

Sustainability in the Loop: AI model development should be multi-objective

Matteo Mugnai, Francesco Pistolesi

Department of Information Engineering
University of Pisa, Italy
matteo.mugnai@ing.unipi.it, francesco.pistolesi@unipi.it

The environmental cost of AI is growing

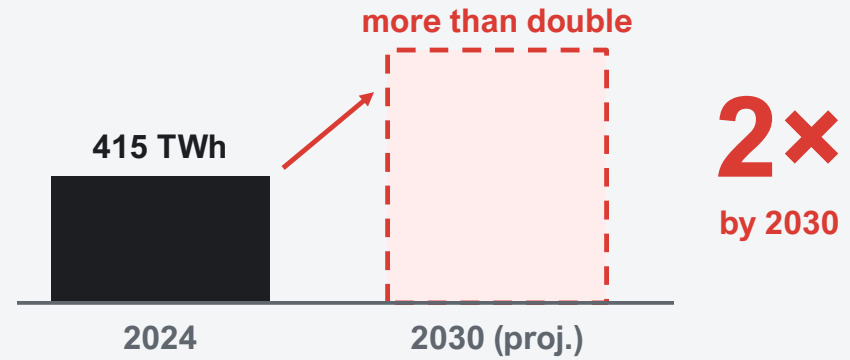
Energy, carbon emissions, and water use **must be part of AI development**, not just its consequences

415 TWh

data center electricity use in 2024

≈ **1.5%** of global electricity demand

AI progress comes at a cost!



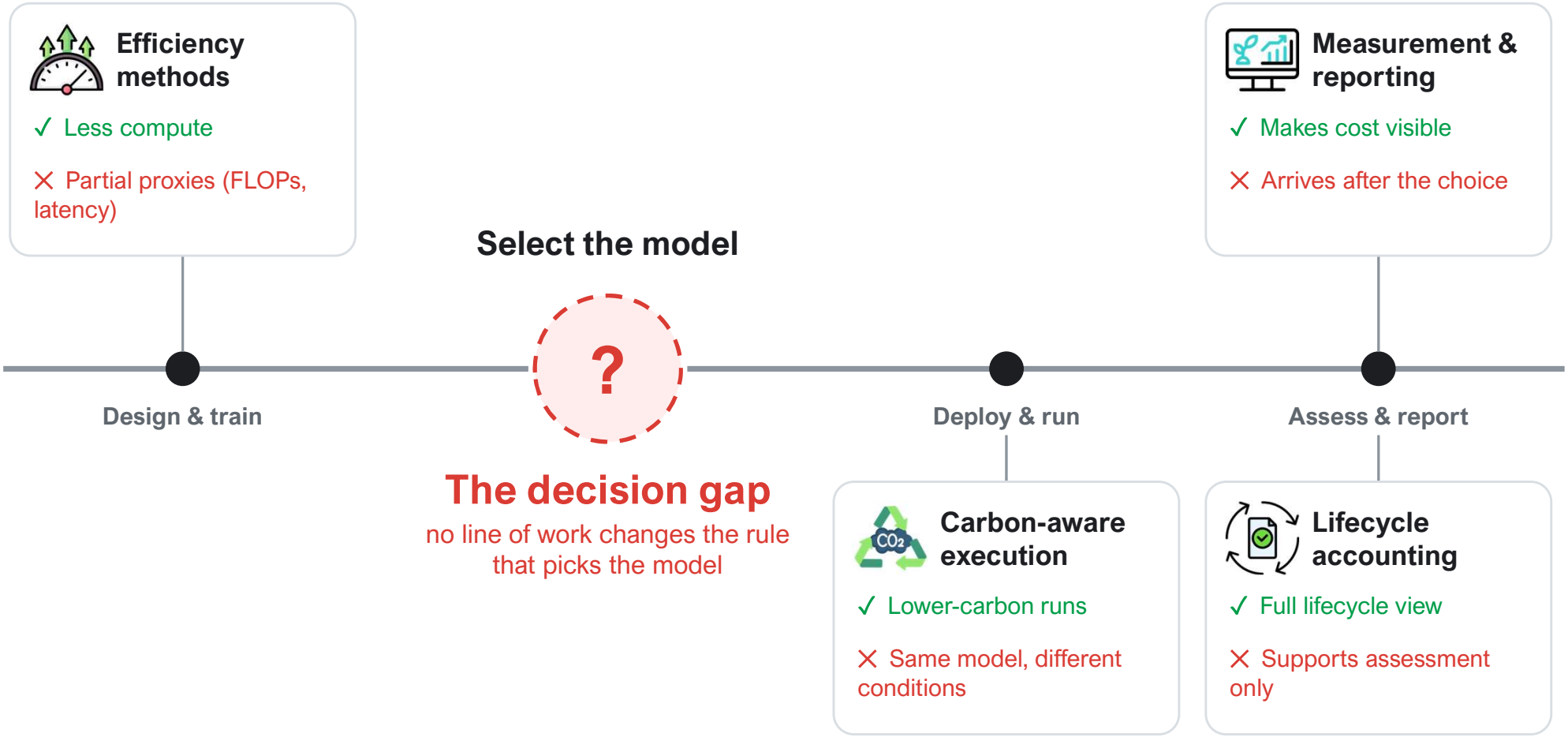
WHAT MODEL SELECTION OPTIMIZES TODAY

Accuracy | Latency | Monetary cost

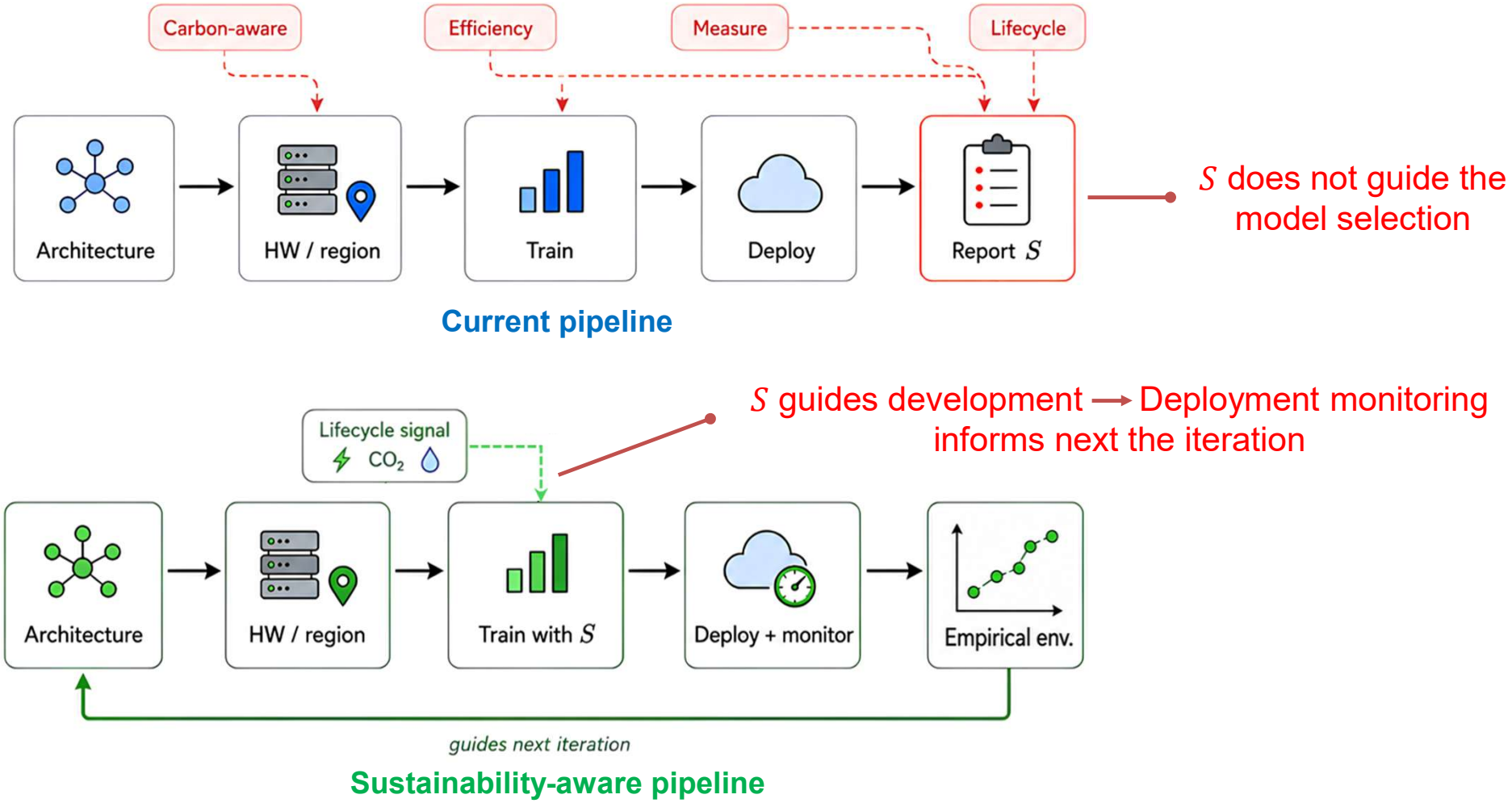
…left outside

- Energy
- CO_2 emissions
- Water use

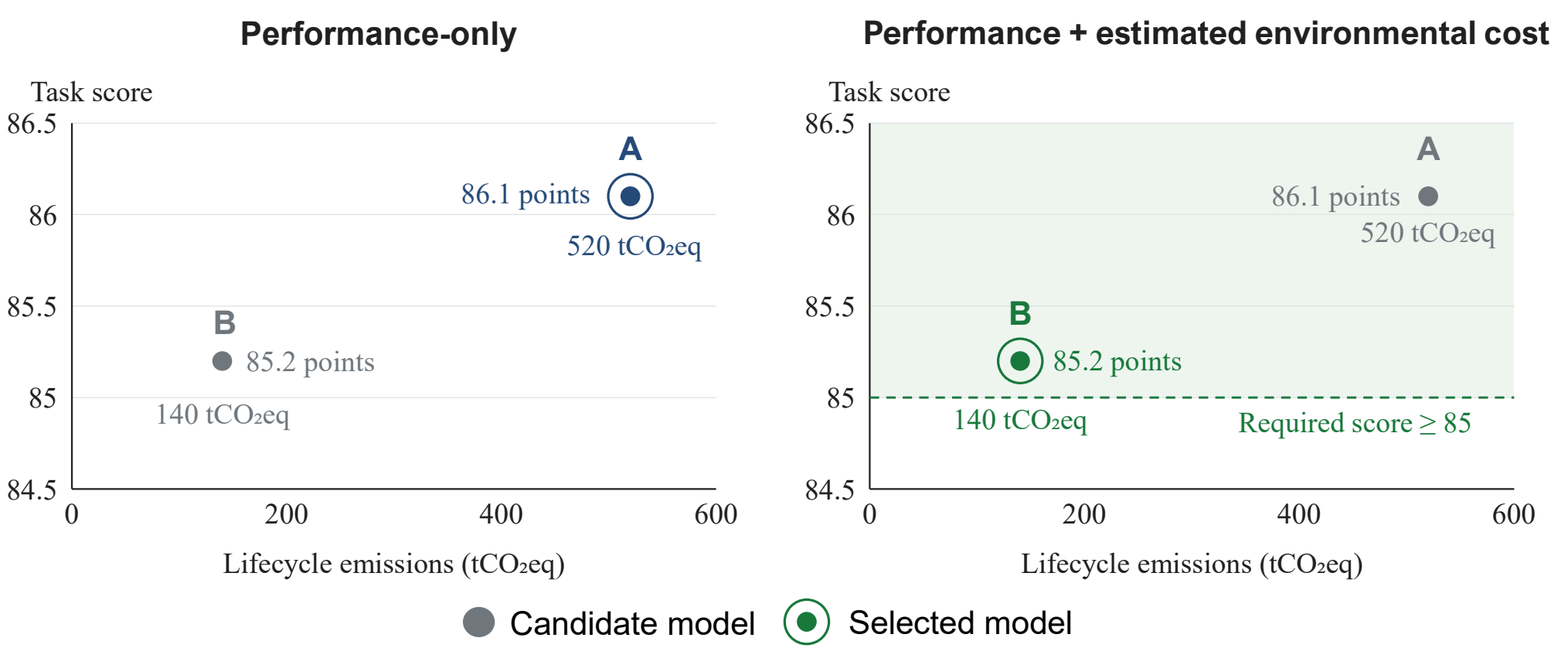
The decision gap



Model development should be multi-objective



What changes in model selection?



tCO_2eq: tonnes of CO_2 equivalent

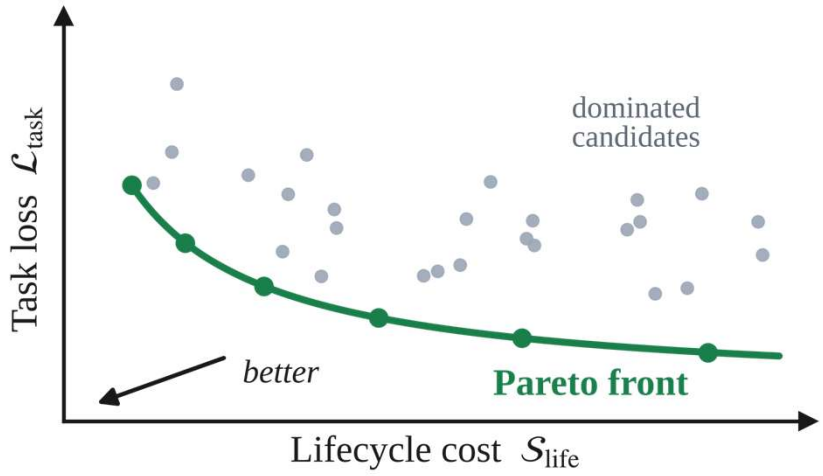
B: −0.9 points, −73% emissions

A possible formulation

$$\mathcal{S}_{\text{life}}(\boldsymbol{\theta}, \boldsymbol{\pi}) = \mathcal{S}_{\text{train}}(\boldsymbol{\theta}, \boldsymbol{\pi}) + V \cdot \mathcal{S}_{\text{query}}(\boldsymbol{\theta}, \boldsymbol{\pi})$$

$$\min_{\boldsymbol{\theta}, \boldsymbol{\pi}} \left(\mathcal{L}_{\text{task}}(\boldsymbol{\theta}, \boldsymbol{\pi}),\ \mathcal{S}_{\text{life}}(\boldsymbol{\theta}, \boldsymbol{\pi}) \right)$$

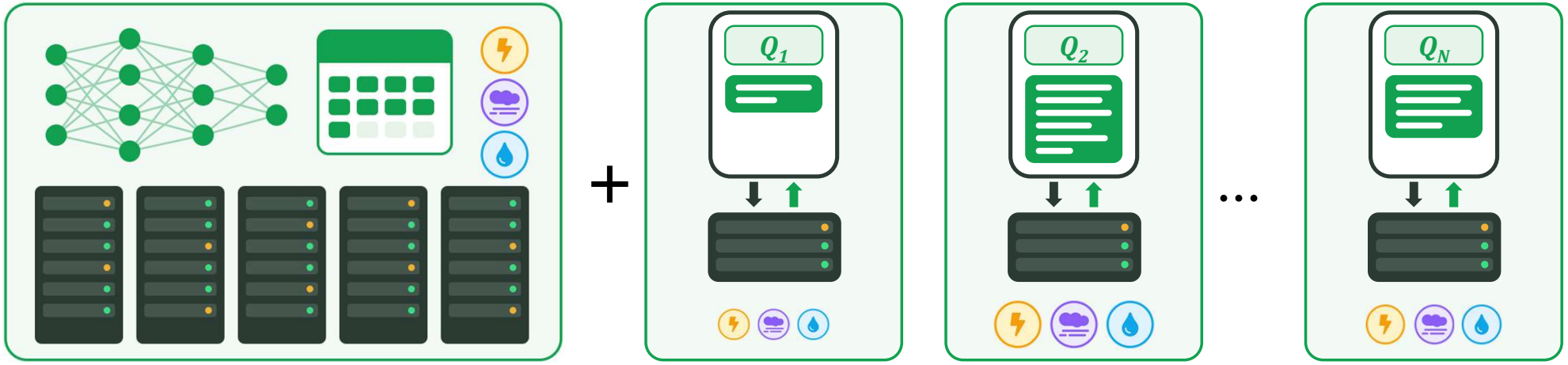
$\boldsymbol{\theta}$: model parameters
$\boldsymbol{\pi}$: controllable design, architecture and deployment choices



A possible formulation

S_{life} is the total environmental cost, from training to the last query

energy CO_2 water



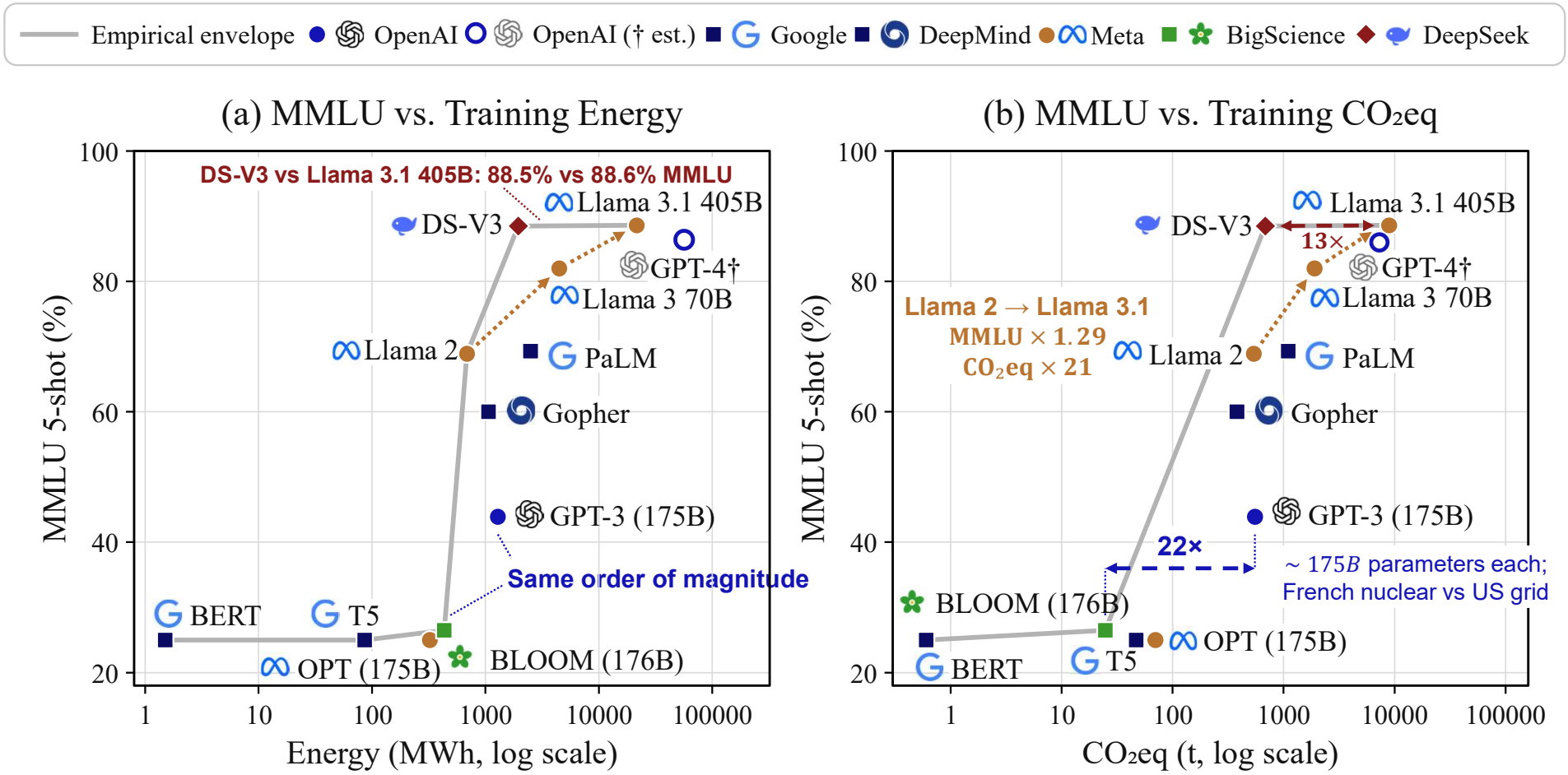
S_{train} + $V \cdot S_{query}$

Training the model
weeks of GPUs, paid once

Answering one question
a small cost, paid at every query

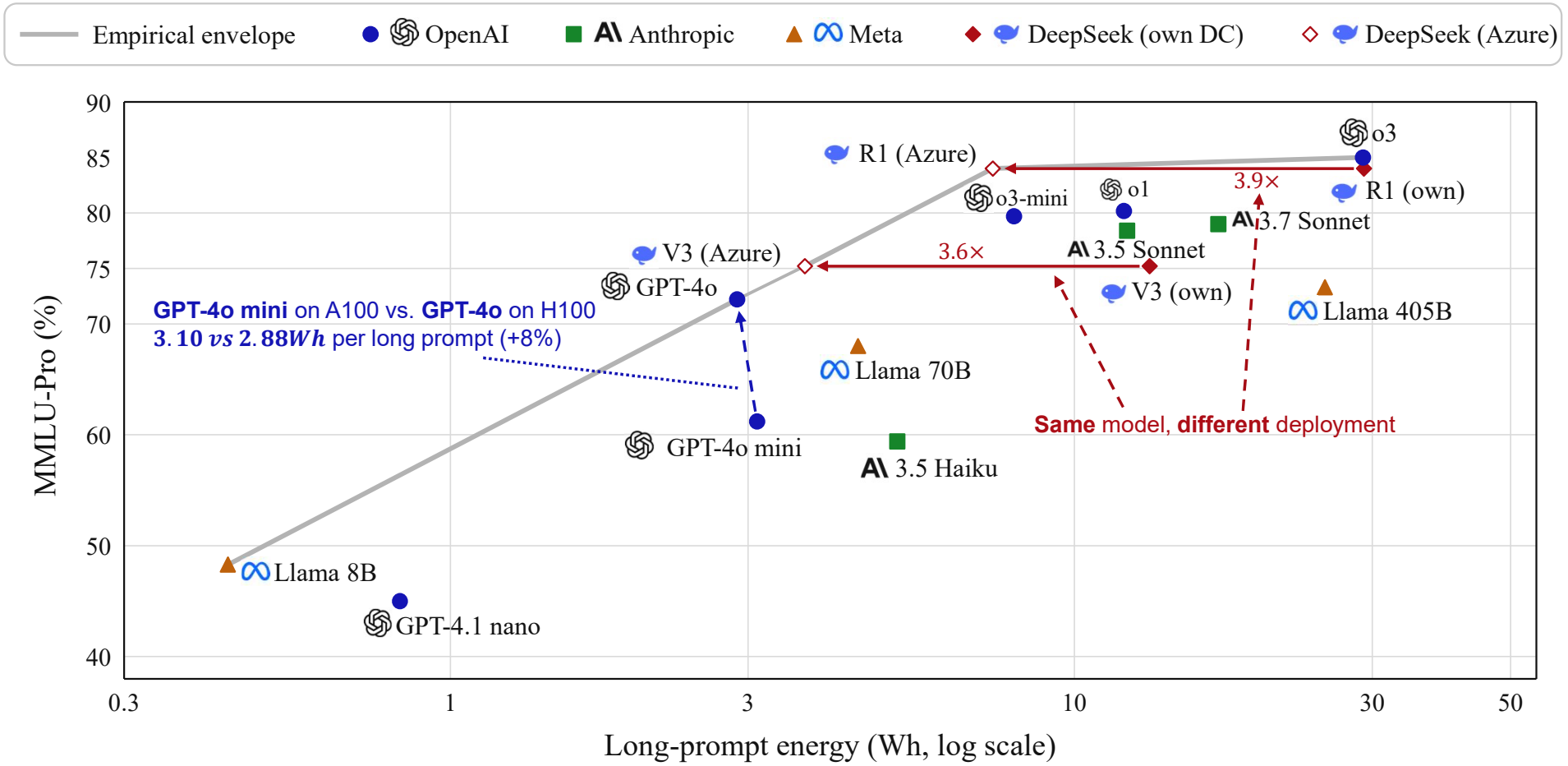
↓ **Optimization:** reduce both terms together with the task loss, acting on θ and π

Lifecycle evidence: Training



† GPT-4: third-party estimate; * Family-level footprint, Llama 2 → Llama 3.1: 539 → 11,390 tCO_2eq; BERT, T5, OPT: MMLU not evaluated in the original release

Lifecycle evidence: Inference



Energy measurements: Jegham et al. [26]. MMLU-Pro scores: model cards and technical reports. Values from Table 3.

Alternative views

REPORTING

"Reporting is enough"

Environmental cost must shape model selection

COMPUTE

"A compute budget is enough"

Compute ≠ lifecycle environmental impact

SYSTEMS

"It is a systems problem"

Improve systems and model choices together

METRICS

"No universal metric exists"

Use separate objectives for energy, carbon, water

UNCERTAINTY

"Costs are too uncertain"

Compare estimates under shared assumptions

QUALITY

"Sustainability lowers quality"

Make the performance–cost trade-off explicit

Conclusions

Performance-only approaches can generate **significant environmental costs**

Similar-scale models can have very **different carbon footprints**

Deployment choices strongly affect inference energy

Environmental cost must become part of how we define **AI progress**