PCA-OS: A Planetary Climate Adaptation Operating System

Chaoyue He
Nanyang Technological University
Singapore, Singapore
cyhe@ntu.edu.sg

Xin Zhou
Nanyang Technological University
Singapore, Singapore
xin.zhou@ntu.edu.sg

Di Wang
Nanyang Technological University
Singapore, Singapore
wangdi@ntu.edu.sg

Hong Xu
Nanyang Technological University
Singapore, Singapore
xuhong@ntu.edu.sg

Wei Liu
Alibaba Group
Hangzhou, China
weiliu.liuwei@alibaba-inc.com

Chunyan Miao*
Nanyang Technological University
Singapore, Singapore
ascymiao@ntu.edu.sg

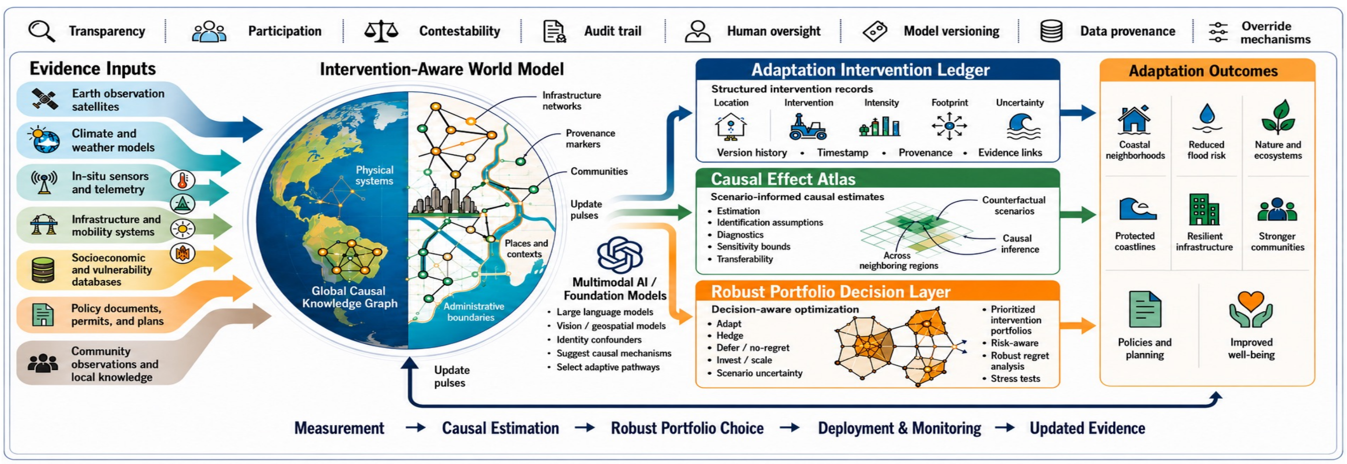

Figure 1: Conceptual infographic overview of PCA-OS.

Abstract

Predictive climate machine learning is increasingly good at forecasting hazards, but hazard maps alone do not decide what to do, where, when, for whom, and under which futures. We argue that climate ML remains insufficient for adaptation unless interventions become first-class, versioned, and auditable objects. Many climate digital twins still prioritize state estimation and simulation, whereas adaptation requires intervention observability, counterfactual effect estimation, and constrained portfolio choice. We propose **PCA-OS** (**P**lanetary **C**limate **A**daptation **O**perating **S**ystem), a decision-support operating abstraction built on an intervention-aware global causal knowledge graph. PCA-OS standardizes schemas, versioned updates, query primitives, and audit interfaces across three core objects: (1) an Adaptation Intervention Ledger recording measurable interventions with provenance and uncertainty; (2) a Causal Effect Atlas storing scenario-indexed, spillover-aware estimands, identification assumptions, diagnostics, and sensitivity bounds; and (3) a Robust Portfolio Decision Layer optimizing intervention portfolios under budget, equity, and no-harm constraints. Foundation models and intervention-aware world models should support, not replace, identification-aware causal analysis by surfacing candidate confounders, mechanisms, and spillover pathways for human review. We also outline AdaptBench, an evaluation suite where systems can fail for inequitable or maladaptive recommendations despite high predictive accuracy. The result is a field-level provocation: move climate ML from read-only hazard intelligence to auditable decision support for adaptation.

CCS Concepts

• **Computing methodologies** → **Machine learning**; **Causal reasoning and diagnostics**; • **Information systems** → **Decision support systems**; *Data mining*; • **Mathematics of computing** → *Mathematical optimization*; • **Applied computing** → *Earth and atmospheric sciences.*

Keywords

climate adaptation; world model; continual learning; causal inference; foundation models; robust optimization; decision support

ACM Reference Format:
Chaoyue He, Xin Zhou, Di Wang, Hong Xu, Wei Liu, and Chunyan Miao. 2026. PCA-OS: A Planetary Climate Adaptation Operating System. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2 (KDD 2026), August 9–13, 2026, Jeju Island, Republic of Korea.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3770855.3818654

*Corresponding author.

KDD 2026, Jeju Island, Republic of Korea

ACM ISBN 979-8-4007-2259-2/2026/08
https://doi.org/10.1145/3770855.3818654

1 Introduction & Provocation

The climate-ML stack is becoming spectacularly good at producing read-only futures. We can now forecast many weather and climate hazards at increasingly high spatiotemporal resolution, even as climate change shatters historical stationarity [37, 58, 64, 69]. Yet adaptation is not won on hazard maps. It is won on interventions: which roofs are cooled, which drainage networks are upgraded, and which communities absorb spillovers. Without first-class, versioned representations of actions and their downstream effects, many climate digital twins remain primarily observation and simulation platforms rather than intervention systems.

This is a structural blind spot in how the KDD community currently frames climate intelligence. Adaptation evidence remains fragmented, highly context-dependent, and thin on causal outcomes [9, 10, 29, 30, 34, 51, 78]. Reported adaptation actions are shaped by enabling conditions, planning capacity, and uneven assessment or reporting systems [14, 60], can generate spatial spillovers [4, 5], and must survive deeply uncertain futures [25, 73]. We have therefore optimized hazard prediction while remaining under-instrumented for intervention choice. We argue that adaptation should be framed as a continual learning and decision loop: **measurement → causal estimation → robust, equity-constrained optimization**.

In this framing, the primary scientific object is no longer the hazard map—it is the *intervention object*. To make hazard prediction actionable, it must be linked to measurable interventions, counterfactual effects, and portfolio trade-offs. PCA-OS directly answers this need by organizing planetary adaptation around three shared infrastructural artifacts: (1) an Adaptation Intervention Ledger that records where and when interventions occur, together with provenance and uncertainty; (2) a Causal Effect Atlas that stores scenario-indexed, spillover-aware causal estimates together with estimands, identification assumptions, diagnostics, and sensitivity bounds; and (3) a Robust Portfolio Decision Layer that turns those estimates into intervention portfolios under explicit budget, equity, and no-harm constraints. Together, these artifacts turn climate intelligence from descriptive monitoring into auditable decision support (Figure 1). Put differently, PCA-OS is an operating abstraction because it standardizes the shared objects, schemas, versioned update rules, query primitives, and audit interfaces through which adaptation evidence becomes intervention decisions.

Terminology. A *climate digital twin* provides state-estimation and simulation; an *intervention-aware world model* is its internal representation, here realized as a global causal knowledge graph over entities, relations, provenance, and updates; and an *adaptation operating system* coordinates interventions, effect objects, portfolio decisions, and audits through shared schemas, versioning, update semantics, and query interfaces. PCA-OS builds on the first two, but its core contribution is this OS layer.

Scope. PCA-OS is a decision-support operating abstraction, not a claim of fully autonomous intervention control; its purpose is to standardize how evidence, assumptions, and recommendations are updated, queried, and audited.

Contributions. This paper introduces PCA-OS, an operating abstraction for climate adaptation in which interventions become first-class, versioned, and contestable objects within climate ML systems. We argue that hazard-centric climate intelligence is structurally insufficient for adaptation unless interventions, their causal effects, and resulting decisions are jointly represented and auditable. To ground this vision, we outline a minimum viable PCA-OS at city scale: detecting parcel-level cool-roof retrofits from Earth observation and permit data, estimating tract-level avoided heat exposure and peak-demand impacts with neighborhood spillovers, and optimizing subsidy rollout under budget, equity-floor, and bounded-harm constraints. This narrow loop operationalizes the coupled measurement–causality–decision stack. **Why Blue Sky, why KDD, why now.** We position this work as a Blue Sky agenda for KDD, enabled by advances in foundation models, planetary-scale data systems, and causal decision methods that make auditable, intervention-centered adaptation science feasible.

Table 1: Comparison of PCA-OS among climate–ML systems, in terms of degree of intervention modeling, causal effect estimation, decision support, and auditability.

System	Interv.	Effects	Decision	Auditability
Hazard/Weather forecasting [58, 69]	no	no	no	limited
Adaptation tracking/assessment [9, 60]	partial	no	no	limited
Climate digital twins [22, 46]	partial	partial	rare	partial
PCA-OS (this work)	**core**	**core**	**core**	**core**

2 Related Work & Positioning

Adaptation tracking already shows why the status quo is inadequate. Prior work argues that adaptation evidence must be consistent, comparable, comprehensive, and coherent [29]; national and global assessments track plans and reported actions [9, 60]; and systematic stocktakes show that outcome evidence remains sparse and heterogeneous [10]. Big-data approaches were proposed to improve adaptation monitoring [30], but most existing pipelines still terminate at inventories, dashboards, or hazard forecasts. PCA-OS advances by making interventions themselves measurable, versioned, and causally queryable.

Climate digital twins show that planetary state estimation is increasingly feasible, but intervention choice is rarely the primary object. Table 1 summarizes the intended distinction. Climate twin efforts such as Destination Earth emphasize environmental states, simulations, and climate services [22, 46]. Existing process-based and engineering simulators already encode intervention-relevant mechanisms, including river/floodplain-routing dynamics, building-energy responses, urban-canopy surface-energy exchange, and coupled human–water dynamics [18, 62, 75, 83]. PCA-OS is therefore not a replacement for such models. It is a higher-level coordination layer that records which interventions were actually deployed, attaches causal and transportability evidence to their effects, and turns those effect objects into auditable portfolio decisions under equity and safety constraints.

The ingredients already exist but remain disconnected: foundation models and geospatial data systems improve planetary measurement [17, 45, 57, 58, 63, 66, 68, 69, 76]; theory-guided and physics-informed learning show how scientific structure can regularize data-driven systems [54, 55]; multimodal and language-grounded models align heterogeneous evidence [2, 13, 33, 52, 72]; causal inference supplies identification-aware estimators under confounding and spillovers [1, 6, 15, 16, 23, 49, 50, 70, 81]; and robust optimization plus fairness-aware learning formalize stress-tested action selection [8, 11, 21, 24, 38, 56, 59]. Here, we fuse these lines into one intervention-centered operating abstraction.

3 PCA-OS: From Architecture to Paradigm Shift

PCA-OS is not merely a simulator. It maintains an intervention-aware world model, implemented as a global causal knowledge graph, with auditable intervention, effect, decision, and governance objects as conditions evolve. By OS, we mean a coordinating layer that manages these shared objects across the measurement → causal estimation → portfolio selection loop through explicit schemas, versioning rules, update semantics, query primitives, and audit trails. At minimum, this means typed ledger, atlas, and portfolio objects with stable identifiers, provenance links, revision histories, and queryable dependency pointers, so both algorithms and human reviewers can inspect what changed, why it changed, which assumptions were used, and which decisions were affected.

Concretely, the OS layer should expose a small object contract rather than a monolithic model: ledger objects support `create`, `revise`, and `supersede`; atlas objects support queries by intervention family, geography, outcome, scenario, estimand, and identification design; portfolio objects expose `trace` links to the exact ledger versions, atlas entries, constraints, and human overrides used to produce a recommendation. These primitives (Figure 2) make the OS claim falsifiable: a PCA-OS prototype fails if a recommendation cannot be traced back to specific intervention evidence, causal assumptions, scenario choices, and governance constraints.

The system follows four design commitments. **First, decision-first:** the primary outputs are intervention, effect, and portfolio objects rather than hazard maps alone. **Second, interventions as first-class objects:** actions are versioned, geolocated, uncertain, and causally queryable. **Third, uncertainty-forward:** measurement, identification, and climate-scenario uncertainty must propagate into the decision layer. **Fourth, normative constraints as infrastructure:** equity, no-harm, and contestability cannot be post-hoc commentary; they must be encoded inside the optimization and the interface. Operationally, this agenda can be bootstrapped from open geospatial standards and cloud-native planetary data planes built around STAC catalogs, NASA HLS, ERA5, the Planetary Computer, OpenStreetMap, and WorldPop [36, 45, 63, 66, 68, 77].

3.1 An Intervention-Aware World Model

We instantiate the underlying world model maintained by PCA-OS at time t as a dynamic global causal knowledge graph (KG):

$$\mathcal{K}_t = (V_t, E_t, \mathcal{R}, X_t), \tag{1}$$

where X_t collects time-varying node and edge attributes, including uncertainty metadata. This is a systems abstraction: a global causal knowledge graph in which intervention, exposure, mechanism, and outcome semantics are represented explicitly, but not every stored edge is itself an identified causal effect. The graph stores entities, relations, topology changes, provenance, uncertainty, and typed causal semantics; identified causal claims are represented separately in atlas entries that attach estimands, assumptions, diagnostics, and sensitivity to selected intervention–outcome relations. This separation keeps the world model updatable without conflating storage with identification.

Nodes can represent spatial units, infrastructure assets, interventions, institutions, communities, and governing documents; edges instantiate relation types such as hydrologic connectivity, ecological adjacency, mobility, service access, ownership, administrative jurisdiction, and intervention–exposure pathways. Recent sustainability knowledge-graph efforts suggest a practical foundation for domain-grounded, updatable global causal knowledge graphs at scale [42, 44, 47]. Each node $u \in V_t$ has an available modality set $\mathcal{M}_{u,t} \subseteq \mathcal{M}$ and carries multimodal observations

$$o_{u,t} = \{m_{u,t}^{(k)} : k \in \mathcal{M}_{u,t}\}, \tag{2}$$

where the open modality registry $\mathcal{M}$ can include EO, climate reanalyses, sensors, text or administrative records, mobility, infrastructure telemetry, and damage or loss signals. A multimodal encoder then produces an adaptation-relevant state embedding $z_{u,t} = f_\theta(o_{u,t})$. Modality-agnostic architectures and aligned representation learning are essential because the relevant evidence is rarely fully paired [33, 52]. Interventions can alter not only node features but also graph topology: a levee attenuates hydrologic edges, urban greening creates ecological corridors, a cooling center creates accessibility edges, and zoning decisions activate governance relations.

We therefore require explicit update operators

$$\mathcal{K}_{t+1} = U_\phi(\mathcal{K}_t, O_{t+1}, \Delta\mathcal{L}_{t+1}, \Delta\mathcal{G}_{t+1}), \tag{3}$$

where O_{t+1} is the batch of new observations arriving between t and $t+1$, $\Delta\mathcal{L}_{t+1}$ is the set of new or revised ledger entries, and $\Delta\mathcal{G}_{t+1}$ is the set of governance or policy updates. These update operators are part of the OS semantics: they specify how new evidence mutates shared state, when revised entries supersede prior versions, and how downstream atlas or portfolio objects are marked for re-estimation or re-optimization. This gives the research community a sharper target: mine, update, and reason over a global causal KG with intervention-aware topologies rather than static hazard layers.

3.2 The Adaptation Intervention Ledger

A planetary adaptation science cannot exist without intervention observability. PCA-OS therefore introduces an **Adaptation Intervention Ledger** $\mathcal{L}$ that records where interventions occurred, when they changed, how intense they were, and what evidence supports those claims. A ledger entry is

$$\ell = (\mathrm{loc}, [t^{\mathrm{start}}, t^{\mathrm{end}}], \mathrm{type}, \mathrm{intensity}, \mathrm{footprint}, p_\ell, \mathrm{provenance}), \tag{4}$$

where loc denotes the affected spatial unit or geometry, and p_ℓ is a joint uncertainty object over type, timing, geometry, and intensity. Each ledger entry should carry a stable identifier, schema-validated fields, and a revision history so downstream atlas and portfolio objects can resolve exactly which intervention state they consumed.

The ledger fuses four evidence streams: **(1) EO change signals** such as albedo shifts, shoreline changes, vegetation dynamics, or

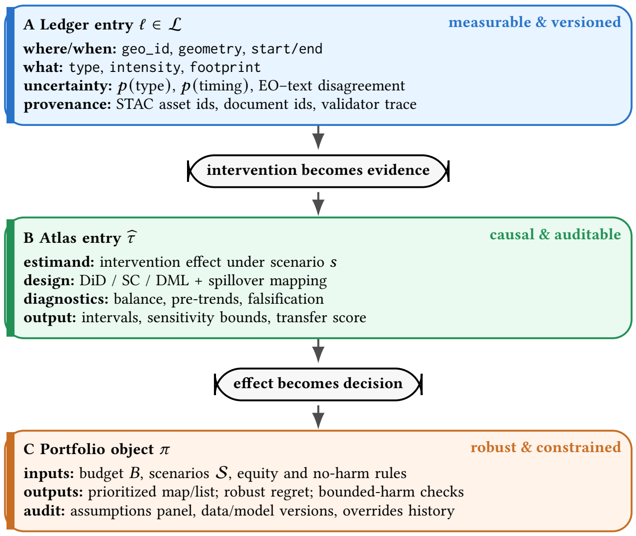

Figure 2: The PCA-OS object flow: interventions become ledger objects, then effect objects, then decision objects; uncertainty, provenance, and assumptions are carried forward.

impervious-surface updates; **(2) SAR and event-response evidence** for flood-related interventions and outcomes [12]; **(3) administrative and documentary text** such as permits, council minutes, engineering standards, procurement records, and adaptation plans; and **(4) operational or participatory traces** such as maintenance logs, sensor anomalies, and community reports. Because many interventions are informal, partially documented, or politically invisible, the ledger must preserve uncertainty, missingness, and multi-view disagreement rather than forcing certainty.

Reliable intervention data is the bottleneck, not a solved input assumption. PCA-OS therefore treats the ledger as a probabilistic and contestable evidence layer: absence from EO is not evidence of no intervention, text-only claims remain provisional until corroborated, and informal community adaptation should be represented with uncertainty rather than discarded. The ledger should expose completeness indicators, source coverage, validator status, and known blind spots so downstream causal estimates can condition on what the system failed to observe.

3.3 The Causal Effect Atlas

Because adaptation placement is non-random, prediction alone is biased. PCA-OS therefore maintains a **Causal Effect Atlas** whose entries are explicit objects: scenario-indexed, design-annotated, spillover-aware estimates with diagnostics and sensitivity bounds. Atlas entries are queryable by intervention family, geography, outcome, climate scenario, estimand, and design, making causal evidence a reusable system object rather than a one-off appendix table. Potential outcomes must admit interference:

$$Y_{i,t}(a_{i,t}, e_{i,t}, s), \qquad e_{i,t} = g(a_{j,t} : j \in \mathcal{N}(i)), \tag{5}$$

where $a_{i,t}$ is local action, $e_{i,t}$ is neighbor exposure, and $s \in \mathcal{S}$ indexes climate scenarios. Each atlas entry stores an estimand, identification strategy, interference model, diagnostics, transportability warning, and the limitations of the outcome proxy. Relevant estimation tools include staggered DiD, synthetic control, heterogeneous treatment effect estimation, double/debiased ML, and off-policy evaluation [1, 15, 16, 23, 81].

Scenario-indexed transportability is deliberately treated as an open scientific problem, connecting climate-scenario stress testing to formal causal transportability across environments [70, 71]. An effect estimated under one climate, infrastructure, or governance regime should not be copied into an unprecedented future as a point estimate. Atlas objects should therefore carry validity envelopes, mechanism notes, and transport warnings: where physical or institutional constraints are violated, the object should downgrade from “estimated effect” to bounds, stress-test input, or unknown. Attribution settings that separate anthropogenic and meteorological influences provide a useful analogue for distinguishing intervention effects from weather and climate confounding [20]. This lets the decision layer see uncertainty growth as portfolios leave the evidence base.

Foundation models and the world model act here as assistive hypothesis engines: they can surface candidate confounders, mechanisms, and spillover pathways for human review, but causal validity still rests on explicit identification assumptions and diagnostics [2, 13, 72].

3.4 The Robust Portfolio Decision Layer

The decision layer chooses portfolios of interventions rather than isolated actions. Let $\mathcal{I}$ index decision locations, let $\pi = (a_i)_{i \in \mathcal{I}}$ denote a portfolio with $a_i \in \mathcal{A}_i$, let $\mathcal{J}$ index protected or potentially harmed groups or locations, let $e_i(\pi)$ denote spillover exposure at location i, let $\widehat{\tau}_i^{(s)}(a_i, e_i(\pi))$ be the estimated normalized benefit under scenario $s \in \mathcal{S}$, and let $w_i \geq 0$ encode policy weights such as vulnerability or priority. Define the scenario-specific welfare of a portfolio as

$$V_s(\pi) = \sum_{i \in \mathcal{I}} w_i \, \widehat{\tau}_i^{(s)}(a_i, e_i(\pi)). \tag{6}$$

The robust portfolio decision then solves

$$\begin{aligned} \max_{\pi \in \prod_{i \in \mathcal{I}} \mathcal{A}_i} \quad & \min_{s \in \mathcal{S}} V_s(\pi) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{I}} \mathrm{cost}_i(a_i) \leq B, \quad \mathrm{Equity}^{(s)}(\pi) \geq \alpha, \qquad \forall s \in \mathcal{S}, \\ & \Delta\mathrm{Risk}_j^{(s)}(\pi) \leq \delta_j, \qquad \forall j \in \mathcal{J},\ \forall s \in \mathcal{S}. \end{aligned} \tag{7}$$

Here, $\mathrm{Equity}^{(s)}(\pi)$ denotes an auditable group-level equity functional under scenario s, and $\Delta\mathrm{Risk}_j^{(s)}(\pi)$ denotes the incremental risk imposed on protected location j relative to a no-action baseline. The equity floor α and bounded-harm thresholds δ_j are policy inputs chosen through stakeholder and governance processes, not values discovered by the model. Environmental-burden settings illustrate why these constraints must track downstream inequity, not only aggregate welfare [61]. To propagate effect-estimation uncertainty into decision making, the atlas can supply conservative lower-confidence values or ambiguity sets for $\widehat{\tau}_i^{(s)}$, so the portfolio layer is robust to both climate scenarios and causal uncertainty. The robust-feasible set is therefore $\Pi = \bigcap_{s \in \mathcal{S}} \Pi_s$. Each portfolio object should retain pointers to the exact ledger versions, atlas versions, constraints, and any human overrides from which it was derived.

A decision-facing score is Robust Decision Regret:

$$\mathrm{RDR}(\pi) = \max_{s \in \mathcal{S}} \left(\max_{\pi' \in \Pi_s} V_s(\pi') - V_s(\pi) \right), \tag{8}$$

where Π_s is the set of portfolios that satisfy the budget, equity, and no-harm constraints under scenario s. Lower is better, and $\mathrm{RDR}(\pi) \geq 0$ by construction [8, 11, 21, 24, 35, 38, 56, 59].

3.5 Foundation Models in PCA-OS

EO models improve intervention detection and outcome proxies such as albedo, heat, and hydrologic context [17, 76]. Weather and climate models provide scenario backbones for effect transport and stress testing [58, 69]. Probabilistic multivariate time-series models can support uncertainty-aware temporal evidence streams for sensors, energy demand, heat exposure, and infrastructure monitoring [26–28]. Language-grounded models align permits, plans, procurement records, and operations text with geospatial evidence [2, 13, 72]. Efficient LVLM and video-LLM inference is relevant when planetary-scale multimodal evidence creates long visual or temporal streams [53, 84]. The open problem is not only representation quality but also binding models and language-agent infrastructure to explicit intervention semantics, runtime controls, identification assumptions, and auditable decision objects [40, 41, 43, 85].

4 Human Interfaces, Governance, & Auditing

Because adaptation is inherently political, PCA-OS must be challengeable by design through four technical pillars. **Transparency:** causal estimands, identification choices, diagnostics, and sensitivity analyses must be inspectable [70]. **Participation:** stakeholders should be able to dispute detections, annotate local constraints, and force explicit handling of missing or contested evidence. **Safety:** optimization must obey no-harm and bounded-harm constraints using safe decision-making principles [31, 80]. **Traceability:** users should be able to inspect what changed, why it changed, and whether recommendation shifts were driven by data, assumptions, or scenario updates. Datasheets, model cards, and glass-box sustainability reports are therefore not peripheral documentation; they are core infrastructure joined to the intervention and audit graph [32, 39, 65]. This stack is also essential for high-spillover interventions, including geoengineering proposals, where detection, transboundary spillovers, and provenance are indispensable [67, 79].

5 Running Exemplars

Minimum viable deployment: cool roofs for extreme heat. A falsifiable city-scale PCA-OS can detect parcel-level cool-roof retrofits from EO albedo/thermal signals and permit text, write them to the ledger, estimate avoided heat and peak-demand effects with spillovers, and optimize subsidy rollout under budget, equity, and no-harm rules (Figure 3) [19, 48, 74]. **Flood defenses with spillovers.** Levees, drainage retrofits, and shoreline protection test interference-aware causal inference because they can redirect water and externalize harm [4]. **Urban greening and nature-based solutions.** Greening can reduce heat but also trigger green gentrification, requiring joint models of cooling, amenity spillovers, and distributive outcomes [3, 5]. **Compound environmental burdens.** The same flow can attach adaptation choices to air-quality and logistics-burden outcomes when heat, ozone, traffic, or delivery systems shift risks unevenly [20, 61, 82].

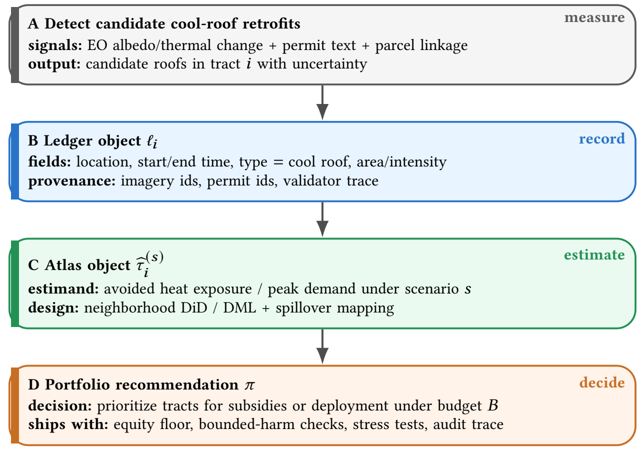

Figure 3: Mini-walkthrough for urban heat: EO and permit evidence become a cool-roof ledger object, a scenario-indexed effect estimate, and an auditable rollout recommendation.

6 Research Agenda & Evaluation

The near-term agenda has three pillars. **Planetary observability** should measure interventions through interoperable ledger schemas aligned with geospatial standards such as STAC [68]. **Causal topologies and transfer** should estimate spillover-aware effects while keeping identification assumptions explicit and scenario transportability uncertain [6, 25, 49, 70, 73]. **Normative optimization and governance** should support equity objectives, robust portfolios, and no-harm constraints [8, 11, 21, 24, 38, 56, 59]. We propose AdaptBench as the evaluation vehicle: systems should be tested on intervention mapping, causal estimation, and portfolio choice, not prediction alone. Appendix A specifies the minimal API contract and guardrails, while Appendices B–C summarize the benchmark and simulator bridge.

Concretely, each AdaptBench task should output an audit record linking evidence, ledger entry, effect object, portfolio decision, and failure labels, so systems can fail for untraceable, maladaptive, or harmful decisions despite accurate hazard prediction [7].

7 Conclusion

PCA-OS reframes climate ML from hazard prediction to auditable adaptation decision support. Its core claim is simple: adaptation needs shared intervention objects, scenario-indexed causal evidence, and robust, equity-constrained portfolio layers that remain contestable under uncertainty while preserving an auditable trace.

Acknowledgments

This research is supported by the RIE2025 Industry Alignment Fund–Industry Collaboration Projects (IAF-ICP) (Award I2301E0026), administered by A*STAR, and by Alibaba Group and NTU Singapore through the Alibaba–NTU Global e-Sustainability CorpLab (ANGEL).

A Supplementary Design Notes

Object contract. A minimal PCA-OS implementation should expose four versioned object types: *intervention evidence* with geometry, timing, intensity, uncertainty, and provenance; *effect evidence* with estimand, design, diagnostics, sensitivity, and validity envelope; *portfolio evidence* with objectives, constraints, selected actions, robust regret, and known harms; and *audit evidence* with data versions, model versions, user overrides, and open disputes. The important point is not a specific database choice, but that every recommendation can be traced from decision back to effect object and ledger object.

Minimal API contract. A minimal PCA-OS prototype should expose five audit operations: retrieve a ledger entry; compare ledger versions; query atlas effects by intervention, outcome, and scenario; trace a portfolio decision back to ledger and atlas objects; and explain active budget, equity, no-harm, and validity constraints. Each response should carry provenance, uncertainty, validation state, and known missingness rather than only a point recommendation.

Transportability checks. Before reusing an atlas effect in a new place or future scenario, the system should compare mechanism, exposure range, infrastructure regime, governance capacity, and observed rollout policy. If any of these differ sharply, the atlas entry should be demoted from a reusable point estimate to a stress-test input, bounded estimate, or unknown. This is the practical meaning of scenario-indexed causal evidence.

Failure modes and guardrails. The strongest early critique of a planetary adaptation OS is not technical infeasibility but false authority: a system could overstate sparse evidence, under-count informal adaptation, optimize away local priorities, or make maladaptation or climate gentrification look efficient [3, 7]. A publishable PCA-OS prototype should therefore surface missingness, source coverage, dispute status, validator identity, privacy sensitivity, and no-harm violations as first-class fields rather than footnotes. Missing informal adaptation, delayed procurement data, and inaccessible operational logs should remain visible metadata, not silent non-interventions. Portfolio recommendations should remain contestable artifacts, not automatic prescriptions.

B AdaptBench Task Card

Task families. A first AdaptBench release can contain: **(1) intervention mapping,** where systems produce calibrated ledger entries with uncertainty and provenance from EO, text, and operational signals; **(2) causal estimation,** where systems recover known or semi-synthetic effects under confounding, targeted rollout, and interference while exposing estimands, identification notes, and diagnostics; and **(3) portfolio choice,** where systems select interventions under budgets, equity floors, and no-harm constraints and are scored by robust regret across scenario ensembles.

Splits and metrics. Evaluation should stress generalization across *space* (held-out regions), *time* (future periods), and *policy regimes* (different rollout logics, budgets, and governance conditions). Because dense labels are rare, the suite should combine real retrospective tasks, semi-synthetic interventions on realistic EO backdrops, and known-mechanism decision tasks. Baselines should span EO segmentation or retrieval and document extraction for mapping, DiD/SC/DML for effect estimation, and robust or fairness-aware optimization for decisions. Metrics include ledger calibration, interval coverage, spillover error, RDR, inequity gaps, and bounded-harm violation rates.

Success criterion. Within a few benchmark cycles, success would include a public ledger schema and reference dataset for at least one intervention family, atlas tasks requiring estimands and diagnostics rather than point effects alone, and portfolio baselines evaluated on robust regret, inequity gaps, and bounded-harm violations.

C Simulator Bridge and Validation States

Simulator bridge. PCA-OS should treat process and engineering simulators as typed evidence services rather than competitors to the OS layer. River/floodplain-routing, building-energy, urban-canopy, and coupled human–water models can serve as typed evidence services that generate mechanism priors, semi-synthetic counterfactuals, and stress tests for atlas entries and portfolio choices [18, 62, 75, 83]. This is where PCA-OS differs from a standalone process simulator: simulators expose mechanisms and counterfactuals, while the OS layer standardizes versioned evidence objects, causal validity envelopes, and decision traces across simulators and observational data. The bridge contract is simple: a simulator call must declare intervention inputs, boundary conditions, scenario assumptions, calibration data, outputs, and validity limits, then write its result as an atlas-supporting evidence object rather than a final policy decision.

Ledger validation states. Each intervention record should move through explicit states: `candidate` when inferred from one source, `provisional` when cross-source support exists but uncertainty remains, `validated` after human or institutional confirmation, `disputed` when sources or stakeholders conflict, and `superseded` when geometry, timing, or intensity is revised. Portfolio queries should expose these states and optionally restrict decisions to validated or sensitivity-weighted records.

Validity envelope. Before transporting an effect, the atlas should compare mechanism, exposure range, rollout policy, infrastructure regime, governance capacity, and spillover topology. A mismatch does not delete evidence; it demotes the effect to a stress-test input or bounded estimate, preserving usefulness without false authority.

References

[1] Alberto Abadie, Alexis Diamond, and Jens Hainmueller. 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association* 105, 490 (2010), 493–505.

[2] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

[3] Isabelle Anguelovski, James JT Connolly, Hamil Pearsall, Galia Shokry, Melissa Checker, Juliana Maantay, Kenneth Gould, Tammy Lewis, Andrew Maroko, and J Timmons Roberts. 2019. Why green "climate gentrification" threatens poor and vulnerable populations. *Proceedings of the national academy of sciences* 116, 52 (2019), 26139–26143.

[4] Abolfazl Hojjat Ansari, Alfonso Mejia, and Raj Cibin. 2024. Flood teleconnections from levees undermine disaster resilience. *npj Natural Hazards* 1, 1 (2024), 2.

[5] Farshid Aram, Ester Higueras García, Ebrahim Solgi, and Soran Mansournia. 2019. Urban green space cooling effect in cities. *Heliyon* 5, 4 (2019).

[6] Peter M Aronow and Cyrus Samii. 2017. Estimating average causal effects under general interference, with application to a social network experiment. (2017).

[7] Jon Barnett and Saffron O'Neill. 2010. Maladaptation. *Global Environmental Change* 20, 2 (2010), 211–213.

[8] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. 2009. *Robust Optimization.* Princeton University Press.

[9] Lea Berrang-Ford, Robbert Biesbroek, James D Ford, Alexandra Lesnikowski, Andrew Tanabe, Frances M Wang, Chen Chen, Angel Hsu, Jessica J Hellmann, Patrick Pringle, et al. 2019. Tracking global climate change adaptation among governments. *Nature Climate Change* 9, 6 (2019), 440–449.

[10] Lea Berrang-Ford, AR Siders, Alexandra Lesnikowski, Alexandra Paige Fischer, Max W Callaghan, Neal R Haddaway, Katharine J Mach, Malcolm Araos, Mohammad Aminur Rahman Shah, Mia Wannewitz, et al. 2021. A systematic global stocktake of evidence on human adaptation to climate change. *Nature climate change* 11, 11 (2021), 989–1000.

[11] Dimitris Bertsimas and Melvyn Sim. 2004. The price of robustness. *Operations research* 52, 1 (2004), 35–53.

[12] Derrick Bonafilia, Beth Tellman, Tyler Anderson, and Erica Issenberg. 2020. Sen1Floods11: A georeferenced dataset to train and test deep learning flood algorithms for sentinel-1. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops.* 210–211.

[13] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.

[14] Tia Brullo, Jon Barnett, Elissa Waters, and Sarah Boulter. 2024. The enablers of adaptation: A systematic review. *NPJ Climate Action* 3, 1 (2024), 40.

[15] Brantly Callaway and Pedro HC Sant'Anna. 2021. Difference-in-differences with multiple time periods. *Journal of econometrics* 225, 2 (2021), 200–230.

[16] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal* 21, 1 (2018).

[17] Yezhen Cong, Samar Khanna, Chenlin Meng, Patrick Liu, Erik Rozi, Yutong He, Marshall Burke, David Lobell, and Stefano Ermon. 2022. Satmae: Pre-training transformers for temporal and multi-spectral satellite imagery. *Advances in Neural Information Processing Systems* 35 (2022), 197–211.

[18] Drury B Crawley, Linda K Lawrie, Frederick C Winkelmann, Walter F Buhl, Y Joe Huang, Curtis O Pedersen, Richard K Strand, Richard J Liesen, Daniel E Fisher, Michael J Witte, et al. 2001. EnergyPlus: creating a new-generation building energy simulation program. *Energy and buildings* 33, 4 (2001), 319–331.

[19] Susan L Cutter, Bryan J Boruff, and WLynn Shirley. 2003. Social Vulnerability to Environmental Hazards. *Social Science Quarterly* 84, 2 (2003), 242–261.

[20] Ruijun Dang, Hong Liao, and Yu Fu. 2021. Quantifying the anthropogenic and meteorological influences on summertime surface ozone in China over 2012–2017. *Science of the Total Environment* 754 (2021), 142394.

[21] Erick Delage and Yinyu Ye. 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research* 58, 3 (2010), 595–612.

[22] Francisco J Doblas-Reyes, Jenni Kontkanen, Irina Sandu, Mario Acosta, Mohammed Hussam Al Turjmam, Ivan Alsina-Ferrer, Miguel Andrés-Martínez, Costanza Anerdi, Leo Arriola, Marvin Axness, et al. 2026. The Destination Earth digital twin for climate change adaptation. *Geoscientific Model Development* 19, 7 (2026), 2821–2848.

[23] Miroslav Dudík, John Langford, and Lihong Li. 2011. Doubly robust policy evaluation and learning. In *Proceedings of the 28th International Conference on International Conference on Machine Learning.* 1097–1104.

[24] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference.* 214–226.

[25] Veronika Eyring, Sandrine Bony, Gerald A Meehl, Catherine A Senior, Bjorn Stevens, Ronald J Stouffer, and Karl E Taylor. 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* 9, 5 (2016), 1937–1958.

[26] Shibo Feng, Chunyan Miao, Ke Xu, Jiaxiang Wu, Pengcheng Wu, Yang Zhang, and Peilin Zhao. 2024. Multi-Scale Attention Flow for Probabilistic Time Series Forecasting. *IEEE Transactions on Knowledge & Data Engineering* 36, 05 (2024), 2056–2068.

[27] Shibo Feng, Chunyan Miao, Zhong Zhang, and Peilin Zhao. 2024. Latent diffusion transformer for probabilistic time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 11979–11987.

[28] Shibo Feng, Peilin Zhao, Liu Liu, Pengcheng Wu, and Zhiqi Shen. 2025. HDT: hierarchical discrete transformer for multivariate time series forecasting. In *Proceedings of the Thirty-Ninth AAAI Conference on Artificial Intelligence and Thirty-Seventh Conference on Innovative Applications of Artificial Intelligence and Fifteenth Symposium on Educational Advances in Artificial Intelligence.* 746–754.

[29] James D Ford and Lea Berrang-Ford. 2016. The 4Cs of adaptation tracking: consistency, comparability, comprehensiveness, coherency. *Mitigation and adaptation strategies for global change* 21, 6 (2016), 839–859.

[30] James D Ford, Simon E Tilleard, Lea Berrang-Ford, Malcolm Araos, Robbert Biesbroek, Alexandra C Lesnikowski, Graham K MacDonald, Angel Hsu, Chen Chen, and Livia Bizikova. 2016. Big data has big potential for applications to climate change adaptation. *Proceedings of the National Academy of Sciences* 113, 39 (2016), 10729–10732.

[31] Javier García and Fernando Fernández. 2015. A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research* 16, 1 (2015), 1437–1480.

[32] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. Datasheets for datasets. *Commun. ACM* 64, 12 (2021), 86–92.

[33] Rohit Girdhar, Alaaeldin El-Nouby, Zhuang Liu, Mannat Singh, Kalyan Vasudev Alwala, Armand Joulin, and Ishan Misra. 2023. Imagebind: One embedding space to bind them all. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* 15180–15190.

[34] Global Commission on Adaptation. 2019. Adapt Now: A Global Call for Leadership on Climate Resilience. https://gca.org/reports/adapt-now-a-global-call-for-leadership-on-climate-resilience/.

[35] Marjolijn HAASNOOT, Jan H KWAKKEL, Warren E WALKER, and Judith TER MAAT. 2013. Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global environmental change* 23, 2 (2013), 485–498.

[36] Mordechai Haklay and Patrick Weber. 2008. Openstreetmap: User-generated street maps. *IEEE Pervasive computing* 7, 4 (2008), 12–18.

[37] Stephane Hallegatte, Colin Green, Robert J Nicholls, and Jan Corfee-Morlot. 2013. Future flood losses in major coastal cities. *Nature climate change* 3, 9 (2013), 802–806.

[38] Moritz Hardt, Eric Price, and Nati Srebro. 2016. Equality of opportunity in supervised learning. *Advances in neural information processing systems* 29 (2016).

[39] Chaoyue He, Xin Zhou, Di Wang, Hong Xu, Wei Liu, and Chunyan Miao. 2026. ESGlass: Glass-Box ESG and Sustainability Reports. *Preprints* (2026).

[40] Chaoyue He, Xin Zhou, Di Wang, Hong Xu, Wei Liu, and Chunyan Miao. 2026. Harness Engineering for Language Agents: The Harness Layer as Control, Agency, and Runtime. *Preprints* (2026).

[41] Chaoyue He, Xin Zhou, Di Wang, Hong Xu, Wei Liu, and Chunyan Miao. 2026. OpenClaw as Language Infrastructure: A Case-Centered Survey of a Public Agent Ecosystem in the Wild. *Preprints* (2026).

[42] Chaoyue He, Xin Zhou, Di Wang, Xinjia Yu, Lei Xiao, Langyue Li, Hong Xu, Wei Liu, and Chunyan Miao. 2026. KG4ESG: The ESG Knowledge Graph Atlas. *Preprints* (2026).

[43] Chaoyue He, Xin Zhou, Yi Wu, Xinjia Yu, Yan Zhang, Lei Zhang, Di Wang, Shengfei Lyu, Hong Xu, Wang Xiaoqiao, et al. 2025. Esgenius: Benchmarking llms on environmental, social, and governance (esg) and sustainability knowledge. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing.* 14612–14653.

[44] Chaoyue He, Xin Zhou, Xinjia Yu, Lei Zhang, Yan Zhang, Yi Wu, Lei Xiao, Liangyue Li, Di Wang, Hong Xu, et al. 2026. SSKG Hub: An Expert-Guided Platform for LLM-Empowered Sustainability Standards Knowledge Graphs. *arXiv preprint arXiv:2603.00669* (2026).

[45] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. 2020. The ERA5 global reanalysis. *Quarterly journal of the royal meteorological society* 146, 730 (2020), 1999–2049.

[46] Jörn Hoffmann, Peter Bauer, Irina Sandu, Nils Wedi, Thomas Geenen, and Daniel Thiemert. 2023. Destination Earth-A digital twin in support of climate services. *Climate Services* 30 (2023), 100394.

[47] Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard De Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, et al. 2021. Knowledge graphs. *ACM Computing Surveys (Csur)* 54, 4 (2021), 1–37.

[48] Angel Hsu, Glenn Sheriff, Tirthankar Chakraborty, and Diego Manya. 2021. Disproportionate exposure to urban heat island intensity across major US cities. *Nature communications* 12, 1 (2021), 2721.

[49] Michael G Hudgens and M Elizabeth Halloran. 2008. Toward causal inference with interference. *Journal of the american statistical association* 103, 482 (2008), 832–842.

[50] Guido W Imbens and Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences.* Cambridge university press.

[51] IPCC. 2022. Climate Change 2022: Impacts, Adaptation and Vulnerability (AR6 Working Group II) — Summary for Policymakers. https://www.ipcc.ch/report/ar6/wg2/chapter/summary-for-policymakers/.

[52] Andrew Jaegle, Sebastian Borgeaud, Jean-Baptiste Alayrac, Carl Doersch, Catalin Ionescu, David Ding, Skanda Koppula, Daniel Zoran, Andrew Brock, Evan Shelhamer, et al. 2021. Perceiver io: A general architecture for structured inputs & outputs. *arXiv preprint arXiv:2107.14795* (2021).

[53] Yicheng Ji, Jun Zhang, Heming Xia, Jinpeng Chen, Lidan Shou, Gang Chen, and Huan Li. 2025. Specvlm: Enhancing speculative decoding of video llms via verifier-guided token pruning. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing.* 7205–7219.

[54] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. 2021. Physics-informed machine learning. *Nature Reviews*

Physics 3, 6 (2021), 422–440.
- [55] Anuj Karpatne, Gowtham Atluri, James H Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar. 2017. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on knowledge and data engineering* 29, 10 (2017), 2318–2331.
- [56] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. 2016. Inherent trade-offs in the fair determination of risk scores. *arXiv preprint arXiv:1609.05807* (2016).
- [57] Alexandre Lacoste, Nils Lehmann, Pau Rodriguez, Evan Sherwin, Hannah Kerner, Björn Lütjens, Jeremy Irvin, David Dao, Hamed Alemohammad, Alexandre Drouin, et al. 2023. Geo-bench: Toward foundation models for earth monitoring. *Advances in Neural Information Processing Systems* 36 (2023), 51080–51093.
- [58] Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, et al. 2023. Learning skillful medium-range global weather forecasting. *Science* 382, 6677 (2023), 1416–1421.
- [59] Robert J. Lempert, Steven W. Popper, and Steven C. Bankes. 2003. *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*. RAND Corporation.
- [60] Alexandra C Lesnikowski, James D Ford, Lea Berrang-Ford, Magda Barrera, and Jody Heymann. 2015. How are we adapting to climate change? A global assessment. *Mitigation and Adaptation Strategies for Global Change* 20, 2 (2015), 277–293.
- [61] Baojie Li, Hong Liao, Ke Li, Jintai Lin, Cheng Gong, Huan Liu, Yan Li, Lei Chen, Yang Yang, Xipeng Jin, et al. 2025. Environmental burden and health inequity in China's road-based express delivery. *Nature Cities* 2, 9 (2025), 825–834.
- [62] David Meyer, Robert Schoetter, Valéry Masson, and Sue Grimmond. 2020. Enhanced software and platform for the Town Energy Balance (TEB) model. *Journal of Open Source Software* 5, 50 (2020), 2008.
- [63] Microsoft. 2021. A Planetary Computer for a Sustainable Future. https://planetarycomputer.microsoft.com/.
- [64] Paul CD Milly, Julio Betancourt, Malin Falkenmark, Robert M Hirsch, Zbigniew W Kundzewicz, Dennis P Lettenmaier, and Ronald J Stouffer. 2008. Stationarity is dead: Whither water management? *Science* 319, 5863 (2008), 573–574.
- [65] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*. 220–229.
- [66] NASA Earthdata. 2025. Harmonized Landsat and Sentinel-2 (HLS). https://www.earthdata.nasa.gov/data/projects/hls.
- [67] Engineering National Academies of Sciences, Medicine, et al. 2021. *Reflecting sunlight: Recommendations for solar geoengineering research and research governance*.
- [68] Open Geospatial Consortium (OGC). 2025. SpatioTemporal Asset Catalog (STAC) Community Standard (OGC 25-004). https://docs.ogc.org/cs/25-004/25-004.html.
- [69] Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, Thorsten Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, et al. 2022. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214* (2022).
- [70] Judea Pearl. 2009. *Causality*. Cambridge university press.
- [71] Judea Pearl and Elias Bareinboim. 2011. Transportability of causal and statistical relations: A formal approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 25. 247–254.
- [72] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International conference on machine learning*. PmLR, 8748–8763.
- [73] Keywan Riahi, Detlef P Van Vuuren, Elmar Kriegler, Jae Edmonds, Brian C O'neill, Shinichiro Fujimori, Nico Bauer, Katherine Calvin, Rob Dellink, Oliver Fricko, et al. 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change* 42 (2017), 153–168.
- [74] Mattheos Santamouris. 2014. Cooling the cities–a review of reflective and green roof mitigation technologies to fight heat island and improve comfort in urban environments. *Solar energy* 103 (2014), 682–703.
- [75] Murugesu Sivapalan, Hubert HG Savenije, Günter Blöschl, et al. 2012. Socio-hydrology: A new science of people and water. *Hydrological Processes* 26, 8 (2012), 1270–1276.
- [76] Daniela Szwarcman, Sujit Roy, Paolo Fraccaro, Orsteinn Elí Gíslason, Benedikt Blumenstiel, Rinki Ghosal, Pedro Henrique De Oliveira, Joao Lucas de Sousa Almeida, Rocco Sedona, Yanghui Kang, et al. 2025. Prithvi-eo-2.0: A versatile multi-temporal foundation model for earth observation applications. *IEEE Transactions on Geoscience and Remote Sensing* (2025).
- [77] Andrew J Tatem. 2017. WorldPop, open data for spatial demography. *Scientific data* 4, 1 (2017), 170004.
- [78] UN Environment Programme (UNEP). 2024. Adaptation Gap Report 2024: Come Hell and High Water - As Fires and Floods Hit the Poor Hardest, it is time for the World to Step up Adaptation Actions. https://wedocs.unep.org/items/b8964bdc-a2f8-4153-96fd-a1e4141bff75.
- [79] David L VanderZwaag and Abdul Hafez Mahamah. 2024. International governance of marine geoengineering: sketchy seascape, foggy future—an essay in honor of Ted L. McDorman. *Ocean Development & International Law* 55, 4 (2024), 624–636.
- [80] Akifumi Wachi, Xun Shen, and Yanan Sui. 2024. A survey of constraint formulations in safe reinforcement learning. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*. 8262–8271.
- [81] Stefan Wager and Susan Athey. 2018. Estimation and inference of heterogeneous treatment effects using random forests. *J. Amer. Statist. Assoc.* 113, 523 (2018), 1228–1242.
- [82] Pinya Wang, Yang Yang, Huimin Li, Lei Chen, Ruijun Dang, Daokai Xue, Baojie Li, Jianping Tang, L Ruby Leung, and Hong Liao. 2022. North China Plain as a hot spot of ozone pollution exacerbated by extreme high temperatures. *Atmospheric Chemistry and Physics* 22, 7 (2022), 4705–4719.
- [83] Dai Yamazaki. 2025. Advancing global river hydrodynamics simulations by catchment-based macro-scale floodplain modeling approach. *Geoscience Letters* 12, 1 (2025), 72.
- [84] Jun Zhang, Yicheng Ji, Feiyang Ren, Yihang Li, Bowen Zeng, Zonghao Chen, Ke Chen, Lidan Shou, Gang Chen, and Huan Li. 2026. Efficient Inference for Large Vision-Language Models: Bottlenecks, Techniques, and Prospects. *arXiv preprint arXiv:2604.05546* (2026).
- [85] Lei Zhang, Xin Zhou, Chaoyue He, Di Wang, Yi Wu, Hong Xu, Wei Liu, and Chunyan Miao. 2025. Mmesgbench: Pioneering multimodal understanding and complex reasoning benchmark for esg tasks. In *Proceedings of the 33rd ACM International Conference on Multimedia*. 12829–12836.