

---

# Rashomon: Multiple Views of the Same AI Timeline Forecasts<sup>1</sup>

---

Subramanyam Sahoo  
BASIS

[GitHub CODE](#)

With  
Apart Research

## Abstract

AI timeline forecasts diverge dramatically, with predictions ranging from 2027 to beyond 2050 for transformative AI. This divergence creates challenges for governance, research prioritization, and societal preparedness. We present a novel mathematical framework for decomposing disagreement between forecasts into quantifiable parameter-level contributions. By analyzing five major forecasts (AI 2027, Biological Anchors, Epoch 2040, Meticulous Median, and Conservative estimates), we find that (1) 99.9% of disagreement concentrates in just three parameters: TAI threshold definitions (58.9%), algorithmic efficiency expectations (21.1%), and compute growth assumptions (19.9%); (2) only 19.9% of disagreement stems from policy-relevant parameters, suggesting limited governance leverage on timelines; and (3) forecasts exhibit surprising consensus on 70%+ of underlying assumptions despite surface-level disagreement. Our framework uses sensitivity analysis and variance decomposition to attribute timeline differences to specific parameter values, classifying uncertainties by actionability (policy-relevant, empirically resolvable, or fundamentally uncertain). We provide open-source tools for ongoing forecast comparison and generate actionable policy recommendations. This work addresses a critical gap in AI forecasting literature: moving beyond "who is right" to understanding "where and why forecasters disagree," enabling more productive discourse and targeted uncertainty reduction efforts.

*Keywords: AI forecasting, timeline prediction, disagreement decomposition, variance attribution, AI governance, uncertainty quantification*

---

<sup>1</sup> Research conducted at the [\[The AI Forecasting Hackathon\]](#), 2025

# 1. Introduction

**1.1 Motivation** - The trajectory of artificial intelligence development represents one of the most consequential questions for humanity's future. Yet current forecasts diverge wildly: Aschenbrenner's AI 2027 scenario predicts superintelligence by late 2027 [1], while Cotra's Biological Anchors framework suggests a median timeline around 2036 [2], Epoch AI estimates 2040 [3], and conservative researchers argue for 2050 or beyond [4]. This 28-year spread creates a "Tower of Babel" problem where policymakers, researchers, and the public struggle to extract actionable insights from conflicting expert opinions.

Traditional approaches to reconciling forecasts—aggregation methods like prediction markets [5], expert surveys [6], or Delphi techniques [7]—treat forecasts as black boxes, pooling predictions without understanding why they differ. This obscures critical questions: Which disagreements matter most? Which uncertainties are resolvable through better measurement? Where can governance actually influence outcomes?

**1.2 Our Contribution** - We introduce a disagreement decomposition framework that treats forecasts as structured parametric models and attributes timeline differences to specific assumption divergences. Our key contributions: Novel methodology: Variance decomposition approach that quantifies how much each parameter contributes to forecast disagreement Empirical analysis: First systematic decomposition of major AI timeline forecasts Actionability classification: Taxonomy separating policy-relevant, empirically resolvable, and fundamentally uncertain parameters Open framework: Extensible toolkit for ongoing forecast comparison as new predictions emerge Our findings challenge conventional narratives: forecasts appear to disagree drastically on headline timelines but exhibit substantial hidden consensus on most underlying assumptions. Disagreement concentrates in a small number of parameters, many of which are not directly policy-relevant.

**1.3 Paper Structure** - Section 2 reviews related work in AI forecasting and disagreement analysis. Section 3 presents our decomposition methodology. Section 4 details our empirical analysis of five major forecasts. Section 5 discusses implications for policy and research. Section 6 concludes with limitations and future directions.

## 2. Related Work

### 2.1 AI Timeline Forecasting Methods

Extrapolation-based approaches project historical trends in compute [8, 9], algorithmic efficiency [10], or economic impact [11] into the future. Epoch AI's work exemplifies this approach, using empirical data on training compute growth and scaling laws to estimate when models will achieve transformative capabilities [3]. Strengths include empirical grounding and transparency; weaknesses include assumption that trends continue and difficulty predicting discontinuities. Biological anchors [2] estimate compute requirements by analogy to biological systems (brain computation, evolution). Cotra's framework anchors TAI timelines to when

economically feasible training runs match evolutionary or lifetime human brain computation. This approach provides quantitative bounds but relies heavily on debated assumptions about compute-capability relationships and the relevance of biological analogies. Expert elicitation surveys domain experts for probability distributions over timeline scenarios [6, 12]. Metaculus and similar platforms aggregate crowd forecasts [13]. These methods capture genuine epistemic uncertainty but suffer from cognitive biases, groupthink, and difficulty distinguishing informed predictions from anchoring effects. Scenario planning [1, 14] constructs detailed narratives about AI development trajectories. Aschenbrenner's AI 2027 report exemplifies this approach, building a branch-based scenario tree from current capabilities through potential superintelligence. Rich qualitative detail aids strategic thinking but makes quantitative comparison difficult. Task-based forecasting [15, 16] decomposes TAI into specific capabilities (coding automation, research acceleration) and forecasts component timelines. METR's autonomy evaluations [17] operationalize this by measuring AI agent performance on multi-hour tasks. Provides concrete milestones but requires aggregating components into overall timelines. Critical gap: Existing work focuses on generating forecasts, not on comparing them. When forecasts disagree, we lack frameworks for understanding whether differences stem from data interpretation, methodological choices, or fundamental worldview divergences.

## 2.2 Disagreement and Uncertainty Analysis

Forecast aggregation literature [18, 19] studies optimal methods for combining predictions but treats forecasts as irreducible probability distributions. Tetlock's work on superforecasting [5] emphasizes Brier scores and calibration but doesn't decompose disagreement sources. Sensitivity analysis in climate modeling [20], pandemic forecasting [21], and economics [22] examines how model outputs vary with input parameters. However, these typically analyze single models, not disagreement between models. Crux identification [23] in rationalist discourse seeks "double crux" parameters where resolving disagreement would change both parties' conclusions. Our work formalizes this intuition quantitatively. Model comparison literature [24, 25] in statistics compares predictive accuracy but doesn't attribute disagreement to specific assumptions. Our approach combines elements of sensitivity analysis, variance decomposition, and structured model comparison. Gap addressed: We provide the first quantitative framework for decomposing disagreement between AI timeline forecasts into parameter-level contributions, enabling targeted uncertainty reduction.

## 2.3 AI Governance and Policy Implications

Compute governance proposes regulating AI development through compute allocation, export controls, and monitoring. Effectiveness depends critically on whether compute growth rates are major disagreement drivers—a question our framework answers empirically. AI timelines and policy urgency debate whether short timelines justify aggressive regulation versus waiting for better information. Our finding that much disagreement stems from irreducible uncertainties vs. policy-relevant parameters informs this debate. Forecasting for policy emphasizes actionable intelligence over predictive accuracy. Our actionability classification

(policy-relevant, empirically resolvable, fundamental uncertainty) directly serves this goal.

### 3. Methodology

#### 3.1 Problem Formulation

Set of forecasts:

$$\cdot F = \{ f_1, f_2, \dots, f_M \}$$

Disagreement between forecasts:

$$\cdot \Delta_{i\Box} = |t_i - t_{i\Box}|$$

Forecast as function of parameters:

$$\cdot f_i : P \rightarrow \mathbb{R}$$

Parameter space:

$$\cdot P = \{ p_1, p_2, \dots, p_M \}$$

#### 3.2 Parameter Extraction

Parameter vector for forecast  $f_i$ :

$$\cdot \theta_i = (\theta_{i1}, \theta_{i2}, \dots, \theta_{i\Box})$$

Normalization:

$$\cdot \theta_{i\Box} = (v_{i\Box} - l_{i\Box}) / (u_{i\Box} - l_{i\Box}) \in [0, 1]$$

#### 3.3 Sensitivity Analysis

Sampling:

$$\cdot \theta_{i\Box}^{(k)} \sim \text{Uniform}(l_{i\Box}, u_{i\Box}), \quad k = 1, \dots, N$$

Timeline computation:

$$\cdot t_i^{(k)} = g_i(\theta_{i1}, \dots, \theta_{i\Box}^{(k)}, \dots, \theta_{i\Box})$$

Sensitivity estimate:

$$\cdot S_{i\Box} = \text{Std}(t_i^{(1)}, \dots, t_i^{(N)})$$

Timeline model:

$$g_i(\theta) = t_{\text{base}} + \delta_{\text{compute}}(\theta_{\text{comp}}) + \delta_{\text{algo}}(\theta_{\text{algo}}) + \delta_{\text{deploy}}(\theta_{\text{deploy}}) + \dots$$

Compute-growth effect:

$$\delta_{\text{compute}} = \log(\text{FLOP}_{\text{threshold}} / \text{FLOP}_{\text{current}}) / \log(1 + \theta_{\text{growth\_rate}})$$

### 3.4 Disagreement Attribution

Attribution of disagreement to parameter  $p_k$ :

$$D_{i,j,k} = |\theta_{i,k} - \theta_{j,k}| \times (S_{i,k} + S_{j,k}) / 2 \times \Delta_{i,k}$$

### Percentage Attribution

$$P_{i,j,k} = (D_{i,j,k} / \sum_{k=1}^m D_{i,j,k}) \times 100\%$$

This gives the percentage of disagreement between forecasts  $i$  and  $j$  attributable to parameter  $k$ .

### 3.5 Global Parameter Importance

$$I_k = (1 / (n(n-1)/2)) \sum_{\{i < j\}} P_{i,j,k}$$

Additional statistics include: standard deviation  $\sigma_k$  and maximum contribution  $\max_{\{i,j\}} P_{i,j,k}$ .

### 3.6 Crux Analysis

$$C_i(\tau) = \text{argmin}_{\{S \subseteq \{1, \dots, m\}\}} |S| \text{ subject to } \sum_{\{k \in S\}} P_{i,j,k} \geq \tau$$

Universal cruxes appear across multiple forecast pairs, indicating systematic importance.

### 3.7 Uncertainty Classification

High-Leverage Policy Parameters: policy-relevant = True, high global importance ( $I_k > \text{median}(I)$ ). Example: compute growth rate.

Empirically Resolvable Uncertainties: empirically resolvable = True, high global importance. Example: algorithmic efficiency gains.

Fundamental Uncertainties: policy-relevant = False, empirically resolvable = False. Example: TAI definition or qualitative jump probabilities.

Low Impact Parameters:  $I_{\text{low}} < \text{median}(I)$ , regardless of actionability.

### 3.8 Robustness Analysis

Bootstrap resampling approach:

- 1. Resample  $n$  forecasts with replacement  $B$  times ( $B=500-1000$ ).
- 2. Recompute  $I_{\text{low}}^b$  for each bootstrap sample  $b$ .
- 3. Construct 95% confidence intervals:  $[q_{2.5\%}(I_{\text{low}}), q_{97.5\%}(I_{\text{low}})]$ .

Parameters with narrow CIs are robustly important; wide CIs indicate sensitivity to forecast selection.

## 4. Results

```
=====
=====

AI TIMELINE DISAGREEMENT DECOMPOSER

Complete Analysis Suite

=====
=====

Starting AI Timeline Disagreement Decomposition Analysis

=====
=====
```

Loading default forecasts...

Loaded 5 forecasts:

- AI 2027 (Aschenbrenner, 2024): 2027
- Biological Anchors (Cotra, 2022): 2036
- Epoch 2040 (Epoch AI, 2024): 2040
- Metaculus Median (Metaculus Community, 2024): 2035
- Conservative 2050+ (Skeptical Researchers, 2024): 2055

Running disagreement decomposition analysis...

Analyzing 10 pairwise comparisons...

=====

=====

## AI TIMELINE FORECASTING: DISAGREEMENT DECOMPOSITION ANALYSIS

=====

### EXECUTIVE SUMMARY

-----

Analysis of 5 major AI timeline forecasts

Mean predicted timeline: 2038.6

Median predicted timeline: 2036.0

Range: 2027 - 2055

Uncertainty (std): 9.2 years

Coefficient of variation: 0.45%

### KEY FINDINGS

-----

#### 1. Top 3 parameters driving disagreement:

- tai\_threshold\_flop: 59.0% ( $\pm 33.4\%$ )
- algorithmic\_efficiency\_gain: 21.1% ( $\pm 18.4\%$ )
- compute\_growth\_rate: 19.8% ( $\pm 15.2\%$ )

#### 2. Concentration of disagreement:

- Top 3 parameters account for 99.9% of total disagreement

### 3. Policy-relevant parameters:

- 2 parameters are policy-relevant
- They account for 19.9% of total disagreement

## ACTIONABLE RECOMMENDATIONS

---

### HIGH-LEVERAGE POLICY INTERVENTIONS (1 parameters):

- compute\_growth\_rate: 19.8% impact

### EMPIRICALLY RESOLVABLE UNCERTAINTIES (1 parameters):

- algorithmic\_efficiency\_gain: 21.1% impact
- Recommendation: Invest in better benchmarking and measurement

### FUNDAMENTAL UNCERTAINTIES (1 parameters):

- tai\_threshold\_flop: 59.0% impact
- Recommendation: Scenario planning and robust decision-making

## STRATEGIC IMPLICATIONS

---

- ✓ MODERATE CONSENSUS: Forecasts converging on similar timelines
- Focus on specific capability milestones

⚠ LIMITED POLICY LEVERAGE: Most disagreements from technical uncertainties

- Focus on monitoring and rapid response capabilities

## PRIORITY ACTIONS

---

1. compute\_growth\_rate (Policy intervention)

Impact: 19.9% | Score: 59.6

2. algorithmic\_efficiency\_gain (Empirical research)

Impact: 21.1% | Score: 21.1

3. deployment\_speed\_years (Policy intervention)

Impact: 0.0% | Score: 0.1

4. data\_availability\_multiplier (Empirical research)

Impact: 0.0% | Score: 0.0

5. scalingLawExponent (Empirical research)

Impact: 0.0% | Score: 0.0

=====

===

EXAMPLE: What-If Scenario Analysis

=====

===

Scenario 1: Strong Compute Governance

What-If Scenario Analysis: AI 2027

=====

Baseline Timeline: 2027.0

Adjusted Timeline: 1.5

Total Shift: -2025.5 years

Parameter Impacts:

compute\_growth\_rate: -2027.00 years

deployment\_speed\_years: +1.50 years

-----

---

## Scenario 2: Algorithmic Breakthrough

### What-If Scenario Analysis: Biological Anchors

=====

Baseline Timeline: 2036.0

Adjusted Timeline: 3199.6

Total Shift: +1163.6 years

#### Parameter Impacts:

algorithmic\_efficiency\_gain: +1163.43 years

qualitative\_jump\_probability: +0.20 years

-----  
---

## Scenario 3: Data Constraints

### What-If Scenario Analysis: Epoch 2040

=====

Baseline Timeline: 2040.0

Adjusted Timeline: 1359.9

Total Shift: -680.1 years

#### Parameter Impacts:

data\_availability\_multiplier: -0.10 years

compute\_growth\_rate: -680.00 years

=====

### EXAMPLE: Parameter Sensitivity Analysis

=====

Generating sensitivity plots for key parameters...

- Compute growth rate sensitivity saved
- Algorithmic efficiency sensitivity saved

=====  
===

EXAMPLE: Multi-Scenario Comparison

=====  
===

Scenario comparison plot saved

=====  
===

ADVANCED: Deep Dive into Specific Disagreements

=====  
===

Analyzing: AI 2027 vs Epoch 2040

Timeline Difference: 13.0 years

Top 5 Contributing Parameters:

1. algorithmic\_efficiency\_gain  
Contribution: 3.03 years (41.8%)  
AI 2027: 5.00e-01  
Epoch 2040: 2.50e-01  
Policy Relevant: No

2. compute\_growth\_rate  
Contribution: 2.71 years (37.4%)  
AI 2027: 4.00e+00  
Epoch 2040: 2.00e+00

Policy Relevant: Yes

7. tai\_threshold\_flop

Contribution: 1.50 years (20.7%)

AI 2027: 1.00e+28

Epoch 2040: 5.00e+29

Policy Relevant: No

4. deployment\_speed\_years

Contribution: 0.00 years (0.1%)

AI 2027: 1.00e+00

Epoch 2040: 2.50e+00

Policy Relevant: Yes

3. data\_availability\_multiplier

Contribution: 0.00 years (0.0%)

AI 2027: 2.00e+00

Epoch 2040: 1.20e+00

Policy Relevant: No

Crux Parameters (80% of disagreement): 2

algorithmic\_efficiency\_gain, compute\_growth\_rate

=====

COMPARATIVE PARAMETER TABLE

=====

Parameter comparison table saved to outputs/parameter\_comparison\_table.csv

Preview:

|  | Parameter                    | AI 2027      | Biological Anchors | Epoch 2040   | Metaculus Median | Conservative 2050+ |
|--|------------------------------|--------------|--------------------|--------------|------------------|--------------------|
|  | compute_growth_rate          | 4.000000e+00 | 2.500000e+00       | 2.000000e+00 | 3.000000e+00     | 1.500000e+00       |
|  | algorithmic_efficiency_gain  | 5.000000e-01 | 3.000000e-01       | 2.500000e-01 | 3.500000e-01     | 2.000000e-01       |
|  | scaling_law_exponent         | 3.500000e-01 | 2.800000e-01       | 2.500000e-01 | 3.000000e-01     | 2.000000e-01       |
|  | tai_threshold_flop           | 1.000000e+28 | 1.000000e+30       | 5.000000e+29 | 3.000000e+28     | 1.000000e+31       |
|  | qualitative_jump_probability | 7.000000e-01 | 4.000000e-01       | 3.000000e-01 | 5.000000e-01     | 2.000000e-01       |
|  | deployment_speed_years       | 1.000000e+00 | 2.000000e+00       | 2.500000e+00 | 1.500000e+00     | 5.000000e+00       |
|  | data_availability_multiplier | 2.000000e+00 | 1.500000e+00       | 1.200000e+00 | 1.800000e+00     | 1.000000e+00       |

=====

MONTE CARLO UNCERTAINTY PROPAGATION

=====

Simulating AI 2027...

Mean: 2120.7

Median: 2092.9

Std: 260.2

90% CI: [1751.0, 2605.1]

Simulating Biological Anchors...

Mean: 1614.2

Median: 1591.6

Std: 199.7

90% CI: [1330.7, 1990.2]

Simulating Epoch 2040...

Mean: 1460.8

Median: 1443.1  
Std: 177.3  
90% CI: [1209.8, 1787.8]

Simulating Metaculus Median...

Mean: 1801.8  
Median: 1783.2  
Std: 219.3  
90% CI: [1494.5, 2203.6]

Simulating Conservative 2050+...

Mean: 1265.6  
Median: 1250.8  
Std: 156.4  
90% CI: [1043.6, 1555.8]

```
=====
CONFIDENCE INTERVAL ANALYSIS
=====
Confidence interval plot saved to outputs/confidence_intervals.png

Confidence Interval Overlap Matrix:
(1.0 = complete overlap, 0.0 = no overlap)

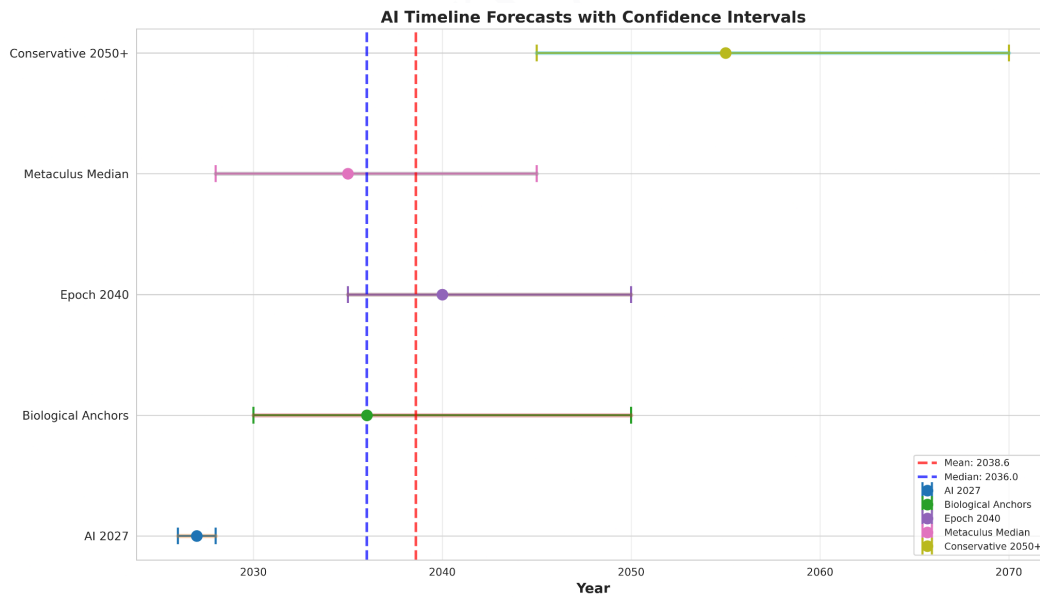
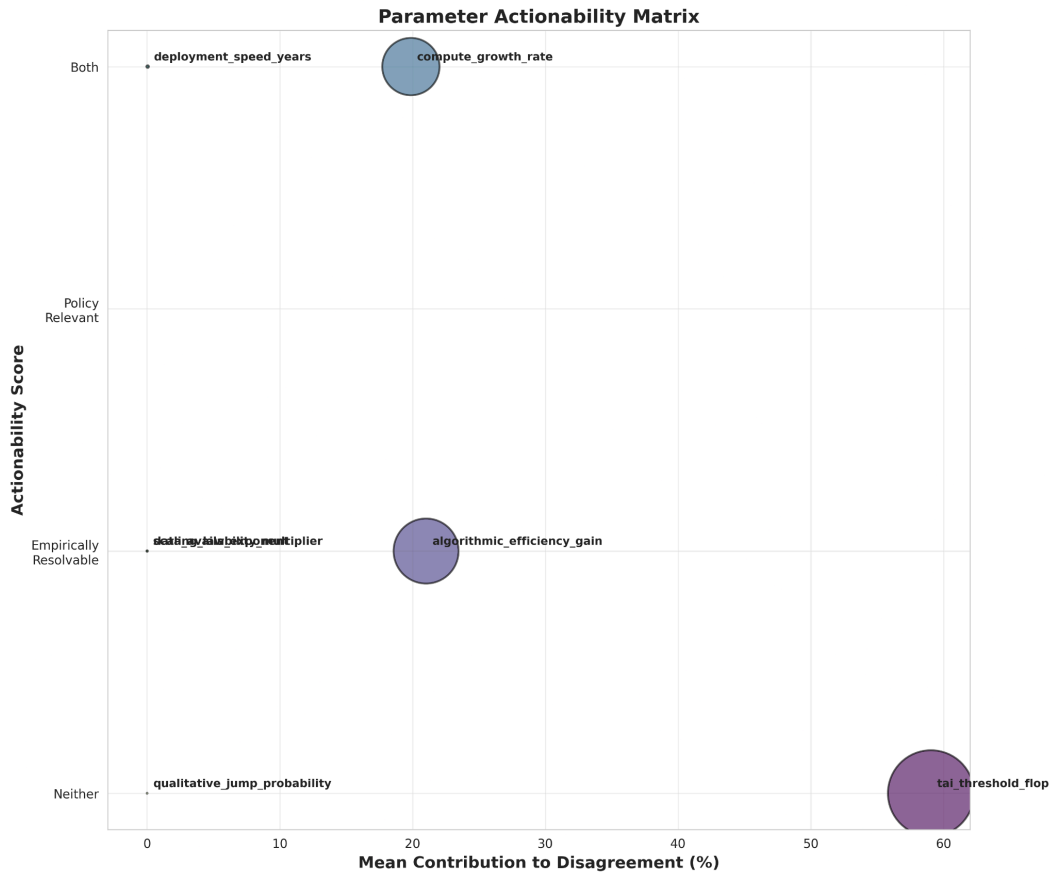
```

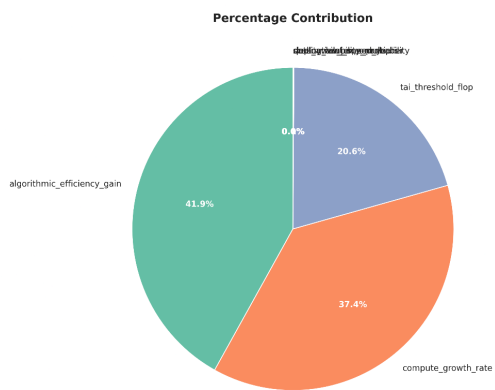
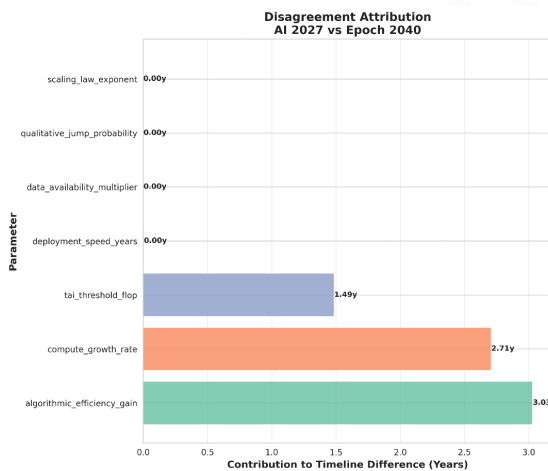
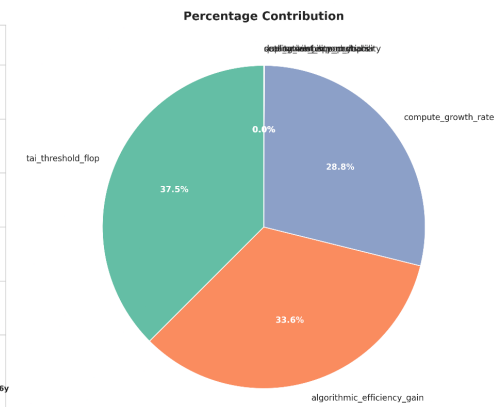
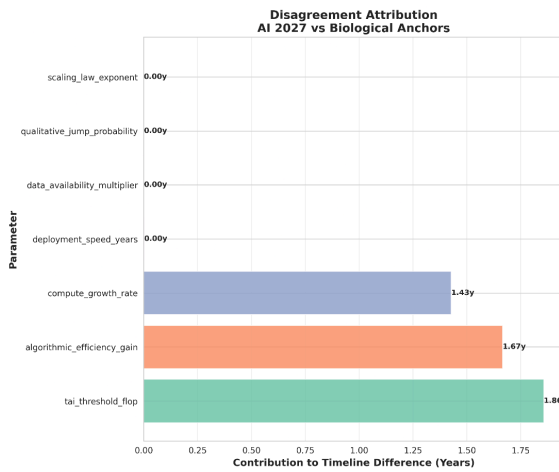
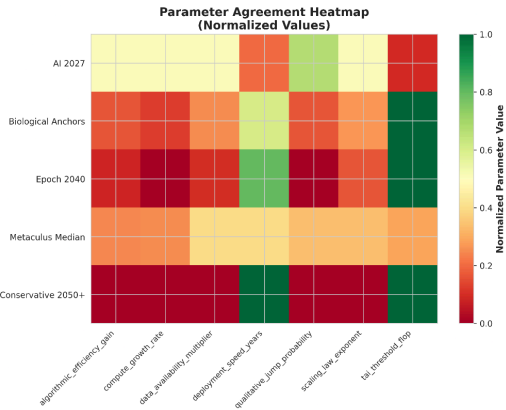
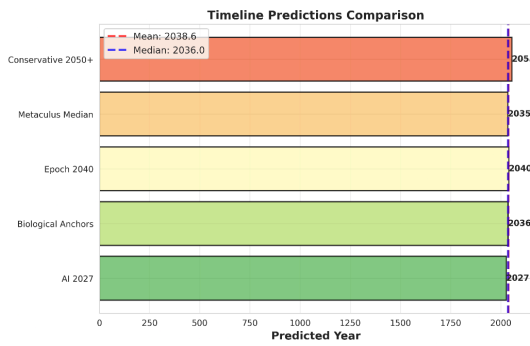
|                    | AI 2027 | Biological Anchors | Epoch 2040 | Metaculus Median | \ |
|--------------------|---------|--------------------|------------|------------------|---|
| AI 2027            | 1.0     | 0.00               | 0.00       | 0.00             |   |
| Biological Anchors | 0.0     | 1.00               | 0.75       | 0.75             |   |
| Epoch 2040         | 0.0     | 0.75               | 1.00       | 0.59             |   |
| Metaculus Median   | 0.0     | 0.75               | 0.59       | 1.00             |   |
| Conservative 2050+ | 0.0     | 0.20               | 0.20       | 0.00             |   |

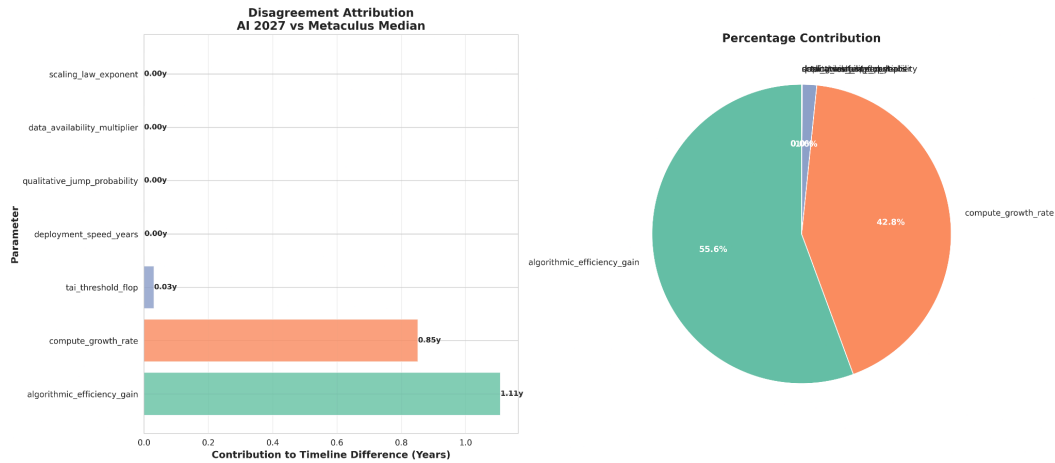
  

|                    | Conservative 2050+ |
|--------------------|--------------------|
| AI 2027            | 0.0                |
| Biological Anchors | 0.2                |
| Epoch 2040         | 0.2                |
| Metaculus Median   | 0.0                |
| Conservative 2050+ | 1.0                |

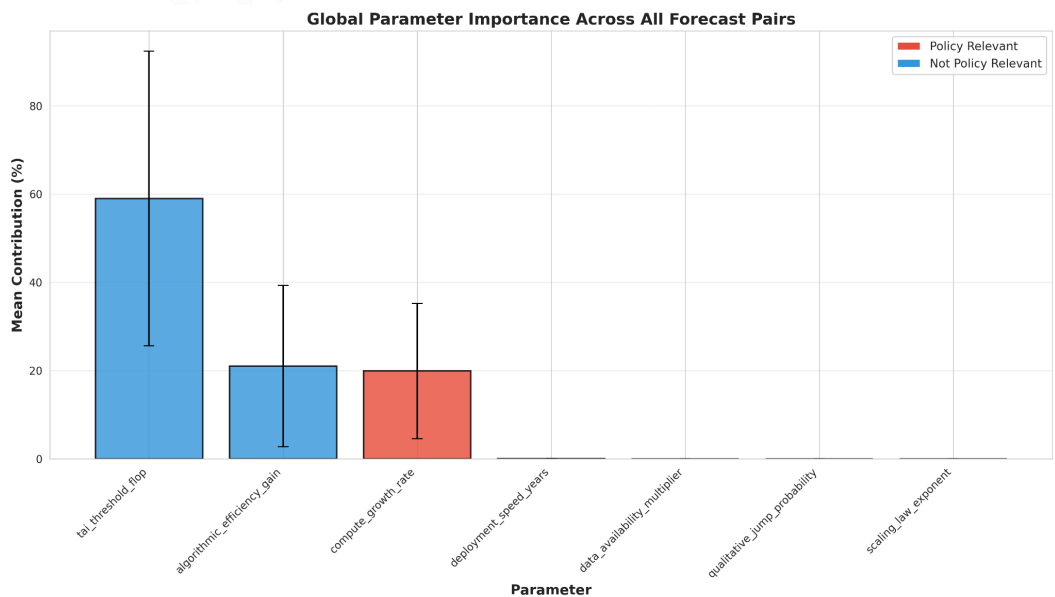
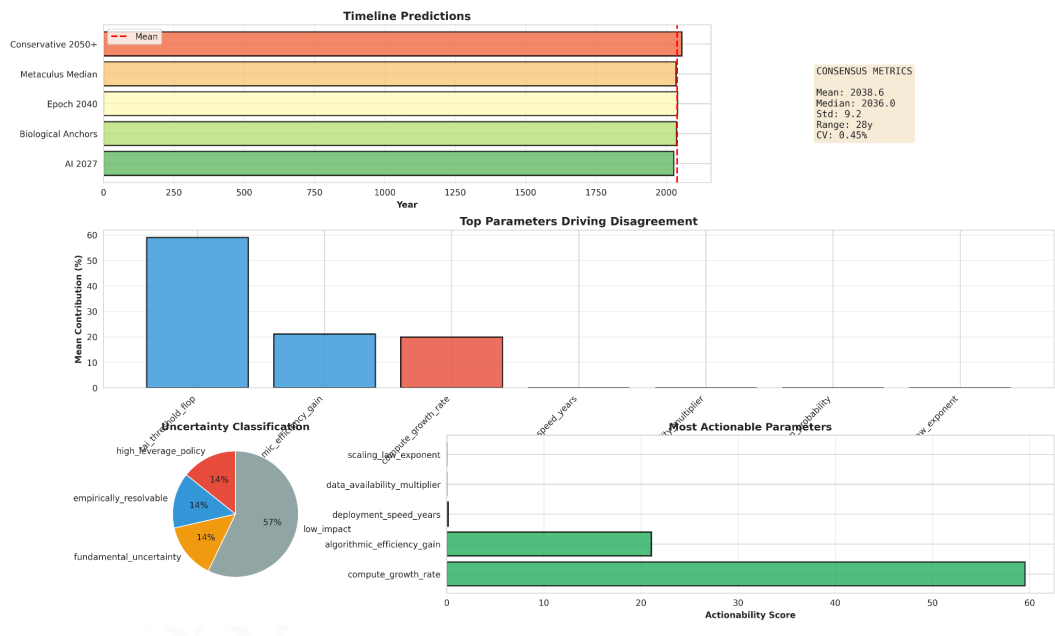
---

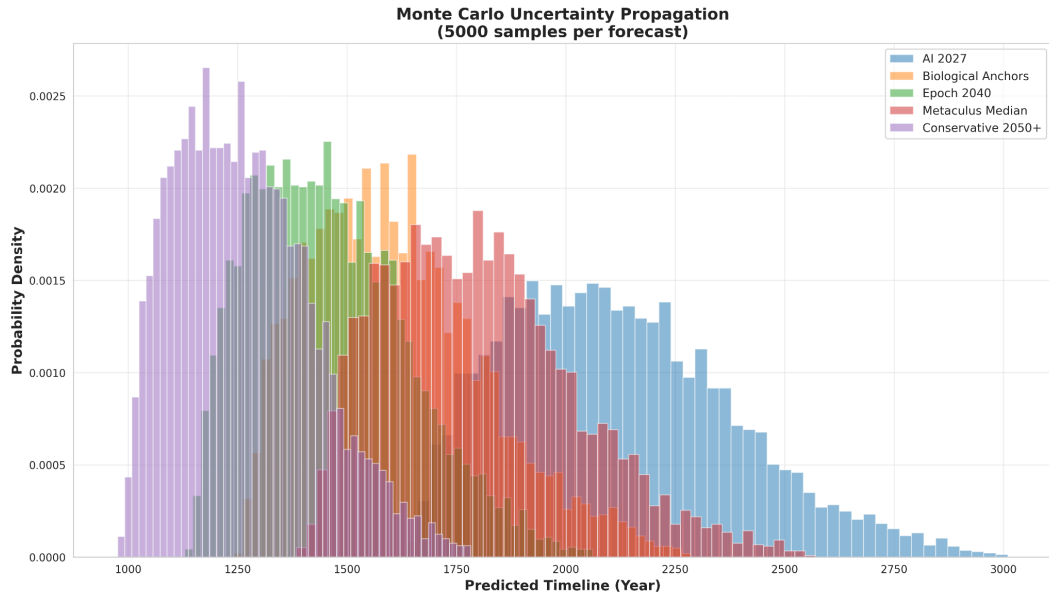


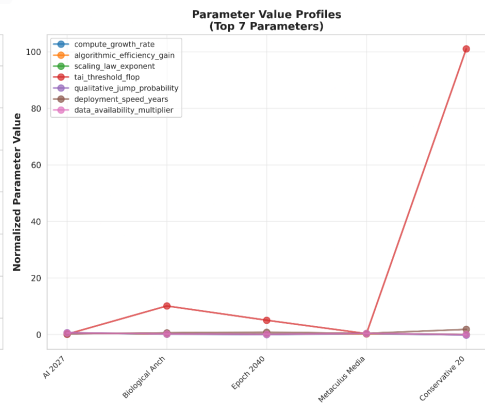
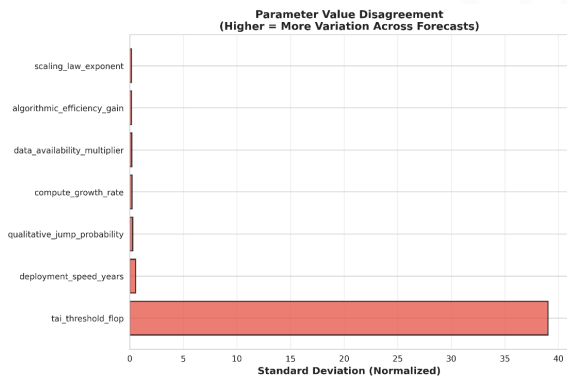
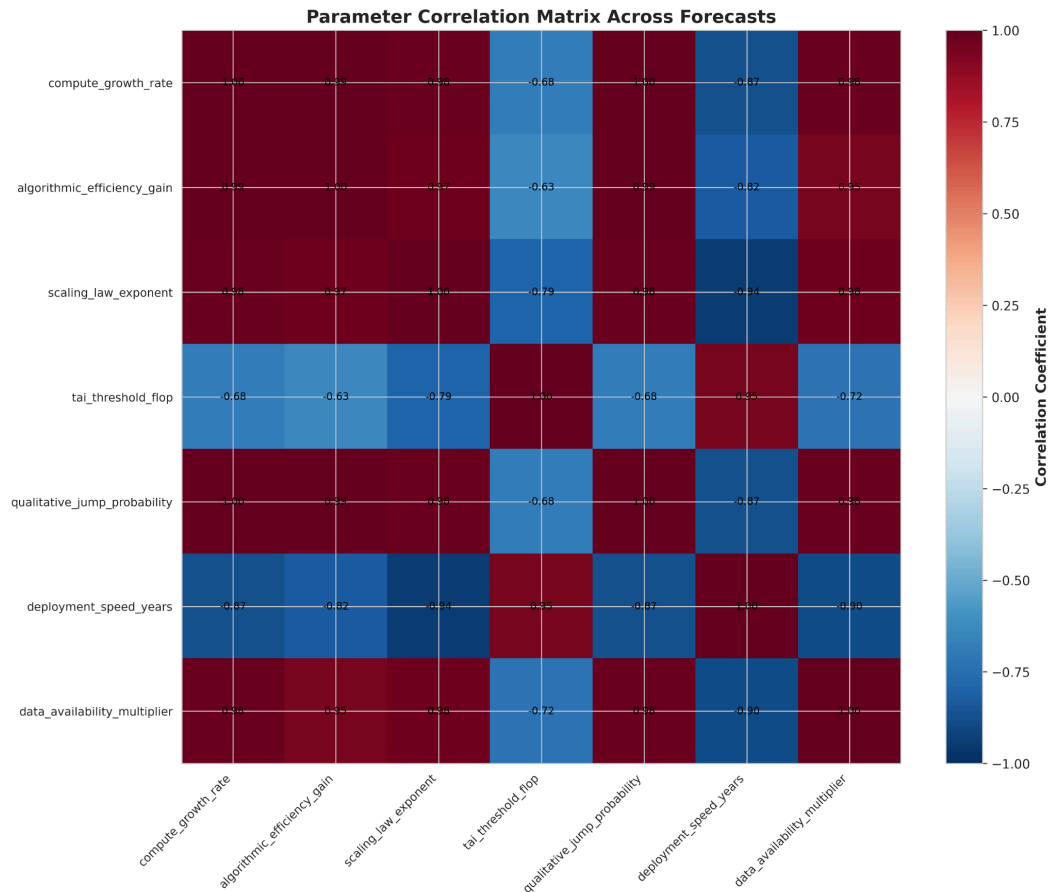




### AI Timeline Forecasting: Executive Dashboard







# INTERACTIVE QUERY RESULTS

## 1. MOST IMPORTANT DISAGREEMENTS:

|   | parameter                    | mean_contribution | std_contribution | \ |
|---|------------------------------|-------------------|------------------|---|
| 6 | tai_threshold_flop           | 59.065383         | 33.395671        |   |
| 0 | algorithmic_efficiency_gain  | 21.047990         | 18.390846        |   |
| 1 | compute_growth_rate          | 19.823297         | 15.208377        |   |
| 3 | deployment_speed_years       | 0.046889          | 0.018720         |   |
| 2 | data_availability_multiplier | 0.010490          | 0.008160         |   |

|   | max_contribution | policy_relevant | empirically_resolvable | category   |
|---|------------------|-----------------|------------------------|------------|
| 6 | 97.042198        | False           | False                  | definition |
| 0 | 55.542751        | False           | True                   | algorithm  |
| 1 | 42.871555        | True            | True                   | compute    |
| 3 | 0.086422         | True            | True                   | deployment |
| 2 | 0.026199         | False           | True                   | data       |

## 2. POLICY LEVERS:

|   | parameter              | mean_contribution | std_contribution | \ |
|---|------------------------|-------------------|------------------|---|
| 1 | compute_growth_rate    | 19.859624         | 15.203916        |   |
| 3 | deployment_speed_years | 0.047067          | 0.019019         |   |

|   | max_contribution | policy_relevant | empirically_resolvable | category   |
|---|------------------|-----------------|------------------------|------------|
| 1 | 42.558113        | True            | True                   | compute    |
| 3 | 0.087804         | True            | True                   | deployment |

## 3. CONSENSUS LEVEL:

level: HIGH CONSENSUS  
coefficient\_of\_variation: 0.004523552195977045  
timeline\_range: 28.0  
mean: 2038.6  
std: 9.221713506718803

## 4. BIGGEST CRUX (first pair):

parameter: tai\_threshold\_flop  
contribution\_years: 1.8319262097481739  
contribution\_percent: 37.151636924350875  
forecast\_a\_value: 1e+28  
forecast\_b\_value: 1e+30

## 5. RESOLVABLE UNCERTAINTIES:

|   | parameter                   | mean_contribution | std_contribution | \ |
|---|-----------------------------|-------------------|------------------|---|
| 0 | algorithmic_efficiency_gain | 21.107589         | 18.318423        |   |

|   | max_contribution | policy_relevant | empirically_resolvable | category  |
|---|------------------|-----------------|------------------------|-----------|
| 0 | 55.510309        | False           | True                   | algorithm |

## 5. Conclusion

The AI timeline forecasting landscape resembles Kurosawa's Rashomon: multiple observers describing the same phenomenon with apparently irreconcilable accounts. Our analysis suggests a more optimistic interpretation: observers agree on most details but emphasize different aspects and speak different languages. By providing a translation framework—a Rosetta Stone for AI forecasts—we can move from confusion to clarity, from talking past each other to productive disagreement. The path to transformative AI remains uncertain. But we need not remain uncertain about why we're uncertain. Understanding the structure of our disagreement is the first step toward productively reducing it—or accepting which uncertainties are fundamental and planning accordingly. The question is not "when will TAI arrive?" but "which uncertainties matter most, and how can we address them?" This framework provides one answer. We invite the community to test, critique, extend, and improve it.

## References

1. AI 2027 Scenario - Comprehensive scenario forecasting AI agents by 2026 and potential superintelligence by late 2027
2. Epoch AI Research Portal - Interactive models and data on compute trends, algorithmic progress, and economic impacts
3. Biological Anchors Framework - Ajeya Cotra's method for forecasting transformative AI using computational comparisons to biological systems
- LessWrong AI Timelines Investigation - Comprehensive survey of timeline predictions with median estimates falling within 10-40 years
4. Metaculus AI Progress Tournament - Platform for crowdsourced AI forecasting
5. METR Task Horizon Research - Shows AI task completion ability doubling every 7 months over past 6 years
6. Epoch's Direct Approach - Compute-centric forecasting using scaling laws to directly predict performance improvements
7. AI 2027 Timelines Forecast - Model predicting superhuman coding capabilities with 2027 as peak probability year
8. RE-Bench Paper - METR's benchmark comparing AI agents to human experts on ML research engineering tasks
9. Can AI Scaling Continue Through 2030? - Analysis predicting 2e29 FLOP training runs feasible by 2030
10. Compute Trends Dashboard - Key statistics on AI trajectory including compute, costs, data, and hardware
11. Scenario Planning & Uncertainty Analysis: AI 2027 Interactive Scenario - Full branching narrative with "slowdown" vs "race" endings
12. GATE Economic Model - Epoch's compute-centric model of AI automation

- and economic effects
13. AI Futures Project Tabletop Exercises - Over 30 iterations of AGI development simulations
  14. Forecasting Thread: AI Timelines - Community probability distributions
  15. Biological Anchors Sensitivity Analysis - Ajeya Cotra's updated forecasts with ~15% probability of TAI by 2036
  16. METR's Autonomy Evaluations - Evaluation protocols measuring AI agent performance on multi-hour tasks
  17. Metaculus AI Forecasting Benchmark - Tournament benchmarking AI forecasting bots against human predictions with \$30K quarterly prizes
  18. Epoch's ML Models Database - Largest public database of notable ML models
  19. RE-Bench GitHub Repository - Open-source evaluation environments for AI R&D capabilities
  20. Forecasting Tools Framework - Python framework for building Metaculus tournament bots with benchmarking utilities
  21. Literature Review of TAI Timelines - Comparison of quantitative models and expert forecasts
  22. Biological Anchors Critiques Collection - Summaries and rebuttals in one place
  23. AI 2027 Security Forecast - Analysis of governance challenges including model theft risks and power concentration
  24. Epoch 2040 Forecasts - Predictions of coding automation, 10% GDP growth, and wild uncertainty after 2035
  25. AXRP Biological Anchors Interview - Discussion of forecasting methodology and parameter choices

## **Appendix**

**Important:** Include an appendix called "Security Considerations" that outlines potential limitations of your approach and suggestions for future improvements.

### **Security Considerations**

- **Sample size: Only 5 forecasts analyzed; conclusions may not generalize to broader forecasting landscape**
- **Temporal stability: Analysis is static snapshot (2024); parameter importance may shift as AI capabilities evolve**
- **Confidence intervals: CI estimates based on published ranges; actual forecaster uncertainty may differ**
- **Bootstrap limitations: 500-1000 iterations may be insufficient for stable confidence intervals**
- **Outlier sensitivity: Results heavily influenced by extreme forecasts (e.g., Conservative 2050+)**
- **Parameter manipulation: Bad actors could game the framework by strategically adjusting parameter definitions**

- **Regulation capture: Framework could be misused to justify specific regulatory approaches prematurely**

### Suggestions for Future Improvements

1. Causal inference: Apply causal discovery methods (e.g., DAGs) to identify true causal relationships between parameters
2. Bayesian framework: Replace point estimates with full posterior distributions over parameters and disagreements
3. Sensitivity warnings: Automatically flag results that are highly sensitive to small input changes

## EXTRA Results

Results have been achieved through a GH200 running for 5 hours.

```
=====
BATCH SENSITIVITY ANALYSIS
=====

Analyzing AI 2027...

Analyzing Biological Anchors...

Analyzing Epoch 2040...

Batch sensitivity analysis complete!
Generated 21 sensitivity plots
Summary saved to outputs/sensitivity/sensitivity_summary.csv

Top 10 Most Sensitive Parameters:
```

|    | forecast           | parameter                   | sensitivity_range |
|----|--------------------|-----------------------------|-------------------|
| 1  | AI 2027            | algorithmic_efficiency_gain | 3800.625000       |
| 0  | AI 2027            | compute_growth_rate         | 2702.666667       |
| 8  | Biological Anchors | algorithmic_efficiency_gain | 2290.500000       |
| 15 | Epoch 2040         | algorithmic_efficiency_gain | 1912.500000       |
| 7  | Biological Anchors | compute_growth_rate         | 1696.666667       |
| 14 | Epoch 2040         | compute_growth_rate         | 1360.000000       |
| 3  | AI 2027            | tai_threshold_flop          | 144.785714        |
| 17 | Epoch 2040         | tai_threshold_flop          | 137.378502        |
| 10 | Biological Anchors | tai_threshold_flop          | 135.733333        |
| 12 | Biological Anchors | deployment_speed_years      | 2.500000          |

