
How should the global south think and act about compute policy?

Bruno De Biase Deo Adre Novais Xavier Rodrigues
Federal University of São Carlos Federal University of São Carlos

With
Apart Research

Abstract

Frontier compute is treated as a governable input to AI, with the prescription that states should fund domestic compute. That guidance was built for the Global North and is untested under the constraints of countries outside the frontier. We build an agent-based model with a 2,000-scenario Monte Carlo layer that turns each country’s AI budget into installable FP64 compute (the TOP500 metric) under chip-access, energy, and endogenous-price constraints, and apply it to Brazil’s AI plan (PBIA). Its two compute targets, funded from one budget, diverge: doubling national compute by 2027 is feasible in 100% of scenarios; a TOP500 top-five placement by 2029 in 0% (about 4% of the bar). The showcase target fails on framing and scale, not chips or energy: those add nothing, and even a tenfold budget reaches only $\sim 7\%$. We conclude that Global-South compute goals should be absolute, budget-anchored targets sized to need, not moving global rankings.

1. Introduction

Compute, defined as the computational power applied to the training and operation of AI models, has become one of the most governable inputs to artificial intelligence. Unlike data and algorithms, which are intangible and readily replicable, frontier compute is detectable, excludable, quantifiable, and produced through a highly concentrated supply chain (Sastry et al., 2024). These properties have led governments to treat investment in domestic compute capacity as a central instrument of AI governance and technological sovereignty, a decision that Sastry et al. classify as one of resource allocation.

This body of work, however, was developed primarily in reference to Global North economies, which operate under stable fiscal conditions and command large compute budgets. It offers little guidance on whether the same allocation strategy is viable under the constraints faced by countries outside the technological frontier, where public budgets are small, currencies are volatile, and access to a

concentrated and inflated hardware market is contested. This gap matters because, for a resource-constrained state, a poorly calibrated compute target is not merely a planning error but a diversion of scarce sovereign capital toward an objective that does not deliver capability. We examine this gap through the Brazilian Artificial Intelligence Plan (PBIA, 2024), which adopts the acquisition of frontier compute as a pillar of technological sovereignty. The PBIA states two compute targets: doubling national AI compute within three years, and placing the Santos Dumont supercomputer among the five largest systems in the global TOP500 ranking. We find that these two targets, despite drawing on the same budget, differ sharply in feasibility, and that the general prescription of the literature holds for one of them but not for the other.

To test this, we build an agent-based model coupled with a Monte Carlo layer of 2,000 scenarios. The model converts each country’s annual budget into installable frontier compute, measured in FP64 Rmax PFLOPs to remain consistent with the TOP500 metric, and advances year by year from 2025 to 2030 under two physical constraints, chip access and energy, together with a chip price that is not fixed but emerges endogenously from aggregate demand against a finite supply. Applied to the PBIA, the model finds that the central target, doubling national compute, is feasible in 100% of scenarios, while the showcase target, a top-five TOP500 placement, is feasible in 0% of them, reaching only about 4% of the required threshold by its deadline. This separation holds across the full plausible range of every model parameter. The difference between the two does not come from the constraints one might expect: counterfactuals that grant Brazil privileged chip access or unlimited energy move the showcase target’s feasibility by zero points, and only budget moves it. What separates the two targets is their design. The central target is absolute, so a fixed budget eventually reaches it. The showcase target is tied to a moving global ranking that recedes at the growth rate of the frontier, so that closing the gap on budget alone would require on the order of ten times the line dedicated to it, roughly the entire plan, concentrated in a single machine, and even that leaves the target met in only about 7% of scenarios.

This work makes three contributions. First, we provide a reusable method for auditing the technical feasibility and effectiveness of national compute targets before expenditure, turning a qualitative governance prescription into a falsifiable, budget-versus-physics test that can be reapplied to any plan by editing a single parameter file. Second, we apply this method to the PBIA and identify a concrete design flaw in its showcase target. The flaw is not simply one of underfunding that more money would fix. Only a small fraction of the plan’s budget is allocated to frontier compute while the majority funds enterprise credit and other pillars that serve their own legitimate ends, so the shortfall cannot be closed by reallocation without sacrificing those objectives, and the total budget is itself fixed. More fundamentally, our counterfactuals show that even a tenfold increase in the dedicated budget leaves the target infeasible, because the problem is the target’s framing rather than its funding level. Third, and specific to the Brazilian case, we identify the design choice that produces this failure. Unlike the other national plans we model, which anchor their goals to absolute capacity, the PBIA pins its showcase target to a position in a moving global ranking, which recedes at the growth rate of the frontier and therefore cannot be reached on a fixed budget. The corrective principle that follows is that compute should be sized to a country’s actual needs, such as running public services, supporting research, and evaluating the models deployed within its borders, rather than pursued as an end in itself in the form of a ranking position. This argument connects directly to AI governance, since the capacity to evaluate, audit, and oversee frontier AI presupposes access to frontier compute, and allocation decisions that spend scarce sovereign capital on an unreachable prestige target weaken, at the source, a country’s ability to govern the systems deployed within it.

2. Related Work

The framework closest to our work analyzes computing power as an instrument of AI governance (Sastry et al., 2024), arguing that compute is uniquely governable because it is detectable, excludable, quantifiable, and produced through a concentrated supply chain, and that it enables three governance capacities: visibility, allocation, and enforcement. We adopt this premise but address a question it leaves open. Its instruments, such as export controls and the monitoring of frontier training, assume an actor that already controls the supply chain or owns the compute to be governed. The prior question for a state outside the frontier, whether it can acquire frontier compute at all under a fixed budget and contested demand, falls outside its scope. Notably, the same authors acknowledge that compute access in the Global South is shaped by additional structural barriers and treat capacity-building there as beyond the paper’s scope, which is precisely the space our work enters.

The quantitative basis for any feasibility claim about compute comes from work on hardware price-performance and the scaling of training compute. Epoch AI estimates that AI hardware performance per dollar has improved by roughly 37% per year (Hobbhahn et al., 2023; Epoch AI, 2024), a trend we adopt directly as a model parameter, and the TOP500 list supplies the standardized FP64 Rmax metric we use for cross-country comparison. These sources characterize the frontier and its incumbents, not the decision problem of a budget-constrained state converting limited public funds into installed capacity while competing for finite supply. A parallel literature names the resulting inequality the compute divide (Ahmed and Wahed, 2020; Thompson, 2024), documenting how access to computational resources concentrates research and capability among a few firms and elite institutions. This work establishes that the divide exists and matters, but it describes the inequality rather than offering a falsifiable, plan-specific test of whether a given national target is reachable under it.

Analyses of national AI strategies, including the PBIA, are largely qualitative, assessing ambition or institutional design without testing whether stated quantitative targets are physically and fiscally achievable. We are not aware of a reusable model that takes a national compute target as written, converts the committed budget into installable compute under explicit physical constraints, and reports the share of scenarios in which the target is met. This is the gap we fill.

3. Methods

3.1 Overview

We model each country as an agent that, in each year from 2025 to 2030, converts its committed AI budget into newly installed frontier compute, subject to the availability and price of chips — which emerge from a finite global supply and from the country’s access tier — and to the energy available to power them. All compute is measured in FP64 Rmax PFLOPs, the metric used by the TOP500 list, so that model output is directly comparable to the ranking the PBIA targets. Brazil is the subject of the analysis; the other countries are included because their demand competes for the same finite chip supply and therefore sets the price Brazil faces. The model is run as a Monte Carlo experiment of 2,000 scenarios, in which uncertain parameters are sampled from documented ranges, and the output is the share of scenarios in which each stated target is met by its deadline. The full model is implemented in Python and is reproducible with a single command under a fixed seed.

3.2 Why an agent-based model

The central design choice is to model countries as interacting agents rather than as independent budget-to-compute calculations. The justification is that the quantity of compute a fixed budget buys is not independent across countries. Global frontier chip supply is finite, so the price per unit of compute emerges from the aggregate demand of all buyers against that supply. The compute Brazil can purchase therefore depends on the appetite of the Gulf states, India, and the rest of the world. This is a system-level effect that parallel per-country spreadsheets cannot produce, and it is the reason the competing countries must be represented explicitly. We are deliberate about the limits of this choice: as shown in Section 4, the emergent price effect is dormant in the central supply regime and becomes material only under tighter supply, so the agent interaction changes the magnitude of Brazil’s outcome rather than the central verdict.

3.3 Agents and their targets

We model six individual agents — Brazil as the base case, India as a peer, Saudi Arabia and the United Arab Emirates as well-funded ceilings, Egypt as an intermediate case, and Nigeria as a control with no compute target — plus one aggregate “Rest of World” agent that enters only as demand in the chip market. Each agent is assigned the compute target its national plan actually states, with its source, and is seeded with the compute it has already installed as of 2025 (Brazil ≈ 13 PFLOP, from Santos Dumont and the SINAPAD network; India $\approx 36,000$ GPUs $\approx 2,400$ PFLOP). It then pursues purchases toward its target, stopping once an absolute target is reached. The PBIA states two targets: doubling national AI compute by 2027, and placing the Santos Dumont supercomputer among the top five of the TOP500 by 2029. Targets defined over total installed capacity, such as India’s 200,000 GPUs, are evaluated against existing plus newly purchased stock; the Gulf agents are seeded at zero, because their stated target is power capacity in gigawatts rather than a compute stock, a simplification we note as a limitation.

3.4 Mechanics of one round

Each round represents one year and proceeds in steps. First, the top-five bar rises as $B_t = B_0 \cdot g^t$, where $B_0 = 561.2$ PFLOP is the fifth-place FP64 Rmax on the June 2025 TOP500 (Eagle, Microsoft Azure) and g is the annual growth rate of the frontier, sampled around $2.2\times$ per year, consistent with the observed TOP500 trend. Second, each agent forms a budget-driven demand: it buys toward its own target — and, for absolute targets, no further — with cost per PFLOP falling at $1.37\times$ per year (Epoch AI). Third, demand meets a finite supply and the chip price emerges endogenously as $\varphi_t = \max(1, (\Sigma \text{demand/supply})^\eta)$, where access tiers enter as a price premium and η is the price elasticity. Fourth, an energy ceiling limits how much of the purchased compute can be brought online, as a function of data-center capacity in megawatts. Fifth, the effective addition is the minimum of desired, affordable, and powerable compute, and comes online the following year. The competitive stock is the sum of additions over the last three years, reflecting the working life of frontier hardware. Finally, each target is evaluated against this stock at its deadline. The full equations and constants are given in Appendix A.

3.5 Monte Carlo layer

Each parameter with genuine uncertainty is sampled across 2,000 scenarios. Physical and economic parameters are drawn from triangular distributions specified by a plausible low, central, and high

value, chosen because they have bounded support and an interpretable mode, avoiding the infinite tails of a normal distribution. The exchange-rate shock applied to budgets denominated in local currency, for Brazil, India, and Egypt, is drawn from a lognormal distribution with mean one (median ≈ 0.99), while the dollar-denominated Gulf budgets are held fixed. The output aggregates, per country and per target, the fraction of scenarios in which the target is met.

3.6 Provenance and honesty discipline

Every input parameter carries a type label indicating whether it is a measured datum, a quantity derived with a shown calculation, or a modeling assumption. This discipline is a deliberate methodological choice: it makes explicit which results rest on hard data and which rest on judgment. The central verdict — that Brazil’s doubling target is feasible and its top-five target is not — is governed entirely by the labeled data and derived quantities, namely budget, cost, and the frontier growth rate, together with the structural fact that an absolute target is fixed while a ranking target moves. The assumptions affect only margins, such as how close the Gulf comes or how much the queue penalty bites, not the qualitative result, and we demonstrate this directly in Section 4. During validation we identified and corrected an error in the competitive-window logic that had previously understated the difficulty of the top-five target; correcting it strengthened, rather than weakened, the central finding.

3.7 Sensitivity and counterfactual analyses

Beyond the headline feasibility figures, we run four robustness analyses. A single-parameter counterfactual on Brazil changes one input at a time — granting privileged chip access, unlimited energy, or a tenfold budget — to isolate which constraint actually binds. A sweep across the full plausible range of the frontier growth rate searches for the threshold at which the top-five target would become feasible; for Brazil no such threshold exists within that range, the target remaining infeasible even at the slowest plausible growth, while the sweep does locate thresholds for the better-funded Gulf plans. A competition experiment compares the chip price Brazil faces with competitors present versus absent, to quantify the competition tax under scarce supply. Finally, a full robustness sweep fixes each uncertain parameter in turn at the low and high end of its range and re-runs all 2,000 scenarios; across every extreme the verdict is unchanged, with the top-five target feasible in 0% of scenarios and the doubling target in 100%, which shows the result is not an artifact of parameter calibration. These analyses are described before results so that the reasoning is transparent.

3.8 Reproducibility

The model is open-source and runs with `python -m abm_ia.run` under a fixed seed (42) and 2,000 scenarios. Every input parameter lives in a single file, and any assumption can be changed, or any other country’s plan substituted, by editing that file alone.

4. Results

4.1 The allocation prescription holds for one target and fails for the other

Observation. Both PBI targets are funded from the same budget, yet the model places them at opposite ends of feasibility (Table 1). Doubling national AI compute by 2027 is met in 100% of

2,000 scenarios; placing Santos Dumont in the TOP500 top five by 2029 is met in 0%, reaching a median of $\sim 4\%$ of the fifth-place bar at its deadline.

Table 1: Feasibility of the two PBIA compute targets (2,000 Monte Carlo scenarios).

Target	Type	Deadline	Feasible (%)	Median attainment
Double national AI compute	absolute	2027	100%	exceeded ($\sim 17\times$)
Santos Dumont, TOP500 top five	ranking	2029	0%	$\sim 4\%$ of the bar

Interpretation. The governance literature’s core prescription — that a state should allocate budget to domestic compute — is implicitly tested here twice on the same money, and it succeeds once and fails once. This is the paper’s first substantive finding about the Global South: the prescription, derived under Global-North conditions, is not unconditional under scarcity. Whether allocating scarce capital to compute builds capability or dissipates it depends on what the compute is sized against. Pinned to an absolute quantity, the budget delivers; pinned to a position in a ranking that recedes as the frontier grows, the same budget delivers almost nothing. Figure 1 makes the mechanism visible: Brazil clears the fixed doubling line early, while the top-five bar climbs geometrically and the gap to it remains about $23\times$ at the deadline.

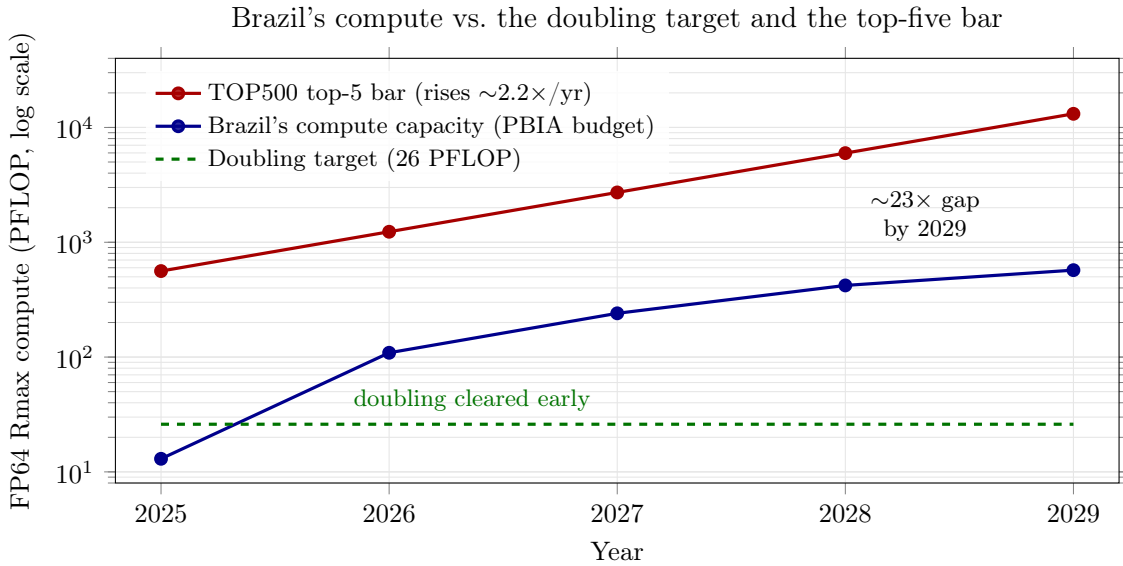


Figure 1: Brazil’s compute capacity (PBIA budget) against the doubling target and the rising top-five bar, 2025–2029, on a log scale. Representative central-parameter scenario; the Monte Carlo median attainment of the bar is $\sim 4\%$ (a $\sim 25\times$ gap). The doubling line (26 PFLOP) is cleared early; the top-five bar recedes at the frontier’s growth rate.

4.2 The shortfall is not the Global-South constraint one expects

Observation. Single-parameter counterfactuals on Brazil isolate what moves the showcase target (Table 2, Figure 2). Privileged chip access $\rightarrow 0\%$; unlimited energy $\rightarrow 0\%$. Only budget moves it at all, and not far: a tenfold annual budget reaches $\sim 7\%$, and raising the compute fraction from 0.08 to 0.60 reaches $\sim 2\%$. The target clears only when budget, access, and energy are raised together to Gulf-like levels ($\rightarrow 100\%$).

Table 2: Single-change counterfactuals on Brazil’s top-five target.

Change applied to Brazil alone	Top-five feasible (%)
Baseline (as declared)	0%
+ privileged chip access	0%
+ unlimited energy	0%
+ budget/yr $\times 10$	$\sim 7\%$
+ compute fraction $0.08 \rightarrow 0.60$	$\sim 2\%$
All combined (Gulf-like)	100%

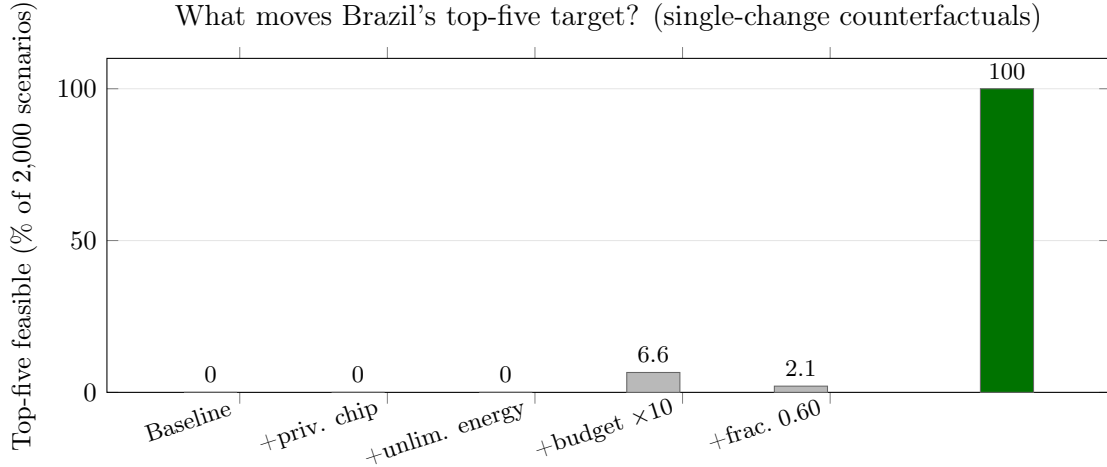


Figure 2: Single-change counterfactuals on Brazil’s top-five target. Chip access and energy each add zero; only budget and the compute fraction move the number at all, and neither makes the target feasible. Only the combined Gulf-like change reaches 100%.

Interpretation. The usual account of the compute divide locates the Global South’s disadvantage outside the state — in contested access to a concentrated hardware market and in energy. The model shows that, for a middle-income economy like Brazil, neither is what binds: removing the access constraint or the energy constraint changes the showcase target’s feasibility by zero. The decisive constraint is internal — how a fixed, scarce budget is allocated against a target. Closing the gap on budget alone would take on the order of ten times the dedicated line, roughly the entire plan, in a single machine, and even that reaches only about 7% of scenarios — capital the state does not have and could not reallocate without cannibalizing other pillars. The compute-governance prescription therefore fails here not because the Global South cannot reach the hardware market, but because the target itself converts sovereign capital into a result the country cannot attain.

4.3 The conditionality is structural, not an artifact

Because the paper’s lesson generalizes beyond Brazil, the divergence must be shown to be a property of the targets rather than of the model’s tuning.

Sampling. At 2,000 scenarios the Monte Carlo standard error of a feasibility share is at most ~ 1.1 percentage points; the headline figures are the extremes 0% and 100%, not estimates near that band.

Invariance. Fixing each uncertain parameter in turn at the low and high end of its range and re-running all scenarios leaves the verdict unchanged — top-five at 0%, doubling at 100% — at every extreme tested (bar anchor, frontier growth rate, cost per PFLOP, chip supply, hyperscaler reserve, competing demand, price elasticity, energy assumptions). The result rests on labeled data and derived quantities — budget, cost, and the frontier growth rate — not on judgment.

No threshold for Brazil. No plausible frontier growth rate makes Brazil’s top-five target feasible; the same sweep does find feasibility thresholds for the far better-funded Gulf plans (UAE at $2.9\times/\text{yr}$, Saudi Arabia at $2.4\times/\text{yr}$) and for India ($4.5\times/\text{yr}$), confirming the model returns success when the budget warrants it.

Interpretation. The conditionality of the prescription is thus structural: an absolute target is reachable on a fixed budget and a ranking target, that makes the compute power comparable to the Global North, is not, for any country at Brazil’s budget scale. This is what licenses generalizing the finding — any Global-South plan that frames a compute goal as a moving rank, trying to achieve compute power comparable with the Global North, inherits the same failure. (For transparency, a validation fix to the competitive-window logic moved Brazil further from the bar, strengthening the finding.)

4.4 The contested market bites, but does not decide the verdict

Observation. The endogenous chip price is the model’s one system-level interaction. Centrally it is inactive — the median price multiplier is $\sim 1.0\times$ and removing all competitors shifts Brazil’s attainment by under 0.2 points (from 3.94% to 4.10% of the bar) — but under a tightened supply regime competitor demand drives the price to $\sim 2.1\times$ and roughly halves the compute Brazil’s fixed budget buys ($\sim 4.1\% \rightarrow \sim 2.1\%$ of the bar).

Interpretation. This is where the introduction’s “concentrated and inflated hardware market” enters as a genuine Global-South handicap: under scarcity, the demand of better-funded buyers raises the price a constrained state pays. But the effect modulates magnitude, never the verdict — even with the market handicap removed, the prestige target fails; even with it at its worst, the doubling target succeeds. The structural design problem dominates the market disadvantage, which we accordingly report as a stress result rather than a central prediction.

4.5 Sized to needs, the same prescription works — and that is the governance point

Observation. Evaluated against their own declared targets, the comparator plans are on track: the Gulf gigawatt build-outs are feasible in 100% of scenarios, and India’s national drive is anchored to an absolute GPU count it is already approaching in reality (Figure 3).

Interpretation. Every plan we model that anchors compute to absolute capacity reaches its goal, including Brazil’s own doubling target. This is the governance stake the introduction sets up, now grounded in the model. In the Global North the compute-governance prescription is met by actors that already sit at the frontier; a resource-constrained state mis-reads it as a mandate to reach that frontier — to place a machine among the world’s five largest, all of them Northern labs and hyperscalers. But the frontier recedes at its own growth rate, so chasing parity with the North on a fixed budget is structurally unwinnable: it would take spending like the North to match the North (our counterfactuals put the gap at roughly ten times the dedicated line, about the entire plan). The cost is not only the missed rank. Under scarcity, pursuing frontier-parity misdirects the same

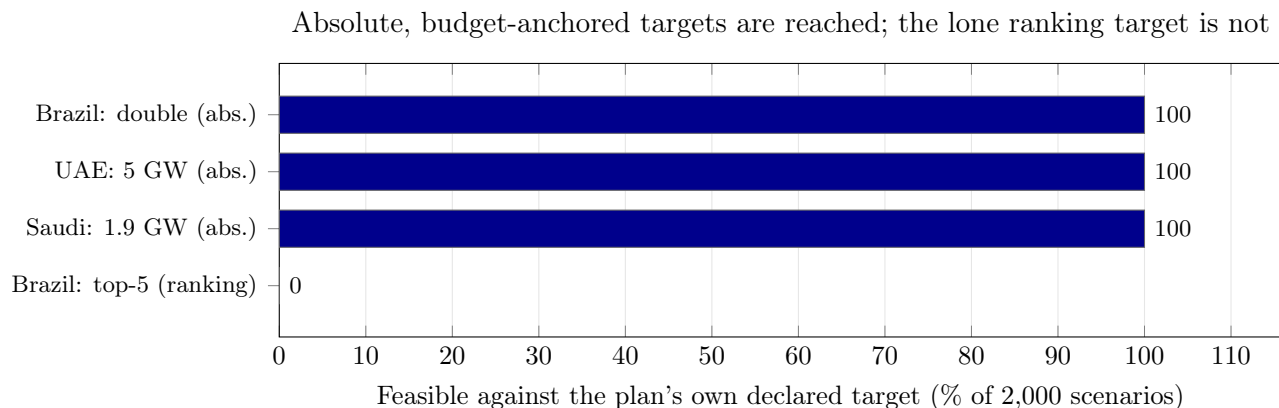


Figure 3: Each plan evaluated against its own declared target. Absolute, budget-anchored targets (Brazil’s doubling, the Gulf gigawatt build-outs) are reached in 100% of scenarios; the only target framed as a position in a moving global ranking, Brazil’s top-five, is reached in 0%.

scarce capital away from the sufficient, needs-based compute that governance actually requires — the capacity to evaluate, audit, and oversee the frontier AI deployed within the country’s borders — and stages a predictable public failure that erodes political support for sovereign compute. The misallocation therefore weakens a country’s ability to govern its own AI at the source, in the budgeting decision itself, before any system is deployed. Sized to its own needs rather than to parity with the North, the same capital builds exactly that governing capacity, which is why the prescription is not wrong for the Global South, only incomplete.

5. Discussion

The result bears directly on AI safety in the Global South. Effective oversight of frontier systems (evaluation, auditing, red-teaming, and the operation of safety-relevant public services) requires compute, but it requires compute sized to those tasks, not compute sized to a global leaderboard. A plan that spends its scarce, dedicated compute line chasing a ranking position both fails to reach the ranking and forgoes the governing capacity the same money could have bought. The misallocation therefore weakens a country’s ability to govern its own AI at the budgeting stage, before any system is deployed.

The broader pattern the model surfaces is that the governability of compute, which the literature treats as an unconditional asset, is conditional on how a target is written. An absolute, budget-anchored target converts a fixed budget into a fixed, reachable quantity. A target defined relative to a moving frontier converts the same budget into a receding goal. For states already at the frontier the distinction is invisible; for states outside it, the distinction decides whether public money builds capability or is dissipated.

5.1 Limitations

Several limitations constrain the scope of our claims, and none favors Brazil. The TOP500 bar measures FP64 (HPL) performance, whereas AI training runs in FP16/FP8 and large industry clusters often do not submit to the list at all; the real AI frontier is therefore above the bar we

measure, which makes the test optimistic. We compare a three-year competitive stock against a single machine, again optimistic, and we use the same three-year window for both a frontier-ranking target and a national-capacity target, even though national capacity arguably should not retire hardware so quickly; over the PBIA’s short horizon this changes little. We treat announced gigawatt capacity as a financed envelope rather than as a budget line competing with chips, and the counterfactual shows this does not change the verdict (energy adds zero for the Gulf). Hyperscalers are exogenous, entering only through the reserve that sets the bar. Finally, the doubling baseline (≈ 13 PFLOP) is an estimate, the 200,000-GPU target mixes public and private capex, and Egypt and Nigeria do not state quantifiable compute targets; where a target is soft we flag it, and the Brazilian verdict is robust to these.

5.2 Future Work

The natural next steps are to model hyperscalers as endogenous market agents rather than an exogenous reserve, which would let scarcity and the emergent price effect arise endogenously rather than only in stress scenarios; to separate power from chips as a competing budget line; and to extend the audit beyond compute to the other PBIA pillars. Because the model reads every assumption and every country’s target from a single parameter file, the most immediate extension is to apply the same budget-versus-physics test to other national AI plans.

6. Conclusion

Using an agent-based model with a 2,000-scenario Monte Carlo layer, we tested whether the two compute targets of Brazil’s national AI plan can be bought with the budget the plan committed. The central target, doubling national AI compute by 2027, is feasible in 100% of scenarios. The showcase target, placing Santos Dumont among the top five of the TOP500 by 2029, is feasible in 0%, reaching a median of about 4% of the required bar. The verdict is invariant across every parameter extreme we tested, and counterfactuals show that the showcase target fails on framing and scale rather than on chip access or energy: those add nothing, and even a tenfold budget reaches only about 7%.

The implication for compute policy in the Global South is that national compute goals should be written as absolute, budget-anchored capability targets, sized to the country’s own needs, rather than as positions in a moving global ranking. A ranking target recedes at the growth rate of the frontier and cannot be reached on a fixed budget, so it converts scarce sovereign capital into a predictable public failure while diverting it from the compute that governance actually requires. Sized to need, the same money builds that capacity.

6.1 Code and Data

- Code repository: <https://github.com/brunodebiased/HackatonGloblSouth>
- The full model (`abm_ia/`), the methodology and validation specifications, and all results (`saida/*.csv`) are in the repository. Reproduce with `python -m abm_ia.run` (seed = 42, 2,000 scenarios).
- Primary source: Plano Brasileiro de Inteligência Artificial (PBIA, MCTI 2024), included in the repository.

References

1. Sastry, G., et al. (2024). *Computing Power and the Governance of Artificial Intelligence*. arXiv:2402.08797.
2. Ahmed, N., & Wahed, M. (2020). *The De-democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research*. arXiv:2010.15581.
3. Hobbhahn, M., et al. / Epoch AI (2023). *Trends in GPU Price-Performance*. Epoch AI. <https://epoch.ai/>
4. Epoch AI (2024). *Data on Machine Learning Hardware*. <https://epoch.ai/data>
5. TOP500 (2025). *TOP500 List, June 2025*. <https://www.top500.org/lists/top500/list/2025/06/>
6. Ministério da Ciência, Tecnologia e Inovação (2024). *Plano Brasileiro de Inteligência Artificial (PBIA) 2024–2028*. Brasília: MCTI.

A. Model equations and constants

Time horizon. One round equals one year, run from $t = 0$ (2025) to 2030. All compute is in FP64 Rmax PFLOPs; all money in US\$ billions; energy in MW.

Per-round dynamics (country i , year index t):

- (1) **Bar.** $B_t = B_0 \cdot g^t$, with $B_0 = 561.2$ PFLOP (TOP500 June 2025, 5th place = Eagle/Azure) and $g \sim \text{tri}(1.5, 2.2, 2.7)$.
- (2) **Budget-driven demand.** $D_{i,t} = \min(\text{resources}_i/c_t, \text{gap-to-target}_i)$, with cost per PFLOP $c_t = c_0/1.37^t$, $c_0 = 0.0005$ US\$bn/PFLOP, and $\text{resources}_i = \text{budget}_i \cdot \text{frac}_i \cdot \text{fx}_i + \text{private}_i$. Absolute targets (GPUs, doubling) impose a ceiling so the agent stops at the target; ranking and capacity targets have no demand-side ceiling.
- (3) **Endogenous chip price.** Available supply $S_t = O_0 \cdot 1.30^t \cdot (1 - 0.72)$ (net of the hyperscaler reserve); per-country demand with tier premium $d_{i,t} = D_{i,t}/q_i$; Rest-of-World demand $R_t = S_t \cdot 0.35$; price $\varphi_t = \max(1, ((\sum_i d_{i,t} + R_t)/S_t)^\eta)$ with $\eta \sim \text{tri}(0.6, 1.0, 1.5)$; chips bought $\text{chip}_{i,t} = d_{i,t}/\varphi_t$.
- (4) **Energy ceiling.** $E_{i,t} = \text{DC}_{i,t} \cdot 0.50 \cdot (55 \cdot 1.15^t)$ PFLOP, where $\text{DC}_{i,t}$ (MW) ramps toward an announced capacity target (Gulf) or grows organically at $1.14\times/\text{yr}$ otherwise.
- (5) **Effective addition.** $\text{add}_{i,t} = \min(D_{i,t}, \text{chip}_{i,t}, E_{i,t} - \text{stock})$, online the following year.
- (6) **Competitive stock.** Sum of additions over the last three years.
- (7) **Target evaluation** at each deadline: top-five $\Rightarrow \text{stock} \geq B_t$; doubling $\Rightarrow \text{base} + \text{stock} \geq 2 \text{ base}$ with $\text{base} \approx 13$ PFLOP; GPUs $\Rightarrow \text{stock}/0.067 \geq N$; gigawatts $\Rightarrow \text{DC}_{i,t}/1000 \geq N$.

Chip access tiers (price premium q): privileged 1.00, open 1.15, queue 1.50 (Brazil), restricted 2.20, blocked ∞ .

Table 3: Input constants. Type: DADO = measured, DERIV = derived (arithmetic shown in repo), SUPOS = modeling assumption, ESTIMATIVA = to confirm.

Constant	Central	MC range	Type	Source / note
TOP500 #5 bar (PFLOP)	561.2	478–793	DADO	TOP500 Jun/2025, 5th = Eagle/Azure
Bar growth ($\times/\text{yr}$)	2.2	1.5–2.7	DADO (FP64)	TOP500 #5 $\sim 2.45\times/\text{yr}$ (2022–24)
Cost per PFLOP (US\$bn)	0.0005	0.00035–0.00075	DERIV	El Capitan/Frontier/Aurora + premium
Price–performance ($\times/\text{yr}$)	1.37	1.20–1.50	DADO	Epoch AI
PFLOP FP64 per GPU	0.067	fixed	DADO	H100 datasheet
PFLOP per MW	55	45–70	DERIV	Rmax/power, El Capitan, Frontier
Energy efficiency ($\times/\text{yr}$)	1.15	1.05–1.25	SUPOS	$\sim$ half of price–perf gain
DC fraction for AI	0.50	0.30–0.70	SUPOS	modeling choice
DC growth ($\times/\text{yr}$)	1.14	1.08–1.25	DADO	McKinsey/Synergy $\sim 14\%$ CAGR
Max DC build rate ($\times/\text{yr}$)	1.7	1.4–2.0	SUPOS	physical build-out limit
World chip supply (PFLOP)	350,000	200k–500k	DERIV	GPU volume $\times$ Rmax FP64/GPU
Supply growth ($\times/\text{yr}$)	1.30	1.15–1.50	DERIV	TSMC/NVIDIA (CoWoS) guidance
Hyperscaler reserve	0.72	0.60–0.85	DERIV	Synergy (2025 DC capex)
Rest-of-World demand	0.35	0.20–0.55	SUPOS	~ 80 -country aggregate; calibrated
Chip price elasticity	1.0	0.6–1.5	SUPOS	price response to scarcity
FX shock σ	0.15	—	SUPOS	emerging-market volatility
Doubling baseline (PFLOP)	13	fixed	ESTIMATIVA	Santos Dumont 5.1 + SINAPAD/ICT
Competitive window (yr)	3	fixed	choice	frontier working life of a cluster

B. Reproduction and additional results

Reproduction. Install `requirements.txt` and run `python -m abm_ia.run` (fixed seed 42, 2,000 scenarios). Inputs live in `abm_ia/params.py` (global constants) and `abm_ia/countries.py` (each agent’s declared target with source). Outputs are written to `saida/*.csv`. To test another country’s plan, edit only its goal in `countries.py`.

Robustness sweep. Fixing each load-bearing parameter at the low and high end of its range (20 configurations: bar anchor, frontier growth rate, cost per PFLOP, price–performance, Rest-of-World demand, price elasticity, DC fraction, energy efficiency, max build rate, world supply, hyperscaler reserve) and re-running all 2,000 scenarios leaves Brazil’s verdict unchanged at every extreme: top-five 0%, doubling 100%.

Table 4: Frontier-growth threshold for a 50% feasible top-five placement, by plan.

Plan	Bar growth ($\times/\text{yr}$) for 50% feasibility
UAE (5 GW)	2.9
Saudi Arabia (1.9 GW)	2.4
India (200k GPUs)	4.5
Brazil	none in the plausible range
Egypt	none

Competition tax. In the central supply regime the median chip-price multiplier is $\sim 1.0\times$, and removing every competitor shifts Brazil’s attainment from 3.94% to 4.10% of the bar (under 0.2 points). Scarcity binds in $\sim 37\%$ of scenarios; in a deliberately tightened supply regime the price reaches $\sim 2.1\times$ and Brazil’s attainment falls to $\sim 2.1\%$ of the bar. The effect changes magnitude, not the verdict.

LLM Usage Statement

We used Claude Opus 4.8 to assist with the more complex mathematical and programming aspects of the project, namely the agent-based model and the Monte Carlo layer, and to help structure the text and with Portuguese-language corrections. We validated and attempted to refute our own thesis ourselves, and additionally used several LLMs, including Claude Opus 4.8 and ChatGPT 5.5, to check for possible flaws in the model and the argument. All results and claims were independently verified by the authors.