

43. The Solland Paradox

An AI-Specific Rebound Hypothesis

Status: Public report

Date: 30 August 2026

Author: Njål Gaute Solland

Executive summary

AI efficiency can reduce the resource required for one unit of intelligence while increasing the total resources consumed by intelligence as a whole.

The mechanism is straightforward:

```
more intelligence per dollar / joule
      ↓
more economically viable use cases
      ↓
more queries, agents, reasoning and automation
      ↓
total demand grows faster than efficiency
      ↓
aggregate compute and energy consumption rises
```

This report calls the AI-specific formulation the **Solland Paradox**.

It does **not** claim that the underlying rebound mechanism is new. It is a direct relative of the Jevons paradox, and McKinsey explicitly applies Jevons to AI infrastructure in its July 2026 analysis of next-generation data center power systems.

The proposed contribution is narrower: a testable formulation for AI in which falling cost per unit of useful intelligence expands both the **number of tasks** and the **intelligence intensity per task**, causing total compute demand to rise even as individual operations become more efficient.

1. The paradox

Suppose one AI task requires 100 units of compute.

An efficiency improvement reduces that to 50 units.

If demand stays constant, total compute use falls by half.

But the lower cost may make four times as many tasks economical.

Then total use becomes:

$$50 \text{ units} \times 4 \text{ tasks} = 200 \text{ units}$$

Per-task consumption fell 50%.

Aggregate consumption doubled.

That is the paradox.

2. The AI-specific extension

AI has an unusually strong rebound channel because efficiency does more than make an existing task cheaper.

It can create new categories of activity.

Lower inference cost enables:

- longer context;
- deeper reasoning;
- more retries;
- multi-agent systems;
- continuous monitoring;
- personalized generation;
- real-time translation;
- autonomous workflows;
- background assistants;
- synthetic data generation;
- simulation;
- machine customers;
- edge inference.

The demand expansion is therefore both **extensive** and **intensive**.

Extensive margin

More tasks use AI.

Intensive margin

Each task can use more AI.

This is stronger than a simple substitution effect.

3. A minimal formal model

Let:

- (E) = useful intelligence produced per unit of physical compute;
- (Q(E)) = total demanded units of useful intelligence;
- (R) = physical compute resources consumed.

Then:

$$[R(E) = \frac{Q(E)}{E}]$$

Efficiency rises when (E) rises.

Total physical resource use rises when demand grows faster than efficiency:

$$[\frac{d \ln Q}{d \ln E} > 1]$$

This is the **Solland condition**.

If demand elasticity with respect to effective AI efficiency exceeds one, aggregate physical resource use rises despite better efficiency.

If it is below one, aggregate resource use falls.

The hypothesis is therefore falsifiable.

4. Why AI may satisfy the condition

AI is not a mature single-purpose technology.

Its demand frontier is still expanding.

Every reduction in effective cost can unlock:

1. a previously uneconomic workflow;
2. a longer reasoning horizon;
3. a higher-frequency execution loop;
4. more users;
5. more autonomous runtime;
6. more machine-to-machine demand.

This means the relevant demand curve is not fixed.

Efficiency changes the shape of the market itself.

5. Evidence consistent with the hypothesis

The current infrastructure forecasts are consistent with strong rebound.

The IEA expects global data center electricity consumption to more than double to around **945 TWh by 2030** in its base case even as hardware and software efficiency continue to improve.

JLL projects global data center capacity to rise from roughly **103 GW in 2025 to 200 GW in 2030**.

McKinsey's 2026 workload model projects AI inference demand to grow rapidly enough to become the dominant AI workload before 2030.

McKinsey also explicitly invokes the Jevons paradox in its July 2026 analysis of 800 VDC data center architecture: more efficient computing enables more applications and economic value, which accelerates adoption and increases compute demand.

These observations do not prove the Solland condition.

They show that the direction of current market forecasts is consistent with it.

6. Why inference matters

Training demand is concentrated and episodic.

Inference demand scales with use.

Every additional user, agent, transaction, sensor, workflow or autonomous process can generate recurring inference.

That makes inference particularly exposed to rebound.

A model that becomes ten times cheaper does not necessarily reduce total spending by 90%.

It can instead move from:

- occasional use → continuous use;
- one response → multi-step reasoning;
- one agent → agent swarms;
- human-triggered → always-on;
- enterprise-only → embedded everywhere.

The unit cost falls while the number of units explodes.

7. The second-order effect: infrastructure scarcity

If the Solland condition holds, efficiency can worsen physical scarcity in the short and medium term.

The sequence is:

```
AI efficiency ↑  
effective price of intelligence ↓  
demand for useful intelligence ↑↑  
aggregate compute ↑  
power-ready capacity tightens  
infrastructure value rises
```

This links the paradox directly to the thesis in Report 42.

Compute becomes abundant in economic terms while power-ready infrastructure remains scarce in physical terms.

8. The third-order effect: efficiency shifts where value accrues

If useful intelligence becomes cheaper, raw compute margins may compress.

At the same time, value can migrate toward constraints that do not improve at semiconductor speed:

- power;
- interconnection;
- land;
- cooling;
- trusted data;
- distribution;
- proprietary workflows;
- authority;
- high-quality outcomes.

The paradox therefore has a strategic implication:

The faster intelligence itself commoditizes, the more valuable non-commoditizing constraints can become.

9. Distinguishing the hypothesis from Jevons

The underlying rebound principle is established economics.

The AI-specific formulation differs in three ways.

9.1 Intelligence intensity is elastic

AI can use more computation to improve reasoning depth, search, verification and personalization.

9.2 AI creates machine demand

Demand does not have to originate from humans. Agents can trigger other agents, tools, simulations and inference loops.

9.3 AI expands the task frontier

Cheaper intelligence can make entirely new activities economical, rather than merely increasing consumption of an existing product.

The claim is not "Jevons is wrong or new."

The claim is:

AI may be an unusually strong, measurable instance of rebound because both the number of tasks and compute intensity per task respond to falling intelligence cost.

10. Measurement

The hypothesis should be tested with four time series.

10.1 Efficiency

Useful output per:

- GPU-second;
- joule;
- dollar;
- rack;
- unit of latency.

10.2 Unit price

Market price per comparable unit of useful intelligence.

Examples:

- cost per million quality-adjusted tokens;
- cost per solved task;
- cost per agent-hour;
- cost per verified completion.

10.3 Demand

Total useful intelligence consumed:

- inference tokens;
- agent runtime;
- task completions;
- reasoning depth;
- model calls;
- active AI workflows.

10.4 Physical consumption

- accelerator-hours;
- data center IT load;
- electricity;
- installed MW/GW;
- network throughput.

The core empirical test is whether:

[\text{growth in useful-intelligence demand}

| \text{growth in efficiency}]

over a meaningful period.

11. Falsification

The Solland Paradox is weakened if:

1. efficiency improves faster than total AI demand;
2. aggregate AI electricity use falls while adoption grows;
3. declining inference cost produces little new usage;
4. model calls and agent runtime saturate quickly;
5. productivity gains substitute for, rather than induce, new compute-intensive activity.

It is strengthened if:

1. unit cost falls while total compute spend rises;
2. energy use grows despite large efficiency gains;
3. inference volume grows faster than hardware efficiency;
4. agentic systems add persistent machine-generated demand;
5. data center scarcity remains high despite generational hardware improvement.

12. Conclusion

The central proposition is simple:

Cheaper intelligence can increase the total amount of intelligence the economy chooses to consume faster than technology reduces the resources required to produce each unit.

That is the Soland Paradox.

The name applies to the AI-specific formulation and test condition, not to the underlying rebound principle.

If the demand response remains above the efficiency response, AI will continue to reduce the cost of intelligence while increasing the total physical infrastructure required to deliver it.

The paradox is therefore not that efficiency failed.

It is that efficiency succeeded so well that demand outran it.

Sources

- International Energy Agency, Energy and AI — Energy demand from AI: <https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai>
- Lawrence Berkeley National Laboratory, United States Data Center Energy Usage Report: 2025 Update: <https://bies.lbl.gov/publications/united-states-data-center-energy-2025>
- JLL, 2026 Global Data Center Market Outlook: <https://www.jll.com/en-us/insights/market-outlook/data-center-outlook>
- McKinsey, The shift to 800-volt DC at data centers: Implications for providers, 30 July 2026: <https://www.mckinsey.com/industries/industrials/our-insights/the-shift-to-800-volt-dc-at-data-centers-implications-for-providers>
- McKinsey, The future of AI workloads, 24 February 2026: <https://www.mckinsey.com/featured-insights/charts/the-future-of-ai-workloads>

This report is research, not investment advice.