\mathrm{MWh}_b} - \frac{\mathrm{Cost}_b \cdot \Delta\mathrm{MWh}(X;b)}{\mathrm{MWh}_b^2}. \tag{24}$$

Equation (24) keeps the *denominator* anchored to the baseline MWh_b, which makes small energy changes explicit and avoids hidden re-scaling.

System-LCOE and value-adjusted metrics. Where integration effects dominate (balancing, reserves, curtailment, profile), plant-level metrics should be supplemented by *system LCOE* so that costs attributable to variability and uncertainty are consistently allocated across technologies [16]. When markets price energy imperfectly, also report a *market value factor* (relative price capture) or a value-adjusted LCOE, but keep plant cash-flows and system metrics separate to avoid double counting [16, 70].

Where deltas originate.

- **Operations (Chs. 4–6):** translate inspection/forecast/control changes into ΔMWh, Δcurtailment, Δreserve, Δstarts, Δimbalance *inside* a UC/ED or bidding model before discounting [9, 10, 54, 55, 92].
- **Soft-costs (Ch. 7):** embed ΔCAC, Δpermitting time, and Δoverheads directly in the CapEx/OpEx lines of the pro forma; monetize timeline shifts via earlier energization and lower cancellations (Section 10.3).

10.3 Cost–benefit under uncertainty: Monte Carlo, global sensitivity, and decision value

Uncertainty propagation. Use Monte Carlo (MC) to propagate parameter uncertainties (detector ROC, degradation rates, reserve prices, CAC variance) and variance-based global sensitivity (Sobol first- and total-effect indices) to attribute output variance to inputs and interactions—standard practice in energy techno-economics [19, 20]. Report MC convergence (e.g., effective sample size), and plot tornado charts for ΔNPV and ΔLCOE drivers.

Proper scoring → decision value. For forecasts, score probabilistic predictions with strictly proper rules (CRPS, logarithmic score) and *then* map improvements into production-cost or revenue changes via UC/ED or market bidding. Economic value depends on both error magnitude and the *correlation of forecast errors with price spreads*; calibrated, price-aware post-processing can outperform equal RMSE reductions that ignore this channel [10, 54, 55, 73, 92].

Value of improved information (VoI). Compute the *operational* value of better forecasts/inspection within the same system model:

$$\mathrm{VoI} = \mathrm{Cost}_{\mathrm{system}}(\mathrm{legacy}) - \mathrm{Cost}_{\mathrm{system}}(\mathrm{improved}), \tag{25}$$

evaluated across multiple weather/price years. Report as $/MWh of system-cost reduction or % decrease in imbalance charges with confidence intervals [9, 54, 55].

Risk and discounting. Publish ΔNPV/ΔLCOE across a *grid* of real WACC values rather than encoding "risk" solely as a higher discount rate. When interventions alter failure risks (e.g., aggressive global MPPT), include actuarial terms in O&M/warranty cash-flows [20].

10.4 Open datasets and open-source models for transparent appraisal

Policy credibility improves when inputs and code are public and citable.

Resource and weather data. ERA5 global reanalysis provides long, consistent meteorology suitable for counterfactuals and stress testing [101]. Renewables.ninja offers validated long-run PV reconstructions with peer-reviewed documentation [102]. For North America, satellite-derived NSRDB may be used provided the *version-specific* methodology paper is cited (not only a website).

Table 17: Illustrative parameter priors for Monte Carlo appraisal. Use site- and programme-specific data where available; ranges below are indicative only.

Quantity	**Suggested distribution (examples)**	**Notes / sources**
Module degradation rate (c-Si)	Lognormal; median 0.5–0.6%/yr, CV 0.4–0.7	Field compendia; climate dispersion [12, 13].
Defect loss severity (EL/IR)	Mixture over classes with class-conditional loss CDFs	Crack/hot-spot functions [34–36].
Detector operating point	Beta priors for TPR/FPR, by class	Cross-plant validation [39, 41].
Reserve price	Empirical price-year bootstrap by zone	Operations data (system operator).
Forecast skill (CRPS)	Empirical distribution by horizon/regime	Out-of-sample skill [51, 92].
CAC reduction	Normal or triangular around measured mean (e.g., 10–20%)	Out-of-sample field tests [11].
Permitting duration	Empirical (jurisdiction bootstrap)	Large-sample panels [60, 62].

Power-system and market data. Open Power System Data (OPSD) curates European time series (load, generation, prices, weather) with a documented data-management pipeline that reduces cleaning ambiguities [103].

PV and system modeling tools. `pvlib` offers validated device-to-plant modeling routines with open tests [104]. `PyPSA` provides open UC/ED and capacity-expansion formulations with transparent datasets [105]. Apply FAIR data principles to all artifacts and publish model/dataset documentation (Model Cards, Datasheets) [83, 84, 106].

10.5 Worked templates (illustrative)

Forecasting → production-cost savings. Calibrate PV to measurements; generate probabilistic day-ahead forecasts scored with CRPS; embed scenarios in UC/ED; report Δproduction cost and Δreserve procurement with CIs and sensitivity to reserve prices [9, 54, 55, 92, 101].

Inspection → ΔLCOE. Combine defect prevalence and class-conditional loss distributions with detector TPR/FPR at a cost-sensitive threshold; simulate actions via break-even rules; compute ΔMWh and Δwarranty, then ΔLCOE, with sensitivity to labor and module prices [34–36, 43].

Soft-cost process reform. Draw permitting duration distributions from large-sample studies; model cancellation elasticity to delay; convert to ΔNPV via earlier energization (time value of energy) and overhead savings; benchmark cost splits to NREL/LBNL reports [60, 62, 64, 65].

10.6 Bottom line

Policy design around AI for PV should look less like a technology showcase and more like *decision analysis with open evidence*: define explicit baselines; use proper scoring and system models to translate technical gains into cash-flow and system-cost deltas; propagate uncertainty with global sensitivity; and publish open data/code so results survive independent replication. The peer-reviewed building blocks—system LCOE accounting [16], sensitivity/uncertainty methods [19, 20], forecast-value mapping [9, 10, 54, 55, 92], and robust open datasets/tools [101–106]—are

Table 18: Minimum reporting for policy-grade valuation. These items enable reproducibility and prevent benefit inflation (compare to Chapter 6).

Item	Expectation
Baseline practice	Hardware, operating rules, tariffs; legacy forecast(s) and reserve rules (UC/ED, imbalance penalties).
Technical KPIs	Proper scores by horizon/regime (CRPS, reliability); ROC/PR for classifiers with independent test periods.
Transfer functions	Equations/parameters linking KPIs to cash-flows; plant vs. system attribution explicit [16].
System model	UC/ED or market-clearing formulation; same engine for baseline and improved cases; scenario method documented.
Uncertainty protocol	MC sample size, parameter priors, Sobol indices; WACC grid; stress years (weather/price).
Valuation outputs	Δproduction cost, Δcurtailment, Δreserve, Δstarts, Δrevenues; then ΔNPV, ΔLCOE (plant & system) with CIs.
Reproducibility	Code, data snapshots (ERA5/OPSD/Ninja extracts), Model Cards/Datasheets; ROPES checklist compliance.

available. Applying these consistently yields policy-grade appraisals that regulators, financiers, and operators can trust.

11 Conclusions and research agenda

11.1 What we know—evidence-based conclusions

This paper quantifies where AI measurably affects the photovoltaics (PV) value chain and how those effects should be valued. Across materials, manufacturing, operations, markets, and soft-costs, the strongest, policy-relevant impacts emerge when *technical gains are mapped to cash-flows through explicit counterfactuals* and *uncertainty is propagated* (Chapter 2).

Materials & interfaces (Ch. 3). AI-assisted search demonstrably compresses iteration cycles in perovskites via Bayesian/active learning and closed-loop labs, with experimentally validated candidates improving interfacial recombination metrics (QFLS/V_{oc}) and stability. These gains are real but bounded by device-stack specificity and the gap between lab-stress and field durability; when translated through field degradation priors for crystalline Si (median ~ 0.5–0.6%/yr with dispersion) and conservative transfer functions, they yield credible, context-dependent ΔLCOE reductions rather than step changes [12, 13].

Manufacturing QA & field inspection (Ch. 4). EL/IR analytics reliably detect high-severity anomalies; field evidence shows a small tail of severe defects drives most losses. Using severity-to-power-loss mappings and cost-sensitive decision thresholds, inspection programs reduce yield loss and warranty outlays *only* when actions are economically justified (e.g., replacement break-even $\sim$10–33% loss under European cost structures). Benefits should not be credited to low-severity anomalies.

Device/plant control (Ch. 5). AI-enhanced MPPT improves tracking on dynamic/shaded profiles by $\sim$0.5–3 percentage points versus tuned conventional baselines on bench tests; after convolution with local irradiance-ramp and shading statistics, gains typically become sub-

percentage to low single-percentage annual ΔMWh for utility plants, higher for complex rooftops. Reliability constraints (hot-spot risk) must be priced into O&M if control is more aggressive.

Forecasting & system integration (Ch. 6). Economic value arises via unit commitment (UC), reserves, redispatch, and bidding—not via MAPE alone. *Probabilistic, calibrated* forecasts embedded in UC/ED and market strategies reduce production cost, reserve procurement, and imbalance charges; value depends on penetration, flexibility, and alignment of error reductions with decision timescales [9, 10, 54, 55].

Soft-costs & adoption (Ch. 7). Machine-learning targeting shows ≈15% CAC reductions ($\sim$ \$0.07/W) under rigorous out-of-sample tests [11]. Large-sample studies link permitting/interconnection reforms to shorter timelines and lower prices; ΔNPV materializes through earlier energization, reduced overhead burn, and fewer cancellations. Distributional analyses document persistent adoption gaps by race/ethnicity, implying fairness guardrails for analytics.

Case studies & meta-analysis (Ch. 8). Across CAISO, Germany/NW Europe, Spain, Japan, China NW, Australia, and U.S. multi-ISO settings, *structure dominates*: flexibility, transmission, and market design set the ceiling for realized AI value. Where intraday liquidity and reserves are well designed, *probabilistic, value-oriented* forecasting and bidding consistently improve revenues and lower system costs; where structural curtailment or cannibalization dominates, AI mainly supports risk-aware damage control.

Risk, governance, and standards (Ch. 9). Proper scoring, reproducible baselines (ROPES), and documented Model Cards/Datasheets are essential to control model risk and bias. Cybersecurity should align to IEEE 1547/1547.3, IEC 62443, NERC CIP, and NIST guidance; value claims must be accompanied by audit trails.

Policy design tools (Ch. 10). Decision-grade appraisal requires ΔNPV, Δ(LCOE) (and *system*-LCOE when integration costs change), Monte Carlo propagation with global sensitivity (Sobol), and fully open datasets/code (ERA5, OPSD, pvlib, PyPSA) for independent replication.

Synthesis. The literature supports cautious optimism: AI delivers *measurable, decision-relevant* gains when it addresses a clear bottleneck, is evaluated against named baselines with proper scores, and its benefits are counted once in cash-flows. The headline opportunities are in *probabilistic forecasting and market operations* at high penetration, *inspection triage* where severe defects are non-trivial, and *soft-cost streamlining* where PII queues and CAC are high. Headline risks include value inflation from lab-only metrics, bias amplification in marketing, and cyber exposure from poorly secured digitalization.

11.2 What we still need—prioritized research agenda

1. **Counterfactual standards and nested baselines.** Regulator-endorsed baseline libraries for each use-case (forecasting, MPPT, inspection, CAC/PII) with site-class distributions (ramps, shading, defect severity, permitting timelines) to support incremental valuation (Ch. 2).

2. **From lab-stress to field durability.** Module-relevant bridges linking QFLS/V_{oc} and passivation metrics to *fielded* degradation trajectories across climates, under preregistered evaluation, to tighten ΔLCOE attribution (Ch. 3).

3. **Severity-aware inspection economics.** Multi-plant EL/IR cohorts with paired IV curves and follow-up to refine severity-to-loss and progression models by climate/vintage; public releases enable transfer validation (Ch. 4).

4. **Value-oriented forecast post-processing.** Optimize calibration and price-spread correlation; report *value-of-information* in /MWh or % imbalance reduction inside UC/ED and bidding simulations (Ch. 6).

5. **Causal field evaluations.** Use modern DiD with staggered rollout and event-study diagnostics to estimate average treatment effects in deployments; consider stepped-wedge trials where feasible (Ch. 2).

6. **Distributional impact audits.** Require subgroup reporting for marketing/PII reforms; include fairness constraints (minimum recall/exposure parity) in procurements so programmes do not widen adoption gaps (Ch. 7).

7. **System-LCOE and market-value integration.** Standardize reporting of system-LCOE deltas and market value factors alongside plant ΔLCOE for interventions affecting reserves/curtailment/profile; include storage/flex demand scenarios (Chs. 6–**??**).

8. **Cybersecurity assurance for DER portfolios.** Operationalize IEEE 1547.3 in procurement and conformance testing (mutual auth, signed firmware, logging); publish anonymized incident post-mortems and testbed results for common PV protocols (Ch. 9).

9. **Open, audited pipelines.** Anchor claims in open-source toolchains (pvlib, PyPSA) with containerized environments and DOI-linked datasets (ERA5, OPSD, Renewables.ninja); require Model Cards/Datasheets, ROPES checklists, and Sobol charts in publicly funded studies (Chs. 9, 10).

10. **Procurement-ready templates.** Publish templates for (i) forecast→UC/ED valuation, (ii) inspection→warranty economics, (iii) MPPT→annual yield, (iv) CAC/PII cost–benefit; each defines baselines, inputs, transfer functions, uncertainty, and minimum reporting sets (Ch. 10).

11.3 Policy implications

- **Target the right levers.** In high-penetration systems with intraday liquidity, prioritize probabilistic forecasting integrated into UC/ED and bidding; where residential soft-costs are high, prioritize PII & CAC reforms.

- **Fund measurement, not marketing.** Focus public support on open datasets, standardized baselines, and field trials that produce transferable knowledge.

- **Codify guardrails.** Tie incentives for AI-enabled improvements to auditability (open artifacts), fairness reporting, and cybersecurity conformance as conditions of eligibility.

Final remark. When AI for PV is evaluated within the discipline framework herein proposed—explicit counterfactuals, proper scoring with economic mapping, quantified uncertainty, governance, and open evidence—the field moves from hopeful claims to *credible, replicable* impacts on cost, reliability, and equity. This agenda aims to make those impacts easier to verify and quantify.

Appendix A Governance and Evidence Package Checklist

A.1 Purpose and scope

This appendix defines a *policy-grade* evidence package for any claim of economic value from AI across the PV value chain. It is structured to be auditable by regulators, lenders, and independent engineers, and to align with: (i) NIST AI Risk Management Framework (AI RMF) [82]; (ii) proper probabilistic evaluation practice [15, 90–92]; (iii) DER and industrial control cybersecurity standards (IEEE 1547/1547.3; ISA/IEC 62443; NERC CIP) [85–89]; (iv) ICS security guidance [107]; and (v) FAIR data principles for reproducibility [106]. Each section provides: (A) a checklist (with ☐ indicating verifiable items), (B) what to submit as evidence, and (C) quantitative acceptance criteria. Where applicable, we reference earlier chapters (e.g., Ch. 2, 4, 5, 6, **??**).

A.2 Evidence bundle manifest (what to submit)

Provide a single, machine-readable inventory (YAML/JSON) that enumerates all artifacts referenced in this appendix and in the main text. Include filenames, SHA-256 hashes, versions/commit IDs, and brief descriptions to enable an auditor to verify integrity and provenance.

Table 19: Submission manifest for an AI value-claim. Use a machine-readable inventory (YAML/JSON) with file hashes.

Component	**Content**	**Verification & references**
☐ Problem definition	Use-case, decision, baseline, counterfactual	Chapter 2; ROPES [90]
☐ Data package	Train/valid/test splits; lineage; licenses; schema	FAIR [106]; Model/Datasheets [83, 84]
☐ Model artifacts	Code (commit hash), weights, post-processing	ROPES [90]; Benchmarks [91, 92]
☐ Eval reports	Proper scores by horizon/regime; reliability plots	[15, 92]
☐ Ops model	UC/ED or bidding model; same engine baseline/new	Ch. 6
☐ Econ mapping	Transfer functions; prices/penalties; WACC grid	Ch. 2
☐ Uncertainty	Monte Carlo set; Sobol indices; sensitivity set	Ch. 2
☐ Cyber SBOM	Firmware list; update policy; comms configs	IEEE 1547.3 [86]; NIST SP 800-82 [107]
☐ Network diagrams	Zones & conduits; ESP perimeters; remote access	IEC 62443 [87]; CIP-005 [88]
☐ IR/EL MSA (if used)	Measurement system analysis; calibration logs	Ch. 4

A.3 Model risk evaluation protocol

This subsection defines minimum expectations for probabilistic skill, calibration, shift robustness, adversarial resilience, and reproducibility. Claims must meet *all* criteria for the intended deployment domain(s).

Proper scoring and reliability. For each horizon/regime (clear, broken, overcast) report CRPS (mean, median, 95% CI), PIT histograms, and reliability diagrams. A claim of "improvement" requires:

- Statistically significant CRPS reduction on the *same* test periods (paired tests or block bootstrap), and
- No reliability degradation in PIT-based coverage (e.g., 90% predictive interval empirical coverage = 90% ± 5 pp) [15, 92].

Distribution shift tests. Evaluate pre/post skill across (i) new weather years, (ii) new plants/sites (hold-out), and (iii) hardware vintages. Report interval coverage by regime and the fraction of out-of-domain inputs. Improvements must persist within 10% of in-domain CRPS; otherwise, gate usage to in-domain envelopes (AI RMF [82]; ROPES [90]).

Adversarial/data-integrity resilience. Apply perturbation tests on key features (irradiance proxies, satellite channels, SCADA) with realistic bounds; document sensitivity of decisions (UC commitment, reserve MW, bids). Require *graceful degradation*: no unstable decision flips under physically implausible but small perturbations; implement plausibility filters and authenticated inputs [93–95].

Table 20: Model-risk acceptance criteria (minimum).

Dimension	**Metric**	**Minimum acceptance**
Calibration	90% interval coverage	90% ± 5 pp across regimes
Sharpness	CRPS vs. baseline	≥ 5–10% reduction with 95% CI non-overlap
Shift robustness	New years/sites	≤ 10% CRPS loss; coverage within bands
Security resilience	Decision sensitivity	No UC/bid instability for $\pm 2\sigma$ sensor noise; authenticated sources
Reproducibility	Seeded reruns	≤ 1 pp variance in CRPS from nondeterminism

A.4 Fairness evaluation protocol (soft-costs and siting)

For soft-cost and siting tools that influence customer acquisition, permitting, or incentives, document intended users and affected populations, measure subgroup performance, and implement guardrails before scale-up.

Documentation. Provide a Model Card and a Datasheet covering intended use, training cohorts, exclusions, subgroup metrics (recall/PPV by income, race/ethnicity, tenure where lawful), and transportability limits [83, 84].

Subgroup performance. Report confusion matrices and calibration curves by subgroup on *held-out* geographies. Require minimum-recall constraints (e.g., subgroup recall $\geq 0.9 \times$ global recall) unless a lawful policy goal justifies otherwise. Track ΔCAC and adoption by subgroup to ensure net benefits do not widen gaps documented in [66, 67].

Prospective guardrails. Before scaling, run a small prospective trial in new jurisdictions and re-estimate subgroup metrics. If subgroup calibration drifts by > 5 pp, retrain with constraint or reweighting; document in the Model Card (AI RMF [82]).

A.5 Cybersecurity controls verification

Submit concrete evidence that devices and networks implicated in the value claim meet minimum segmentation, authentication, update, and incident-response practices for DER and ICS contexts.

Architecture & segmentation. Submit diagrams of network zones and conduits, VLANs, and data flows; identify Electronic Security Perimeters (ESPs) for assets in scope (CIP-005-6) [88]. For plants/factories, apply IEC 62443 zones and documented conduits with firewall rules [87].

Authenticated/secure control channels. Show evidence of mutual authentication and encryption (cert inventories, TLS configs, cipher suites) for IEEE 2030.5, DNP3-SA, or SunSpec/Modbus over TLS [86, 108, 109]. Demonstrate replay protection and rate limiting at DER gateways (NIST SP 800-82 [107]).

Firmware updates and SBOM. Provide signed-firmware policy, rollback protection, update cadence, and a software bill of materials (SBOM). Demonstrate verification logs for updates in the last 12 months (IEEE 1547.3 [86]).

Incident response. Provide runbooks and evidence of a tabletop exercise in the last 12 months (CIP-008-6) [89]. Include alarm integration for unusual command sequences and failed authentications (NIST SP 800-82 [107]).

Table 21: Minimum cybersecurity evidence and tests.

Control	Evidence	Standard(s)
Zones/Conduits	Diagrams; firewall configs	IEC 62443 [87]
ESP/Remote access	Jump-host & MFA logs	CIP-005-6 [88]
Secure comms	Cert inventories; TLS checks	IEEE 1547.3 [86]; NIST 800-82 [107]
Signed updates	SBOM; signature logs	IEEE 1547.3 [86]
IR playbooks	Drill reports	CIP-008-6 [89]

A.6 Audit trail and reproducibility

Package artifacts so an independent engineer can reproduce skill and value estimates deterministically (modulo hardware nondeterminism), and document any deviations.

Artifact packaging. Provide a single `.zip/.tar.zst` with a machine-readable manifest (file names, SHA-256 hashes), Dockerfile or Conda `environment.yml`, seed settings, and README with exact commands. Follow FAIR and community benchmarking practice [90–92, 106].

Independent rerun. An independent engineer should be able to (i) re-evaluate skill on the provided test split; (ii) run the UC/ED or bidding model with both baselines and improved forecasts; and (iii) reproduce Δproduction cost, Δreserve, and Δrevenues within the stated CIs. Report any nondeterminism (GPU kernels, multi-threading).

Change management. Maintain a change log linking model versions to deployment dates and operating periods used in the valuation (AI RMF governance function [82]). If models changed during the study period, segment results by version.

A.7 Quantitative acceptance criteria for value claims

The following acceptance criteria must be satisfied *jointly* for any economic value claim.

- **Skill → Value:** No claim of economic value without (i) proper-score improvement with CI, and (ii) end-to-end valuation inside the decision model (UC/ED, reserve, or bidding) [15, 92].
- **Double counting:** Use nested baselines when multiple tools apply (e.g., inspection+forecasting) per Chapter 2; document the order of application and incremental deltas.
- **Uncertainty:** Provide Monte Carlo distributions for ΔNPV/ΔLCOE and variance-based sensitivity (Sobol totals); highlight dominant drivers.
- **Security gating:** If secure comms and signed updates are absent for devices involved in the claim, flag the value as *conditional* pending remediation (IEEE 1547.3 [86], NIST SP 800-82 [107]).
- **Fairness:** For soft-cost claims, report subgroup outcomes (conversion, CAC, cancellations) and certify that constraints or mitigations are in place if gaps exceed thresholds [66, 67, 83, 84].

A.8 Templates (author-facing)

Use these concise templates to ensure consistent, auditable documentation.

Model Card (abstract). Intended use; decision and baseline; training domains; evaluation splits and horizons; proper scores with CIs; failure modes; shift and adversarial tests; subgroup metrics (if applicable); change log [83].

Datasheet (abstract). Data provenance; collection period; licensing; known biases; missingness; transformations; privacy considerations; intended exclusions [84].

Cyber bill of materials (SBOM). Vendor SKUs, firmware versions, signing status, last update date, supported secure protocols (IEEE 2030.5, DNP3-SA, SunSpec/TLS), certificate expiry, and next scheduled patch window [86, 108, 109].

Decision mapping table. Explicit equations and parameter values linking skill to value (transfer functions); price/penalty inputs; reserve rules; discount rates; with citations and ranges (Chapter 2).

Appendix B One-page submission checklist (fillable)

Instructions. This single page accompanies the full evidence package in Appendix A. Check each item and supply a pointer (file name, link, or page # in the manifest). Leave comments if any acceptance criterion is not met (cf. Tables 20 and 21). This checklist summarizes, not replaces, the full evidence package in Appendix A.

Checklist

Item	Requirement (summary)	Pointer (file/link/page)
☐ Baseline	Decision, legacy practice, explicit counterfactual; scope/time/region aligned	
☐ Data	Train/valid/test splits; lineage; licenses; schema; access method	
☐ Model	Code commit, weights, post-processing; seeds; environment file	
☐ Skill	Proper scores by horizon/regime (CRPS, reliability); out-of-sample	
☐ Shift	New years/sites/vintages; coverage within ± 5 pp; $\leq$10% CRPS loss	
☐ Adversarial	Perturbation tests; plausibility filters; stable UC/bid decisions	
☐ Ops model	Same UC/ED/bidding engine baseline vs. improved; reproducible runs	
☐ Mapping	Transfer functions; prices/penalties; WACC grid; nested baselines	
☐ Uncertainty	Monte Carlo distributions; Sobol totals; tornado/sensitivity plots	
☐ IR/EL MSA	Measurement system analysis & calibration (if imaging used)	
☐ Fairness	Model Card + Datasheet; subgroup metrics (if applicable)	
☐ SBOM	Firmware list; signing/rollback; update cadence	
☐ Network	Zones & conduits diagram; ESPs; remote-access controls	
☐ Secure comms	TLS/mutual auth evidence (2030.5/DNP3-SA/SunSpec+TLS)	
☐ Incident	CIP-008-style runbook; last tabletop drill evidence	

Acceptance criteria (tick and provide numeric evidence)

Tick	Criterion	Value / Evidence
☐	CRPS reduction $\geq$ 5–10% vs. baseline (95% CI non-overlap)	
☐	Predictive interval coverage 90% $\in [85, 95]$ pp across regimes	
☐	Shift robustness: $\leq$ 10% CRPS loss on new years/sites	
☐	No UC/bid instability under $\pm 2\sigma$ sensor noise	
☐	Nested baselines applied (no double counting)	
☐	Monte Carlo & Sobol reported; dominant drivers identified	
☐	TLS/mutual auth in place; signed firmware + rollback	
☐	Model Card & Datasheet attached; subgroup metrics OK	

Variance decomposition (top three drivers)

List variable name $\rightarrow$ total Sobol index. #1: _____ #2: _____ #3: _____

Economic outputs (with 95% CI)

Metric	Reported value	Notes (prices/penalties, years)
Δ Production cost		
Δ Reserve procurement		
Δ Curtailment		
Δ Starts / cycling		
Δ Revenues (bidding)		
Δ O&M / Warranty		
Δ LCOE / System LCOE		

Sign-off

Name _____ Role _____

Institution _____ Date _____

Signature _____ Contact _____

Appendix C Glossary of Abbreviations

Notes on usage:
(i) Several abbreviations have context-dependent meanings. In particular, **PR** denotes either *Precision–Recall* (machine vision metrics) or *Performance Ratio* (PV plant KPI); we disambiguate as *PR (precision–recall)* vs. *PR (performance ratio)* where needed.
(ii) Symbols such as Δ indicate differences (e.g., ΔLCOE is the change in LCOE).
(iii) For probabilistic forecasts, **CRPS** refers to the Continuous Ranked Probability Score; **CDF** is a cumulative distribution function.

Abbreviation	**Meaning / Context (where used)**
AC	Alternating Current (grid-side quantities; Chs. 5–6).
AGC	Automatic Grid/Generation Control (real-time regulation; Ch. 6).
AI	Artificial Intelligence (statistical learning/optimization across PV value chain; Chs. 1–9).
AI RMF / RMF	NIST Artificial Intelligence Risk Management Framework (governance; App. A, Ch. 9).
ANL / ANN	Artificial Neural Network (forecasting, MPPT, image analytics; Chs. 4–6).
ANSI	American National Standards Institute (standards references; Ch. 9).
AS	Ancillary Services (reserves, regulation; Ch. 6).
AUC / AUROC	Area Under ROC Curve (classifier performance; Ch. 4).
BES	Bulk Electric System (scope for NERC CIP; App. A.5).
BESS	Battery Energy Storage System (PV+storage operations; Chs. 5–6, 8).
BoS	Balance of System (non-module costs; Ch. 3).
CAC	Customer Acquisition Cost (soft-costs; Ch. 7).
CAISO	California Independent System Operator (case study; Ch. 8).
CapEx	Capital Expenditure (project finance; Chs. 2, 7).
CDF	Cumulative Distribution Function (probabilistic forecasts; Ch. 6).
CIP	Critical Infrastructure Protection (NERC CIP standards; Ch. 9).
CI(s)	Confidence Interval(s) (uncertainty reporting; Chs. 2, 4–8).
CRPS	Continuous Ranked Probability Score (proper score; Chs. 2, 6, 9).
CV (stats)	Coefficient of Variation (dispersion parameter in priors; Ch. 10).
DA	Day-Ahead (market horizon; Chs. 6, 8).
DC	Direct Current (DC-side controls; Chs. 5–6).
DER	Distributed Energy Resource (smart inverter fleets, rooftop PV; Chs. 6–7, 9).
DERMS	Distributed Energy Resource Management System (aggregation/control; Ch. 9).
DFI / FDI	False Data Injection (cyber risk; Ch. 9).
DFT	Density Functional Theory (materials modeling; Ch. 3).
DG	Distributed Generation (behind-the-meter or small front-of-meter PV; Chs. 6–8).
DJ (phase)	Dion–Jacobson phase (2D perovskite interlayers; Ch. 3).
DNP3-SA	Secure Authentication for DNP3 (hardened SCADA protocol; Ch. 9).
DOE	U.S. Department of Energy (DER cybersecurity guidance; Ch. 9).
ED	Economic Dispatch (operations; Chs. 2, 6).
EL	Electroluminescence (imaging modality; Ch. 4).
ERA5	ECMWF global reanalysis dataset (long-run weather; Ch. 10).
ESP	Electronic Security Perimeter (NERC CIP-005; Ch. 9).

Abbreviation	Meaning / Context (where used)
FAIR (data)	Findable, Accessible, Interoperable, Reusable (data principles; Ch. 3, App. A).
F1	F1-score (classifier metric; Ch. 4).
FPR	False Positive Rate (classifier operating point; Ch. 4).
GHI/DNI/DHI	Global/Direct/Diffuse Horizontal Irradiance (context; Ch. 8 via climate).
GSA	Global Sensitivity Analysis (variance-based; Ch. 2).
HIL	Hardware-in-the-Loop (controls testing; Ch. 5).
HOIP	Hybrid Organic–Inorganic Perovskite (materials; Ch. 3).
HVAC	Heating, Ventilation, Air Conditioning; incidental.
ICS	Industrial Control System (cyber/plant networks; Ch. 9).
ID	Intraday (market horizon; Chs. 6, 8).
IEC	International Electrotechnical Commission (standards; Ch. 9).
IEA PVPS	International Energy Agency Photovoltaic Power Systems Programme (Task 13 reports; Chs. 2, 4).
IEEE	Institute of Electrical and Electronics Engineers (standards; Ch. 9).
IEEE 1815 (DNP3)	Distributed Network Protocol 3 for SCADA; see also DNP3-SA (Ch. 9).
IEEE 2030.5	Smart Energy Profile 2.0 over IP (DER communications; Ch. 9).
INC (MPPT)	Incremental Conductance (conventional MPPT; Ch. 5).
I–V	Current–Voltage (device/array characteristic; Chs. 3–5).
IR	Infrared (thermography; Ch. 4).
IR/EL	Combined Infrared/Electroluminescence (inspection stack; Ch. 4).
ISO	Independent System Operator (market/operator; Chs. 6, 8).
ISO-NE	ISO New England (study context in earlier references; Ch. 1).
JSON	JavaScript Object Notation (data artifact; App. A).
KPI	Key Performance Indicator (technical-to-economic mapping; Chs. 2–7).
LBNL	Lawrence Berkeley National Laboratory (Tracking the Sun; Ch. 7).
LCOE	Levelized Cost of Energy (plant metric; Chs. 2, 3, 4, 5, 7, 8).
LCOE (system)	System LCOE (includes integration costs; Chs. 2, 6).
MAE	Mean Absolute Error (forecast metric; Ch. 8 discussion).
MAPE/RMSE	Mean Absolute Percentage Error / Root Mean Square Error (deterministic forecast errors; Chs. 6, 8).
MFA	Multi-Factor Authentication (remote-access control; App. A.5).
MHP	Metal–Halide Perovskite (materials; Ch. 3).
MIBEL	Iberian Electricity Market (Spain/Portugal; Ch. 8).
MPP	Maximum Power Point (device operating point; Chs. 3, 5).
MPPT	Maximum Power Point Tracking (controls; Ch. 5).
MSA	Measurement System Analysis (metrology QA; Ch. 4).
MVF	Market Value Factor (relative price of PV; Ch. 8).
NEM	National Electricity Market (Australia; Ch. 8).
NERC	North American Electric Reliability Corporation (bulk system standards; Ch. 9).
NIST	National Institute of Standards and Technology (US; Ch. 9).
NPV	Net Present Value (finance; Chs. 2–8).
NREL	National Renewable Energy Laboratory (US; Chs. 4, 7, 8, 9).
NSRDB	National Solar Radiation Database (U.S. solar resource; Ch. 10).
NWP	Numerical Weather Prediction (forecast input; Chs. 6, 8).
O&M	Operations and Maintenance (cost category; Chs. 2, 4, 7).
OPSD	Open Power System Data (European time-series; Ch. 10).
P&O (MPPT)	Perturb and Observe (conventional MPPT; Ch. 5).

Abbreviation	Meaning / Context (where used)
PCE	Power Conversion Efficiency (cell/module; Ch. 3).
PDF (stats)	Probability Density Function (probabilistic modeling; Ch. 6 context).
PID	Potential-Induced Degradation (failure mode; Ch. 4).
PII	Permitting–Inspection–Interconnection (soft-cost process; Ch. 7).
PIT	Probability Integral Transform (reliability check for probabilistic forecasts; App. A.3).
PNAS	Proceedings of the National Academy of Sciences (venue; Ch. 7/8 refs).
PPV (classifier)	Positive Predictive Value (precision analog; Ch. 4).
PR (metrics)	Precision–Recall (classifier operating characteristic; Ch. 4).
PR (perf.)	Performance Ratio (energy-normalized plant KPI; Ch. 4).
PV	Photovoltaics (domain; all chapters).
PVPS	See IEA PVPS (programme; Chs. 2, 4).
PyPSA	Python for Power System Analysis (UC/ED/capacity expansion; Ch. 10).
pvlib	Open PV modeling library (Python/MATLAB; Ch. 10).
QFLS	Quasi-Fermi Level Splitting (device photophysics; Ch. 3).
QA	Quality Assurance (manufacturing/field; Ch. 4).
RBAC	Role-Based Access Control (device/SCADA access; Ch. 9).
Renewables.ninja	Public PV/wind reconstructions platform (long-run series; Ch. 10).
ROC	Receiver Operating Characteristic (classifier curve; Ch. 4).
ROC/PR	ROC and Precision–Recall (joint reporting; Ch. 4).
ROPES	Reproducible, Operational, Probabilistic/Physically-based, Ensemble, Skill (solar-forecasting guideline; Chs. 6, 9).
RT	Real-Time (market/operations; Ch. 6).
SBOM	Software Bill of Materials (cyber–supply chain; Ch. 9).
SCADA	Supervisory Control and Data Acquisition (plant control system; Ch. 9).
SCED / SCUC	Security-Constrained Economic Dispatch / Unit Commitment (implied in UC/ED; Ch. 6).
SoC (bat.)	State of Charge (storage; Chs. 6, 8).
Sobol index	Variance-based sensitivity index (global sensitivity; Ch. 2).
SunSpec	SunSpec Modbus (DER/inverter protocol; Ch. 9).
TLS	Transport Layer Security (secure comms; Ch. 9).
TPR	True Positive Rate (classifier recall; Ch. 4).
TPR_d	Class-conditional True Positive Rate for defect class d (inspection valuation; Ch. 4).
TSO	Transmission System Operator (EU context; Chs. 6, 8).
TTI	Time-to-Innovation (program-level metric in materials; Ch. 3).
UAV	Unmanned Aerial Vehicle (IR surveys; Ch. 4).
UC	Unit Commitment (operations; Chs. 2, 6).
UC/ED	Unit Commitment / Economic Dispatch (joint operations model; Chs. 2, 6).
VLAN	Virtual Local Area Network (network segmentation; App. A.5).
VoI	Value of Information (operational value of improved forecasts; Ch. 10).
V_oc (V_{oc})	Open-Circuit Voltage (cell metric; Ch. 3).
Volt/VAR, Volt/Watt	Grid-support functions of inverters (AC-side control; Ch. 5).
WACC	Weighted Average Cost of Capital (discount rate; Chs. 2, 6–8).
Wh / MWh	Watt-hour / Megawatt-hour (energy units; all chapters).
YAML	YAML Ain't Markup Language (artifact manifests; App. A.6).

Symbols and operators (select)

Symbol	Meaning
Δ	Difference/change operator (e.g., ΔNPV, ΔLCOE).
$\Pr(\cdot)$	Probability (e.g., $\Pr(d)$ is prevalence of defect class d).
$\mathbb{E}[\cdot]$	Expectation operator (used in valuation equations; Chs. 2, 4–6).
η_{trk}	MPPT tracking efficiency (Ch. 5).

Standards and guidance (acronyms used)

Abbrev.	Expanded form / role
IEEE 1547	DER interconnection/interoperability standard (Ch. 9).
IEEE 1547.3	Guide for DER cybersecurity (Ch. 9).
IEC 62443	Industrial automation/control systems security (zones/conduits; Ch. 9).
NERC CIP-005/008	ESP/Incident response requirements for BES assets (Ch. 9).
NIST (Inverter Guide)	NIST guidance for smart inverter cybersecurity (2024; Ch. 9).

References

[1] Benjamin Sanchez-Lengeling et al. Scaling deep learning for materials discovery. *Nature*, 2023. doi: 10.1038/s41586-023-06735-9. URL `https://doi.org/10.1038/s41586-023-06735-9`.

[2] J. Xu et al. Anion optimization for bifunctional surface passivation in perovskite solar cells. *Nature Materials*, 2023. doi: 10.1038/s41563-023-01705-y. URL `https://doi.org/10.1038/s41563-023-01705-y`.

[3] Zhenzhen Ren et al. Machine learning for perovskite materials design and discovery. *npj Computational Materials*, 7:95, 2021. doi: 10.1038/s41524-021-00495-8. URL `https://doi.org/10.1038/s41524-021-00495-8`.

[4] Y. Liu, Y. Wu, Y. Yuan, and L. Zhao. Deep learning-based method for defect detection in electroluminescent images of polycrystalline silicon solar cells. *Optics Express*, 32(10): 17295–17317, 2024. doi: 10.1364/OE.517341. URL `https://doi.org/10.1364/OE.517341`.

[5] Z. Qu et al. A photovoltaic cell defect detection model capable of topological knowledge extraction. *Scientific Reports*, 14:21904, 2024. doi: 10.1038/s41598-024-72717-0. URL `https://doi.org/10.1038/s41598-024-72717-0`.

[6] Lukas Koester et al. Review of photovoltaic module degradation, field inspection techniques and techno-economic assessment. *Renewable and Sustainable Energy Reviews*, 165: 112616, 2022. doi: 10.1016/j.rser.2022.112616. URL `https://doi.org/10.1016/j.rser.2022.112616`.

[7] M. Kermadi and E. M. Berkouk. Artificial intelligence-based maximum power point tracking controllers for photovoltaic systems: Comparative study. *Renewable and Sustainable Energy Reviews*, 69:369–386, 2017. doi: 10.1016/j.rser.2016.11.125. URL `https://doi.org/10.1016/j.rser.2016.11.125`.

[8] M. Khan et al. Conventional and artificial intelligence-based maximum power point tracking techniques for photovoltaic systems: A comparative study. *Engineering Reports*, 6(7): e12963, 2024. doi: 10.1002/eng2.12963. URL `https://doi.org/10.1002/eng2.12963`.

[9] C. B. Martinez-Anido et al. The value of day-ahead solar power forecasting improvement. *Solar Energy*, 129:192–203, 2016. doi: 10.1016/j.solener.2016.01.049. URL https://doi.org/10.1016/j.solener.2016.01.049.

[10] Omkar Gandhi et al. The value of solar forecasts and the cost of their errors: A review. *Renewable and Sustainable Energy Reviews*, 189:113915, 2024. doi: 10.1016/j.rser.2023.113915. URL https://doi.org/10.1016/j.rser.2023.113915.

[11] Cheng Dong et al. Machine learning reduces soft costs for residential solar photovoltaics. *Scientific Reports*, 13:7213, 2023. doi: 10.1038/s41598-023-33014-4. URL https://doi.org/10.1038/s41598-023-33014-4.

[12] Dirk C. Jordan and Sarah R. Kurtz. Photovoltaic degradation rates—an analytical review. *Progress in Photovoltaics: Research and Applications*, 21(1):12–29, 2013. doi: 10.1002/pip.1182. URL https://doi.org/10.1002/pip.1182.

[13] Dirk C. Jordan, Sarah R. Kurtz, Katherine VanSant, and John Newmiller. Compendium of photovoltaic degradation rates. *Progress in Photovoltaics: Research and Applications*, 24(7):978–989, 2016. doi: 10.1002/pip.2744. URL https://doi.org/10.1002/pip.2744.

[14] Mahmoud Dhimish, Vicky Holmes, Behzad Mehrdadi, and Mark Dales. The impact of cracks on photovoltaic power performance. *Journal of Science: Advanced Materials and Devices*, 2(2):199–209, 2017. doi: 10.1016/j.jsamd.2017.05.005. URL https://doi.org/10.1016/j.jsamd.2017.05.005.

[15] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437. URL https://doi.org/10.1198/016214506000001437.

[16] Falko Ueckerdt, Lion Hirth, Gunnar Luderer, and Ottmar Edenhofer. System lcoe: What are the costs of variable renewables? *Energy*, 63:61–75, 2013. doi: 10.1016/j.energy.2013.10.072. URL https://doi.org/10.1016/j.energy.2013.10.072.

[17] Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021. doi: 10.1016/j.jeconom.2021.03.014. URL https://doi.org/10.1016/j.jeconom.2021.03.014.

[18] Brantly Callaway and Pedro H. C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021. doi: 10.1016/j.jeconom.2020.12.001. URL https://doi.org/10.1016/j.jeconom.2020.12.001.

[19] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output: design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270, 2010. doi: 10.1016/j.cpc.2009.09.018. URL https://doi.org/10.1016/j.cpc.2009.09.018.

[20] Thang T. D. Tran and Adrian D. Smith. Incorporating performance-based global sensitivity and uncertainty analysis into lcoe calculations for emerging renewable energy technologies. *Applied Energy*, 216:157–171, 2018. doi: 10.1016/j.apenergy.2018.02.024. URL https://doi.org/10.1016/j.apenergy.2018.02.024.

[21] IEA PVPS Task 13. Degradation and failure modes in new photovoltaic cell and module technologies. Technical Report IEA-PVPS T13-30:2025, International Energy Agency, 2025. URL https://iea-pvps.org/wp-content/uploads/2025/02/IEA-PVPS-T13-30-2025-REPORT-Degradation-and-Failure.pdf.

[22] Andy Walker, Eric Lockhart, Jigar Desai, Kristen Ardani, Geoffrey Klise, et al. Model of operation-and-maintenance costs for photovoltaic systems. Technical Report NREL/TP-5C00-74840, National Renewable Energy Laboratory, 2020. URL https://www.nrel.gov/docs/fy20osti/74840.pdf.

[23] Henry C. Herbol et al. Efficient search of compositional space for hybrid organic–inorganic perovskites via bayesian optimization. *npj Computational Materials*, 4:51, 2018. doi: 10.1038/s41524-018-0106-7. URL https://doi.org/10.1038/s41524-018-0106-7.

[24] Shijing Sun et al. A data fusion approach to optimize compositional stability of halide perovskites. *Matter*, 3(3):939–955, 2021. doi: 10.1016/j.matt.2021.01.008. URL https://doi.org/10.1016/j.matt.2021.01.008.

[25] Nathan J. Szymanski et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature*, 623:322–329, 2023. doi: 10.1038/s41586-023-06734-w. URL https://doi.org/10.1038/s41586-023-06734-w.

[26] Shiyou You et al. Bifunctional hole-shuttle molecule for improved interfacial energetics in perovskite solar cells. *Nature Energy*, 8:1253–1263, 2023. doi: 10.1038/s41560-023-01249-0. URL https://doi.org/10.1038/s41560-023-01249-0.

[27] Cheng Zhi, Shuang Wang, Shijing Sun, et al. Machine-learning-assisted screening of interface passivation materials for perovskite solar cells. *ACS Energy Letters*, 8(3):1424–1433, 2023. doi: 10.1021/acsenergylett.2c02818. URL https://doi.org/10.1021/acsenergylett.2c02818.

[28] Michael Saliba, Taisuke Matsui, Krzysztof Domanski, Jangwon Y. Seo, Asli Ummadisingu, Shaik M. Zakeeruddin, Juan-Pablo Correa-Baena, Wolfgang R. Tress, Antonio Abate, Anders Hagfeldt, and Michael Grätzel. Incorporation of rubidium cations into perovskite solar cells improves photovoltaic performance. *Science*, 354(6309):206–209, 2016. doi: 10.1126/science.aah5557. URL https://doi.org/10.1126/science.aah5557.

[29] Tsz-Lung Leung et al. Stability of 2d and quasi-2d perovskite materials and devices. *Communications Materials*, 3:85, 2022. doi: 10.1038/s43246-022-00285-9. URL https://doi.org/10.1038/s43246-022-00285-9.

[30] Feng Zhang et al. Metastable dion–jacobson 2d structure enables efficient and stable perovskite solar cells. *Science*, 375(6584):71–76, 2022. doi: 10.1126/science.abj2637. URL https://doi.org/10.1126/science.abj2637.

[31] Erkan Aydin et al. Pathways toward commercial perovskite/silicon tandem solar modules. *Science*, 384(6692):eadh3849, 2024. doi: 10.1126/science.adh3849. URL https://doi.org/10.1126/science.adh3849.

[32] Tossi J. Jacobsson, Ebrahim H. Anaraki, Emil Nordman, et al. An open-access database and analysis tool for perovskite solar cells based on the fair data principles. *Nature Energy*, 7:107–115, 2022. doi: 10.1038/s41560-021-00941-3. URL https://doi.org/10.1038/s41560-021-00941-3.

[33] M. Köntges, I. Kunze, S. Kajari-Schröder, X. Breitenmoser, and B. Bjørneklett. The risk of power loss in crystalline silicon based photovoltaic modules due to micro-cracks. *Solar Energy Materials and Solar Cells*, 95(4):1131–1137, 2011. doi: 10.1016/j.solmat.2010.10.034. URL https://doi.org/10.1016/j.solmat.2010.10.034.

[34] Mahmoud Dhimish, Vicky Holmes, Paul Mather, et al. Rapid testing on the effect of cracks on solar cells output power and hotspot development. *Scientific Reports*, 12:12961, 2022. doi: 10.1038/s41598-022-16546-z. URL https://doi.org/10.1038/s41598-022-16546-z.

[35] Shereen Hassan and Mahmoud Dhimish. Broad-scale electroluminescence analysis of 5 million+ photovoltaic cells for defect detection and degradation assessment. *Renewable Energy*, 237:121868, 2024. doi: 10.1016/j.renene.2024.121868. URL https://doi.org/10.1016/j.renene.2024.121868.

[36] Mahmoud Dhimish, Paul Mather, and Vicky Holmes. Evaluating power loss and performance ratio of hot-spotted photovoltaic modules. *IEEE Transactions on Electron Devices*, 65(12):5419–5427, 2018. doi: 10.1109/TED.2018.2877806. URL https://doi.org/10.1109/TED.2018.2877806.

[37] S. A. Rahaman et al. Investigate the impact of environmental and operating conditions on ir imaging-based pv defect detection. *Solar Energy*, 243:70–82, 2022. doi: 10.1016/j.solener.2022.06.136. URL https://doi.org/10.1016/j.solener.2022.06.136.

[38] Wei Luo, Yong Shen, Dongxiang Wang, and Deqing Yang. Potential-induced degradation in photovoltaic modules: A critical review. *Energy & Environmental Science*, 10:43–68, 2017. doi: 10.1039/C6EE02271E. URL https://doi.org/10.1039/C6EE02271E.

[39] Stefan Deitsch, Vincent Christlein, Sebastian Berger, Claudia Buerhop-Lutz, Andreas Maier, Florian Gallwitz, and Christian Riess. Automatic classification of defective photovoltaic module cells in electroluminescence images. *Solar Energy*, 185:455–468, 2019. doi: 10.1016/j.solener.2019.02.067. URL https://doi.org/10.1016/j.solener.2019.02.067.

[40] X. Chen, T. Karin, and A. Jain. Automated defect identification in electroluminescence images of solar modules. *Solar Energy*, 242:20–29, 2022. doi: 10.1016/j.solener.2022.06.031. URL https://doi.org/10.1016/j.solener.2022.06.031.

[41] Lukas Bommes, Markus Hoffmann, Claudia Buerhop-Lutz, Thomas Pickel, Jens Hauch, Christoph Brabec, Andreas Maier, and Irene M. Peters. Anomaly detection in ir images of pv modules using supervised contrastive learning. *Progress in Photovoltaics: Research and Applications*, 30(12):1617–1634, 2022. doi: 10.1002/pip.3518. URL https://doi.org/10.1002/pip.3518.

[42] Muhammad Waqas Akram, Guohui Li, Yan Jin, Xiang Chen, Chuan Zhu, and Asif Ahmad. Automatic detection of photovoltaic module defects in infrared images with isolated and develop-model transfer deep learning. *Solar Energy*, 208:692–705, 2020. doi: 10.1016/j.solener.2020.08.080. URL https://doi.org/10.1016/j.solener.2020.08.080.

[43] Vegard Lofstad-Lie, Benedicte L. Aarseth, Nick Roosloot, Erik S. Marstein, and Torbjørn Skauli. Modeling cost-effectiveness of pv module replacement based on quantitative assessment of defect power loss. *Solar*, 4(4):728–743, 2024. doi: 10.3390/solar4040034. URL https://doi.org/10.3390/solar4040034.

[44] Trishan Esram and Patrick L. Chapman. Comparison of photovoltaic array maximum power point tracking techniques. *IEEE Transactions on Energy Conversion*, 22(2):439–449, 2007. doi: 10.1109/TEC.2006.874230. URL https://doi.org/10.1109/TEC.2006.874230.

[45] Fangrui Liu, Shanxu Duan, Fei Liu, Bangyin Liu, and Yong Kang. A variable step size inc mppt method for pv systems. *IEEE Transactions on Industrial Electronics*, 55(7): 2622–2628, 2008. doi: 10.1109/TIE.2008.920550. URL https://doi.org/10.1109/TIE.2008.920550.

[46] Xiaomeng Liu, Shanxu Duan, Tao Cai, Cheng Chen, and Hongbing Hu. A review on maximum power point tracking techniques for photovoltaic systems under partial shading conditions. *Renewable and Sustainable Energy Reviews*, 41:436–453, 2015. doi: 10.1016/j.rser.2014.08.038. URL https://doi.org/10.1016/j.rser.2014.08.038.

[47] Matthew Lave, Jan Kleissl, and Ernesto Arias-Castro. High-frequency irradiance fluctuations and geographic smoothing. *Solar Energy*, 86(8):2190–2199, 2012. doi: 10.1016/j.solener.2012.04.020. URL https://doi.org/10.1016/j.solener.2012.04.020.

[48] Matthew Lave, Matthew J. Reno, and Robert J. Broderick. Characterizing local high-frequency solar variability and its impact to distribution studies. *Solar Energy*, 118: 327–337, 2015. doi: 10.1016/j.solener.2015.05.028. URL https://doi.org/10.1016/j.solener.2015.05.028.

[49] J. W. Bishop. Computer simulation of the effects of electrical mismatches in photovoltaic cell interconnection circuits. *Solar Cells*, 25(1):73–89, 1988. doi: 10.1016/0379-6787(88)90059-2. URL https://doi.org/10.1016/0379-6787(88)90059-2.

[50] Francisco Martínez-Moreno, Javier Muñoz, and Eduardo Lorenzo. Experimental model to estimate shading losses on pv arrays. *Solar Energy*, 84(9):1805–1817, 2010. doi: 10.1016/j.solener.2010.07.011. URL https://doi.org/10.1016/j.solener.2010.07.011.

[51] Javier Antonanzas, Noelia Osorio, Ramón Escobar, Raquel Urraca, Francisco J. Martinez-de Pison, and Fernando Antonanzas-Torres. Review of photovoltaic power forecasting. *Solar Energy*, 136:78–111, 2016. doi: 10.1016/j.solener.2016.06.069. URL https://doi.org/10.1016/j.solener.2016.06.069.

[52] L. M. Aguiar, B. Pereira, M. David, F. Díaz, and P. Lauret. Use of satellite data to improve solar radiation forecasting with bayesian artificial neural networks. *Solar Energy*, 122:1309–1324, 2015. doi: 10.1016/j.solener.2015.10.041. URL https://doi.org/10.1016/j.solener.2015.10.041.

[53] L. M. Aguiar, B. Pereira, P. Lauret, F. Díaz, and M. David. Combining solar irradiance measurements, satellite-derived data and an nwp model to improve intra-day solar forecasting. *Renewable Energy*, 97:599–610, 2016. doi: 10.1016/j.renene.2016.06.018. URL https://doi.org/10.1016/j.renene.2016.06.018.

[54] Qi Wang, Ning Zhang, Zhi Li, Chongqing Kang, Qing Xia, and Ying Du. Quantifying the value of probabilistic forecasting for power system operation planning. *Applied Energy*, 343:121254, 2023. doi: 10.1016/j.apenergy.2023.121254. URL https://doi.org/10.1016/j.apenergy.2023.121254.

[55] Luuk R. Visser, Taha A. AlSkaif, Ahsan Khurram, Jan Kleissl, and Wilfried van Sark. Probabilistic solar power forecasting: An economic and technical evaluation of an optimal market bidding strategy. *Applied Energy*, 370:123573, 2024. doi: 10.1016/j.apenergy.2024.123573. URL https://doi.org/10.1016/j.apenergy.2024.123573.

[56] Miguel A. Matos and Ricardo J. Bessa. Setting the operating reserve using probabilistic wind power forecasts. *IEEE Transactions on Power Systems*, 26(2):594–603, 2012. doi: 10.1109/TPWRS.2012.2191978. URL https://doi.org/10.1109/TPWRS.2012.2191978.

[57] Pierre Pinson, Christophe Chevallier, and Georges Kariniotakis. Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems*, 22(3):1148–1156, 2007. doi: 10.1109/TPWRS.2007.901117. URL https://doi.org/10.1109/TPWRS.2007.901117.

[58] Bryan Bollinger and Kenneth Gillingham. Peer effects in the diffusion of solar photovoltaic panels. *Marketing Science*, 31(6):900–912, 2012. doi: 10.1287/mksc.1120.0727. URL https://doi.org/10.1287/mksc.1120.0727.

[59] Jonathan E. Hughes and Michael Podolefsky. Getting green with solar subsidies: Evidence from the california solar initiative. *Journal of the Association of Environmental and Resource Economists*, 2(2):235–275, 2015. doi: 10.1086/681131. URL https://doi.org/10.1086/681131.

[60] Eric O'Shaughnessy, Galen Barbose, and Ryan Wiser. Patience is a virtue: A data-driven analysis of rooftop solar pv permitting timelines in the united states. *Energy Policy*, 144: 111615, 2020. doi: 10.1016/j.enpol.2020.111615. URL https://doi.org/10.1016/j.enpol.2020.111615.

[61] Jesse Burkhardt, Ryan Wiser, Naim Darghouth, Changgui Dong, and Jason Huneycutt. Exploring the impact of permitting and local regulatory processes on residential solar prices in the united states. *Energy Policy*, 78:102–112, 2015. doi: 10.1016/j.enpol.2014.12.020. URL https://doi.org/10.1016/j.enpol.2014.12.020.

[62] Eric O'Shaughnessy, Shu Dong, Jeffrey J. Cook, Jesse Cruce, Kristin Ardani, Elizabeth Fekete, and Robert Margolis. Effects of local permitting and interconnection requirements on solar pv installation durations. *Energy Policy*, 161:112734, 2022. doi: 10.1016/j.enpol.2021.112734. URL https://doi.org/10.1016/j.enpol.2021.112734.

[63] Jeffrey Cook, Emily Dalecki, et al. Solarapp+ performance review (2023 data). Technical Report NREL/TP-6A20-89618, National Renewable Energy Laboratory, 2024. URL https://docs.nrel.gov/docs/fy24osti/89618.pdf.

[64] Vignesh Ramasamy et al. U.s. solar photovoltaic system and energy storage cost benchmark: Q1 2023. Technical Report NREL/TP-7A40-87303, National Renewable Energy Laboratory, 2023. URL https://docs.nrel.gov/docs/fy23osti/87303.pdf.

[65] Galen Barbose et al. Tracking the sun: 2024 edition. Technical report, Lawrence Berkeley National Laboratory, 2024. URL https://emp.lbl.gov/sites/default/files/2024-10/Tracking%20the%20Sun%202024_Report.pdf.

[66] Deborah A. Sunter, Sergio Castellanos, and Daniel M. Kammen. Disparities in rooftop photovoltaics deployment in the united states by race and ethnicity. *Nature Sustainability*, 2:71–76, 2019. doi: 10.1038/s41893-018-0204-z. URL https://doi.org/10.1038/s41893-018-0204-z.

[67] Fedor A. Dokshin and Brian C. Thiede. Revised estimates of racial and ethnic disparities in rooftop photovoltaic deployment in the united states. *Nature Sustainability*, 2023. doi: 10.1038/s41893-023-01134-4. URL https://doi.org/10.1038/s41893-023-01134-4.

[68] Rebecca R. Hernandez et al. Solar energy development impacts on land cover change and protected areas. *Proceedings of the National Academy of Sciences*, 112(44):13579–13584, 2015. doi: 10.1073/pnas.1517656112. URL https://doi.org/10.1073/pnas.1517656112.

[69] Paul Denholm, Matthew O'Connell, Greg Brinkman, and Jennie Jorgenson. Overgeneration from solar energy in california: A field guide to the duck chart. *Energy Policy*, 127:537–549, 2019. doi: 10.1016/j.enpol.2018.12.064. URL https://doi.org/10.1016/j.enpol.2018.12.064.

[70] Lion Hirth. The market value of variable renewables: The effect of solar and wind power variability on their relative price. *Energy Economics*, 38:218–236, 2013. doi: 10.1016/j.eneco.2013.02.004. URL https://doi.org/10.1016/j.eneco.2013.02.004.

[71] Maik Eising, Hendrik Hobbie, and Dominik Möst. Future wind and solar power market values in germany—evidence of spatial and technological dependencies? *Energy Economics*, 86:104638, 2020. doi: 10.1016/j.eneco.2019.104638. URL https://doi.org/10.1016/j.eneco.2019.104638.

[72] Joachim Seel, Galen Barbose, and Ryan Wiser. Why are residential pv prices in the united states so high? a comparative assessment of installation prices in the u.s. and germany. *Energy Policy*, 69:467–475, 2014. doi: 10.1016/j.enpol.2014.01.043. URL https://doi.org/10.1016/j.enpol.2014.01.043.

[73] Oliver Ruhnau, Philipp Hennig, and Reinhard Madlener. Economic implications of forecasting electricity generation from variable renewable energy sources. *Renewable Energy*, 161:1318–1327, 2020. doi: 10.1016/j.renene.2020.06.110. URL https://doi.org/10.1016/j.renene.2020.06.110.

[74] Michel David, John Boland, Lorenzo Cirocco, Philippe Lauret, and Cyril Voyant. Value of deterministic day-ahead forecasts of pv generation in pv+storage operation for the australian electricity market. *Solar Energy*, 224:672–684, 2021. doi: 10.1016/j.solener.2021.06.011. URL https://doi.org/10.1016/j.solener.2021.06.011.

[75] Sheila Marie G. Dumlao and Keiichi N. Ishihara. Reproducing solar curtailment with fourier analysis using japan dataset. *Energy Reports*, 6(S2):199–205, 2020. doi: 10.1016/j.egyr.2019.11.063. URL https://doi.org/10.1016/j.egyr.2019.11.063.

[76] Ning Tang, Yong Zhang, Yanan Niu, and Xinying Du. Solar energy curtailment in china: Status quo, reasons and solutions. *Renewable and Sustainable Energy Reviews*, 97:509–528, 2018. doi: 10.1016/j.rser.2018.07.021. URL https://doi.org/10.1016/j.rser.2018.07.021.

[77] Juan Ignacio Peña, Raquel Rodríguez, and Silvia Mayoral. Cannibalization, depredation, and market remuneration of power plants. *Energy Policy*, 167:113086, 2022. doi: 10.1016/j.enpol.2022.113086. URL https://doi.org/10.1016/j.enpol.2022.113086.

[78] Patrick R. Brown and Francis M. O'Sullivan. Spatial and temporal variation in the value of solar power across united states electricity markets. *Renewable and Sustainable Energy Reviews*, 121:109594, 2020. doi: 10.1016/j.rser.2019.109594. URL https://doi.org/10.1016/j.rser.2019.109594.

[79] Setareh Goodarzi, Harsha N. Perera, and Derek Bunn. The impact of renewable energy forecast errors on imbalance volumes and electricity spot prices. *Energy Policy*, 134:110827, 2019. doi: 10.1016/j.enpol.2019.06.035. URL https://doi.org/10.1016/j.enpol.2019.06.035.

[80] Thomas Huld, Ralph Gottschalg, Hans-Georg Beyer, and Marko Topič. Mapping the performance of pv modules, effects of module type and data uncertainty. *Solar Energy*, 84(2):324–338, 2010. doi: 10.1016/j.solener.2010.01.005. URL https://doi.org/10.1016/j.solener.2010.01.005.

[81] Eftihia Skoplaki and John A. Palyvos. On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. *Solar Energy*, 83(5):614–624, 2009. doi: 10.1016/j.solener.2008.10.008. URL https://doi.org/10.1016/j.solener.2008.10.008.

[82] National Institute of Standards and Technology. Artificial intelligence risk management framework (ai rmf 1.0). Technical report, 2023. URL https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf.

[83] Margaret Mitchell, Simone Wu, et al. Model cards for model reporting. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, 2019. doi: 10.1145/3287560.3287596. URL https://doi.org/10.1145/3287560.3287596.

[84] Timnit Gebru et al. Datasheets for datasets. *Communications of the ACM*, 64(12):86–92, 2021. doi: 10.1145/3458723. URL https://doi.org/10.1145/3458723.

[85] IEEE Standards Association. Ieee std 1547-2018: Standard for interconnection and interoperability of distributed energy resources, 2018. URL https://standards.ieee.org/standard/1547-2018.html.

[86] IEEE Standards Association. Ieee 1547.3-2023: Guide for cybersecurity of distributed energy resources interconnected with electric power systems, 2023. URL https://webstore.ansi.org/standards/ieee/ieee15472023.

[87] ISA/IEC. Isa/iec 62443 series of standards: Industrial automation and control systems security, 2025. URL https://www.isa.org/standards-and-publications/isa-iec-62443-series-of-standards.

[88] North American Electric Reliability Corporation. Cip-005-6 cyber security—electronic security perimeter(s), 2019. URL https://www.nerc.com/pa/Stand/Reliability%20Standards/CIP-005-6.pdf.

[89] North American Electric Reliability Corporation. Cip-008-6 cyber security—incident reporting and response planning, 2024. URL https://www.nerc.com/pa/stand/reliability%20standards/cip-008-6.pdf.

[90] Dazhi Yang. A guideline to solar forecasting research practice (ropes). *Journal of Renewable and Sustainable Energy*, 11:022701, 2019. doi: 10.1063/1.5087462. URL https://doi.org/10.1063/1.5087462.

[91] Kyle Doubleday, Victoria Van Scyoc-Hernandez, and Bri-Mathias Hodge. Benchmark probabilistic solar forecasts: Characteristics and recommendations. *Solar Energy*, 206:52–67, 2020. doi: 10.1016/j.solener.2020.05.051. URL https://doi.org/10.1016/j.solener.2020.05.051.

[92] Tilmann Gneiting. Probabilistic solar forecasting: Benchmarks, post-processing, and evaluation. *Solar Energy*, 257:106–124, 2023. doi: 10.1016/j.solener.2023.04.067. URL https://doi.org/10.1016/j.solener.2023.04.067.

[93] Xiangyu Zhou. Load forecasting with adversarial attacks in power systems. In *ACM e-Energy Workshop*, 2019. doi: 10.1145/3314058.3317729. URL https://doi.org/10.1145/3314058.3317729.

[94] Zhiyuan Fan et al. Adversarial attacks on deep neural networks for time-series forecasting. In *International Conference on Computational Data and Social Networks*, 2022. doi: 10.1145/3485314.3485316. URL https://doi.org/10.1145/3485314.3485316.

[95] Ali Moradzadeh et al. Electric load forecasting under false data injection attacks using deep learning. *Energy Reports*, 8:140–152, 2022. doi: 10.1016/j.egyr.2021.12.030. URL https://doi.org/10.1016/j.egyr.2021.12.030.

[96] Joseph Marron et al. Cybersecurity for smart inverters: Guidelines for residential and light-commercial solar energy systems, 2024. URL https://www.nist.gov/publications/cybersecurity-smart-inverters-guidelines-residential-and-light-commercial-solar-energy.

[97] Fouzi Harrou et al. Cybersecurity of photovoltaic systems: challenges, threats, and mitigation strategies—a short survey. *Frontiers in Energy Research*, 11:1274451, 2023. doi: 10.3389/fenrg.2023.1274451. URL https://doi.org/10.3389/fenrg.2023.1274451.

[98] Rubén J. Rodríguez et al. A toolkit to facilitate security auditing of ics devices using plcs. *Computers & Security*, 128:103148, 2023. doi: 10.1016/j.cose.2023.103148. URL https://doi.org/10.1016/j.cose.2023.103148.

[99] Esteban D. Gutiérrez Mlot et al. Towards a testbed for critical industrial systems: Sunspec modbus case study. In *2022 IEEE ETFA*, 2022. doi: 10.1109/ETFA52439.2022.9921522. URL https://doi.org/10.1109/ETFA52439.2022.9921522.

[100] U.S. DOE and NREL. Cybersecurity standards for distributed energy resources, 2025. URL https://www.nrel.gov/security-resilience/cybersecurity-standards.

[101] Hans Hersbach et al. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. doi: 10.1002/qj.3803. URL https://doi.org/10.1002/qj.3803.

[102] Stefan Pfenninger and Iain Staffell. Long-term patterns of european pv output using the renewables.ninja dataset. *Energy*, 114:1251–1265, 2016. doi: 10.1016/j.energy.2016.08.060. URL https://doi.org/10.1016/j.energy.2016.08.060.

[103] Frauke Wiese, Ingmar Schlecht, et al. Open power system data—frictionless data for electricity system modelling. *Energy Strategy Reviews*, 23:100399, 2019. doi: 10.1016/j.esr.2019.100399. URL https://doi.org/10.1016/j.esr.2019.100399.

[104] William F. Holmgren, Clifford W. Hansen, and Mark A. Mikofski. pvlib python: a python package for modeling solar energy systems. *Journal of Open Source Software*, 3(29):884, 2018. doi: 10.21105/joss.00884. URL https://doi.org/10.21105/joss.00884.

[105] Tom Brown, Jonas Hörsch, and David Schlachtberger. Pypsa: Python for power system analysis. *Journal of Open Research Software*, 6(1):4, 2018. doi: 10.5334/jors.188. URL https://doi.org/10.5334/jors.188.

[106] Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, and et al. The fair guiding principles for scientific data management and stewardship. *Scientific Data*, 3: 160018, 2016. doi: 10.1038/sdata.2016.18. URL https://doi.org/10.1038/sdata.2016.18.

[107] Keith Stouffer, Victoria Lightman, Victoria Pillitteri, Michael Abrams, and Adam Hahn. Guide to industrial control systems (ics) security, sp 800-82 rev. 3. Technical report, National Institute of Standards and Technology, 2024. URL https://csrc.nist.gov/publications/detail/sp/800-82/rev-3/final.

[108] IEEE Standards Association. Ieee std 2030.5-2018: Smart energy profile application protocol, 2018. URL https://standards.ieee.org/standard/2030_5-2018.html.

[109] IEEE Standards Association. Ieee std 1815-2012 (dnp3): Standard for electric power systems communications, 2012. URL https://standards.ieee.org/standard/1815-2012.html.

[110] ISO/IEC. Iso/iec 27019:2017 information technology — security techniques — information security controls for the energy utility industry, 2017. URL https://www.iso.org/standard/68095.html.