CLIMATE MOONSHOT PORTFOLIO

Billions to save trillions

Global Climate Model Experimental Research Initiative

Proposed $1 Billion / Five-Year R&D Program

12 PROGRAM AREAS

By Glenn Weinreb

Developed with assistance from ChatGPT

OPEN SOURCE

Executive Summary

Purpose. This proposed $1 billion, five-year R&D initiative is designed to make global climate models more decision-ready for questions about how much sunlight reflection would be required, where and when it might be applied, and what regional risks or benefits could follow. It combines observations, field campaigns, laboratory measurements, natural perturbations, high-resolution modeling, data assimilation, and independent validation. The program is research infrastructure for climate understanding and risk assessment; it is not authorization for SAI or MCB deployment.

Scientific hypothesis. The largest practical errors for a reflectivity decision are not captured by global-mean temperature alone. Clouds, aerosols, Earth's radiation budget, ocean heat uptake, AMOC and freshwater pathways, sea ice, deep convection, and regional precipitation can compensate for one another in a model. The initiative therefore tests the physical links separately and jointly, with traceable measurements and held-out observations, so that a model must reproduce the climate for the right reasons rather than because offsetting errors happen to cancel.

Five-year work. The portfolio allocates the $1 billion across eleven scientific and infrastructure areas plus a centrally controlled Funding-Source reserve: Earth radiation-budget and albedo closure; anthropogenic aerosol forcing; marine low clouds; mixed-phase and high-latitude clouds; deep convection and precipitation; AMOC, ocean heat and freshwater observations; rapid-response natural perturbations; storm-resolving and large-eddy modeling; model evaluation, data assimilation and uncertainty quantification; a shared observing/calibration backbone; international participation and open-data infrastructure; and the central reserve. Field measurements are linked directly to model-development and prospective prediction tests.

Success criterion. The objective is measurable reduction and transparent propagation of uncertainty in the quantities policymakers would actually need: required radiative forcing, spatial and seasonal intervention patterns, regional temperature and precipitation effects, ocean-circulation response, monitoring thresholds, failure signals, and the consequences of stopping or changing an intervention. Success means multiple independent models can predict protected observations and natural perturbations better without degrading energy, water, or regional-climate constraints elsewhere. Remaining uncertainty and model disagreement are reported explicitly rather than hidden by averaging.

Climate Moonshot Portfolio. We define "Climate Moonshot" as conducting R&D to the extent required to solve the climate problem. This would require a portfolio of R&D initiatives, an example of which is at www.APlanToSaveThePlanet.org. The document you are reading is one initiative within that portfolio. It uses a Common Framework that establishes shared rules for governance, procurement, data rights, independent review, and oversight. All documents within this portfolio are open-source and may be freely viewed, used, and modified.

Program at a Glance

The proposed program aims to improve the accuracy of global climate models.

Item	Program design
Total funding / duration	$1.0B over five years
Mission	Make global climate models more decision-ready for evaluating sunlight-reflection quantity, location, timing, monitoring, and risk.
Scientific chain	Radiation -> clouds/aerosols -> ocean heat and circulation -> temperature/precipitation -> regional impacts -> intervention strategy.
Program architecture	11 scientific/infrastructure areas plus a centrally controlled Funding-Source reserve.
Principal uncertainties	Radiation budget, aerosol forcing, low clouds, mixed-phase clouds, convection, ocean heat uptake, AMOC, and regional precipitation.
Field strategy	Traceable observations, natural perturbations, laboratory work, multi-platform field campaigns, and only bounded deliberate perturbations that pass separate authorization.
Modeling strategy	LES and storm-resolving models -> scale-aware parameterizations -> multiple structurally different global models.
Evidence standard	Preregistered predictions, protected observations, simultaneous budget closure, independent evaluation, and explicit model disagreement.
Decision quantities	Required radiative forcing; spatial/seasonal pattern; regional temperature and precipitation; ocean/AMOC response; monitoring and termination risk.
Stage-gate logic	Fund the next increment only when the preceding evidence still justifies it; redesign or stop weak paths.
Five-year endpoint	Independent decision-readiness assessment identifying what improved, what remains uncertain, what should continue, and what should stop.
Boundary	Research and decision support only; completion does not authorize SAI or MCB deployment.

Table of Contents

Executive Summary 2
Program at a Glance 3
Table of Contents 4
What Is a Global Climate Model, and Why Improve It? 6
What a Reflectivity Decision System Must Calculate 7
Program Architecture and $1 Billion Budget 11
Program Crosswalk 13
Relationship to SAI and MCB Experimental Research 14
Stage Gates and Go/No-Go Criteria 15
Model Validation, Acceptance, and Decision Readiness 16
Area 1. Earth Radiation Budget and Planetary-Albedo Benchmark - $100M 20
Area 2. Anthropogenic Aerosol Radiative-Forcing Closure - $175M 25
Area 3. Marine Low-Cloud and Cloud-Feedback Experiments - $145M 30
Area 4. Mixed-Phase, High-Latitude Cloud and Surface-Albedo Experiments - $100M 35
Area 5. Deep Convection, Precipitation and Tropical-Organization Experiments - $90M 39
Area 6. AMOC, Ocean Heat Uptake, Freshwater and Sea-Ice Observations - $120M 43
Area 7. Rapid-Response Natural-Perturbation Program - $55M 48
Area 8. Storm-Resolving, Large-Eddy and Process-Model Experiments - $70M 52
Area 9. Model Evaluation, Data Assimilation and Uncertainty Quantification - $55M 56
Area 10. Observing Backbone, Calibration and Autonomous Platforms - $45M 60
Area 11. International Participation, Governance and Open-Data Infrastructure - $25M 64
Area 12. Central Reserve and Rapid-Response Contingency - $20M 68
Policymaker Questions and Program Boundaries 72
Termination Shock 74
Campaign Development and Execution Phases 79
Purchase Orders 83
Recipients of Purchase Orders 91
Project Management 95
Funding Source Requirements 96
Top GCM Science Papers 97

Top Researchers and Organizations 103
Scientific Advisory Panel 107
Management Advisory Panel 108
Agreements 109
GCM Aircraft, Instruments and Airborne Measurement Architecture 111
SAI / MCB / GCM Aircraft and Instrument Overlap 117
Existing Infrastructure to Leverage 119

What Is a Global Climate Model, and Why Improve It?

A global climate model is a numerical representation of the atmosphere, oceans, land, sea ice and their interactions. It divides Earth into three-dimensional grid cells, calculates the movement of energy, water and momentum, and represents processes that occur below the grid scale through parameterizations. Modern Earth-system models also include atmospheric chemistry, aerosols, vegetation, the carbon cycle and other components.

These models are indispensable, but they are not perfect replicas of Earth. They reproduce many large-scale features and broad historical trends, yet important quantities still differ among models and from observations. The largest practical uncertainties for sunlight-reflection decisions include clouds, aerosol-cloud interactions, the preindustrial aerosol baseline, ocean heat uptake, regional precipitation, sea-ice feedbacks and circulation responses.

For ordinary climate projection, an ensemble of imperfect models can still provide useful risk information. A reflectivity program, however, would ask a more demanding control question:

> Given a proposed intervention, can we predict the resulting change in top-of-atmosphere radiation, global temperature, regional temperature gradients, precipitation, sea ice, ocean circulation and ecological risks with enough accuracy to choose a quantity, location and schedule?

That question cannot be answered by tuning a model to match global-mean temperature alone. Different combinations of cloud feedback, aerosol forcing, ocean heat uptake and natural variability can produce a similar temperature history. A model may appear correct because errors compensate. The research program must therefore test the physical links separately and jointly.

The central scientific chain

The proposed program would connect observations to decisions through the following chain:

> Energy entering Earth system → reflection by clouds, aerosols and the surface → absorption by the atmosphere and ocean → heat transport and storage → temperature and precipitation response → tipping-element and regional impacts → reflectivity control strategy

A reliable decision system must close the energy budget at the top of the atmosphere, explain where the energy goes, and reproduce how the climate responds over days, seasons, decades and centuries.

Why observations alone are not enough

Measurements reveal what happened in the real atmosphere and ocean, but the historical record contains only one realization of climate and does not include a deliberate global sunlight-reflection program. Observations also contain sampling gaps and instrument discontinuities. Models are needed to separate causes, fill gaps, run counterfactual scenarios and estimate future outcomes. The goal is therefore not to replace models with observations; it is to use carefully designed experiments and observations to make models more physically constrained, independently testable and decision-ready.

Why modeling alone is not enough

Model intercomparisons can identify robust behavior and disagreement, but models often share assumptions, parameterizations and training datasets. Agreement among related models is not the same as validation. The proposed fund would emphasize field campaigns, laboratory measurements, long-term observing systems, natural experiments and prospective predictions made before observations are revealed.

What success would mean

The program would not claim that climate can be predicted perfectly. Success would mean that the dominant uncertainty sources are measured, bounded and propagated into decisions. A credible output would state, for example:

- How much additional reflected shortwave radiation is required to reduce a specified global warming rate, with an uncertainty range.
- How alternative latitude, season and geographic patterns affect regional temperature, precipitation and circulation.
- How much of the historical temperature trend is attributable to greenhouse gases, anthropogenic aerosols, natural variability and changes in ocean heat uptake.
- Whether an intervention designed to reduce one risk, such as extreme heat, increases another risk, such as regional drought or disruption of ocean circulation.
- Which observations would reveal early that the model prediction is failing and trigger adjustment or termination.

Climate Science Videos

```
The Science of Global Warming [#7]
https://www.youtube.com/watch?v=Por9aWKLdc4

The Uncertainty of Climate Change [#8]
https://www.youtube.com/watch?v=HoqX7uBaeKU
```

What a Reflectivity Decision System Must Calculate

Prompt to AI: What should an improved climate-model system calculate before anyone decides how much sunlight to reflect, from where, and when?

Response From AI:

A reflectivity decision system should be formulated as a constrained, uncertain control problem. It must connect an intervention pattern to radiative forcing, climate response, harms, benefits and monitoring signals. It should not output one deceptively precise number.

1. Required radiative forcing

The system should estimate how much additional reflected shortwave energy is required to meet an explicit objective, such as reducing the decadal global warming rate, limiting peak temperature, slowing ice loss or reducing the probability of an AMOC transition. It should report watts per square meter, uncertainty and the time dependence of the required forcing.

2. Spatial distribution

The same global-mean forcing can have different consequences depending on latitude, longitude and altitude. The model system should compare multiple spatial patterns and quantify effects on hemispheric temperature balance, equator-to-pole gradients, monsoons, storm tracks, tropical rainfall and ocean circulation.

3. Seasonal timing

Seasonal interventions could affect sea ice, snow, clouds and circulation differently from a year-round forcing. The system should test whether timing can reduce specific risks while avoiding large precipitation or ecological penalties.

4. Intervention physics

For SAI, models must translate material, particle size, altitude, latitude and release schedule into aerosol evolution and radiative forcing. For MCB, they must translate sea-salt emission, particle size and meteorological conditions into cloud droplet number, cloud brightness, lifetime, precipitation and regional radiation. The climate model should receive realistic forcing from validated process models rather than an idealized reduction in the solar constant alone.

5. Climate objectives and constraints

Decision quantity	What should be calculated
Global mean surface temperature	Level and rate of change; land-ocean and day-night contrasts.
Temperature gradients	Interhemispheric and equator-to-pole differences that influence circulation.
Top-of-atmosphere radiation	Shortwave reflection, longwave emission and net planetary imbalance.
Precipitation and evaporation	Global hydrological sensitivity, monsoons, drought and extreme rainfall.
Ocean heat content	Heat uptake by basin and depth, including deep-ocean commitment.
AMOC and ocean circulation	Overturning strength, freshwater transport, convection and threshold indicators.
Sea ice, ice sheets and glaciers	Mass balance, albedo feedback and regional sea-level implications.
Ozone, air quality and deposition	Intervention-specific chemistry and surface exposure.
Ecosystems and food systems	Direct/diffuse sunlight, temperature, rainfall, crop and marine impacts.
Equity and regional risk	Distribution of benefits, harms and uncertainty among countries and communities.

6. Counterfactual baseline

Every intervention outcome must be compared with a credible no-intervention pathway. The baseline should include uncertainty in emissions, greenhouse-gas concentrations, aerosol emissions, land use, volcanic activity and natural variability. Otherwise the apparent benefit or harm may be an artifact of the chosen baseline.

7. Uncertainty and probability

The system should propagate observational, parametric, structural, scenario and internal-variability uncertainty. It should present probability distributions and robust ranges rather than simply averaging models. Where models share code or assumptions, they should not be counted as statistically independent.

8. Monitoring and adaptive control

A real program would need feedback: observe the climate, compare it with predictions, update the model and adjust or stop the intervention. The model system should identify which signals emerge fastest and which instruments are needed to detect divergence from predictions.

- Top-of-atmosphere reflected shortwave radiation.
- Aerosol or cloud-property changes attributable to the intervention.
- Global and regional temperature gradients.

- Precipitation and circulation indicators.
- Ocean heat-content and AMOC precursor indicators.
- Ozone, air-quality and ecological warning signals.

9. Stress tests and failure modes

The decision system should be tested against abrupt volcanic eruptions, El Niño and La Niña, wildfire smoke, unexpected aerosol-emission changes, satellite gaps, instrument drift, model bias and intervention interruption. It should show what happens when a forecast is wrong, when a delivery system underperforms or when international participation changes.

10. A decision-ready output

A final model assessment should produce a transparent portfolio rather than a single optimal plan. Each option should include the expected climate outcome, uncertainty, regional tradeoffs, monitoring plan, reversibility, termination risk and conditions under which the option should not proceed.

Program Architecture and $1 Billion Budget

Prompt to AI: If a funding source wants to spend $1 billion over five years on experiments that improve global climate-model accuracy for sunlight-reflection decisions, what program should it consider?

Response From AI:

A $1 billion, five-year program should concentrate on the physical processes that most strongly limit estimates of Earth's energy imbalance, cloud and aerosol forcing, ocean heat uptake and regional climate response. It should combine measurements, controlled perturbations, natural experiments, laboratory work, high-resolution modeling and independent global-model evaluation.

The program should not be judged by the number of model runs or publications. It should be judged by whether it reduces uncertainty in a defined set of decision quantities: required radiative forcing, spatial and seasonal intervention pattern, temperature and precipitation response, ocean-circulation response, and the probability of crossing or avoiding major climate thresholds.

Illustrative $1 billion portfolio

Program	Five-year budget	Main purpose
Area 1. Earth radiation budget and planetary-albedo benchmark	$100M	Measure and reconcile incoming and outgoing radiation, albedo trends and cloud radiative effects.
Area 2. Anthropogenic aerosol radiative-forcing closure	$175M	Measure direct and cloud-mediated reflection and absorption from pollution and its removal.
Area 3. Marine low-cloud and cloud-feedback experiments	$145M	Constrain subtropical low-cloud feedbacks and cloud susceptibility relevant to MCB.
Area 4. Mixed-phase, high-latitude cloud and surface-albedo experiments	$100M	Improve Arctic and Southern Ocean cloud, sea-ice and snow feedbacks.
Area 5. Deep convection, precipitation and tropical-organization experiments	$90M	Constrain convective aggregation, anvil clouds, humidity and rainfall response.
Area 6. AMOC, ocean heat uptake, freshwater and sea-ice observations	$120M	Measure circulation, mixing and thresholds that control heat uptake and AMOC risk.
Area 7. Rapid-response natural-perturbation program	$55M	Observe volcanic eruptions, wildfire smoke, shipping-fuel changes and major pollution shifts.
Area 8. Storm-resolving, large-eddy and process-model experiments	$70M	Translate observations into improved parameterizations and scale-aware physics.
Area 9. Model evaluation, data assimilation and uncertainty quantification	$55M	Run held-out tests, causal attribution, model weighting and decision-focused ensembles.

Area 10. Observing backbone, calibration and autonomous platforms	$45M	Maintain common radiometry, aircraft, ship, buoy, balloon and satellite-calibration systems.
Area 11. International participation, governance and open-data infrastructure	$25M	Support Global South leadership, independent review, engagement and permanent archives.
Area 12. Contingency and rapid-response reserve	$20M	Respond to rare events, failed instruments, ship or aircraft delays and unexpected findings.
Total	$1.000B	

These are order-of-magnitude planning estimates in 2026 dollars, not vendor quotations. Some facilities and satellite missions would be contributed through government or international agreements rather than conventional purchase orders.

Five-year sequence

Period	Principal work
Year 1	Governance, data audit, calibration, baseline models and campaign design.
Year 2	Begin radiation, aerosol, low-cloud and AMOC campaigns; activate rapid response and prospective tests.
Year 3	Major cloud, convection and aerosol campaigns; storm-resolving simulations and parameterization development.
Year 4	Replicate contrasting regimes; test revised physics in multiple models; run reflectivity-control ensembles.
Year 5	Independent replication, held-out validation, integrated energy-budget assessment and public decision-readiness report. Completion does not authorize deployment.

Program boundaries

The fund should explicitly prohibit:

- Any experiment intended to create detectable regional or global cooling.
- Using the climate-model program as an implicit authorization for SAI or MCB deployment.
- Tuning only to global-mean temperature while ignoring precipitation, radiation, ocean heat content or regional patterns.
- Withholding measurement data, model code or negative results for proprietary advantage.
- Treating agreement among models as validation when the models share parameterizations or calibration data.
- Continuing an expensive campaign after the decision-relevant uncertainty has already been resolved or the measurement cannot meet its required accuracy.

Program Crosswalk

This crosswalk links each budget Area to the uncertainty it is intended to reduce, the evidence it must produce, and the decision quantity that evidence supports. Purchase-order references later in the proposal are mappings to this same budget, not additional funding.

Area	Budget	Principal uncertainty	Main evidence	Decision impact / gate
1	$100M	Planetary radiation / albedo	Radiometric closure plus ocean-heat reconciliation	Required forcing; G1-G5
2	$175M	Anthropogenic aerosol forcing	Source-to-aerosol-to-cloud-to-flux closure	Historical forcing; SAI/MCB process constraints; G2-G5
3	$145M	Marine low-cloud feedback / MCB susceptibility	Field campaigns, LES, held-out cloud-response tests	MCB effectiveness; cloud feedback; G2-G5
4	$100M	Polar mixed-phase clouds / surface albedo	Cloud-phase and surface-energy closure	Polar response, sea ice, regional forcing; G2-G5
5	$90M	Convection / precipitation / organization	Life-cycle convection and water/energy-budget closure	Monsoons and regional hydrology; G2-G5
6	$120M	Ocean heat, freshwater, AMOC	Arrays, process experiments, initialized predictions	AMOC risk and delayed climate response; G1-G6
7	$55M	Natural / policy perturbation response	Rapid observations plus prospective forecasts	Model transferability and forcing response; event gates
8	$70M	Scale transfer and parameterization	LES/CRM/storm-resolving intercomparisons	Scale-aware physics for global models; G2-G5
9	$55M	Model credibility and uncertainty	Protected tests, assimilation, uncertainty engine	All decision quantities; all gates
10	$45M	Measurement comparability	Calibration, modular payloads, intercomparisons	Evidence quality and traceability; G1-G5
11	$25M	Regional validity and reproducibility	Regional leadership, replication, usable open science	Regional risk, independent checking; all gates
12	$20M	Unplanned high-value need	Funding-Source-approved reserve action with accepted deliverable	Program resilience; formal reserve approval

Common evidence chain for every Area

Baseline uncertainty -> experiment or observation -> measured evidence -> model or parameterization change -> held-out validation -> revised uncertainty -> decision impact.

Relationship to SAI and MCB Experimental Research

The GCM initiative should integrate evidence from the separate SAI and MCB experimental programs without paying twice for the same release system, aircraft block, observing platform, or field campaign. The division below keeps one technical owner for each activity while preserving independent model validation.

Program	Owns	Boundary / interface
GCM Experimental Research	Climate-response modeling; common benchmarks; uncertainty propagation; natural perturbations; cross-program model validation; energy, water, ocean, AMOC, regional-impact and termination analysis.	May fund incremental validation measurements or independent analysis on an SAI/MCB campaign, but not duplicate the source system or baseline campaign.
SAI Experimental Research	Deliberate stratospheric release experiments; SAI-specific plume, chemistry, microphysics, optical forcing, ozone/heating and forcing-decay measurements.	Provides measured release-to-forcing evidence to GCM; GCM owns the broader coupled-climate response assessment.
MCB Experimental Research	Sea-salt source, plume delivery, cloud activation, cloud response, radiative closure and MCB-specific continuity/decay measurements.	Provides measured source-to-cloud-to-radiation evidence to GCM; GCM owns broader regional and coupled-climate response.
Shared assets	Monitoring aircraft, calibration standards, radiometers, lidar, aerosol/cloud instruments, common data systems and selected autonomous platforms.	Fund the common capability once. Other initiatives pay only incremental adaptation, operating time, consumables or unique requirements.

Shared-asset rule

The Program Management Office should maintain one cross-program asset register identifying owner, funding source, availability, calibration status, utilization, incremental charges, and the initiative responsible for each campaign. This implements the Common Framework shared-asset rule and prevents duplicate capitalization.

Stage Gates and Go/No-Go Criteria

These are program-level gates. They implement the Common Framework evidence-before-scale principle for this initiative. Campaign-Level Gates C0-C7 in the later Campaign Development and Execution Phases chapter remain subordinate operational gates for individual field campaigns. Passing a science gate never authorizes a deliberate perturbation; any release requires separate environmental, legal, ethics and safety approval, public disclosure, source limits, monitoring, stop conditions, and may be cancelled while the surrounding research Area continues.

Gate	Timing	Required evidence	Decision enabled
G0 - Mission and baseline	Month 6	Decision quantities, baseline models, benchmark datasets, uncertainty map, alternatives, legal/safety boundaries, and initial value-of-information ranking.	Freeze or revise mission and research priorities.
G1 - Observing-system readiness	Month 12	Traceable calibration, protected datasets, preregistered metrics, shared-asset plan, campaign designs, model baselines, permits strategy, and independent review.	Authorize major field campaigns and protected prediction tests.
G2 - First scientific discrimination	Months 18-24	Early field/process evidence shows major uncertainties are observable, competing explanations can be separated, and proposed model changes remain plausible.	Continue, down-select, redesign, or stop weak paths before larger spending.
G3 - Mid-program integration	~Month 30	Revised physics improves protected process/regime tests in multiple models without material compensating degradation; updated uncertainty and cost/schedule review.	Authorize major replication, Year-4 campaigns, and integrated strategy testing.
G4 - Cross-regime replication	Months 40-42	Improvements transfer across regions/seasons and independent implementations; data quality, conservation, and no-compensating-error tests pass.	Proceed to integrated decision-focused ensembles or return for redesign.
G5 - Reflectivity-strategy readiness	~Month 50	Multiple models can evaluate physically realistic SAI/MCB forcing patterns with explicit regional, ocean, monitoring, and termination uncertainty.	Authorize final strategy comparisons and independent assessment.
G6 - Five-year decision-readiness assessment	Month 60	Accepted archive, audited program evidence, protected-test results, quantified remaining disagreement, monitoring requirements, unresolved risks, and explicit research recommendations.	Proceed with targeted follow-on research, redirect, or stop specific paths. No deployment authorization is implied.

Model Validation, Acceptance, and Decision Readiness

Prompt to AI: How should the program determine whether climate models have improved enough to support decisions about reflecting sunlight?

Response From AI:

The program should not declare victory because a new model has finer resolution, more processes or better agreement with the historical global temperature curve. It should use a formal Model Validation, Acceptance, and Decision Readiness framework with preregistered tests, held-out observations, cross-variable constraints and explicit uncertainty targets.

Model Validation Ladder

Level	Required evidence
Level 1: Numerical and conservation integrity	The model conserves energy, water, mass and tracers; results are reproducible across machines and documented versions.
Level 2: Process fidelity	Laboratory and field-scale processes—cloud microphysics, aerosol aging, turbulence, convection, mixing—match observations within stated uncertainty.
Level 3: Regime and regional fidelity	The model reproduces cloud regimes, radiation, precipitation, ocean transports and variability in multiple regions not used for tuning.
Level 4: Historical causal attribution	The model explains observed changes in temperature, radiation, aerosols, clouds and ocean heat without implausible compensating errors.
Level 5: Prospective prediction	The model successfully predicts withheld campaigns, natural perturbations or later periods before seeing the data.
Level 6: Intervention transferability	Validated process models translate realistic SAI or MCB operations into forcing, and global models reproduce responses across independent implementations.
Level 7: Decision readiness	Uncertainty in required forcing, spatial pattern, timing and major risks is narrow enough for a specified decision and monitoring system.

Core benchmark suite

Benchmark family	Examples
Planetary energy budget	Reflected shortwave, outgoing longwave, net imbalance and trend consistency with ocean heat uptake.
Cloud radiative effect	By low-cloud, deep-convective, mixed-phase, polar and transition regimes.

Benchmark family	Examples
Aerosol direct forcing	Optical depth, absorption, vertical profile, composition and clear-sky flux response.
Aerosol-cloud forcing	CCN-to-droplet-to-cloud-to-radiation chain under controlled meteorological conditions.
Hydrological cycle	Global and regional precipitation, evaporation, water vapor, monsoons and extremes.
Ocean heat uptake	Basin and depth distribution, mixing, air-sea flux and heat-content trends.
AMOC and transports	Mean state, variability, heat/freshwater transport and response to freshwater/heat forcing.
Sea ice and snow	Seasonal cycle, thickness, albedo, melt and coupling to clouds and ocean.
Natural perturbations	Volcanic aerosol, wildfire smoke, ENSO and rapid emissions changes.
Intervention experiments	Idealized and physically realistic SAI/MCB scenarios, including spatial patterns, termination and monitoring.

No-compensating-error rule

A model should not pass because a positive cloud bias cancels a negative aerosol-forcing bias, or because excessive ocean heat uptake hides excessive climate sensitivity. The evaluation unit should diagnose component budgets and causal relationships. Every claimed improvement should include a “what got worse?” report across the full benchmark suite.

Prospective prediction challenges

Challenge	Design
Campaign forecast	Before a field campaign, predict cloud, aerosol, radiation or ocean variables for specified cases and archive the prediction.
Natural-event forecast	When a volcano, wildfire or emissions perturbation occurs, issue predictions before the full observational record is available.
Time-split test	Develop on an earlier period; test on later years with independent instruments.
Geographic transfer	Develop in one region; test in a physically similar but independent region.
Variable holdout	Tune on state variables but test on radiative flux, precipitation, transport or other withheld responses.
Cross-model replication	Implement the same physical change in at least three structurally different models.

Reflectivity Decision Readiness Levels

Readiness level	Meaning
RDRL 1 — Conceptual	Models can represent an idealized forcing, but observational constraints on key processes remain broad.
RDRL 2 — Process constrained	Relevant aerosol/cloud/radiative processes are measured in several regimes and process models reproduce them.
RDRL 3 — Global response constrained	Multiple global models with validated process links agree on broad response and explain disagreement.
RDRL 4 — Strategy comparison	Models can compare quantity, location and timing with quantified regional tradeoffs, monitoring and termination risk.
RDRL 5 — Decision support	For a specified limited decision, uncertainty and early-warning systems meet agreed thresholds; independent panels accept the evidence.

The $1 billion fund should aim to move priority questions from RDRL 1–2 toward RDRL 3–4. It should not presume that RDRL 5 will be achieved in five years, and reaching RDRL 5 for one objective would not authorize deployment or imply readiness for all regions and risks.

Illustrative quantitative targets

Decision metric	Illustrative target
Top-of-atmosphere annual mean net flux	Bias and trend uncertainty small enough to detect intervention-scale changes; exact threshold set by metrology review.
Historical anthropogenic aerosol forcing	Meaningful narrowing of the assessed uncertainty range, with regional and cloud-regime decomposition.
Marine low-cloud radiative response	Observed susceptibility reproduced across multiple regimes and held-out cases.
Southern Ocean cloud radiative bias	Substantial reduction without degrading precipitation, circulation or sea-ice performance.
Ocean heat uptake	Closed heat budgets by basin/depth and improved initialized prediction.
AMOC	Reduced disagreement in mean transports, variability and forcing response; monitoring variables tied to model thresholds.
Reflectivity requirement	Independent models produce an uncertainty distribution for required forcing and explain differences mechanistically.
Regional precipitation	Tradeoffs and uncertainty quantified for major monsoon and vulnerable regions; no single-model claims.

Final five-year assessment

At the end of Year 5, an independent international assessment should answer:

1. Which physical uncertainties were materially reduced?
2. Which model improvements passed held-out and cross-variable tests?
3. Which apparent improvements were rejected because of overfitting or compensation?
4. How much uncertainty remains in required reflected sunlight, location and timing?
5. What do models agree and disagree about for AMOC, precipitation, sea ice and regional harms?
6. Which observations would be required for adaptive monitoring?
7. Which controlled experiments, if any, have sufficient value to justify a future proposal?
8. Which research paths should be stopped?
9. What information is still too uncertain for a responsible decision?
10. How should the open observing and model-evaluation infrastructure continue after the fund ends?

The final report should distinguish scientific readiness from political, ethical and legal authorization. Better models can inform decisions; they cannot make the decision by themselves.

Area 1. Earth Radiation Budget and Planetary-Albedo Benchmark - $100M

Prompt to AI: Describe a five-year, $100 million experimental program that establishes a decision-grade benchmark for Earth's radiation budget and planetary albedo.

Response From AI:

This chapter proposes a metrology-centered program for determining how much solar energy Earth reflects, how much infrared energy it emits, how the net imbalance changes through time, and which clouds, aerosols and surface changes cause the observed variations. Climate models reproduce many broad features of the energy budget, but sunlight-reflection decisions require a more demanding standard: the measured change in reflected shortwave radiation must be distinguishable from instrument drift, orbital sampling, cloud-regime shifts and changes in ocean heat uptake.

The program would not attempt to replace existing satellite missions. It would create a common calibration, validation and reconciliation system around them. The principal product would be a versioned global benchmark that links top-of-atmosphere fluxes to surface radiation, cloud properties, aerosol optical properties and ocean heat-content change, with uncertainty stated by region, season and timescale.

Why this program is decision-critical

A reflectivity strategy is specified in watts per square metre, yet the quantity to be adjusted is not observed by one perfect instrument. Broadband radiometers measure outgoing flux, imagers classify clouds, polarimeters infer particle and droplet properties, surface networks measure downwelling and upwelling radiation, and the ocean integrates the remaining imbalance as heat. Each system has different sampling and calibration errors. A decision-grade benchmark must reconcile them rather than select whichever record best supports a preferred conclusion.

The program should answer two distinct questions. First, what is the present planetary energy imbalance and its trend? Second, what physical changes produced that trend: cloud amount, cloud optical depth, aerosol loading, snow and sea ice, water vapor, surface temperature or circulation? The second question matters because a model can match the net flux for the wrong reason. Such compensating errors would make estimates of required SAI or MCB forcing unreliable.

Core scientific questions

- Can independent satellite and ocean records close the global energy budget within a predeclared uncertainty range?

- How rapidly is planetary albedo changing, and which cloud regimes, aerosol layers and surface regions account for the change?
- What calibration stability is required to detect a deliberate reflectivity change that is small relative to ordinary weather variability?
- Where do three-dimensional cloud effects, angular sampling and spectral assumptions bias broadband flux retrievals?
- Can models reproduce shortwave and longwave fluxes simultaneously without tuning away errors in cloud phase, humidity or surface albedo?
- Which observing variables provide the earliest reliable indication that a reflectivity intervention is producing a different forcing than predicted?

Experimental portfolio

A. Absolute radiometric calibration and cross-calibration

Establish a traceable calibration chain for broadband shortwave and longwave measurements. The work would combine laboratory standards, on-orbit calibration maneuvers, lunar and stellar observations, deep-space views, stable terrestrial targets and overlapping satellite records. Independent calibration teams would quantify instrument drift and produce correction histories rather than allowing each mission to define its own reference. Aircraft and high-altitude balloon radiometers would underfly selected satellite overpasses to test the entire measurement chain under known atmospheric conditions.

B. Spectral and angular closure campaigns

Conduct repeated aircraft campaigns over representative scenes: clean ocean, marine stratocumulus, broken cumulus, desert, snow, sea ice, smoke above cloud and polluted continental boundary layers. Stacked flight legs above and below cloud would measure broadband and spectral flux, cloud geometry, aerosol extinction and surface reflectance. Multi-angle polarimetry and lidar would characterize the anisotropy and vertical structure that convert radiance into flux. The acceptance test would be whether measured scene properties predict the observed flux divergence within stated uncertainty.

C. Surface-to-space energy-budget supersites

Upgrade a small network of marine, continental and high-latitude supersites with redundant shortwave and longwave radiometers, spectral instruments, radar, lidar, microwave radiometers, aerosol and cloud-condensation-nuclei measurements, radiosondes and surface-flux systems. The sites would provide continuous records for satellite validation and model process evaluation. At least one site should be located in a remote ocean regime and one in a high-albedo region where small cloud errors strongly affect the net flux.

D. Ocean heat-content reconciliation

Use Argo, deep-Argo, hydrographic sections, satellite altimetry and ocean state estimation to compare the accumulated ocean heat gain with the top-of-atmosphere imbalance. The goal is not to force exact agreement by adjustment, but to identify where sampling gaps, deep-ocean uncertainty, freshwater effects or calibration offsets prevent closure. Independent teams would perform the reconciliation with different methods and publish the residual budget.

E. Albedo attribution and benchmark dataset

Create a causal attribution system that decomposes reflected-shortwave changes by cloud regime, aerosol type, surface class, latitude and season. The dataset would preserve instrument-level uncertainty and sampling information rather than only providing monthly means. A protected subset of years and regions would be withheld from model developers so that revised models can be tested prospectively.

Required measurements and products

Measurement domain	Required observation	Decision use
Top of atmosphere	Broadband and spectral reflected shortwave, outgoing longwave and net flux	Defines the forcing that must be changed and detects intervention-scale anomalies.
Clouds	Fraction, optical depth, phase, top height, water path, droplet or ice properties	Attributes albedo changes and exposes compensating cloud errors.
Aerosols	Optical depth, absorption, size, composition, vertical profile and humidity growth	Separates aerosol reflection and absorption from cloud and surface effects.
Surface	Ocean state, snow and sea-ice albedo, vegetation, desert reflectance and emissivity	Closes the lower boundary and tracks high-latitude feedbacks.
Ocean heat	Heat content by basin and depth, steric change and air-sea flux	Independent integral check on the net planetary imbalance.
Calibration	Traceability, drift, overlap and uncertainty covariance	Determines whether apparent trends are physical or instrumental.

Five-year execution and budget

The first year would emphasize standards, data audits and observing-system simulations. Years two and three would conduct the principal closure campaigns and supersite upgrades. Year four would repeat the most informative cases under contrasting conditions. Year five would complete independent reconciliation and deliver a benchmark suitable for held-out model tests.

Work package	Five-year allocation	Principal deliverable
Radiometric standards and satellite cross-calibration	$22M	Traceable calibration chain and drift-corrected overlapping records.
Aircraft and balloon closure campaigns	$26M	Approximately 8–12 intensive deployments across key scene types.
Surface supersites and reference radiometers	$18M	Continuous surface-to-space closure datasets in diverse regimes.
Ocean heat-content reconciliation	$12M	Independent global and basin energy-budget assessments.
Albedo attribution, data engineering and benchmark archive	$14M	Versioned dataset, uncertainty model and protected validation cases.
Independent audit and contingency	$8M	Metrology audit, replication and replacement observations.
Total	$100M	

Deliverables and success criteria

- A public, version-controlled Earth energy-budget benchmark with raw, calibrated and derived products linked to provenance.
- A quantified trend in reflected shortwave radiation with regional, seasonal and instrument components of uncertainty.
- A cloud-regime and surface-regime attribution of albedo change that can be reproduced by independent teams.
- Prospective model tests showing whether revised models predict withheld flux and ocean-heat observations.
- Detection thresholds and monitoring requirements for identifying a specified SAI or MCB forcing anomaly.
- A documented residual budget that openly states where closure has not been achieved.

Risks and boundaries

The main program risk is false precision. Combining many datasets can produce a narrow number while retaining shared calibration or retrieval assumptions. Therefore, uncertainty covariance, independent processing and disagreement among methods must be preserved. A second risk is optimizing only the global annual mean. The benchmark must retain regional and seasonal structure because intervention decisions depend on gradients and circulation.

No instrument or institution should be declared the reference by reputation alone. The reference should emerge from traceable standards, overlap, closure and independent audit. The program should publish negative closure results and should not adjust observations merely to make the energy budget agree.

Area 2. Anthropogenic Aerosol Radiative-Forcing Closure - $175M

Prompt to AI: Describe a five-year, $175 million experimental program that measures the direct and cloud-mediated radiative forcing of anthropogenic air pollution.

Response From AI:

Anthropogenic aerosols cool and warm the climate through several pathways. Sulfate, nitrate, organics, black carbon, dust mixtures and industrial particles scatter or absorb sunlight; they also change cloud-droplet number, droplet size, precipitation, cloud water and cloud fraction. The net historical forcing is difficult to infer because emissions, meteorology and clouds change together and because the preindustrial aerosol baseline is poorly observed.

This program would build a chain of experiments from known emissions to aerosol properties, cloud response and measured radiative flux. It would combine regional gradients, policy-driven emissions changes, ship and industrial plumes, laboratory studies, satellite observations and carefully governed small perturbations only where natural variability cannot answer the question. The objective is not one global number obtained by statistical fitting. It is a physically decomposed forcing estimate that models can reproduce across regimes.

Why this program is decision-critical

Historical aerosol cooling affects estimates of climate sensitivity, the warming rate expected as air pollution declines, and the amount of sunlight reflection that might be required to bend the temperature curve. If aerosol cooling has been stronger than assumed, future clean-air improvements could unmask more warming; if it has been weaker, models may be compensating with cloud or ocean errors. For MCB, the same aerosol-cloud physics determines whether added sea-salt particles brighten clouds or trigger adjustments that offset the initial effect.

The program should separate direct forcing from cloud-mediated forcing and rapid adjustments. It should also distinguish source sectors and particle types, because sulfur controls, black-carbon reductions, shipping rules and wildfire smoke do not produce interchangeable perturbations. Each flagship campaign should close the sequence: emissions or perturbation amount, plume transport and aging, aerosol optical properties, activation into cloud droplets, cloud macrophysical response, and shortwave and longwave flux change.

Core scientific questions

- What is the radiative forcing per unit emission for major anthropogenic aerosol precursors and absorbing particles?
- How does aerosol hygroscopicity, mixing state and vertical placement change direct scattering and absorption?
- Under which cloud and meteorological regimes does added aerosol increase droplet number, suppress drizzle, thicken cloud, thin cloud or change cloud fraction?
- How much of observed aerosol-cloud correlation is causal rather than a response to shared meteorology?
- What forcing was present in the preindustrial atmosphere, and how sensitive is the historical forcing estimate to that baseline?
- Can multiple models reproduce the same measured plume-to-flux chain without regime-specific retuning?

Experimental portfolio

A. Regional pollution-gradient closure campaigns

Conduct paired campaigns in regions where polluted continental outflow transitions to cleaner marine air. Aircraft would sample upwind emissions, aging plumes, cloud inflow, in-cloud microphysics and above-cloud radiation while ships and surface stations measure boundary conditions. Repeated Lagrangian sampling would follow the same air mass for hours to days. Contrasting regions should include sulfate-dominated, carbonaceous, mixed industrial and biomass-burning environments.

B. Shipping and industrial-plume natural experiments

Use shipping corridors, fuel-rule changes, port disruptions, power-plant closures and industrial operating schedules as quasi-controlled perturbations. The program would preregister expected plume locations and cloud regimes, acquire before-and-after satellite and aircraft observations, and use adjacent unaffected regions as controls. Ship tracks are valuable but should not be treated as globally representative; the work must identify the meteorological conditions under which visible tracks appear, disappear or produce surrounding cloud adjustments.

C. Absorbing aerosol above cloud

Repeat and extend southeast Atlantic and other smoke-over-cloud campaigns with polarimetry, lidar, in-situ absorption, cloud probes and flux radiometers. The forcing sign depends on aerosol absorption, cloud brightness, layer separation and cloud response. The experiment should measure all four simultaneously and test whether satellite retrievals can distinguish absorption from cloud and surface effects.

D. Cloud activation and drizzle-transition experiments

In clean and moderately polluted marine clouds, perform repeated stacked legs below cloud, at cloud base, in cloud and above cloud. Measure aerosol size and composition, cloud-condensation-nuclei spectra, cloud-base vertical velocity, droplet spectra, liquid water, drizzle and radiative flux. These observations would constrain activation functions and identify when additional particles cease to increase droplet number or begin to alter precipitation and entrainment.

E. Metered perturbation option

Where natural experiments cannot establish causality, the program could consider small, independently governed releases of an environmentally acceptable tracer or sea-salt aerosol into a limited marine cloud field. The release would be designed to measure the plume-to-cloud-to-flux response, not to create regional cooling. It would proceed only after modeling, environmental review, public disclosure and a finding that the same question cannot be answered from existing emissions changes or observations.

F. Preindustrial baseline and laboratory constraints

Combine ice-core and sediment proxies, remote-region measurements, emissions reconstruction, cloud-chamber studies and global aerosol modeling to constrain natural aerosol and nucleation. Laboratory work should determine optical properties, hygroscopic growth, cloud activation and ice-nucleating behavior for representative mixtures rather than idealized pure particles. The global forcing calculation should propagate uncertainty from the baseline instead of hiding it in model tuning.

Measurement chain

Link in causal chain	Minimum measurements	Closure test
Source and emissions	Mass, composition, timing, stack or ship conditions and co-emitted gases	Can the emitted aerosol and precursors be reconstructed independently?
Transport and aging	Wind, turbulence, chemistry, size distribution, mixing state and humidity growth	Does the model predict plume location and transformed particle properties?
Direct radiation	Extinction, scattering, absorption, single-scattering albedo and layer height	Do measured aerosol properties reproduce above- and below-layer flux changes?
Cloud activation	CCN spectrum, updraft distribution, droplet number and effective radius	Can activation be predicted from aerosol and dynamics?

Link in causal chain	Minimum measurements	Closure test
Cloud adjustment	Liquid water, fraction, thickness, entrainment, drizzle and organization	Are persistence and precipitation responses causal and regime dependent?
Net forcing	Shortwave and longwave flux above and below cloud, plus satellite context	Does the complete measured state reproduce the observed radiative perturbation?

Five-year execution and budget

Work package	Five-year allocation	Principal deliverable
Four regional pollution-gradient campaigns	$48M	Multi-regime plume, cloud and flux closure datasets.
Shipping, fuel-rule and industrial natural experiments	$28M	Before/after and matched-control forcing estimates.
Absorbing aerosol above-cloud campaigns	$22M	Constrained forcing sign and satellite retrieval validation.
Cloud activation, drizzle and adjustment experiments	$30M	Parameterizations and susceptibility curves across regimes.
Metered marine perturbation option	$15M	Governed causal test, exercised only if stage gates are met.
Preindustrial baseline, laboratory and synthesis work	$18M	Natural baseline distribution and particle-property database.
Independent replication, data systems and contingency	$14M	Blind model challenges, open archive and replacement operations.
Total	$175M	

Deliverables and success criteria

- A decomposed probability distribution for direct aerosol forcing and aerosol-cloud forcing by region, source class and cloud regime.
- Observed forcing efficiencies linked to emissions and particle properties rather than inferred solely from model residuals.
- Activation and cloud-adjustment relationships that improve held-out predictions in at least three structurally different models.
- A quantified preindustrial aerosol baseline with transparent proxy and model uncertainty.
- A public database of emissions events, plume observations, cloud response and radiative closure.

- Clear evidence showing which uncertainties remain irreducible without larger or longer observations.

Risks and boundaries

Meteorology can create convincing but noncausal aerosol-cloud correlations. Every campaign should therefore use matched controls, Lagrangian sampling, causal designs and prospective model predictions. Another risk is selecting only visually obvious ship tracks or strong pollution episodes, which biases susceptibility upward. Sampling must include cases where added aerosol produces no visible track or where cloud adjustments offset the initial brightening.

Any deliberate perturbation must remain far below a scale intended to affect regional climate and must be optional, not a spending obligation. The program should not equate a narrower global forcing range with complete knowledge of regional health, precipitation or ecosystem effects.

Area 3. Marine Low-Cloud and Cloud-Feedback Experiments - $145M

Prompt to AI: Describe a five-year, $145 million experimental program that improves the representation of marine low clouds, cloud feedback and MCB susceptibility in global climate models.

Response From AI:

Subtropical stratocumulus, trade cumulus and the transition between them exert strong control over planetary albedo. Their response to warming is a major source of uncertainty in climate sensitivity, and their response to additional cloud-condensation nuclei is central to marine cloud brightening. Global models cannot explicitly resolve many of the turbulent eddies, entrainment processes, drizzle cells and mesoscale organizations that determine cloud amount and reflectivity.

The proposed program would combine three flagship field campaigns, persistent ocean observing sites, large-eddy simulations, satellite retrieval evaluation and parameterization teams embedded in multiple global models. The goal is to predict cloud radiative response from measurable environmental conditions, not to derive one universal susceptibility number.

Why this program is decision-critical

MCB effectiveness depends on whether added particles reach cloud base, activate into droplets and increase reflectivity without causing compensating changes in liquid water, entrainment, drizzle or cloud fraction. The same physical processes govern low-cloud feedback under greenhouse warming. A model that brightens clouds by prescribing droplet number but misrepresents boundary-layer dynamics may overstate both MCB cooling and ordinary climate feedback.

The program should therefore characterize a multidimensional response surface: cloud radiative effect as a function of aerosol, cloud-base updraft, inversion strength, free-tropospheric humidity, sea-surface temperature, surface flux, precipitation state and mesoscale organization. Observations must follow cloud systems through time because an initially brightened cloud can later thin, reorganize or dissipate.

Core scientific questions

- What controls the transition among solid stratocumulus, open cells, closed cells, broken cumulus and clear sky?
- How do entrainment and free-tropospheric humidity change liquid water path and cloud fraction as the climate warms?
- How does cloud-base vertical-velocity variability control aerosol activation and droplet number?

- When does drizzle suppression increase cloud lifetime, and when do dynamical adjustments offset the Twomey brightening?
- How rapidly does a localized aerosol perturbation spread, dilute and interact with mesoscale circulation?
- Which variables can a global model use to predict susceptibility without resolving every turbulent eddy?

Experimental portfolio

A. Eastern-boundary stratocumulus campaign

Conduct an intensive campaign in a persistent stratocumulus deck using two research aircraft, a ship, autonomous surface vehicles and coordinated satellite overpasses. Flight stacks would sample sub-cloud aerosol, cloud-base activation, in-cloud droplet and drizzle spectra, cloud-top entrainment and above-cloud flux. Lagrangian sequences would follow the same cloud field from morning through afternoon and, where feasible, across multiple days. The campaign should intentionally sample clean, moderately polluted and high-aerosol conditions.

B. Stratocumulus-to-cumulus transition campaign

Follow air masses moving from cool upwelling regions toward warmer open ocean. The transition reorganizes cloud fraction and may dominate regional albedo response. Aircraft and autonomous platforms would measure surface flux, inversion weakening, free-tropospheric humidity, drizzle, cold pools and mesoscale cellular organization. Models would be required to predict the location and timing of breakup before seeing the final observations.

C. Trade-cumulus campaign

Study shallow cumulus in a region where cloud fields are sparse, rapidly evolving and strongly coupled to circulation. Dense radiosonde arrays, scanning radar, lidar, aircraft and surface flux measurements would constrain mass flux, cloud-core properties, detrainment and organization. This campaign is essential because MCB applied outside solid stratocumulus could have a smaller, more variable or even counterproductive response.

D. Cloud susceptibility and metered-aerosol cases

First exploit natural aerosol gradients, ship plumes and changing wind direction to estimate susceptibility. If a controlled case passes governance gates, release a small metered sea-salt plume from a vessel and sample it with stacked aircraft legs and an adjacent control region. Measure delivery efficiency to cloud base, droplet response, drizzle, cloud water, cloud fraction and reflected flux. The test should distinguish immediate microphysical brightening from later dynamical adjustment.

E. Persistent low-cloud observatories

Maintain at least two ocean or island supersites with cloud radar, lidar, microwave radiometry, broadband flux, aerosol and CCN instruments, precipitation sensors and regular soundings. Autonomous surface vehicles would extend observations along Lagrangian trajectories. These long records would place short campaigns in climatological context and provide independent cases for satellite and model evaluation.

F. LES-to-global-model parameterization strike teams

For each flagship campaign, run ensembles of large-eddy simulations initialized with observed meteorology and aerosol. Teams would identify which processes explain observed cloud evolution, then develop scale-aware parameterizations for turbulence, entrainment, shallow convection, microphysics and precipitation. At least three global models would implement the changes and test them on withheld regimes. A change would not be accepted if it improves low-cloud radiation while degrading precipitation, circulation or energy conservation.

Observation and modeling matrix

Domain	Key variables	Purpose
Aerosol and activation	Size, composition, hygroscopicity, CCN spectrum and cloud-base updraft	Predict how many particles become droplets.
Cloud microphysics	Droplet number, effective radius, spectral width, liquid water and drizzle	Quantify immediate brightening and precipitation response.
Boundary-layer dynamics	Turbulence, inversion strength, entrainment, humidity and surface flux	Explain cloud thickness, breakup and persistence.
Mesoscale organization	Open/closed cells, cold pools, cloud fraction and spatial covariance	Determine whether local brightening changes surrounding cloud.
Radiation	Above- and below-cloud shortwave and longwave flux, spectral reflectance	Close the measured cloud state to radiative effect.
Model hierarchy	LES, regional storm-resolving and global models	Translate process understanding across scales.

Five-year execution and budget

Work package	Five-year allocation	Principal deliverable
Eastern-boundary stratocumulus campaign	$30M	Radiative and process closure in persistent low clouds.
Transition-zone campaign	$24M	Observed and predicted cloud breakup and organization.
Trade-cumulus campaign	$24M	Mass-flux, organization and radiative-response constraints.
Susceptibility and optional metered-aerosol cases	$18M	Causal response curves and delivery-efficiency measurements.
Persistent observatories and autonomous platforms	$18M	Multi-year regime statistics and satellite validation.
LES, parameterization and global-model implementation	$23M	Tested scale-aware physics in multiple models.
Independent replication and contingency	$8M	Held-out evaluations, audit and replacement operations.
Total	$145M	

Deliverables and success criteria

- Cloud-susceptibility distributions conditioned on meteorology and regime, not one global constant.
- Observed closure from aerosol and dynamics through droplet response, cloud adjustment and reflected flux.
- LES cases and forcing files that independent groups can reproduce.
- Global-model improvements that predict withheld low-cloud regimes and preserve other climate metrics.
- A practical map of where MCB is likely to be effective, ineffective or unusually uncertain.
- Monitoring variables that indicate when a brightening operation is departing from the predicted response.

Risks and boundaries

Short campaigns can mistake weather sampling for climate feedback. The program must combine intensive process observations with multi-year statistics and repeated regimes. A second risk is designing experiments around conditions most favorable to MCB. Site selection should include marginal and unfavorable regimes so that the resulting model does not extrapolate optimistic cases globally.

Any aerosol release should be small, temporary and scientifically justified. The program should not test regional cooling, should not develop deployment-scale spray systems, and should report cases in which cloud response is weak, negative or dominated by dynamical adjustment.

Area 4. Mixed-Phase, High-Latitude Cloud and Surface-Albedo Experiments - $100M

Prompt to AI: Describe a five-year, $100 million experimental program for mixed-phase clouds, high-latitude radiation, sea ice, snow and surface-albedo feedbacks.

Response From AI:

High-latitude climate is governed by tightly coupled cloud, radiation, sea-ice, snow, ocean and boundary-layer processes. Mixed-phase clouds can contain supercooled liquid water and ice for long periods, yet models differ in how quickly they glaciate, precipitate and dissipate. These differences contribute to persistent shortwave and longwave radiation biases over the Southern Ocean and Arctic and can be masked by compensating errors in surface albedo or heat transport.

The proposed program would observe the coupled column rather than treating cloud and surface as separate topics. It would combine aircraft and ship campaigns, year-round supersites, ice-tethered and autonomous platforms, laboratory ice-nucleation work, satellite validation and coupled model experiments. The output would improve both ordinary climate projections and estimates of how sunlight reflection interacts with polar amplification and AMOC-related risks.

Why this program is decision-critical

A reflectivity intervention may cool the globe while producing different high-latitude responses depending on season, cloud phase and surface state. Additional reflected sunlight has little direct effect during polar night, but summer cooling can alter melt, snow cover and ocean stratification. Cloud longwave warming can dominate in winter, while cloud shortwave cooling dominates over dark open water in summer. Models must capture these sign changes to evaluate targeted strategies.

Southern Ocean cloud errors also influence the amount of sunlight absorbed by the ocean and therefore affect heat uptake, circulation and climate sensitivity. Correcting the top-of-atmosphere shortwave bias by artificially changing surface albedo or cloud amount can yield the right global temperature for the wrong reason. The program should require simultaneous closure of cloud phase, precipitation, surface radiation and ocean response.

Core scientific questions

- What maintains supercooled liquid water in mixed-phase clouds, and what triggers glaciation or precipitation?

- How do ice-nucleating particles, secondary ice production, turbulence and aerosol removal affect cloud lifetime?
- How do cloud errors interact with sea-ice concentration, thickness, snow cover and melt ponds?
- What is the seasonal net cloud radiative effect over open ocean, marginal ice zones, pack ice and snow-covered land?
- How do Southern Ocean frontal systems and cold-air outbreaks organize cloud and precipitation?
- Can coupled models reproduce radiation, cloud phase and surface energy balance without compensating errors?

Experimental portfolio

A. Southern Ocean frontal and cold-air-outbreak campaign

Operate a research vessel and one or two aircraft across frontal zones and post-frontal cold-air outbreaks. Measure cloud phase, supercooled liquid, ice number and habit, precipitation, aerosol and ice-nucleating particles, surface fluxes and above-cloud radiation. Repeat transects through comparable meteorological structures to separate systematic model errors from case-to-case variability. Satellite underflights would evaluate cloud-phase and radiation retrievals in conditions where passive sensors often struggle.

B. Arctic seasonal transition campaign

Conduct coordinated observations during spring melt onset, summer open water and autumn freeze-up. Aircraft, ice-tethered buoys, autonomous surface vehicles, radar, lidar and frequent soundings would measure boundary-layer structure, cloud phase, radiation, sea-ice state and ocean heat flux. The campaign should follow the same region through transitions rather than sampling one season in isolation.

C. Year-round polar supersites

Strengthen at least two long-duration sites, one Arctic and one Southern Ocean or Antarctic, with redundant cloud radar and lidar, microwave radiometers, broadband and spectral flux instruments, aerosol and ice-nucleating-particle measurements and surface energy-budget systems. The sites would provide statistics on cloud persistence, phase occurrence and radiative sign by season and surface type.

D. Ice-nucleation and microphysics laboratory program

Collect representative sea spray, mineral dust, biological particles and mixed aerosols from field sites and test them in chambers spanning relevant temperature, humidity and cloud-processing histories. The work would evaluate primary and secondary ice production, riming, aggregation and aerosol scavenging.

Parameterizations should be tested against observed cloud evolution, not only laboratory freezing spectra.

E. Coupled cloud-sea-ice-ocean process experiments

Use local and regional coupled models to test how cloud radiative errors alter melt ponds, snow persistence, upper-ocean stratification and ice formation. Observed surface heat budgets would constrain the models. Perturbation experiments would separate the effects of cloud phase, precipitation, ocean mixing and surface albedo. Results would then be implemented in multiple global models and evaluated on withheld seasons and regions.

Measurement matrix

Component	Required measurements	Model constraint
Cloud phase and microphysics	Liquid and ice water, droplet and crystal spectra, habits, temperature and vertical velocity	Persistence, glaciation, precipitation and optical properties.
Aerosol and INP	Number, size, composition, ice-nucleating activity and scavenging	Aerosol control of mixed-phase cloud lifetime.
Radiation	Surface and above-cloud SW/LW flux, spectral albedo and cloud radiative effect	Seasonal energy-budget closure.
Surface state	Sea-ice concentration/thickness, snow depth, melt ponds, ocean temperature and waves	Albedo and conductive/latent heat feedbacks.
Boundary layer	Turbulence, inversions, humidity, surface flux and advection	Cloud maintenance and coupling to surface.
Precipitation	Snowfall, drizzle, radar Doppler velocity and accumulation	Water removal and cloud adjustment.

Five-year execution and budget

Work package	Five-year allocation	Principal deliverable
Southern Ocean campaign	$26M	Frontal and cold-air-outbreak cloud/radiation closure.
Arctic seasonal transition campaign	$22M	Coupled melt, open-water and freeze-up observations.
Year-round polar supersites	$18M	Multi-season cloud-phase and surface-energy records.
Laboratory ice-nucleation and	$10M	Field-linked microphysical

Work package	Five-year allocation	Principal deliverable
microphysics		parameterizations.
Coupled process and global-model experiments	$16M	Reduced Southern Ocean and Arctic radiation biases without compensation.
Independent validation and contingency	$8M	Satellite validation, blind tests and replacement operations.
Total	$100M	

Deliverables and success criteria

- Seasonal cloud-phase and radiative-effect benchmarks over major polar surface regimes.
- Field-tested parameterizations for mixed-phase persistence, ice formation and precipitation.
- A closed surface energy budget linking cloud, snow, sea ice and upper-ocean heat.
- Reduced Southern Ocean shortwave bias and Arctic seasonal radiation bias in held-out tests.
- Quantified sensitivity of polar amplification and sea-ice response to alternative reflectivity patterns.
- Clear attribution of remaining model disagreement to observable processes.

Risks and boundaries

Polar field work is vulnerable to weather, icing, darkness, sea-ice hazards and limited aircraft or ship availability. The program should use redundant platforms, multi-year windows and contingency sites. It should not interpret a single unusual season as a representative climate response.

The program must avoid tuning cloud phase solely to improve top-of-atmosphere radiation. Improvements must also reproduce precipitation, boundary-layer structure, surface flux and sea-ice evolution. It should distinguish research on climate response from any proposal to deliberately alter polar clouds or ice.

Area 5. Deep Convection, Precipitation and Tropical-Organization Experiments - $90M

Prompt to AI: Describe a five-year, $90 million experimental program that improves deep convection, tropical organization, precipitation and hydrological response in global climate models.

Response From AI:

Deep convection transports heat, water, momentum and trace gases through the tropical atmosphere. It creates anvils, cold pools and organized storm systems that affect both reflected sunlight and outgoing infrared radiation. Most global models still represent important convective processes through parameterizations, and even kilometre-scale models require uncertain microphysics and turbulence assumptions.

This program would observe tropical convection as an evolving system from initiation through organization, anvil production and precipitation. It would pair field campaigns with convection-permitting regional and global simulations, water-budget closure and prospective prediction challenges. The objective is to improve estimates of regional rainfall and circulation response to both greenhouse forcing and spatially patterned sunlight reflection.

Why this program is decision-critical

Solar forcing and greenhouse forcing do not affect the hydrological cycle identically. A reflectivity intervention that restores global-mean temperature can still alter evaporation, precipitation, monsoons and the position of the intertropical convergence zone. These changes depend on how convection responds to surface flux, atmospheric stability, humidity and temperature gradients. A model that is adequate for global temperature may therefore be inadequate for rainfall and regional risk.

The program should connect local convective physics to large-scale circulation. It must measure the distribution of storm sizes, lifetimes and organization rather than only average rain. Anvil-cloud radiative effects, convective aggregation and cold-pool interactions can feed back on the environment and affect the sensitivity of the tropical atmosphere to forcing.

Core scientific questions

- What environmental conditions control convective initiation, inhibition and transition to deep storms?
- How do cold pools, boundary-layer moisture and mesoscale circulations organize convection?
- What controls anvil area, ice microphysics, lifetime and net shortwave/longwave effect?

- How does convective aggregation change humidity, radiation and precipitation efficiency?
- Can models close column water and energy budgets during active and suppressed periods?
- How do alternative latitude and seasonal patterns of sunlight reflection shift monsoons and tropical rain belts?

Experimental portfolio

A. Tropical ocean convection campaign

Deploy a research vessel, aircraft, radar network, autonomous surface vehicles and dense radiosondes in a tropical ocean region that samples suppressed, developing and organized convection. Aircraft would measure boundary-layer moisture, cloud microphysics, anvil outflow and radiative flux. Doppler radar and satellite observations would map storm populations and mesoscale organization. The campaign should last long enough to capture multiple large-scale environmental states.

B. Continental monsoon and land-convection campaign

Conduct a complementary campaign in a monsoon or tropical continental region where surface heterogeneity, aerosols and orography affect initiation and precipitation. Measure soil moisture, surface flux, boundary-layer growth, aerosol, updrafts, lightning, precipitation microphysics and cold pools. The comparison with the ocean campaign would test whether parameterizations transfer across fundamentally different environments.

C. Anvil-cloud and upper-tropospheric humidity experiment

Use high-altitude aircraft and satellite underflights to measure anvil ice, particle size and habit, cloud-top temperature, outflow humidity and broadband radiation. Track anvils after convective cores decay to determine lifetime and radiative effect. Models would predict anvil coverage and humidity evolution using only the observed large-scale environment and convective history.

D. Convective organization and cold-pool arrays

Create temporary dense arrays of scanning radar, surface meteorology, disdrometers, microwave radiometers and unmanned systems to resolve cold-pool propagation and storm interactions. The goal is to derive observable organization metrics that can be used to evaluate kilometre-scale models and inform coarser parameterizations.

E. Storm-resolving model intercomparison

Run coordinated regional and global storm-resolving simulations at multiple grid spacings using standardized initial conditions and observations. Require teams to submit predictions before selected periods are released. Diagnose whether differences arise from dynamics, microphysics, turbulence, radiation or surface coupling. Parameterization teams would then test reduced representations in conventional global models.

F. Reflectivity-pattern hydrology experiments

Use validated models to compare uniform solar reduction with latitude-, hemisphere- and season-dependent forcing patterns. Analyze monsoon rainfall, ITCZ position, tropical cyclone environment, drought and extreme precipitation. The program should present probability distributions and model disagreement, not select a single optimal pattern from one model.

Measurement and closure framework

Process	Measurements	Closure objective
Initiation	Profiles of temperature/humidity, inhibition, surface flux, aerosol and boundary-layer convergence	Predict when and where convection begins.
Dynamics and organization	3-D winds, cold pools, updrafts, mesoscale circulations and storm tracking	Predict storm size, lifetime and aggregation.
Microphysics	Liquid/ice spectra, mixed-phase processes, precipitation and latent heating	Predict rain efficiency and anvil production.
Radiation	Anvil optical properties, SW/LW flux and upper-tropospheric humidity	Close convective radiative effect.
Water and energy budgets	Moisture convergence, evaporation, storage, precipitation and radiative/sensible flux	Test conservation and hydrological sensitivity.
Large-scale response	Monsoon indices, ITCZ, circulation and extremes	Translate process improvements to regional risk.

Five-year execution and budget

Work package	Five-year allocation	Principal deliverable
Tropical ocean campaign	$22M	Full life-cycle convection and water-budget dataset.
Continental monsoon campaign	$18M	Land-aerosol-convection and precipitation constraints.
Anvil and upper-troposphere experiment	$12M	Anvil lifetime, humidity and radiative closure.

Work package	Five-year allocation	Principal deliverable
Cold-pool and organization arrays	$10M	Observed organization metrics and process statistics.
Storm-resolving intercomparison and parameterization	$18M	Prospective simulations and transferable physics.
Reflectivity-pattern hydrology ensembles	$6M	Regional precipitation tradeoff distributions.
Independent validation and contingency	$4M	Audit, held-out tests and weather reserve.
Total	$90M	

Deliverables and success criteria

- Observed convective life-cycle cases with closed water, energy and radiative budgets.
- Metrics for initiation, organization, anvil response and precipitation efficiency that models can predict prospectively.
- Improved storm statistics and tropical rainfall in multiple model frameworks without degrading global energy balance.
- Quantified hydrological differences between greenhouse warming and alternative sunlight-reflection patterns.
- Regional monsoon and ITCZ uncertainty distributions suitable for governance and risk evaluation.
- Open simulation outputs, forcing files and diagnostics that allow independent reproduction.

Risks and boundaries

Convection is noisy and field campaigns can miss the intended regime. Observing-system simulations, long deployment windows and multiple regions are necessary. The program should not claim that kilometre-scale resolution automatically produces truth; microphysics, turbulence and surface coupling remain uncertain.

A model improvement should be rejected if it matches rainfall by violating energy or water conservation or by compensating errors in cloud radiation. Reflectivity-pattern experiments should be framed as sensitivity analysis, not as authorization to manipulate monsoons or tropical circulation.

Area 6. AMOC, Ocean Heat Uptake, Freshwater and Sea-Ice Observations - $120M

Prompt to AI: Describe a five-year, $120 million experimental program that reduces uncertainty in AMOC behavior, ocean heat uptake, freshwater pathways and sea-ice coupling.

Response From AI:

The ocean stores most of the excess heat associated with Earth's energy imbalance and redistributes heat and freshwater through currents and mixing. The Atlantic Meridional Overturning Circulation influences North Atlantic climate, sea level, ecosystems and atmospheric circulation. Its future behavior depends on processes that are incompletely observed: deep-water formation, boundary currents, overflows, interior mixing, freshwater storage and export, Greenland meltwater pathways and remote Southern Ocean changes.

This program would strengthen existing observing arrays rather than replace them. It would add targeted measurements in formation, overflow and freshwater-control regions, expand autonomous coverage, conduct process experiments and assimilate the results into multiple ocean and coupled models. The objective is to connect observable warning indicators to physically plausible AMOC responses and to determine how alternative reflectivity patterns affect ocean heat and freshwater forcing.

Why this program is decision-critical

A sunlight-reflection strategy could reduce surface warming quickly while the deep ocean retains heat accumulated over decades. The spatial pattern of cooling may alter air-sea heat loss, sea ice, precipitation and freshwater transport. These effects could strengthen or weaken particular components of overturning even if global-mean temperature is held fixed. Models therefore need a correct ocean state and credible process sensitivity, not merely a global surface-temperature target.

Claims about preventing or causing an AMOC collapse are currently much less certain than global temperature response. A five-year program cannot observe a century-scale tipping event, but it can reduce uncertainty in present transports, freshwater feedbacks, mixing and model drift. It can also establish which measurements would reveal movement toward dangerous states and how quickly an intervention could be adjusted.

Core scientific questions

- What are the mean, variability and uncertainty of overturning, heat and freshwater transports across key Atlantic sections?

- Which regions and processes control deep-water formation and the transformation of water masses?
- How do boundary currents, overflows, eddies and interior mixing transmit anomalies?
- Where does Greenland and Arctic freshwater travel, mix and accumulate?
- How much ocean heat uptake occurs by basin and depth, and which processes set the uptake efficiency?
- Which observable indicators distinguish temporary variability from a persistent weakening or threshold approach?
- How do different spatial and seasonal reflectivity patterns alter air-sea flux, freshwater and overturning in multiple models?

Experimental portfolio

A. Enhanced trans-basin transport arrays

Add targeted moorings, bottom-pressure instruments, gliders and repeat hydrography to existing subtropical and subpolar observing systems. Measurements should resolve boundary currents, interior geostrophic transport, Ekman transport, heat and freshwater flux. Common calibration and processing would allow independent estimates and identify structural differences among array methods.

B. Deep-water formation and overflow process studies

Conduct winter and shoulder-season observations in key convection and overflow regions using ships, gliders, profiling floats, moorings and turbulence instruments. Measure surface heat loss, restratification, mixed-layer depth, water-mass transformation, entrainment and overflow mixing. Tracer-release or deliberate inert-tracer studies could be considered where they provide unique mixing constraints and pass environmental review.

C. Greenland and Arctic freshwater pathways

Deploy autonomous platforms and shelf-to-basin sections to trace meltwater, river input, sea-ice melt and freshwater storage. Use salinity, oxygen isotopes or other tracers, velocity, sea-surface height and hydrographic structure to distinguish sources and pathways. The goal is to determine how much freshwater reaches sensitive formation regions, on what timescale and through which boundary currents or interior routes.

D. Ocean heat-content and abyssal uptake

Expand deep-Argo and repeat full-depth hydrography in regions that dominate deep heat uptake and uncertainty. Combine these observations with altimetry, gravimetry and air-sea flux products.

Independent state-estimation teams would close basin heat budgets and quantify the residual between observed heat storage and modeled transport and mixing.

E. Sea-ice-ocean coupling

Observe ice concentration, thickness, drift, snow, brine rejection, upper-ocean stratification and air-sea flux in the Labrador, Nordic and Arctic gateways. These measurements would constrain how changing ice cover affects winter heat loss and freshwater input. Coupled process models would test whether cloud and surface-radiation biases are contaminating AMOC sensitivity.

F. Initialized prediction and threshold experiments

Assimilate observations into several ocean and coupled models and issue seasonal-to-decadal predictions before later observations are available. Perform standardized freshwater, heat-flux and wind perturbations to diagnose feedbacks and threshold behavior. The purpose is not to assign a single collapse date, but to identify robust mechanisms, model-dependent assumptions and monitoring variables.

Observing system and model links

Domain	Measurements	Decision relevance
Overturning and transport	Velocity, density, pressure, wind and hydrographic sections	Current AMOC strength, variability and heat/freshwater transport.
Formation and mixing	Surface flux, mixed layers, turbulence, tracers and water-mass transformation	Sensitivity of dense-water production and overflow.
Freshwater	Salinity, source tracers, sea-ice export, meltwater and river inputs	Location and timing of freshwater forcing.
Heat uptake	Full-depth temperature, steric change, air-sea flux and state estimates	Ocean commitment and delayed response after surface cooling.
Sea ice	Thickness, concentration, snow, drift and brine processes	Coupling among radiation, freshwater and winter heat loss.
Prediction	Initialized ensembles and standardized perturbations	Early-warning skill and intervention-pattern sensitivity.

Five-year execution and budget

Work package	Five-year allocation	Principal deliverable
Enhanced transport arrays and repeat	$30M	Improved AMOC, heat and

Work package	Five-year allocation	Principal deliverable
sections		freshwater transport estimates.
Formation, overflow and mixing process studies	$24M	Field-constrained transformation and mixing physics.
Greenland and Arctic freshwater pathways	$18M	Source-attributed freshwater budgets and transit times.
Deep-ocean heat uptake and deep-Argo	$18M	Basin/depth heat-budget closure.
Sea-ice-ocean coupling observations	$10M	Winter flux and freshwater coupling constraints.
Assimilation, prediction and threshold experiments	$14M	Prospective forecasts and standardized sensitivity ensembles.
Independent audit and contingency	$6M	Sensor loss reserve, calibration and replication.
Total	$120M	

Deliverables and success criteria

- A reconciled, uncertainty-qualified estimate of AMOC, heat and freshwater transports across key sections.
- Improved parameterizations for overflow, mixing, restratification and water-mass transformation validated against process data.
- A source-to-basin map of Greenland and Arctic freshwater pathways and timescales.
- Closed basin and depth-resolved ocean heat budgets connected to top-of-atmosphere energy imbalance.
- Prospective initialized forecasts and documented skill limits.
- A set of observable AMOC risk indicators tied to model mechanisms rather than simple trend extrapolation.
- Multi-model sensitivity of AMOC and ocean heat uptake to alternative reflectivity patterns.

Risks and boundaries

Five years is short relative to low-frequency ocean variability, and severe conditions can cause instrument loss or sampling gaps. The program should use long-lived arrays, historical records and state estimation to place process campaigns in context. It should publish uncertainty and avoid interpreting one temporary weakening as evidence of an approaching collapse.

The program cannot guarantee that a particular reflectivity strategy would prevent an AMOC tipping event. Its proper outcome is narrower process uncertainty, better monitoring and clearer model disagreement. Any claimed prevention benefit should be presented probabilistically and alongside regional precipitation, sea-ice and termination risks.

Area 7. Rapid-Response Natural-Perturbation Program - $55M

Prompt to AI: Describe a five-year, $55 million rapid-response program that studies natural and policy-driven atmospheric and ocean perturbations before their information is lost.

Response From AI:

Rare events often create stronger and more revealing perturbations than deliberate climate experiments could ethically produce. Volcanic eruptions inject sulfur and ash, wildfire smoke can reach the upper troposphere or stratosphere, abrupt emissions changes alter aerosol and cloud properties, and exceptional ocean or sea-ice events expose mixing and feedback processes. The scientific value is frequently lost because aircraft, instruments, permits, data agreements and teams are not ready when the event occurs.

This program would maintain a standing capability that can move from detection to measurement within days or weeks. It would fund readiness, not only deployments. Preapproved protocols, modular payloads, satellite tasking, balloon kits, autonomous platforms, model forecast teams and reserve agreements with aircraft and ships would allow the program to observe an event from its initial state through chemical, microphysical and climate evolution.

Why this program is decision-critical

Natural perturbations provide scale tests for climate models. Volcanic aerosols test stratospheric transport, particle growth, radiation and temperature response. Wildfire smoke tests absorbing aerosol, cloud interaction and long residence times. Shipping-fuel changes and industrial shutdowns provide quasi-experiments for anthropogenic aerosol forcing. Extreme freshwater, convection or sea-ice events test ocean and cloud parameterizations under conditions outside the calibration mean.

Rapid response also supports prospective prediction. A model team can issue forecasts immediately after an event using only early observations, then compare them with later measurements. This is a stronger validation than explaining the event after all data are known. The program should preserve unsuccessful forecasts and missed deployments because they reveal observing and model limitations.

Event classes and activation criteria

- Volcanic eruptions with material reaching the upper troposphere or stratosphere.
- Pyroconvective wildfire events that produce long-lived smoke layers.

- Large dust, pollution or biomass-burning episodes with unusual vertical placement or cloud interaction.
- Major shipping, fuel-standard, plant-closure or transport disruptions that create a measurable emissions discontinuity.
- Exceptional cloud-regime transitions, marine heat waves, sea-ice loss or freshwater anomalies.
- Instrument or satellite opportunities that provide a unique overlap or calibration event.

Operational architecture

A. Standing event-detection and triage center

Maintain automated monitoring of satellite aerosol, fire, cloud, radiation, ocean and sea-ice products; volcano and emissions alerts; and model forecasts. A rotating scientific duty team would score events against preregistered criteria: uniqueness, expected information value, observability, safety, cost and relevance to program uncertainties. Activation should require documented approval, but the process must be fast enough to capture early evolution.

B. Modular aircraft and balloon payloads

Keep a core set of calibrated instruments in deployable modules: particle size and composition, SO2 and trace gases, ozone, water vapor, radiation, lidar, cloud probes and meteorology. Maintain balloon packages for aerosol, ozone and water vapor profiles and agreements with operators for launches from multiple latitudes. Instrument teams would perform regular readiness exercises so that a nominally available payload is actually flight ready.

C. Satellite tasking and ground networks

Establish advance agreements for rapid satellite observations where tasking is possible and rapid processing of routine sensors. Coordinate surface sun photometers, lidars, radiation sites, radar and air-quality networks along predicted transport paths. Common formats and near-real-time quality control would allow daily assimilation and flight planning.

D. Ocean and cryosphere response kits

Maintain autonomous floats, gliders, surface vehicles, expendable profilers and mooring spares that can be deployed during unusual freshwater, mixing, heat-content or sea-ice events. These assets would complement atmospheric deployments and help determine whether an event changes ocean heat uptake or surface coupling.

E. Prospective model challenge

Within a specified time after activation, independent model teams would archive predictions of plume transport, aerosol evolution, radiative forcing, cloud response or ocean evolution. Later observations would remain withheld until predictions are fixed. The comparison would diagnose which model components failed and whether new observations improve forecasts.

F. Event archive and synthesis

Each event would receive a complete package: initial conditions, measurement plan, raw and calibrated data, model forecasts, uncertainty, operational lessons and a final report. Cross-event synthesis would identify which parameterizations generalize across volcanoes, wildfire smoke, pollution changes and other perturbations.

Readiness levels

Readiness element	Requirement before an event	Activation deliverable
Governance and permits	Preapproved protocols, contacts and decision authority	Documented event-specific authorization.
Platforms	Framework agreements and availability windows	Confirmed aircraft, ship, balloon or autonomous assets.
Instruments	Calibrated modular payloads and spares	Flight-ready configuration and uncertainty record.
Forecasting	Operational trajectories and data-assimilation pipelines	Daily event-location and sampling products.
Data	Common formats, repositories and rapid QC	Open preliminary data and metadata on a fixed schedule.
Model validation	Registered teams and protected observations	Archived prospective predictions and later scorecard.

Five-year budget

Work package	Five-year allocation	Principal deliverable
Standing center, readiness exercises and permits	$10M	Continuous event triage and annual deployment rehearsals.
Aircraft and instrument availability reserve	$16M	Rapid access to calibrated airborne measurements.
Balloon, lidar and surface-network reserve	$7M	Distributed vertical and transport observations.

Work package	Five-year allocation	Principal deliverable
Satellite processing and data assimilation	$6M	Near-real-time event products and targeting.
Ocean and cryosphere response assets	$5M	Autonomous deployment capability.
Prospective model challenges and archive	$6M	Forecast scorecards and permanent event datasets.
Unallocated event contingency	$5M	Additional flights, ship days or instrument replacement.
Total	$55M	

Deliverables and success criteria

- A documented ability to deploy a useful first measurement package within the event-specific scientific window.
- At least several fully observed natural or policy perturbations, while recognizing that event occurrence cannot be guaranteed.
- Prospective model predictions archived before the mature observations are released.
- Measured evolution from source through transport, chemistry or mixing to radiative or climate effect.
- Open event archives that can be reused by modeling groups worldwide.
- A readiness assessment showing which missing platforms or agreements prevented successful response.

Risks and boundaries

The program may experience years with few qualifying events, followed by several simultaneous events. Funds should therefore be held in reserve and not spent merely to meet an annual target. Scientific triage must prevent deployment to spectacular but low-information events.

Rapid response does not waive environmental, aviation, safety or international requirements. The program would observe naturally occurring or independently caused events; it would not create a large perturbation to ensure that the reserve is used. Any small deliberate component would be governed under the relevant experiment program, not activated through an emergency shortcut.

Area 8. Storm-Resolving, Large-Eddy and Process-Model Experiments - $70M

Prompt to AI: Describe a five-year, $70 million modeling-experiment program that converts field observations into improved, scale-aware physics for global climate models.

Response From AI:

Field observations do not automatically improve global models. Measurements are local and intermittent, while global models operate on grid cells far larger than individual clouds, plumes, eddies and turbulent layers. A hierarchy of models is needed to connect scales: laboratory and parcel models, plume and aerosol microphysics, large-eddy simulations, cloud-resolving regional models, global storm-resolving models and conventional Earth-system models.

This program would treat high-resolution simulations as experiments with preregistered hypotheses, controlled perturbations and validation data. Its purpose is not to produce the largest possible simulation. It is to identify which unresolved processes control decision-relevant responses and to create parameterizations that work across grid spacing, climate state and intervention scenario.

Why this program is decision-critical

The amount of reflected sunlight required from SAI or MCB depends on processes that begin at centimetre-to-kilometre scales: particle formation, plume dilution, aerosol activation, cloud turbulence, drizzle, convection and ocean mixing. Global models must represent the aggregate effect. If the parameterization is calibrated only to present-day averages, it may fail when aerosols, temperature gradients or cloud regimes are deliberately changed.

High-resolution models can reveal mechanisms, but they are not automatically correct. They share uncertain microphysics, numerical diffusion and boundary conditions. The program should therefore use observations as constraints, multi-model ensembles to expose structural uncertainty and held-out cases to test transferability. The desired outcome is a documented chain from process evidence to global response.

Core scientific questions

- Which subgrid processes dominate uncertainty in radiation, precipitation, aerosol forcing and ocean heat uptake?
- At what resolution do key structures emerge, and which behaviors remain parameterization dependent?

- Can process models reproduce complete observed budgets rather than one selected variable?
- Which reduced variables and closures transfer across regimes and climate states?
- How should parameterizations change with grid spacing so that model behavior is scale aware?
- Do revised process links alter estimates of required reflectivity, regional impacts or AMOC risk?

Experimental portfolio

A. Benchmark-case library

Convert field campaigns into standardized cases containing initial and boundary conditions, forcing, observations, uncertainty and protected validation variables. Cases would cover low clouds, mixed-phase clouds, deep convection, aerosol plumes, stratospheric aerosol evolution, ocean mixing and overflows. The library should include ordinary and extreme conditions and should retain failed closure cases that expose missing physics.

B. Large-eddy simulation ensembles

Run multi-model LES ensembles for low-cloud, boundary-layer and mixed-phase cases. Vary grid spacing, microphysics, turbulence, radiation and aerosol assumptions in a designed experiment. Compare full distributions and budgets, not only mean cloud fraction. Use the ensemble to identify robust mechanisms and to quantify how uncertainty in observations propagates into inferred parameterizations.

C. Convection-permitting regional experiments

Simulate tropical ocean, monsoon and cold-air-outbreak campaigns at kilometre scale with common forcing. Evaluate storm populations, precipitation extremes, organization, cloud radiation and water budgets. Nested experiments would examine how errors enter from large-scale boundary conditions versus local physics.

D. Global storm-resolving intercomparison

Run coordinated short global simulations at several kilometre grid spacing for selected seasons and perturbations. Because such simulations are expensive, the program would focus on controlled comparisons and process diagnostics rather than century-scale projections. Teams would archive forecasts before evaluation periods and examine whether explicit convection reduces or merely relocates uncertainty.

E. Aerosol and intervention process models

Develop coupled near-field plume, aerosol microphysics and radiative models for SAI, and plume-to-cloud LES for MCB. These models would translate release conditions into particle size, cloud activation and radiative effect. Field measurements from the aerosol and cloud programs would test the first minutes and hours that global models cannot resolve.

F. Parameterization strike teams and global implementation

Small teams combining field scientists, process modelers and global-model developers would convert findings into code. Every proposed change would include a science traceability statement, conservation tests, computational cost, expected regime and failure conditions. At least three structurally different global models would implement the change and test it on data not used in development.

Model-experiment standards

Requirement	Implementation	Reason
Preregistration	Archive hypotheses, configurations, metrics and protected variables before runs	Separates prediction from retrospective explanation.
Budget closure	Track energy, water, aerosol mass and tracers	Prevents improvement through hidden conservation errors.
Multi-model design	Use independent dynamical cores and physics packages	Exposes structural rather than parametric uncertainty.
Resolution sweep	Repeat at several grid spacings	Determines convergence and scale dependence.
Observation uncertainty	Sample instrument and retrieval uncertainty in ensembles	Avoids treating observations as exact truth.
Transfer tests	Evaluate new regions, seasons and warmer states	Tests whether the closure generalizes.

Five-year budget

Work package	Five-year allocation	Principal deliverable
Benchmark-case library and data engineering	$9M	Standardized, uncertainty-aware cases and protected tests.
LES and mixed-phase ensembles	$13M	Process sensitivities and transferable closures.
Regional convection-permitting experiments	$11M	Storm, precipitation and organization benchmarks.

Work package	Five-year allocation	Principal deliverable
Global storm-resolving intercomparison	$15M	Coordinated global high-resolution experiments.
SAI/MCB plume and microphysics models	$8M	Release-to-forcing process links.
Parameterization strike teams and implementation	$10M	Tested code in multiple global models.
Independent replication and computing contingency	$4M	Reproduction, code audit and failed-run reserve.
Total	$70M	

Deliverables and success criteria

- A public benchmark library with observations, uncertainty, forcing files and reference diagnostics.
- Multi-model process ensembles that identify the source of disagreement rather than only its magnitude.
- Scale-aware parameterizations linked to measured mechanisms and explicit failure conditions.
- Prospective improvement on held-out campaigns and regions in multiple global models.
- A tested release-to-radiative-forcing chain for physically realistic SAI and MCB scenarios.
- Quantified changes in reflectivity requirements and regional outcomes caused by the revised physics.

Risks and boundaries

The principal risk is equating resolution with accuracy. High-resolution models can produce realistic-looking clouds while retaining biased microphysics or energy budgets. Visual realism is not an acceptance criterion. Models must predict measured distributions, fluxes and causal responses.

The program should avoid a single-model software monoculture. Shared libraries are useful, but independent implementations are needed to reveal structural uncertainty. Computational spending should be tied to specific information goals and stop if additional resolution does not change the decision-relevant result.

Area 9. Model Evaluation, Data Assimilation and Uncertainty Quantification - $55M

Prompt to AI: Describe a five-year, $55 million program for rigorous climate-model evaluation, data assimilation, causal attribution and uncertainty quantification.

Response From AI:

A climate model can reproduce global temperature while misrepresenting clouds, aerosol forcing, ocean heat uptake or regional precipitation. Conventional evaluation often reports many metrics but does not determine whether the model can answer a specific decision. This program would create a validation and uncertainty system focused on the quantities needed to calculate how much sunlight to reflect, where and when, and how to monitor the response.

The program would combine held-out tests, prospective forecasts, causal attribution, perturbed-physics ensembles, multi-model structural comparisons, data assimilation and decision-focused emulators. It would explicitly penalize compensating errors and would report uncertainty as a distribution traceable to observations, model structure, scenario and internal variability.

Why this program is decision-critical

Agreement among models is not validation when models share code, parameterizations, tuning targets or observational products. Conversely, disagreement does not necessarily mean every model is equally plausible. A decision framework needs evidence-based weighting that rewards successful out-of-sample prediction and process fidelity without collapsing uncertainty prematurely.

Data assimilation can produce the best estimate of the present climate state and can connect sparse observations to complete fields, but it can also inherit model bias. The program should therefore compare multiple assimilation systems and preserve the distinction between direct observation, model-constrained estimate and forecast. Uncertainty should be propagated into intervention quantities rather than added as an afterthought.

Core scientific questions

- Which model errors matter most for required reflectivity and regional climate risk?
- Can models predict new campaigns, natural perturbations and later time periods before seeing the data?
- How much apparent skill results from tuning to the evaluation dataset?

- Where do models match observations through compensating errors?
- How should observations and model skill be used to weight ensembles without discarding plausible structural alternatives?
- How do uncertainty sources propagate into forcing quantity, spatial pattern, timing and monitoring thresholds?

Program components

A. Decision-focused benchmark suite

Define a small, stable suite of benchmarks for top-of-atmosphere radiation, cloud regimes, aerosol forcing, precipitation, ocean heat uptake, AMOC, sea ice and natural perturbations. Each benchmark would include uncertainty, independence, temporal and geographic coverage, and whether it was used for model development. The suite should be versioned slowly so that models cannot improve merely because the target changes.

B. Protected and prospective tests

Hold back selected campaign periods, regions and variables. Require model teams to archive predictions before release. Additional tests would use time splits, geographic transfer and natural events. Scores should include calibration, bias, variability, extremes, causal response and conservation. Poor performance must be published even when it contradicts a favored model.

C. No-compensating-error diagnostics

Develop joint metrics that require simultaneous agreement in related quantities. For example, a model should match reflected shortwave, cloud amount, cloud optical properties and aerosol loading together; match ocean heat content, air-sea flux and transport together; and match precipitation, evaporation and atmospheric moisture convergence together. An apparent success created by offsetting biases would be flagged rather than averaged away.

D. Multi-system data assimilation and state estimation

Run several atmospheric and ocean data-assimilation systems with common observations and document their priors, bias corrections and spread. Compare reconstructed fluxes, aerosol fields, cloud regimes, ocean transports and heat content. Observation impact experiments would identify which new measurements most reduce uncertainty in the state and in short-term prediction.

E. Causal attribution and emergent constraints

Use process-informed statistical methods to attribute observed changes in radiation, temperature, aerosols, clouds and ocean heat. Emergent constraints would be accepted only when the physical mechanism is clear, the relationship is robust across model families and it predicts data not used to discover the relationship. The program should publish failed constraints and sensitivity to model genealogy.

F. Decision uncertainty engine

Build an open system that propagates observation, parameter, model-structure, scenario and internal-variability uncertainty into quantities such as required reflected flux, injection or spray patterns, temperature and precipitation response, AMOC risk and termination consequences. Fast emulators may be used, but they must be trained on transparent ensembles and checked against full models.

Evaluation ladder

Level	Question	Evidence required
1. Integrity	Does the code conserve and reproduce its own result?	Versioned code, exact configurations and budget tests.
2. Process	Does the model reproduce measured local physics?	Laboratory and field closure with uncertainty.
3. Regime	Does it work in multiple regions and seasons?	Held-out geographic and regime tests.
4. Historical cause	Does it explain observed changes for the right reasons?	Joint attribution of radiation, temperature, aerosol and ocean heat.
5. Prospective prediction	Can it predict later observations?	Archived forecasts and independent scorecards.
6. Intervention transfer	Can it translate physical SAI/MCB operations into climate response?	Validated process chain and cross-model replication.
7. Decision readiness	Is uncertainty narrow enough for a specified choice?	Predeclared thresholds accepted by independent review.

Five-year budget

Work package	Five-year allocation	Principal deliverable
Benchmark suite and protected-data governance	$8M	Stable decision-focused benchmarks and holdouts.
Prospective model challenges	$10M	Campaign, event and time-split prediction scorecards.

Work package	Five-year allocation	Principal deliverable
Joint metrics and no-compensating-error diagnostics	$7M	Process-linked evaluation tools.
Multi-system data assimilation and state estimation	$12M	Atmosphere-ocean state ensembles and observation-impact tests.
Causal attribution and model weighting	$8M	Mechanism-based constraints and genealogy-aware weights.
Decision uncertainty engine and emulators	$7M	Open probability distributions for reflectivity decisions.
Independent audit and contingency	$3M	Statistical review, replication and protected-data security.
Total	$55M	

Deliverables and success criteria

- A public benchmark and scoring system that distinguishes development data from independent validation data.
- Prospective scorecards for major field campaigns and natural events.
- Model weights and uncertainty ranges that are transparent, sensitivity tested and not based on consensus alone.
- A present-state estimate that clearly labels observed and model-constrained quantities.
- An open uncertainty engine connecting physical inputs to reflectivity quantity, pattern, timing and risk.
- Evidence that selected model changes improve decision metrics without compensating degradation elsewhere.

Risks and boundaries

Metrics can be gamed if they become the sole target. The benchmark suite should therefore include physically linked quantities, rotating protected tests and expert review. Model ranking should not imply that one model is best for every question.

The program should not compress deep uncertainty into a deceptively precise probability. Where model structure cannot be assigned a defensible probability, results should present scenario ranges and qualitative confidence. Data assimilation should not be described as observation when the estimate depends strongly on model priors.

Area 10. Observing Backbone, Calibration and Autonomous Platforms - $45M

Prompt to AI: Describe a five-year, $45 million shared observing and calibration backbone that supports the climate-model experimental fund.

Response From AI:

The flagship experiments depend on common capabilities: calibrated radiometers, aerosol and cloud instruments, aircraft and ship integration, autonomous ocean platforms, balloon systems, data synchronization, communications, maintenance and permanent archives. Rebuilding these capabilities separately for every campaign would waste money and produce incompatible measurements.

This program would fund a shared backbone that any approved experiment can use. It would not purchase a deployment fleet or duplicate entire government observing systems. Instead, it would close critical gaps, maintain modular payloads, provide traceable calibration and ensure that data from satellites, aircraft, ships, ground sites and autonomous platforms can be combined with known uncertainty.

Why this program is decision-critical

Many apparent scientific disagreements originate in measurement inconsistency: different inlets lose different particles, radiometers drift, cloud probes shatter ice, clocks are misaligned, ship exhaust contaminates aerosol samples and autonomous sensors develop biofouling or pressure offsets. A shared backbone can reduce these errors and allow observations from different campaigns to be compared directly.

The backbone also improves responsiveness. Modular instruments and integration drawings shorten the time between scientific approval and field operations. Common data formats and real-time quick-look systems help teams adapt sampling while the event is occurring. Long-term stewardship prevents valuable measurements from becoming unreadable after the original investigators move on.

Core capabilities

A. Radiometry and optical metrology laboratory

Maintain traceable calibration for broadband and spectral shortwave and longwave radiometers, sun photometers, lidar and polarimetric instruments. The facility would provide pre- and post-campaign calibration, field transfer standards, angular-response characterization and uncertainty budgets. Intercomparisons would identify drift and instrument-to-instrument offsets.

B. Modular airborne instrument suites

Develop standardized racks, power, timing, telemetry and inlets for aerosol, trace-gas, cloud, radiation and meteorological instruments. The suites would be configurable for high-altitude aerosol, marine cloud, mixed-phase cloud or convection missions. Aircraft-specific airworthiness approval would still be required, but reusable mechanical and electrical interfaces would reduce repeated engineering.

C. Ship, buoy and autonomous surface packages

Create robust packages for broadband radiation, meteorology, sea-surface temperature, waves, aerosol and cloud-base observations on research vessels, buoys and unmanned surface vehicles. Designs would address motion correction, salt contamination, exhaust avoidance, biofouling, power and satellite communications. Reference deployments alongside research ships would quantify autonomous-system bias.

D. Balloon and profiling systems

Maintain balloon payloads for aerosol, ozone, water vapor, radiation and meteorology, together with launch kits and data systems. Expand expendable and reusable ocean profilers for temperature, salinity, oxygen, turbulence and optical properties. Standard operating procedures would allow rapid deployment by partner institutions in multiple countries.

E. Calibration field sites and intercomparison exercises

Conduct annual exercises where satellite, aircraft, surface and autonomous instruments observe the same scene. Include clear sky, aerosol layers, low cloud and mixed-phase cloud when feasible. The objective is to quantify cross-platform bias before a flagship campaign relies on small differences between perturbed and control regions.

F. Common data and operations infrastructure

Provide synchronized time standards, mission planning, real-time displays, data transfer, automated quality checks, metadata templates and persistent storage. Each instrument would have a machine-readable calibration and configuration history. The system should support embargo-free operational metadata while allowing a short, fixed period for quality control of raw science data.

Backbone service catalog

Shared service	Users	Acceptance criterion
Radiometric calibration	Radiation-budget, cloud and aerosol campaigns	Traceable uncertainty and pre/post drift record.

Shared service	Users	Acceptance criterion
Airborne payload integration	Aerosol, cloud, convection and rapid-response teams	Approved interfaces, synchronized data and airworthiness documentation.
Autonomous ocean platforms	Low-cloud, AMOC, polar and rapid-response programs	Reference comparison, uptime and calibrated data return.
Balloon/profiling kits	Stratosphere, polar, aerosol and event programs	Deployment-ready package and standardized processing.
Intercomparison exercises	All observing programs	Published bias and uncertainty across platforms.
Data operations	All recipients	Common formats, provenance, quick-look and permanent archive.

Five-year budget

Work package	Five-year allocation	Principal deliverable
Radiometry and optical metrology	$8M	Traceable calibration and transfer standards.
Modular airborne suites and integration	$10M	Reusable payload interfaces and instrument modules.
Ship, buoy and autonomous surface systems	$8M	Validated remote observing packages.
Balloon and ocean profiling systems	$6M	Rapid, distributed vertical measurements.
Annual field intercomparisons	$5M	Cross-platform bias and uncertainty records.
Common data and mission operations	$6M	Synchronized processing, quick-look and archive.
Spares and contingency	$2M	Replacement sensors and urgent calibration.
Total	$45M	

Deliverables and success criteria

- A catalog of shared instruments and platforms with readiness, calibration and availability status.
- Traceable uncertainty and configuration records for every measurement used in a major closure claim.
- Reduced integration time and cost for repeated campaigns.

- Cross-platform intercomparison datasets linking satellite, aircraft, ship, surface and autonomous observations.
- Common metadata and time synchronization sufficient to combine measurements at process timescales.
- Permanent archives that remain usable without proprietary software or undocumented expert knowledge.

Risks and boundaries

A central facility can become a bottleneck or suppress instrument innovation. The backbone should support external instruments through documented interfaces and should competitively select calibration and integration providers. Readiness claims should be tested through exercises rather than accepted from inventory lists.

The program should not purchase expensive platforms without a utilization plan and life-cycle budget. Equipment ownership, maintenance, access and disposition must be specified in advance. Shared infrastructure should serve measurement and validation, not become an undeclared SAI or MCB deployment capability.

Area 11. International Participation, Governance and Open-Data Infrastructure - $25M

Prompt to AI: Describe a five-year, $25 million program for international participation, independent governance, public engagement and open scientific infrastructure.

Response From AI:

Sunlight-reflection research has global implications, while many of the most consequential climate risks occur in regions with limited access to high-performance computing, aircraft campaigns and major observing systems. A research program dominated by a few institutions could produce technically sophisticated results that lack legitimacy, omit regional priorities or cannot be independently checked.

This program would finance substantive scientific leadership from low- and middle-income countries, independent governance and ethics review, public and community engagement, multilingual communication, data stewardship and open tools. Participation would not mean asking representatives to endorse SAI or MCB. It would provide the resources needed to define questions, analyze regional impacts, challenge assumptions and decline unacceptable experiments.

Why this program is decision-critical

Models used to optimize global temperature can underrepresent monsoons, small islands, mountain hydrology, African rainfall, South American river basins or local ocean circulation. Regional scientists often have the observations and contextual knowledge needed to detect such errors but lack stable funding or access to model ensembles. International leadership improves both scientific accuracy and governance.

Open data alone is insufficient if files are too large, undocumented or dependent on expensive software. The program should fund usable access: cloud-hosted subsets, analysis-ready formats, training, computing grants, reproducible notebooks and long-term repositories. Governance bodies must be independent of organizations that receive large research awards.

Program components

A. Regional science hubs and investigator grants

Support competitively selected hubs in Africa, South and Southeast Asia, Latin America, small-island regions and other underrepresented areas. Hubs would lead regional model evaluation, observation analysis and impact questions relevant to monsoons, food systems, water, heat and ocean circulation.

Funding would cover principal investigators, students, data access, computing and participation in field campaigns.

B. Independent scientific and safety review

Fund panels that are institutionally and financially separate from the program-management and experiment teams. Reviews would assess scientific value, environmental risk, ethics, transboundary implications, data rights and whether a controlled perturbation is necessary. Dissenting opinions, conflicts and recusals would be published.

C. Public and affected-community engagement

For field campaigns and any proposed perturbation, provide accessible descriptions, local-language materials, meetings, response-to-comment documents and ongoing incident communication. Engagement should begin while alternatives remain open, not after a design is fixed. Communities and jurisdictions along likely observation or plume paths should have clear contacts and grievance procedures.

D. Open data, software and hardware access

Maintain repositories for raw and calibrated measurements, model code, configurations, analysis workflows, CAD files and reports. Use open licenses, persistent identifiers and machine-readable metadata. Provide cloud-based analysis environments and reduced-size products so that participation does not require downloading petabytes or owning a supercomputer.

E. Training and independent replication

Fund summer schools, fellowships, visiting scientists and replication grants. A meaningful share of validation work should be performed by teams that did not design the experiment or model. Training would cover instrument uncertainty, causal inference, high-performance computing, climate-model evaluation and research governance.

F. International legal and institutional analysis

Support comparative analysis of aviation, marine, environmental, research and data rules that affect campaigns in different jurisdictions. The work would clarify notification, permitting, liability, transboundary consultation and information-sharing requirements. It would not attempt to create a private substitute for intergovernmental decision-making.

Participation and openness requirements

Area	Minimum requirement	Evidence

Area	Minimum requirement	Evidence
Scientific leadership	Regional investigators define and lead funded work	Named leadership, budget authority and authorship.
Access	Data, code and models usable without proprietary barriers	Open licenses, documentation and cloud-access options.
Independence	Review bodies separated from award recipients	Conflict disclosures, recusals and separate funding line.
Engagement	Potentially affected communities consulted before irreversible decisions	Meeting record, response document and grievance channel.
Replication	Important results checked by unaffiliated teams	Independent code, methods and published comparison.
Language and communication	Core materials available in relevant languages and nontechnical form	Public repository and communication audit.

Five-year budget

Work package	Five-year allocation	Principal deliverable
Regional science hubs and investigator grants	$10M	Locally led regional evaluation and impact research.
Independent scientific, safety and ethics review	$4M	Published decisions, dissent and conflict records.
Public and community engagement	$3M	Accessible consultation and response processes.
Open repositories and analysis environments	$4M	Usable data, code and computing access.
Training, fellowships and replication grants	$3M	Expanded independent analytical capacity.
Legal/institutional analysis and contingency	$1M	Jurisdictional guidance and rapid consultation support.
Total	$25M	

Deliverables and success criteria

- Regional research products that materially change model evaluation, experiment design or risk framing.
- Independent panels with published reasoning, dissent and conflict management.
- Open data and code that can be reproduced by teams outside the original consortium.

- Demonstrated access through cloud-hosted tools, documentation and training, not merely file posting.
- Public records showing how comments and regional concerns altered or constrained research plans.
- A durable international network capable of continuing independent assessment after the five-year fund ends.

Risks and boundaries

Token participation is a serious risk. The program should measure who controls research questions, budgets, data and publication, not only the number of countries represented. Another risk is using engagement as public-relations support for a predetermined intervention. The process must allow disagreement and cancellation.

Open access must respect narrow legitimate limits for personal data, aircraft security, export controls and pre-existing proprietary technology, but exceptions should be written, minimal and time limited. The program should not promise that philanthropic governance can authorize global deployment; its authority is limited to research funding and research conduct.

Area 12. Central Reserve and Rapid-Response Contingency - $20M

Prompt to AI: Describe how a five-year, $20 million contingency reserve should be structured and governed within the climate-model experimental fund.

Response From AI:

Atmospheric and ocean field research is exposed to weather, aircraft and ship delays, lost instruments, failed satellite overpasses, calibration problems, geopolitical disruptions and rare scientific opportunities. A contingency reserve is therefore necessary, but an unrestricted pool can become a mechanism for avoiding competition or expanding weak projects. The reserve should be centrally controlled, stage gated and publicly accounted for.

The $20 million reserve would not be allocated evenly across years or programs. It would remain uncommitted until a documented event or failure creates a high-value need. Unused funds would return to the Funding Source or be reauthorized through the same governance process as other program funds. Prepriced replacement flights, ship days, instrument spares, and other operating options already included in an approved campaign budget are baseline operating provisions, not draws on this central reserve. The reserve remains under Funding Source control and is not allocated pro rata to recipients or charged overhead while unused.

Purposes of the reserve

- Replace or repair a critical instrument whose failure would invalidate an approved campaign.
- Purchase limited replacement aircraft, ship or field days after weather or safety cancellations.
- Respond to a rare natural event that falls outside the standing rapid-response allocation.
- Extend observations when an unexpected result creates a time-limited, high-value scientific opportunity.
- Address a calibration or data-integrity problem discovered after field operations.
- Support transition, data rescue or closeout if a recipient or Managing Organization cannot complete its obligations.

What the reserve should not fund

The reserve should not cover ordinary cost overruns, optimistic budgeting, late deliverables, unapproved scope growth or a desire to spend the full $1 billion. It should not be used to bypass competitive procurement, rescue a project that has failed its scientific gate, or finance an atmospheric perturbation that lacks independent approval. A recipient should bear costs caused by negligence or failure to meet contractual requirements where the agreement permits.

Reserve categories

A. Platform and weather replacement

Provide options for replacement sorties, ship days, balloon launches or temporary relocation when the intended meteorological regime is missed. The original contract should price these options where possible. Reserve approval would require evidence that the replacement still answers the registered question and that the campaign has not already collected sufficient information.

B. Instrument and calibration failure

Fund replacement detectors, pumps, inlets, electronics, calibration work or independent reanalysis when a critical measurement fails. The program should first determine whether another instrument or existing dataset can answer the question more cheaply. Funds may also support a forensic review when data quality is uncertain.

C. Rare scientific opportunity

Support additional observations of an unexpected volcano, wildfire, cloud transition, freshwater event or satellite overlap when the information value is unusually high. The request should include a rapid science case, cost, readiness, safety review and data-release plan. The reserve is not a substitute for the standing rapid-response program, but it can expand an activated response.

D. Data rescue and reproducibility

Protect irreplaceable data if a storage system, contractor or institution fails. Eligible actions include copying raw data, reconstructing metadata, recovering calibration records, converting proprietary formats and completing an open archive. This category can also fund independent replication when a major result is disputed and a timely test would resolve the issue.

E. Program transition and closeout

If a major recipient or the Managing Organization is replaced, the reserve can finance orderly transfer of equipment, records, contracts, staff knowledge and data. It should not pay termination penalties caused by avoidable misconduct unless necessary to protect the research assets, after which the program should pursue recovery where appropriate.

Approval process

1. The Project Manager submits a written reserve request identifying the triggering event, alternatives, amount, schedule and decision value.
2. Scientific staff confirm that the request preserves or materially increases information value.
3. Finance and procurement staff verify price reasonableness, contract authority and whether another party is responsible for the cost.
4. Safety, environmental and governance review is repeated if the change affects operations or perturbation conditions.
5. The Agreement Oversight Board approves material requests and the Funding Source approves requests above the reserved threshold.
6. The program publishes the request, decision, amount and later outcome, subject only to narrow temporary safety restrictions.

Illustrative allocation ceilings

Reserve category	Illustrative ceiling	Control
Platform/weather replacement	$6M	Prepriced options and scientific reauthorization.
Instrument/calibration failure	$4M	Technical failure report and responsibility review.
Rare scientific opportunity	$5M	Rapid value-of-information and readiness decision.
Data rescue/independent replication	$3M	Open-archive or dispute-resolution plan.
Program transition/closeout	$2M	Oversight Board and Funding Source approval.
Total	$20M	Ceilings may be rebalanced by formal board action.

Payment and accountability

Reserve actions should use amendments or task-order options with measurable deliverables. Payment should remain milestone based, and the final holdback should depend on accepted data, documentation and open-access obligations. The public financial dashboard should show the original event, approved amount, recipient, deliverable, payments and whether the reserve action succeeded.

Deliverables and success criteria

- Critical campaigns are not invalidated by foreseeable platform or instrument failures when a cost-effective remedy exists.
- Reserve decisions are fast enough for time-sensitive opportunities but retain documented scientific and financial review.
- No reserve expenditure is used to conceal routine overruns or failed stage gates.
- Every activation produces a defined scientific or data-protection deliverable and a public after-action report.
- Unused balances remain identifiable and are returned or explicitly reauthorized.
- Annual independent audit confirms that reserve spending followed thresholds and approval rules.

Risks and boundaries

The central risk is moral hazard: teams may underbudget or take avoidable risks because they expect the reserve to rescue them. Contracts should distinguish unavoidable events from recipient responsibility and should require reasonable insurance, spares and contingency planning before reserve access.

A second risk is urgency being used to bypass governance. No reserve request should relax scientific, environmental, safety, open-data or conflict-of-interest requirements. If a rare opportunity cannot be pursued responsibly within the available time, the program should decline it rather than create a precedent for unreviewed experimentation.

Policymaker Questions and Program Boundaries

This proposal supports a specific part of climate policy: improving the evidence needed to evaluate sunlight reflection. It is not, by itself, a complete climate strategy.

What this GCM program is intended to answer

- What is Earth's present energy imbalance, and which cloud, aerosol, surface, and ocean processes explain it?
- Can climate models reproduce those processes for the right reasons and predict protected observations and natural perturbations prospectively?
- If a physically realistic SAI or MCB forcing were considered, how much radiative forcing would be needed, where and when, and with what uncertainty?
- What would alternative strategies do to regional temperature, precipitation, sea ice, ocean heat uptake, AMOC, monitoring requirements, and termination risk?

What this program does not decide

- How rapidly emissions reduction, clean-energy deployment, adaptation, and carbon removal should proceed.
- Whether SAI or MCB should ever be deployed, who would have authority to decide, or what international liability and compensation rules should apply.
- Which global climate-policy portfolio minimizes total human and ecological harm across mitigation, adaptation, carbon removal, SAI, and MCB.
- Whether a model result that is scientifically decision-ready is politically, ethically, legally, or institutionally acceptable.

Portfolio context

- SAI research asks whether candidate stratospheric pathways can be predicted, measured, controlled, and operated safely enough to justify further research.
- MCB research asks whether sea-salt delivery can brighten real clouds predictably across relevant regimes and whether practical equipment can do so reliably.
- GCM research integrates those measured forcing pathways to estimate how much forcing might be needed, where and when, and what regional and coupled-climate consequences follow.

The appropriate policy comparison is not sunlight-reflection risk versus zero risk. It is the risk of any intervention portfolio versus the evolving no-intervention or mitigation-only counterfactual, with regional consequences and uncertainty stated explicitly.

Termination Shock

Prompt to AI: If SAI or MCB were used for years or centuries and then reduced or stopped while greenhouse-gas concentrations remained high, what should the climate-model program calculate? Include illustrative 10-year and 200-year cases, differences between SAI and MCB, regional risks, gradual phaseout, monitoring, and research requirements.

Response From AI:

There is no single permanent "new warming rate" after sunlight reflection stops. Termination shock is a transient response: the imposed cooling forcing weakens, the underlying greenhouse forcing remains, and the climate moves toward the warmer state associated with the accumulated greenhouse gases. The initial warming can be much faster than the pre-termination warming rate, especially when a long-running intervention has been masking several degrees of warming.

SAI and MCB determine how the forcing is produced and how quickly it disappears; the GCM program must calculate what that forcing change does to the coupled atmosphere, ocean, cryosphere, hydrological cycle, and ecosystems.

Illustrative Scenarios

The calculations below are deliberately simplified arithmetic examples used to show how termination risk scales with masked warming and intervention duration. They are not forecasts. The program should replace them with multi-model ensembles spanning greenhouse-gas pathways, ocean states, intervention patterns, forcing-decay times, and abrupt versus gradual phaseout.

Let's examine two scenarios. In Scenario #1, we reflect sunlight for 10 years to offset the 0.3°C/decade global warming rate to 0.0°C/decade, and then stop reflecting. Alternatively, in Scenario #2 we reflect for 200 years and then stop when the concentration of CO_2 in atmosphere is 1200 ppm, instead of the current 430 ppm. In both cases, we assume Earth Climate Sensitivity (ECS) is 4°C and global warming is 0.0°C/decade while reflecting.

Scenario #1 – Stop Reflecting after 10-years

Without SAI, temperature would have risen approximately: $0.3^{\circ}\text{C/decade} \times 1\ \text{decade} = 0.3^{\circ}\text{C}$

SAI has therefore created approximately a **0.3°C temperature gap** between the geoengineered world and the world without SAI. After termination, two effects occur simultaneously:

- The normal underlying warming continues at approximately 0.3°C per decade.
- Earth begins recovering the approximately 0.3°C of warming that SAI had suppressed.

Climate-model experiments indicate that approximately 38%–61%—roughly half—of the equilibrium temperature response to an abrupt forcing change occurs during the first decade. A simple estimate of warming for the first decade after termination is therefore: $0.30 + 0.5(0.30) \approx 0.45°\text{C/decade}$

After this initial one decade catch-up, the warming rate would move back toward approximately 0.3°C per decade, assuming the underlying greenhouse forcing continued increasing at the same rate. This would be noticeable termination warming, but it would be small compared with terminating a centuries-long program.

Scenario #2 – Stop Reflecting after 200-years

Using ECS = 4°C and a preindustrial CO_2 concentration in atmosphere of 280 ppm:

$$T_{\text{eq}}(C) = 4\log_2\left(\frac{C}{280}\right)$$

At 1,200 ppm:

$$T_{\text{eq}}(1200) = 4\log_2\left(\frac{1200}{280}\right) \approx 8.4°\text{C}$$

That is the approximate equilibrium warming above preindustrial conditions associated with 1,200 ppm under the assumed ECS. If the program had held global temperature near today's level—approximately 1.4°C above preindustrial—SAI would be masking approximately 7.0°C (8.4 – 1.4). A slightly narrower calculation that counts only the additional CO_2 increase from 430 to 1,200 ppm gives:

$$4\log_2\left(\frac{1200}{430}\right) \approx 5.9°\text{C}$$

The difference is whether one includes the warming already committed but not yet realized at 430 ppm. Since your program is assumed to hold temperature completely flat for 200 years, the more appropriate masked amount is probably approximately 7°C. If roughly half emerged during the first decade, the average temperature of the planet would increase roughly 3.5°C (50% × 7°C). Continuing greenhouse-gas growth might add roughly another 0.2–0.3°C per decade, producing approximately 3–4°C of global warming during the first decade after abrupt termination. An approximate trajectory might be:

Period after termination	Approximate warming
First decade	3 to 4°C
Second decade	1 to 2°C
Following decades	Progressively slower
Long-run rate after catch-up	Returns toward the underlying CO_2-driven rate

SAI and MCB do not terminate on the same timescale

For sulfate SAI, aerosol already in the stratosphere continues to exert a cooling influence after injection stops and decays on roughly a year-scale. The radiative forcing therefore ramps down over months to several years rather than vanishing in a day. For MCB, sea-salt aerosol and cloud perturbations persist for days, so the deliberate forcing can disappear much more rapidly after source shutdown or a change in suitable cloud conditions.

That difference affects operational warning time, but it does not change the basic climate problem: if either system has been masking large greenhouse forcing, removal exposes the climate to rapid catch-up warming. The GCM program should therefore model SAI and MCB separately rather than applying one generic forcing-decay curve.

What the model ensemble must calculate

Termination experiments should be a standard part of every reflectivity strategy evaluated by the program. At minimum, the ensemble should vary:

- intervention duration, from a few years to many decades or centuries;
- the amount of warming or radiative forcing being masked at termination;
- greenhouse-gas concentration and emissions trajectories before and after termination;
- SAI aerosol lifetime and injection pattern, or MCB cloud-regime and forcing persistence;
- abrupt shutdown, partial outage, regional outage, and controlled ramp-down schedules;
- global and regional temperature rate, extremes, precipitation, monsoons, drought, sea ice, snow, and storm statistics;
- ocean heat uptake, upper-ocean stratification, AMOC and other circulation indicators; and
- ecological and agricultural exposure to unusually rapid rates of change, not only final temperature.

Gradual phaseout is a control problem, not a fixed schedule

A long-running intervention should not be evaluated without an exit path. The simplest arithmetic illustrates the scale: spreading 1°C of rebound warming over 50 years corresponds to about 0.2°C per decade of additional warming from phaseout. A program masking 4°C would require a much longer or more complicated transition if the same rate limit were desired. This arithmetic is useful for intuition, but the actual safe ramp depends on greenhouse-gas decline, carbon removal, ocean state, regional impacts, and the forcing technology.

The model program should therefore optimize phaseout jointly with emissions cuts and carbon removal. If CO_2 concentrations are falling, required reflected sunlight can decline while limiting rebound. If CO_2

remains high or continues rising, an orderly exit becomes progressively harder. Termination planning is thus inseparable from the long-term greenhouse-gas pathway.

Regional and tipping-element risks matter as much as the global mean

A globally averaged warming rate can hide the most important consequences. Termination experiments should track regional temperature gradients, precipitation and monsoon shifts, sea-ice loss, snow and glacier response, ocean heat uptake, AMOC and freshwater pathways, and ecosystems sensitive to the rate of climate change. A strategy that produces an acceptable global-mean phaseout could still create an unacceptable regional transition.

The program should also distinguish threshold risk from rate risk. Some systems may be harmed because a threshold is crossed; others may be harmed because temperature or precipitation changes too quickly for ecosystems, infrastructure, or agriculture to adapt. Both need explicit metrics.

Monitoring and adaptive control

A credible termination plan requires measurements that reveal loss of forcing and climate response early enough to act. The observing system should identify the fastest reliable indicators and the uncertainty in each:

- top-of-atmosphere reflected shortwave and net radiation;
- SAI aerosol or MCB cloud-property changes attributable to the intervention;
- global and regional temperature gradients and upper-ocean heat content;
- precipitation, monsoon, circulation and extreme-event indicators;
- sea-ice and snow response; and
- AMOC, ozone, air-quality and ecological warning signals where relevant.

The model should translate each observed deviation into a decision: continue, compensate, redistribute forcing, begin a controlled ramp-down, or stop a particular pathway. Monitoring without a predefined control response is not sufficient.

Failure scenarios must be modeled explicitly

Termination is not limited to a deliberate political decision to stop. The program should stress-test partial and complete failures: grounded SAI aircraft, loss of sulfur supply or operating bases, MCB vessel or generator failures, unsuitable cloud regimes, cyber or communications outages, instrument gaps, satellite loss, international participation changes, sanctions or conflict, and simultaneous natural perturbations such as major volcanic eruptions or El Niño.

Each scenario should report how much forcing is lost, how quickly, what redundancy remains, whether other regions or technologies can compensate, and what climate response occurs before repair or replacement. This connects the GCM work directly to the reliability requirements in the SAI and MCB equipment programs without duplicating their engineering designs.

Validation standard

Climate-scale termination cannot be experimentally reproduced for validation. The program must therefore build confidence from multiple lines of evidence: conservation and process tests, observed responses to volcanic and other natural radiative perturbations, historical aerosol changes, ocean heat uptake, prospective prediction, and coordinated model intercomparisons. Termination conclusions should be rejected if they depend on one model family or on tuning only global-mean temperature.

A key benchmark is Jones (2013), "The Impact of Abrupt Suspension of Solar Radiation Management—Termination Effect—in GeoMIP Experiment G2" (https://doi.org/10.1002/jgrd.50762). Eleven models showed rapid warming, increased precipitation, and declining sea ice after abrupt suspension. The proposed program should update this logic with more realistic SAI and MCB forcing, improved process constraints, modern models, regional diagnostics, and explicit phase-down scenarios.

Portfolio division of responsibility

The GCM initiative owns the portfolio-wide climate-response analysis: warming rates, regional impacts, ocean and cryosphere response, tipping and rate risks, greenhouse-gas pathways, monitoring thresholds, and abrupt versus gradual phaseout. The SAI Experimental Research proposal retains SAI-specific aerosol lifetime, forcing decay, release-control and continuity measurements. The MCB Experimental Research proposal retains MCB-specific aerosol/cloud decay, weather dependence, source control and continuity measurements.

Central conclusion

Termination risk scales primarily with how much warming is being masked and how abruptly the cooling forcing is lost. A short research intervention and a two-century climate-control program are therefore not versions of the same risk. A decision-ready reflectivity system must include the endgame from the beginning: a quantified failure envelope, redundant monitoring, technology-specific forcing-decay models, greenhouse-gas and carbon-removal pathways, and a credible gradual phaseout. If a strategy has no plausible exit path, that is itself a decision-relevant reason not to pursue it.

Campaign Development and Execution Phases

Prompt to AI: For the proposed GCM experiments, what phases should occur first, and how much might representative campaigns cost?

Response From AI:

A field experiment should be treated as a complete scientific campaign, not merely as a flight, ship cruise or instrument deployment. Each campaign should begin with a decision question and end with an independent test of whether the new model prediction improved. The following costs are illustrative planning estimates in 2026 dollars.

A. Anthropogenic aerosol radiative-forcing closure campaign - approximately $65M

Assumed campaign: six to eight weeks in a polluted-to-clean marine transition region, with two research aircraft, one research vessel or autonomous surface fleet, satellite coordination, ground sites, repeated plume and cloud transects, and model analysis extending through one year.

Phase	Principal work	Estimated cost
1. Decision question and science design	Define causal chain, required uncertainty, sites, meteorological regimes and stop criteria.	$2.5M
2. Existing-data audit and target selection	Analyze emissions, satellites, reanalysis and previous campaigns; choose locations and dates.	$2.0M
3. Governance, environmental and regulatory review	Permits, consultation, environmental review and any perturbation authorization.	$2.0M
4. Instrument development and calibration	Aerosol, CCN, cloud microphysics, radiation, trace gases, lidar and radar.	$9.0M
5. Aircraft, vessel and site integration	Payload installation, airworthiness, ship systems, synchronization and rehearsals.	$5.5M
6. Baseline observations	Clean and polluted background, vertical profiles, surface flux and cloud-regime mapping.	$5.0M
7. Intensive field operations	Aircraft sorties, vessel operations, cloud/plume transects and satellite underflights.	$18.0M
8. Natural or metered perturbation component	Observe policy/emissions change or conduct a small, governed perturbation if justified.	$4.0M
9. Laboratory and sample analysis	Composition, hygroscopicity, ice nucleation, morphology and optical properties.	$4.0M

Phase	Principal work	Estimated cost
10. Data assimilation and radiative closure	Reconstruct emissions-to-aerosol-to-cloud-to-flux causal chain.	$5.0M
11. Model challenge and parameterization testing	LES/CRM and global-model predictions using withheld observations.	$4.5M
12. Independent audit, archive and publication	Calibration audit, uncertainty review, open repository and final report.	$2.0M
13. Contingency	Weather, missed regimes, instrument failures and replacement operations.	$1.5M
Total		$65.0M

The core acceptance test should be a closed, quantified estimate of aerosol effective radiative forcing for the observed cases, with uncertainty decomposed among emissions, aerosol properties, cloud response, meteorology and radiative transfer.

B. Marine low-cloud feedback and radiative-closure campaign - approximately $55M

Assumed campaign: four to six weeks in a subtropical low-cloud region, approximately 40–60 research-aircraft sorties, a vessel or autonomous surface platforms, radar and lidar, satellite coordination and extensive large-eddy simulation.

Phase	Principal work	Estimated cost
1. Science and decision design	Specify cloud-feedback and MCB-susceptibility questions, regimes and validation metrics.	$2.0M
2. Site and season selection	Use satellite climatology, forecasts and model disagreement to select targets.	$1.5M
3. Instrument and platform preparation	Cloud probes, aerosol/CCN, turbulence, radiation, radar, lidar and surface flux systems.	$8.0M
4. Integration and rehearsal	Aircraft certification, vessel installation, communications and coordinated test flights.	$4.0M
5. Baseline and synoptic mapping	Characterize inversion, aerosol, cloud thickness, drizzle and mesoscale organization.	$4.0M
6. Intensive aircraft and surface operations	Stacked legs below, within and above clouds; repeated Lagrangian sampling.	$15.0M
7. Cloud-susceptibility observations	Natural gradients, ship tracks and any approved metered aerosol cases.	$3.0M

Phase	Principal work	Estimated cost
8. LES/CRM simulations	Case simulations, ensemble perturbations and process attribution.	$5.0M
9. Parameterization strike teams	Develop and implement improved cloud, turbulence and precipitation physics.	$4.5M
10. Global-model testing	Run revised physics in multiple models and compare with withheld regimes.	$3.5M
11. Independent closure and archive	Radiative-budget audit, reproducibility package and open data.	$2.0M
12. Contingency	Weather, aircraft availability and failed instruments.	$2.5M
Total		$55.0M

The expensive elements are not sea salt or aircraft fuel alone. They are obtaining simultaneous measurements of aerosol, cloud microphysics, turbulence, precipitation and radiation; sampling the same evolving air mass repeatedly; and then testing whether models predict the observed transition.

C. AMOC and ocean-heat-uptake campaign - approximately $60M

Assumed campaign: a five-year regional program strengthening an existing trans-basin or high-latitude array, adding autonomous platforms, conducting two major research cruises and assimilating the data into multiple ocean and coupled models.

Phase	Principal work	Estimated cost
1. Observing-system design	Identify AMOC decision metrics, observing gaps, array design and forecast experiments.	$2.5M
2. Mooring and sensor procurement	Current meters, temperature/salinity, pressure, acoustics and telemetry.	$11.0M
3. Autonomous platforms	Gliders, Argo/deep-Argo floats, surface vehicles and profiling systems.	$8.0M
4. Deployment cruise	Install moorings, deploy floats/gliders and conduct hydrographic/tracer survey.	$7.5M
5. Continuous operations and maintenance	Telemetry, calibration, recoveries, redeployments and quality control.	$6.0M
6. Process and mixing experiment	Tracer, turbulence, overflow and convection measurements.	$6.5M
7. Freshwater and Greenland pathway observations	River, meltwater, shelf and basin measurements and source attribution.	$4.0M
8. Recovery and repeat cruise	Recover assets, repeat sections and close mass/heat/freshwater budgets.	$5.5M

Phase	Principal work	Estimated cost
9. Data assimilation and state estimation	Multi-model reconstruction of overturning, heat and freshwater transports.	$4.0M
10. Prediction challenge	Prospective seasonal-to-decadal forecasts and coupled-model sensitivity tests.	$2.5M
11. Independent audit and open archive	Sensor audit, uncertainty, documentation and final assessment.	$1.5M
12. Contingency	Lost instruments, ship delays and severe weather.	$1.0M
Total		$60.0M

Campaign-Level Gates C0-C7

Gate	Required finding
C0 - Relevance	The uncertainty materially affects a reflectivity or climate-risk decision.
C1 - Existing evidence	The question cannot be answered adequately from existing data and published analysis.
C2 - Observability	Instruments and sampling can achieve the required accuracy and distinguish competing hypotheses.
C3 - Readiness	Platforms, calibration, permits, logistics, safety and data systems are ready.
C4 - Operations	Forecast conditions are suitable and the campaign can capture the intended regime.
C5 - Data quality	Measurements pass calibration, completeness and uncertainty requirements.
C6 - Model test	Predictions are compared with held-out data and improvements are demonstrated without degrading other metrics.
C7 - Decision impact	The result changes an uncertainty range, model weighting, parameterization or future research decision.

A failed gate should lead to redesign, postponement or cancellation. The program should not treat spending the approved amount as a measure of success.

Purchase Orders

Prompt to AI: For representative climate-model experiments, what purchase orders, subawards or task orders might be issued?

Response From AI:

The term "purchase order" is used broadly below. Research universities would usually receive subawards or sponsored-research agreements. Government aircraft and ships may operate through reimbursable agreements. Commercial engineering, instruments, logistics and computing would use procurement contracts or task orders.

The following structures show how a project manager could convert a scientific campaign into measurable deliverables. Costs are illustrative and should be competed or independently evaluated.

A. Aerosol radiative-forcing closure campaign - $65M

PO	Procurement package	Deliverable	Estimate
A1-01	Integrated science-design subaward	Science traceability matrix, hypotheses, required uncertainty, sampling plan and decision rules.	$1.2M
A1-02	Observing-system simulation contract	Synthetic experiments showing whether proposed sampling distinguishes aerosol direct and cloud-mediated forcing.	$0.8M
A1-03	Independent design review	Published review and proceed/redesign decision.	$0.5M
A2-01	Emissions and satellite data audit	Regional emissions inventory, plume climatology, satellite retrieval assessment and target atlas.	$1.0M
A2-02	Meteorological regime analysis	Probability of target cloud and transport conditions by site and season.	$0.5M
A2-03	Prior-campaign synthesis	Harmonized dataset and gap analysis from relevant field campaigns.	$0.5M
A3-01	Environmental and regulatory services	Permits, environmental review, airspace and maritime plan.	$0.8M
A3-02	Public and jurisdictional engagement	Meetings, materials, comments and response record.	$0.7M
A3-03	Independent governance panel	Conflict disclosures and authorization decision for any controlled perturbation.	$0.5M
A4-01	Aerosol composition and size package	Calibrated measurements from nanometres through coarse particles, with composition and mixing state.	$2.3M

PO	Procurement package	Deliverable	Estimate
A4-02	CCN and hygroscopicity package	Cloud-condensation-nuclei spectra and water-uptake characterization.	$1.2M
A4-03	Absorption and optical-properties package	Scattering, absorption, extinction and single-scattering albedo.	$1.2M
A4-04	Cloud microphysics package	Droplet number, size, liquid water, drizzle and ice measurements.	$1.4M
A4-05	Radiometry and remote-sensing package	Broadband/spectral flux, lidar and radar measurements.	$1.7M
A4-06	Calibration and intercomparison services	Traceable calibrations and side-by-side instrument comparisons.	$1.2M
A5-01	Research-aircraft block 1	Primary in-situ aircraft, crew and approximately 25–35 sorties.	$6.5M
A5-02	Research-aircraft block 2	Remote sensing and vertical-context aircraft, approximately 20–25 sorties.	$4.5M
A5-03	Vessel or autonomous surface operations	Surface flux, aerosol source, cloud radar/lidar and trajectory support.	$4.0M
A5-04	Ground supersite and mobile network	Continuous aerosol, cloud and radiation observations.	$1.5M
A5-05	Satellite coordination and rapid products	Overpass planning and campaign-quality retrievals.	$1.0M
A6-01	Baseline flight and surface series	Clean-background and polluted-background characterization.	$3.0M
A6-02	Trajectory and forecast center	Daily plume, cloud-regime and flight-target products.	$1.0M
A6-03	Formation and sampling rehearsals	Coordinated aircraft/vessel test and source-reconstruction exercise.	$1.0M
A7-01	Intensive operations task order	Coordinated sorties and surface operations meeting the sampling matrix.	$10.0M
A7-02	Plume and Lagrangian follow-up options	Repeat sampling of the same air mass and cloud system.	$3.0M
A7-03	Field logistics and maintenance	Deployment, hangar, shipping, spares, lodging and communications.	$3.0M
A7-04	Weather and replacement-flight options	Reserved platform availability for missed cases.	$2.0M
A8-01	Natural-policy experiment analysis	Quantify radiative response to a documented emissions change.	$1.5M
A8-02	Metered perturbation option	Small approved release and synchronized measurement package, only if gated.	$2.0M

PO	Procurement package	Deliverable	Estimate
A8-03	Independent source verification	Mass, timing, location and composition audit.	$0.5M
A9-01	Filter/sample laboratory analysis	Composition, morphology, isotopes and source markers.	$1.5M
A9-02	Optical and hygroscopic laboratory analysis	Measured refractive index, absorption and water uptake.	$1.0M
A9-03	Cloud-active-particle analysis	CCN/ice-nucleating properties and uncertainty.	$0.8M
A9-04	Laboratory quality assurance	Blanks, standards, chain of custody and interlaboratory comparison.	$0.7M
A10-01	Campaign data system	Synchronized aircraft, vessel, satellite and ground dataset.	$1.3M
A10-02	Causal plume reconstruction	Emissions-to-aerosol transport and aging history.	$1.0M
A10-03	Cloud response attribution	Meteorology-controlled estimate of aerosol effect on cloud properties.	$1.2M
A10-04	Radiative-closure analysis	Top-of-atmosphere and surface forcing with uncertainty budget.	$1.5M
A11-01	LES/CRM model challenge	Prospective simulations compared with withheld cloud and radiation data.	$1.5M
A11-02	Global-model parameterization tests	Implementation in at least three independent global models.	$1.5M
A11-03	Structural uncertainty comparison	Evaluate alternative aerosol-cloud formulations and emergent constraints.	$0.8M
A11-04	Independent model evaluation	Firewalled evaluation of performance and compensating errors.	$0.7M
A12-01	Data-quality and metrology audit	Calibration, completeness, exclusions and uncertainty audit.	$0.7M
A12-02	Open repository and reproducibility package	Raw and calibrated data, code, metadata and persistent identifiers.	$0.6M
A12-03	International review workshop	External review and final decision-focused report.	$0.7M
A13-01	Program-manager reserve	Unforeseen technical and scheduling requirements.	$0.7M
A13-02	Instrument and platform contingency	Repairs, replacements and additional operations.	$0.8M

B. Marine low-cloud campaign - $55M

PO	Procurement package	Deliverable	Estimate

PO	Procurement package	Deliverable	Estimate
C1-01	Integrated campaign design	Cloud-regime matrix, decision metrics, flight patterns and validation plan.	$1.0M
C1-02	Observing-system simulation	Sampling sensitivity to cloud variability and mesoscale organization.	$0.6M
C1-03	Independent design review	Proceed/redesign report.	$0.4M
C2-01	Site and season atlas	Satellite and reanalysis assessment of target regimes.	$0.7M
C2-02	Forecast-system development	Operational cloud-regime and trajectory forecasts.	$0.8M
C3-01	Cloud probe and turbulence suite	Calibrated droplet, drizzle, liquid-water, temperature and turbulence measurements.	$2.5M
C3-02	Aerosol/CCN suite	Below-cloud aerosol size, composition, CCN and sea-salt characterization.	$1.6M
C3-03	Radiation suite	Broadband and spectral flux above, within and below cloud.	$1.4M
C3-04	Radar, lidar and imaging suite	Cloud structure, drizzle, boundary-layer depth and mesoscale context.	$1.8M
C3-05	Surface flux and autonomous platforms	Sensible/latent heat, aerosol source, SST and boundary-layer profiles.	$0.7M
C4-01	Primary aircraft integration	Payload, inlets, power, airworthiness and data systems.	$1.5M
C4-02	Secondary aircraft integration	Remote sensing and coordinated operations.	$1.0M
C4-03	Vessel/autonomous integration	Radar, lidar, radiometry, aerosol and communications.	$0.8M
C4-04	Calibration and rehearsal	Side-by-side flights and surface intercomparison.	$0.7M
C5-01	Baseline mapping flights	Cloud deck, transition and cumulus environmental mapping.	$2.0M
C5-02	Surface baseline period	Continuous cloud, aerosol, radiation and flux measurements.	$1.2M
C5-03	Mission planning center	Daily targeting, trajectories and quick-look products.	$0.8M
C6-01	Primary-aircraft sortie block	Approximately 30–40 stacked and Lagrangian sorties.	$7.0M
C6-02	Secondary-aircraft sortie block	Approximately 15–20 remote-sensing and context sorties.	$3.5M
C6-03	Vessel and autonomous operations	Continuous surface and boundary-layer observations.	$2.0M
C6-04	Field logistics, spares and standby	Deployment support and replacement cases.	$2.5M
C7-01	Ship-track and natural-gradient analysis	Cloud response as aerosol and meteorology evolve.	$1.2M
C7-02	Metered sea-salt option	Governed small perturbation with source and cloud-response measurements.	$1.3M

PO	Procurement package	Deliverable	Estimate
C7-03	Independent attribution analysis	Separate aerosol signal from meteorological variability.	$0.5M
C8-01	Case LES ensemble	High-resolution simulations of all priority cases.	$2.0M
C8-02	Uncertainty and process attribution	Entrainment, drizzle, organization and aerosol sensitivity.	$1.5M
C8-03	Held-out flight prediction	Prospective LES predictions for selected withheld periods.	$1.5M
C9-01	Cloud parameterization team 1	Improved turbulence/entrainment treatment.	$1.4M
C9-02	Cloud parameterization team 2	Improved microphysics/drizzle treatment.	$1.4M
C9-03	Scale-aware coupling team	Transfer physics from LES to global-model resolutions.	$1.0M
C9-04	Independent code review	Numerical, conservation and reproducibility audit.	$0.7M
C10-01	Multi-model implementation	At least three global models with revised physics.	$1.5M
C10-02	Out-of-sample climate tests	Evaluate cloud radiative effect, feedback and precipitation in other regions.	$1.2M
C10-03	Reflectivity sensitivity tests	Compare MCB-like forcing using physically constrained cloud response.	$0.8M
C11-01	Radiative and water-budget closure	Independent campaign closure report.	$0.8M
C11-02	Public archive and reproducibility	Complete data/code package.	$0.6M
C11-03	Review workshop and assessment	External review and future-decision recommendations.	$0.6M
C12-01	Contingency reserve	Weather, aircraft and instrument failures.	$2.5M

C. AMOC and ocean heat uptake - $60M

PO	Procurement package	Deliverable	Estimate
O1-01	Decision and observing-system design	Target variables, thresholds, array gaps and forecast experiments.	$1.2M
O1-02	Array optimization modeling	Observing-system simulation and uncertainty reduction per instrument.	$0.8M
O1-03	Independent design review	Proceed/redesign recommendation.	$0.5M
O2-01	Mooring instrument packages	Current, temperature, salinity, pressure and oxygen sensors.	$5.5M
O2-02	Mooring frames, lines and releases	Deep-ocean hardware and recovery systems.	$2.5M

PO	Procurement package	Deliverable	Estimate
O2-03	Telemetry and data systems	Near-real-time data return and system-health monitoring.	$1.2M
O2-04	Calibration and spares	Pre/post calibrations and replacement inventory.	$1.8M
O3-01	Glider fleet	Long-duration sections across formation and boundary-current regions.	$2.5M
O3-02	Argo and deep-Argo floats	Temperature, salinity, oxygen and deep-ocean profiles.	$2.0M
O3-03	Autonomous surface vehicles	Air-sea flux, meteorology and upper-ocean observations.	$1.5M
O3-04	Specialty profiling and turbulence systems	Mixing, microstructure and boundary-current observations.	$2.0M
O4-01	Deployment research cruise	Mooring installation, hydrography, tracer and platform deployment.	$5.5M
O4-02	Cruise science and laboratories	Shipboard chemistry, CTD, tracers and data processing.	$1.2M
O4-03	International berthing and permits	Port, diplomatic and maritime coordination.	$0.8M
O5-01	Array operations center	Telemetry, QA, instrument health and data release.	$2.0M
O5-02	Glider/autonomous operations	Piloting, maintenance, recoveries and redeployments.	$1.5M
O5-03	Annual servicing options	Targeted ship time and replacement deployments.	$1.5M
O5-04	Data quality and calibration	Continuous quality control and cross-platform comparison.	$1.0M
O6-01	Tracer-release experiment	Approved tracer release and multi-month sampling.	$2.5M
O6-02	Turbulence and overflow survey	Mixing, entrainment and dense-water pathway measurements.	$2.0M
O6-03	Convection process campaign	Wintertime air-sea flux and deep-mixing observations.	$2.0M
O7-01	Greenland meltwater pathway study	Source markers, shelf transport and basin penetration.	$1.5M
O7-02	Arctic freshwater export analysis	Strait transports, storage and attribution.	$1.2M
O7-03	Freshwater-budget closure	Integrated freshwater inventory and uncertainty.	$1.3M
O8-01	Recovery and repeat cruise	Recover instruments, repeat sections and redeploy selected assets.	$4.2M
O8-02	Final calibrations and asset closeout	Post-deployment calibration and equipment disposition.	$1.3M
O9-01	Ocean state estimation	Dynamically consistent heat, salt and transport reconstruction.	$1.5M
O9-02	Multi-model assimilation comparison	Independent estimates and structural uncertainty.	$1.3M

PO	Procurement package	Deliverable	Estimate
O9-03	Heat and freshwater budget analysis	Closed regional and basin budgets.	$1.2M
O10-01	Seasonal prediction challenge	Prospective forecasts of convection and transport anomalies.	$0.8M
O10-02	Decadal prediction challenge	Initialized forecasts and hindcast evaluation.	$0.8M
O10-03	Reflectivity-response experiments	Coupled-model tests of spatial/seasonal cooling and AMOC response.	$0.9M
O11-01	Independent instrument and data audit	Calibration, loss, drift and uncertainty assessment.	$0.5M
O11-02	Open ocean-data archive	Standardized data, code and documentation.	$0.5M
O11-03	International assessment workshop	Decision-focused AMOC and model-evaluation report.	$0.5M
O12-01	Lost-instrument and ship-delay reserve	Replacement instruments, ship time and severe-weather response.	$1.0M

How to contract the program efficiently

Issuing hundreds of unrelated contracts would create excessive transaction cost and fragmented accountability. The Managing Organization should establish approximately ten master agreements with campaign-specific task orders:

Master agreement	Task orders covered
Aircraft and airborne science	Aircraft access, payload integration, airworthiness, sorties and field support.
Research vessels and autonomous ocean platforms	Cruises, moorings, gliders, floats, surface vehicles and maritime logistics.
Atmospheric instrumentation and calibration	Aerosol, cloud, radiation, radar, lidar, gases and metrology.
Satellite products and calibration/validation	Overpass planning, retrievals, underflights and long-term data continuity.
Field-campaign logistics	Sites, shipping, facilities, communications and safety.
High-resolution and global modeling	LES, CRM, storm-resolving and Earth-system model development.
Data assimilation and scientific software	Common data systems, state estimation, visualization and HPC.
Scientific-analysis consortium	University and laboratory research subawards.
Environmental, legal and engagement services	Review, permits, consultation, public process and ethics.

Master agreement	Task orders covered
Independent evaluation and open science	Audits, blinded challenges, replication, archives and assessment workshops.

Each task order should specify a decision-relevant deliverable, measurable acceptance criteria, uncertainty requirements, open-data and code obligations, safety and stop-work clauses, payment milestones, change-control procedures and options that are exercised only when scientifically justified.

Recipients of Purchase Orders

Prompt to AI: Who might receive the purchase orders, subawards and task orders described above?

Response From AI:

The fund should use competitive selection and should not assume that any named organization would participate. The examples below identify plausible capabilities rather than preselected winners. Government participation would often require cooperative or reimbursable agreements rather than ordinary commercial purchase orders.

Recommended program consortium

Role	Plausible organizations
Fiduciary and program-management prime	RTI International, Battelle, or another independent nonprofit research institute with large-program procurement and audit systems.
Atmospheric-science integration	UCAR/NSF NCAR, working with NOAA laboratories, universities and international partners.
Model evaluation and intercomparison	Program for Climate Model Diagnosis and Intercomparison (PCMDI), GeoMIP/CMIP networks, NCAR and independent model centers.
Radiation and satellite calibration	NASA Langley, NASA Goddard, NOAA, DOE ARM, national metrology organizations and university remote-sensing groups.
Aerosol and cloud observations	NOAA Chemical Sciences Laboratory, DOE ARM, NASA airborne-science teams, NCAR Earth Observing Laboratory and university laboratories.
Ocean and AMOC integration	NOAA Atlantic Oceanographic and Meteorological Laboratory, Woods Hole Oceanographic Institution, Scripps, University of Miami, UK National Oceanography Centre and international array partners.
Independent environmental review	AECOM, Ramboll, RTI or another competitively selected environmental assessment organization separated from experiment leadership.
Public engagement and conflict resolution	Consensus Building Institute, Keystone Policy Center or comparable international organizations.
Independent scientific review	National Academies or a separately appointed international panel with disclosed conflicts.

A. Aerosol radiative-forcing closure campaign

Procurement area	Plausible recipients
Science design and synthesis	NOAA CSL/CIRES, University of Washington, University of Leeds, University of Oxford, NASA GISS or NCAR-led consortium.
Aerosol composition and optical instruments	NOAA CSL instrument teams, Aerodyne Research, TSI, Droplet Measurement Technologies, Handix Scientific and university aerosol laboratories.
CCN and hygroscopicity	NOAA, DOE ARM, NCAR and university cloud-aerosol groups, with commercial instrument suppliers.
Cloud microphysics	NCAR EOL, DOE ARM, NASA airborne teams and experienced university cloud laboratories.
Airborne radiation and remote sensing	NASA Langley, NASA Goddard, NOAA and university lidar/radar groups.
Research aircraft	NASA airborne-science programs, NSF NCAR aircraft, NOAA Aircraft Operations Center, SAFIRE or other currently verified international operators.
Research vessel and autonomous surface platforms	NOAA fleet, UNOLS institutions, Saildrone or comparable autonomous-platform operators.
Satellite products	NASA, NOAA, EUMETSAT, ESA and university retrieval teams.
LES and process modeling	NCAR, University of Washington, Max Planck Institute for Meteorology, University of Hamburg, UK Met Office and selected university groups.
Independent model evaluation	PCMDI, a firewalled NCAR team, national metrology laboratories or an independent statistics consortium.

B. Marine low-cloud campaign

Procurement area	Plausible recipients
Campaign science leadership	University of Washington cloud group, NCAR, NOAA, DOE ARM, University of Leeds, University of Miami and international low-cloud experts.
Aircraft platforms	NSF NCAR C-130 or University of Wyoming King Air for in-cloud work; NSF NCAR Gulfstream V for above-cloud/context work; NASA, NOAA, SAFIRE or other verified aircraft depending on payload and altitude.
Ship and surface observations	UNOLS research vessels, NOAA vessels, DOE ARM Mobile Facility, autonomous surface-vehicle operators.
Cloud radar and lidar	DOE ARM, NCAR EOL, NOAA/NASA groups and university radar laboratories.
Large-eddy simulation	University of Washington, NCAR, Max Planck Institute for Meteorology, University of Hamburg, Delft University of Technology and UK Met Office groups.

Procurement area	Plausible recipients
Parameterization development	GFDL, NCAR, NASA GISS, UK Met Office, ECMWF, MPI-M and university teams, through open competitive calls.
Independent closure	PCMDI, DOE ARM data quality office, national metrology institutes and independent radiative-transfer experts.

C. AMOC and ocean heat uptake

Procurement area	Plausible recipients
Array and cruise leadership	WHOI, NOAA AOML, Scripps Institution of Oceanography, University of Miami, UK National Oceanography Centre and existing RAPID/OSNAP partners.
Mooring systems	WHOI, Scripps, NOC, NOAA laboratories and qualified ocean-instrument manufacturers.
Gliders, floats and autonomous vehicles	Scripps, WHOI, NOAA, Teledyne Webb Research, Saildrone and international Argo partners.
Tracer and mixing experiments	WHOI, Scripps, University of Washington, University of Southampton/NOC and specialist ocean-mixing groups.
Greenland and Arctic freshwater	WHOI, University of Washington, NOC, GEOMAR, Danish/Greenland institutions and Arctic research consortia.
Ocean state estimation	MIT, NASA JPL, NOAA/GFDL, NCAR, ECMWF and international ocean-assimilation centers.
Coupled-model experiments	GFDL, NCAR, NASA GISS, UK Met Office, MPI-M, IPSL and other CMIP centers.
Independent review	An international panel administered separately from array operators and model developers.

Selection principles

1. Separate design from evaluation. The team that builds a parameterization should not be the only team deciding whether it improved the model.
2. Use existing infrastructure. Strengthen proven aircraft, vessel, satellite, ARM, Argo and AMOC programs rather than duplicating them.
3. Fund international replication. At least one major campaign and model challenge should include independently led non-U.S. teams.
4. Avoid institutional capture. No single model center, university or laboratory should control the observing plan, analysis, funding decisions and final assessment.

5. Require open deliverables. Recipients should provide raw and calibrated data, code, configuration files, uncertainty calculations and documentation.
6. Preserve critical expertise. Long-term instrument, calibration and data-stewardship teams should receive stable support rather than one-off grants that disappear after a campaign.

Project Management

Prompt to AI: What project-management provisions are specific to this Global Climate Model Experimental Research initiative after incorporating the Climate Moonshot Portfolio Common Framework?

Response From AI:

Portfolio-wide governance, Managing Organization selection, fiduciary controls, procurement, financial controls, reserve governance, independent review, audit, conflicts, open-source/data rights, and transition rules are specified in the Moonshot Common Framework. This chapter states only the additional technical-management structure required by the GCM experimental-research program.

Initiative-specific management requirements

- Maintain one accountable Program Director and integrated science/operations baseline spanning the twelve research areas, shared observing backbone, field campaigns, model hierarchy, data assimilation, held-out validation, rapid response, and final decision-readiness assessment.
- Maintain distinct atmospheric-science, ocean/AMOC, model-evaluation, field-operations, metrology/calibration, data/software, and international-participation leads. No model-development team should be the sole evaluator of its own improvement.
- The independent model-evaluation function must control protected observations, preregistered scorecards, model genealogy, cross-variable/no-compensating-error tests, prospective predictions, and the final evidence record used to claim improved decision readiness.
- Any metered aerosol or tracer perturbation remains subject to a separate scientific, environmental, legal, ethics, and safety authorization. The program-management chain cannot convert an observational or model-development award into a larger perturbation activity by change order alone.
- Aircraft, vessel, satellite, ARM, Argo, AMOC-array, autonomous-platform, laboratory, and computing resources should be scheduled through an integrated campaign calendar with common calibration, data, and readiness interfaces so shared assets are not double-funded.
- Phase II or any deployment-oriented sunlight-reflection program requires separate authorization; completion of this research program does not authorize SAI or MCB deployment.

Funding Source Requirements

We assume the Funding Source aims to maximize climate benefit per dollar spent, as opposed to maximize financial return on an investment. This is sometimes referred to as "climate money", and it differs from the traditional investment portfolio. Sources of climate money include high-net-worth individuals, foundations, and governments.

The portfolio-wide Funding Source requirements are specified in the Moonshot Common Framework. The additions below address the program's climate-model validation mission, protected observations, controlled perturbations, international participation, measurement traceability, and decision-focused model evaluation.

Program-specific record additions

- Predeclare the decision quantities, scientific hypotheses, uncertainty targets, campaign success/failure criteria, model versions, protected variables, and stopping rules for each major experiment or model challenge.
- Preserve traceable calibration, uncertainty covariance, platform configuration, raw and calibrated observations, exclusion/quality flags, model source/configuration, forcing files, analysis code, and prospective predictions needed to reproduce every major closure or validation claim.
- Maintain protected/held-out observations under controlled access until preregistered predictions are frozen; release the protected data afterward under the common open-source/data policy.
- Require independent evaluation across energy, water, aerosol mass, clouds, precipitation, ocean heat, AMOC, sea ice, and regional climate so an improvement cannot be claimed solely from better global-mean temperature or one favored metric.
- Any deliberate atmospheric or ocean perturbation must remain scientifically bounded, independently authorized, environmentally reviewed, publicly documented, and materially below a scale intended to create regional or global climate change.
- Fund substantive participation and analytical capacity in underrepresented regions where monsoon, water, food, island, ocean, or other regional risks are decision-relevant; participation must include authority over research questions and analysis rather than symbolic consultation.

Decision-readiness metric

Do not measure success by publications, model resolution, or number of simulations. Report whether the program materially narrows uncertainty in required radiative forcing, spatial/seasonal intervention patterns, regional temperature and precipitation response, ocean-circulation risk, monitoring thresholds, and failure/termination consequences, while explicitly preserving unresolved structural disagreement.

Top GCM Science Papers

Prompt to AI: Summarize top papers that are especially useful for designing a research fund to improve climate models for clouds, aerosols, radiation, ocean circulation and sunlight-reflection decisions.

Response From AI:

"Top" is necessarily subjective. The list below emphasizes historical influence, scientific importance and direct usefulness for the proposed program. It is not a citation-count ranking and is not a complete literature review.

Climate Model Evaluation, Feedbacks, and Sensitivity

Bony (2006), "How Well Do We Understand and Evaluate Climate Change Feedback Processes?"

A foundational review of feedback diagnosis. It explains why clouds dominate uncertainty in climate sensitivity and why model evaluation must examine physical mechanisms rather than global averages alone.

https://doi.org/10.1175/JCLI3819.1

Gleckler (2008), "Performance Metrics for Climate Models"

Introduces systematic multi-variable metrics for comparing models with observations. It is central to building transparent scorecards while avoiding reliance on a single metric.

https://doi.org/10.1029/2007JD008972

Stevens (2013), "What Are Climate Models Missing?"

Argues for closer integration of observations, theory and models, with particular attention to clouds and convection. The paper motivates process-focused experiments rather than simply increasing model complexity.

https://doi.org/10.1126/science.1237554

Bony (2015), "Clouds, Circulation and Climate Sensitivity"

Frames cloud feedback as inseparable from atmospheric circulation. It helped organize modern research around specific cloud regimes and dynamical controls.

https://doi.org/10.1038/ngeo2398

Eyring (2016), "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6)"

Defines the coordinated experimental framework used by global modeling centers. CMIP demonstrates the value—and limits—of common protocols for separating scenario, model and internal-variability effects.

https://doi.org/10.5194/gmd-9-1937-2016

Sherwood (2020), "An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence"

Integrates process understanding, historical change, paleoclimate and emergent constraints. It provides a model for combining evidence without treating any one line as decisive.

https://doi.org/10.1029/2019RG000678

Zelinka (2020), "Causes of Higher Climate Sensitivity in CMIP6 Models"

Identifies stronger positive cloud feedback—especially from low clouds—as a major reason some CMIP6 models are more sensitive. It shows how parameterization choices alter decision-relevant warming estimates.

https://doi.org/10.1029/2019GL085782

Palmer (2019), "The Scientific Challenge of Understanding and Estimating Climate Change"

Calls for a step change in computing, observations and physical modeling. It argues that climate science needs prediction systems whose uncertainty is connected to resolved processes.

https://doi.org/10.1073/pnas.1906691116

Clouds and Cloud Feedback

Ceppi (2017), "Cloud Feedback Mechanisms and Their Representation in Global Climate Models"

A comprehensive review of low, high and mixed-phase cloud feedbacks. It maps which physical processes most need field and high-resolution constraints.

https://doi.org/10.1002/wcc.465

Bodas-Salcedo (2014), "Origins of the Solar Radiation Biases over the Southern Ocean in CFMIP2 Models"

Shows how cloud-phase and microphysical errors create major Southern Ocean radiation biases. It is a direct rationale for mixed-phase cloud campaigns and process-oriented satellite simulators.

https://doi.org/10.1175/JCLI-D-13-00169.1

Tan (2016), "Observational Constraints on Mixed-Phase Clouds Imply Higher Climate Sensitivity"

Demonstrates how the representation of supercooled liquid water influences cloud feedback and climate sensitivity. It also illustrates the opportunities and limitations of observational constraints.

https://doi.org/10.1126/science.aad5300

Brient (2016), "Constraints on Climate Sensitivity from Space-Based Measurements of Low-Cloud Reflection"

Links observed low-cloud behavior to model feedback and sensitivity. It is relevant to designing held-out satellite and field tests for subtropical clouds.

https://doi.org/10.1175/JCLI-D-15-0897.1

Stevens (2021), "EUREC4A" campaign overview"

Describes a major international field campaign designed around shallow trade-wind clouds, circulation and climate feedback. EUREC4A is a strong organizational and scientific precedent for the proposed low-cloud campaign.

https://doi.org/10.5194/essd-13-4067-2021

Aerosols and Aerosol-Cloud Forcing

Bellouin (2020), "Bounding Global Aerosol Radiative Forcing of Climate Change"

Provides a comprehensive assessment of aerosol direct, cloud and surface effects. It highlights the persistent uncertainty in anthropogenic aerosol forcing and the need for causal closure experiments.

https://doi.org/10.1029/2019RG000660

Myhre (2013), “Radiative Forcing of the Direct Aerosol Effect from AeroCom Phase II Simulations”

Compares global aerosol models and observations. The spread demonstrates how emissions, optical properties, transport and humidity produce uncertainty even before cloud interactions are included.

https://doi.org/10.5194/acp-13-1853-2013

Ghan (2013), “Technical Note: Estimating Aerosol Effects on Cloud Radiative Forcing”

Clarifies how to diagnose aerosol-cloud radiative effects while separating changes in clear-sky and cloudy-sky radiation. It is useful for defining campaign and model output requirements.

https://doi.org/10.5194/acp-13-9971-2013

Malavelle (2017), “Strong Constraints on Aerosol-Cloud Interactions from Volcanic Eruptions”

Uses a volcanic natural experiment to test aerosol-cloud responses. It demonstrates both the power of large perturbations and the difficulty of extrapolating from one event or regime.

https://doi.org/10.1038/nature22974

Christensen (2022), “coordinated analyses of the 2020 marine-fuel sulfur change and ship-track response”

Studies the global emissions perturbation created by low-sulfur shipping fuel. This class of work illustrates how policy changes can constrain aerosol-cloud forcing without deliberate climate-scale experiments.

https://doi.org/10.5194/acp-22-641-2022

Storm-Resolving and High-Resolution Modeling

Stevens (2019), “DYAMOND: The Dynamics of the Atmospheric General Circulation Modeled on Non-hydrostatic Domains”

Introduces coordinated global storm-resolving simulations. DYAMOND shows how kilometre-scale models can expose differences in convection and cloud organization that parameterized models hide.

https://doi.org/10.1186/s40645-019-0304-z

Satoh (2019), "review of global cloud-resolving modeling"

Summarizes the development, promise and computational limitations of global storm-resolving models. It supports using them as experimental laboratories and parameterization guides, not automatically as truth.

https://doi.org/10.1007/s40641-019-00131-0

Earth Energy Budget

Stephens (2012), "An Update on Earth's Energy Balance in Light of the Latest Global Observations"

Synthesizes top-of-atmosphere, atmospheric and surface energy flows and their uncertainty. It shows why radiative closure and instrument calibration are prerequisites for detecting small changes in planetary albedo.

https://doi.org/10.1038/ngeo1580

Loeb (2021), "Satellite and Ocean Data Reveal Marked Increase in Earth's Heating Rate"

Combines satellite radiation and ocean heat observations to diagnose changes in Earth's energy imbalance. It underscores the importance of independent observing systems and careful attribution.

https://doi.org/10.1029/2021GL093047

Ocean Circulation and AMOC

Lozier (2019), "A Sea Change in Our View of Overturning in the Subpolar North Atlantic"

Uses the OSNAP array to show that overturning variability is strongly influenced by the eastern subpolar gyre rather than only the Labrador Sea. It demonstrates how sustained observations can overturn model assumptions.

https://doi.org/10.1126/science.aau6592

Frajka-Williams (2019), "Atlantic Meridional Overturning Circulation: Observed Transport and Variability"

Reviews AMOC observing arrays, methods and variability. It is a guide to what is measured, where uncertainty remains and how observations should constrain coupled models.

https://doi.org/10.3389/fmars.2019.00260

Weijer (2020), “Stability of the Atlantic Meridional Overturning Circulation: A Review and Synthesis”

Synthesizes theory, paleoclimate, observations and models concerning AMOC stability and thresholds. It is essential for defining what a reflectivity strategy intended to reduce AMOC risk must predict and monitor.

https://doi.org/10.1029/2019JC015083

Ten papers to read first

For program design, I would begin with Bony 2015; Ceppi 2017; Bellouin 2020; Stevens 2021 (EUREC4A); Stevens 2019 (DYAMOND); Palmer and Stevens 2019; Loeb 2021; Lozier 2019; Frajka-Williams 2019; and Weijer 2020. Together they connect the program’s central needs: cloud processes, aerosol forcing, radiative closure, high-resolution modeling, sustained observations and ocean-circulation risk.

Bibliographic details and paper selections should receive a final literature review by the Scientific Advisory Panel before publication or procurement use.

Top Researchers and Organizations

Prompt to AI: Who are leading researchers and organizations relevant to this climate-model experimental research program?

Response From AI:

There is no authoritative ranking. The examples below emphasize scientific influence, current leadership and relevance to clouds, aerosols, radiation, model evaluation, storm-resolving modeling, ocean heat uptake, AMOC and decision-focused climate prediction. Affiliations should be verified before outreach.

Leading researchers

Researcher	Current or recent base	Principal relevance
Björn Stevens	Max Planck Institute for Meteorology	Clouds, convection, climate sensitivity, field campaigns and storm-resolving modeling.
Sandrine Bony	LMD/IPSL and Sorbonne Université/CNRS	Cloud feedbacks, circulation, climate sensitivity and EUREC4A leadership.
Robert Wood	University of Washington	Marine boundary-layer clouds, aerosol-cloud interactions, field observations and cloud susceptibility.
Piers Forster	University of Leeds	Earth radiation budget, radiative forcing, climate sensitivity and assessment leadership.
Trude Storelvmo	University of Oslo	Aerosol-cloud interactions, mixed-phase clouds and climate feedback.
Mark Zelinka	Lawrence Livermore National Laboratory	Cloud feedback, climate sensitivity and model diagnosis.
Johannes Quaas	Leipzig University	Aerosol-cloud forcing, satellite constraints and climate-model evaluation.
Olivier Boucher	IPSL/Sorbonne Université	Aerosols, radiative forcing, climate modeling and assessments.
Jim Haywood	University of Exeter / UK Met Office community	Aerosol-radiation interactions, observations, volcanic analogues and climate intervention.
Graham Feingold	NOAA Chemical Sciences Laboratory	Cloud-aerosol interactions, process modeling and observational analysis.
Zhanqing Li	University of Maryland	Aerosol, cloud, radiation and surface-observation networks.

Researcher	Current or recent base	Principal relevance
Graeme Stephens	NASA JPL / Colorado State University community	Clouds, precipitation, satellite observations and Earth energy balance.
Norman Loeb	NASA Langley Research Center	CERES radiation measurements, Earth energy imbalance and calibration.
Christian Jakob	Monash University	Cloud and convection parameterization, model evaluation and process-oriented development.
Tapio Schneider	California Institute of Technology	Climate dynamics, cloud feedback and computational climate modeling.
Tim Palmer	University of Oxford	Ensembles, uncertainty, numerical prediction and high-resolution climate modeling.
Ruby Leung	Pacific Northwest National Laboratory	Earth-system modeling, convection, precipitation extremes and model development.
David Randall	Colorado State University	Cloud parameterization, atmospheric modeling and model hierarchy.
Laure Zanna	New York University	Ocean circulation, machine learning, parameterization and climate prediction.
Susan Lozier	Georgia Institute of Technology	AMOC observations and OSNAP leadership.
Eleanor Frajka-Williams	National Oceanography Centre / University of Southampton community	AMOC observing systems, transport and variability.
M. Susan Lozier / OSNAP consortium leaders	International consortium	Subpolar overturning, sustained arrays and ocean-observation strategy.
John Marshall	MIT	Ocean circulation, climate dynamics and state estimation.
Raffaele Ferrari	MIT	Ocean mixing, circulation and heat uptake.
Marika Holland	NSF NCAR	Sea ice, polar climate and coupled-model feedbacks.
Clara Deser	NSF NCAR	Internal variability, large ensembles and attribution.
Ben Kravitz	Indiana University	GeoMIP, multi-model climate intervention experiments and control strategies.
Douglas MacMartin	Cornell University	Climate intervention control, uncertainty and decision design.
Simone Tilmes	NSF NCAR	Chemistry-climate modeling, SAI scenarios and Earth-system impacts.
Daniele Visioni	Cornell University	Stratospheric aerosols, intervention scenarios and model uncertainty.

Top organizations and networks

Organization or network	Principal strength
NSF National Center for Atmospheric Research (NCAR) and UCAR	Integrated atmosphere, ocean, chemistry and Earth-system modeling; aircraft; field campaigns; supercomputing; data stewardship.
NOAA Chemical Sciences Laboratory	Aerosol, cloud, radiation, atmospheric chemistry, high-altitude and airborne measurements.
NOAA Geophysical Fluid Dynamics Laboratory	Global coupled-model development, clouds, aerosols, ocean circulation and climate projections.
NASA Langley Research Center	CERES radiation, airborne remote sensing, cloud/aerosol instruments and satellite validation.
NASA Goddard Space Flight Center	Satellite missions, Earth-system observations, data assimilation and global modeling.
DOE Atmospheric Radiation Measurement user facility	Long-term ground and mobile observations of clouds, aerosols, radiation and precipitation.
Lawrence Livermore National Laboratory / PCMDI	CMIP infrastructure, model diagnostics, benchmarks and intercomparison.
Max Planck Institute for Meteorology	Clouds, convection, field campaigns, Earth-system modeling and global storm-resolving simulation.
LMD/IPSL	Cloud feedback, circulation, aerosols, Earth-system modeling and international assessment.
UK Met Office Hadley Centre	Operational and climate prediction, coupled modeling, convection and aerosol/cloud research.
European Centre for Medium-Range Weather Forecasts	Data assimilation, numerical prediction, reanalysis and Earth-system model development.
University of Washington atmospheric science groups	Marine low clouds, aerosol-cloud interaction and field campaigns.
University of Leeds	Radiative forcing, aerosols, clouds, climate sensitivity and assessments.
EUREC4A consortium	International field-campaign model for shallow clouds, circulation and open data.
DYAMOND consortium	Coordinated global storm-resolving model intercomparison.
Cloud Feedback Model Intercomparison Project (CFMIP)	Common diagnostics and experiments for cloud feedback and model evaluation.
Aerosol Comparisons between Observations and Models (AeroCom)	Global aerosol-model intercomparison and forcing analysis.
Woods Hole Oceanographic Institution	AMOC arrays, ocean mixing, floats, moorings, state estimation and ocean process studies.

Organization or network	Principal strength
Scripps Institution of Oceanography	Argo, ocean observations, gliders, mixing and climate dynamics.
NOAA Atlantic Oceanographic and Meteorological Laboratory	AMOC and Atlantic observing systems, ocean heat and operational monitoring.
UK National Oceanography Centre	RAPID array, AMOC observations, autonomous systems and ocean modeling.
OSNAP consortium	Subpolar North Atlantic overturning observations and international coordination.
Argo and Deep Argo programs	Global and deep-ocean temperature, salinity and biogeochemical profiles.
Geoengineering Model Intercomparison Project (GeoMIP)	Common sunlight-reflection scenarios and multi-model comparison.
Degrees Initiative and Global South research networks	Independent regional research capacity, equity and broad participation in SRM assessment.

Best scientific core for the proposed fund

Function	Suggested scientific core
Earth radiation budget and calibration	NASA Langley, NOAA, NASA Goddard, DOE ARM and metrology partners.
Aerosol forcing	NOAA CSL, AeroCom, Leeds, Leipzig, IPSL and university observation groups.
Marine low clouds	University of Washington, NCAR, DOE ARM, MPI-M, LMD/IPSL and EUREC4A teams.
Mixed-phase and polar clouds	NCAR, DOE ARM, University of Oslo, UK groups and Southern Ocean/Arctic campaign teams.
Storm-resolving modeling	DYAMOND centers, MPI-M, NCAR, ECMWF, UK Met Office, JMA and selected university groups.
Global-model evaluation	PCMDI, CFMIP, CMIP, independent statistics/metrology teams and satellite-observation groups.
AMOC and ocean heat	WHOI, NOAA AOML, Scripps, NOC, OSNAP, RAPID, Argo and coupled-model centers.
Reflectivity-control scenarios	GeoMIP, Cornell, NCAR, UChicago and independent regional-impact teams.
Global participation	Degrees Initiative, IISc, African, Latin American, Pacific and other regional research institutions.

No single organization has all required capabilities. The strongest program would deliberately distribute authority among observation teams, model developers, oceanographers, high-resolution modelers, model evaluators, regional-impact researchers and independent critics.

Scientific Advisory Panel

The initiative should use a technology-specific Scientific Advisory Panel while portfolio-wide governance, fiduciary oversight, procurement, audit, and management-board functions remain in the Common Framework. Individual names in the preceding Top Researchers and Organizations chapter are examples of the expertise pool, not preselected appointees. Members must recuse from awards or protected-data decisions involving their institutions or close collaborators; model-development teams may not be the sole evaluators of their own changes. The panel may recommend proceed, redesign, replication, pause, or stop, but it does not itself authorize controlled perturbations, negotiate contracts, approve invoices, or replace the independent safety/governance process.

Seat	Expertise	Primary role
Chair / systems science	Broad climate-science and assessment leadership	Integrate evidence across radiation, atmosphere, ocean, uncertainty, and decision metrics.
Earth radiation / metrology	Broadband and spectral radiation, calibration	Challenge energy-budget closure and detection limits.
Aerosol forcing	Aerosol optics, chemistry, aerosol-cloud forcing	Review Area 2 and intervention-process evidence.
Marine low clouds	Boundary-layer clouds, aerosol activation, field campaigns	Review Area 3 and MCB-relevant susceptibility.
Mixed-phase / polar clouds	Cloud phase, sea ice, Southern Ocean / Arctic processes	Review Area 4 and polar radiative closure.
Convection / hydrology	Deep convection, precipitation, monsoons	Review Area 5 and regional hydrological response.
Ocean / AMOC	Overturning, freshwater, heat uptake, observing arrays	Review Area 6, initialized prediction, and ocean risk.
High-resolution modeling	LES, storm-resolving models, parameterization	Review Area 8 scale transfer and computational claims.
Model evaluation / statistics	Held-out testing, UQ, causal inference, genealogy	Protect independent evaluation and no-compensating-error rules.
SAI / MCB intervention modeling	Realistic forcing, GeoMIP/control strategies	Connect measured forcing evidence to global response without controlling release authorization.
Regional / Global South climate	Monsoon, water, food, island and regional impacts	Ensure consequential regional errors and priorities are not hidden by global averages.
Independent critic	Methods skeptical of favored pathways	Provide structured challenge and support credible stop/redirection recommendations.

Management Advisory Panel

The Moonshot Common Framework governs the panel charter, independence, audit access, conflicts, escalation rights, reserve review, corrective action, and Managing Organization oversight. This initiative adds several seats or equivalent capabilities because it combines expensive international field campaigns, government research platforms, protected validation data, high-performance computing, and potentially sensitive atmospheric perturbation research.

Initiative-specific oversight capabilities

- Large multi-platform field-campaign operations: aircraft/vessel scheduling, airworthiness and maritime readiness, weather risk, instrument integration, remote-site logistics, and rapid-response deployment.
- Climate-model and HPC program delivery: model-version control, computing procurement, protected validation data, reproducibility, software transition, and avoidance of model-center capture.
- International research governance: cross-border permits and agreements, regional participation, public engagement, data access, and independent review of controlled perturbations or transboundary activities.
- Scientific metrology and data stewardship: calibration traceability, long-lived archives, open formats, cyber/data integrity, and preservation of raw observations and model configurations through contractor or institution changes.

All other management-panel seats and responsibilities are supplied by the common framework. Technology-specific scientific review remains with the Scientific Advisory Panel in this proposal.

Agreements

Prompt to AI: Which agreement provisions should remain inside this proposal after the portfolio-wide Funding Source-to-Managing Organization and Managing Organization-to-Recipient master terms are moved to the common framework?

Response From AI:

The generic master agreement terms are specified in the Moonshot Common Framework and are not repeated here. Final contracts would use those common terms plus the GCM-specific schedules below, adapted by counsel to the parties, jurisdictions, government-facility rules, environmental requirements, and the actual procurement instrument.

Program-Specific Schedules to the Common Agreements

Scientific traceability and protected validation schedule

Identify the decision question, hypotheses, required uncertainty, instrument/model configuration, protected observations, preregistered metrics, release trigger, independent evaluator, and evidence needed for acceptance.

Field-platform and calibration schedule

For aircraft, vessels, balloons, satellites, autonomous systems, supersites, and laboratory instruments, specify readiness, calibration, timing/location accuracy, data completeness, safety/airworthiness or maritime responsibility, weather cancellation, replacement options, and accepted data products.

Controlled-perturbation schedule

Any metered aerosol or tracer release requires experiment-specific source limits, material, location, duration, monitoring, environmental/legal approvals, public disclosure, stop conditions, incident reporting, and independent source verification. The common agreements do not themselves authorize a perturbation.

Model/data/reproducibility schedule

Specify source-code repositories, exact configurations, forcing and boundary files, compiler/library information, data formats, uncertainty calculations, provenance, release dates, cloud/HPC transition rights, and the package required for an unaffiliated team to reproduce major figures and scorecards.

International and government-facility schedule

Where work uses government aircraft/ships, international arrays, national laboratories, or foreign field sites, document reimbursable/cooperative authorities, permits, data rights, equipment disposition, publication constraints, liability/insurance boundaries, and any narrow lawful exceptions to open release.

No duplicate master terms

General procurement, milestone payment, audit, conflicts, open-source/data rights, background-IP disclosure, reserve, changes/claims, remedies, termination, step-in, succession, insurance, and routine reporting provisions are governed by the common framework unless a written GCM-specific schedule expressly adds or overrides a term.

GCM Aircraft, Instruments and Airborne Measurement Architecture

Prompt to AI: The Global Climate Model (GCM) experimental research program relies on aircraft in many of its campaigns. Tell me more about the airplanes it needs. What parameters should each type measure? What instruments are needed? Which existing research-aircraft operators might provide the aircraft? Do we need new aircraft, modifications, or new airborne instrumentation? How should the program organize aircraft so that the same capability can serve radiation, aerosol, cloud, convection, polar and rapid-response experiments?

Response From AI:

The GCM program needs the broadest aircraft portfolio of the sunlight-reflection research programs because it is not studying one plume or one cloud regime. It must measure Earth radiation, anthropogenic aerosol forcing, marine low clouds, mixed-phase/polar clouds, deep convection, upper-tropospheric humidity and rare natural perturbations. The efficient solution is not to buy one universal airplane. It is to maintain a small set of interoperable aircraft classes and modular payloads, with common calibration, data, timing and mission-planning systems.

Six GCM airborne roles

Airborne role	Decision-relevant parameters	Representative instruments	Typical experiments
G1. High-altitude radiation / satellite validation	Up/down SW/LW flux; spectral radiance/reflectance; polarization; cloud/aerosol optical depth; layer height; surface/cloud albedo; scene geometry.	Broadband/spectral radiometers; multi-angle polarimeter/imager; HSRL/lidar; precision GPS/INS; calibrated cameras.	Area 1 radiation-budget closure; satellite underflights; above-cloud aerosol/radiation cases.
G2. High-altitude in-situ aerosol and chemistry	Particle number/size/composition; scattering/absorption/extinction; SO_2, ozone, NOy, H_2O and trace gases; UTLS meteorology; volcanic/wildfire material; anvil ice.	CPC/fast sizers; aerosol mass spectrometry; absorption/nephelometry; trace-gas spectrometers; hygrometers; cloud/ice probes; gust package.	Area 2 aerosol forcing; Area 5 anvil/upper troposphere; Area 7 volcano/smoke rapid response.
G3. Boundary-layer and cloud microphysics	Aerosol size/composition; CCN; cloud-base updraft; droplet/ice spectra; LWC/IWC; drizzle/precipitation; turbulence; inversion strength; T/RH and surface coupling.	CCN counter; optical aerosol sizers; cloud droplet/imaging probes; hot-wire water; precipitation probes; fast 3-D winds and thermodynamics.	Area 2 aerosol-cloud closure; Area 3 marine low clouds; Area 4 mixed-phase clouds.

Airborne role	Decision-relevant parameters	Representative instruments	Typical experiments
G4. Cloud/storm remote sensing	3-D cloud/precipitation structure; Doppler velocity; storm organization; cold pools; cloud top/base; anvil structure; precipitation microphysics.	Doppler/dual-polarization radar; cloud radar; HSRL/lidar; microwave radiometer; imagers; dropsondes.	Area 3 organization; Area 4 polar systems; Area 5 convection and precipitation.
G5. Polar / mixed-phase aircraft	Cloud phase; supercooled liquid; ice number/habit; snowfall/drizzle; aerosol/INP; radiation; boundary-layer coupling; sea-ice/surface context.	Cloud phase/imaging probes; LWC/IWC; INP/aerosol samplers; radar/lidar; radiometers; IR/surface imagers; turbulence package.	Southern Ocean and Arctic campaigns in Area 4.
G6. Rapid-response modular aircraft	Event-specific plume location, aerosol evolution, gases, cloud response, radiation and meteorology.	Deployable combinations of G1-G5 payloads plus satellite/trajectory mission system.	Volcanoes, pyroconvection, wildfire smoke, abrupt emissions changes and unique calibration events in Area 7.

Measurement packages should be modular

Shared airborne pallet	Core measurements	Why it matters
Radiative-closure pallet	Up/down broadband SW/LW, spectral radiometry, polarimetry/imaging, attitude/GPS and traceable calibration.	Connects clouds/aerosols/surface state to the actual energy-budget quantity that GCMs must reproduce.
Aerosol/chemistry pallet	Particle number/size/composition, scattering, absorption, CCN where relevant, SO2, ozone, NOy, water vapor and related gases.	Constrains direct aerosol forcing, activation, natural perturbations and aerosol-cloud coupling.
Cloud-microphysics pallet	Droplet and ice spectra, LWC/IWC, drizzle/precipitation, phase and temperature.	Constrains low-cloud, mixed-phase, anvil and precipitation parameterizations.
Dynamics/turbulence pallet	3-D winds, vertical velocity, T/RH/P, turbulence, surface/boundary-layer flux proxies.	Separates aerosol/cloud causality from meteorology and constrains entrainment and convection.
Active remote-sensing pallet	HSRL/lidar, cloud radar, Doppler/dual-pol radar where practical, microwave radiometry.	Supplies vertical and 3-D context that in-situ flight legs cannot provide alone.
Mission/data pallet	Common clocks, GPS/INS, aircraft state, telemetry, quick-look processing and aircraft-ship-ground-satellite synchronization.	Makes multi-platform closure and prospective model tests possible.

Illustrative current aircraft facilities

Aircraft / operator	Illustrative current capability	Best GCM roles
NASA ER-2	NASA lists about 70,000 ft maximum altitude, 8 hour duration and about 2,900 lb useful payload.	G1 radiation/remote sensing, satellite calibration/validation, high-altitude lidar/imaging.
NASA WB-57	NASA lists roughly 6.5 hour duration and up to about 8,800 lb payload, with operation above 60,000 ft.	G2 high-altitude in-situ aerosol/chemistry, volcanic/smoke events, upper-tropospheric/stratospheric sampling.
NSF/NCAR HIAPER Gulfstream V	About 10 hour endurance, 5,000 nmi range, 5,600 lb payload and altitude above 45,000 ft; multiple optical ports and external mounts.	G1 radiation, G2 upper-tropospheric chemistry, above-cloud remote sensing, polar/long-range and convection context.
NSF/NCAR C-130	Middle/lower troposphere; typical research configuration about 13,000 lb payload and 8-9 hour endurance.	G3 cloud/aerosol in-situ work and G4 precipitation/convection campaigns with heavy payloads.
NSF / University of Wyoming King Air 350i	About 3-5 hour endurance, 2,100 nmi range, 33,000-35,000 ft ceiling and roughly 2,970 lb maximum payload.	Agile G3 low/mid-level cloud, aerosol and turbulence sampling; radar/lidar missions.
NOAA WP-3D Orion	Severe-weather and atmospheric-research platform with tail/lower-fuselage radar, dropsondes and long missions.	G4 deep convection, precipitation, storm organization, aerosol/gas and ocean-atmosphere campaigns.
NOAA Gulfstream IV-SP	NOAA describes 45,000 ft cruise altitude and about 4,000 nmi range, with dropsonde and radar missions.	Wide-area upper-air context, atmospheric rivers, storm environment and large-scale control-region mapping.
SAFIRE ATR42 (France)	Medium-troposphere atmospheric laboratory for chemistry, aerosol/cloud microphysics, remote sensing and turbulence; up to about 9 scientists and 2 tonnes of instruments.	International G3/G4 campaigns and independent replication.

These are examples of operators/user facilities rather than preselected vendors. Government and university aircraft are commonly accessed by facility proposals, cooperative or reimbursable agreements and campaign-specific task orders. The program should contract for scientific capability - aircraft, integration, airworthiness, crew, instrument teams, calibration and data - rather than treat the airplane as an ordinary charter.

How aircraft map to the GCM research areas

GCM area	Primary aircraft need	Secondary / complementary aircraft
Area 1 - Earth radiation budget	ER-2 or GV-class high-altitude radiation/remote sensing and satellite underflights.	King Air/C-130 or surface/balloon platforms for below-cloud and scene

GCM area	Primary aircraft need	Secondary / complementary aircraft
		closure.
Area 2 - Anthropogenic aerosol forcing	GV/C-130/King Air-class in-situ aerosol, CCN, cloud and radiation sampling.	High-altitude remote-sensing aircraft for above-cloud aerosol and radiation.
Area 3 - Marine low clouds	C-130 or King Air primary in-cloud aircraft.	GV/King Air remote-sensing aircraft above cloud plus vessel/autonomous platforms.
Area 4 - Mixed-phase/polar	C-130, King Air, GV or comparable aircraft selected for icing, range and campaign altitude.	High-altitude remote sensing/satellite underflight where useful.
Area 5 - Deep convection	C-130/WP-3D-class storm and precipitation aircraft plus remote sensing.	GV/ER-2-class aircraft for anvil, upper-troposphere and radiative measurements.
Area 7 - Rapid response	WB-57/ER-2/GV/C-130 chosen to match event altitude and payload.	Balloon, surface, satellite and autonomous assets with the same data system.
Area 10 - Observing backbone	Maintains modular interfaces, calibration and payload readiness rather than owning every aircraft.	Framework agreements and annual cross-platform intercomparisons.

Airborne instruments and suppliers

Many core instruments already exist commercially or within research facilities. Droplet Measurement Technologies supplies airborne CCN and cloud/aerosol probes such as CCN-100/200, CDP, PCASP and CAPS-family systems. Aerodyne supplies AMS systems that provide real-time non-refractory submicron aerosol composition and can be deployed on aircraft, while TSI supplies CPC and mobility-sizing families for particle number and size. Radar, lidar, specialized trace-gas, polarimetric and high-altitude instruments are often facility- or laboratory-built by NCAR, NOAA, NASA, DOE ARM and university teams. The procurement unit should therefore buy integrated measurement capability and instrument-team support, not merely boxes of hardware.

What should be developed rather than a new airframe?

- Standard mechanical racks, external mounts, probe interfaces, optical-port documentation, power, cooling, Ethernet/data, timing, GPS/INS and telemetry that can move between participating aircraft.
- Characterized aerosol and trace-gas inlets with size-dependent transmission, humidity, heating/anti-icing and contamination documentation.
- A common radiometry and optical-metrology system with pre/post-flight calibration, attitude correction and cross-aircraft intercomparison.

- Reusable cloud/aerosol/chemistry payloads configured for low cloud, mixed phase, high-altitude aerosol and rapid-response missions.

- A shared mission computer and data-fusion system that merges aircraft, radar/lidar, ship, surface, balloon and satellite measurements in near real time.

- A high-value GCM-specific development program for rapid 3-D cloud and precipitation remote sensing. An electronically scanned dual-polarization airborne phased-array radar could materially improve deep-convection and storm-organization measurements if cost, weight, power and airworthiness are acceptable.

A new general-purpose GCM airframe should not be a first-line investment. Existing fleets cover most altitude/payload combinations. A purpose-built aircraft should be considered only after an observing-system analysis demonstrates a persistent unmet combination of altitude, payload, endurance, severe-weather/icing capability or operating cost and shows that the information gain justifies lifecycle ownership.

Rapid-response readiness is an aircraft problem as well as an instrument problem

Volcanic eruptions, pyroconvective smoke and unusual pollution events may be scientifically valuable for only days or weeks. The GCM fund should maintain framework agreements, pre-engineered payload configurations, instrument readiness exercises, export/customs plans, spares and standing scientific triage. A nominally available aircraft is not a rapid-response capability if payload integration takes six months.

Current-status correction for international aircraft

The proposal should not list the former UK FAAM BAe-146 as a current aircraft source. FAAM states that the Airborne Laboratory formally closed and all operational activity ended on March 31, 2026. International aircraft should be listed generically or by verified current facilities such as SAFIRE, with availability checked immediately before procurement.

Recommended program structure

Area 10 should function as a Climate Moonshot Airborne Science Facility: maintain the shared payload standards, instrument pools, calibration, mission software, integration drawings, airworthiness records and aircraft-access agreements. Flagship Areas 1-7 then pay their campaign-specific aircraft hours, field integration and unique instruments. This avoids buying duplicate radar, lidar, radiometers and aerosol systems for each scientific topic while preserving independent operators and competing instrument teams.

The main conclusion is that the GCM program needs a fleet architecture, not one airplane. High-altitude remote sensing, high-altitude in-situ chemistry, cloud penetration, storm radar and rapid response are distinct jobs; common interfaces and modular instruments are the technology that makes the fleet act like one observing system.

SAI / MCB / GCM Aircraft and Instrument Overlap

Most monitoring aircraft and instruments should be shared across the three atmospheric initiatives. The programs differ mainly in altitude, release-source interfaces, and payload emphasis. Common infrastructure lowers cost and makes measurements directly comparable.

Shared aircraft role	GCM	SAI	MCB
High-altitude remote sensing / radiation	Satellite underflights; radiation closure; high-cloud context	Stratospheric plume layer height, optical depth, radiation	Above-cloud MCB context and radiative closure
High-altitude in situ	Volcano/smoke, UTLS aerosol/chemistry, anvils	Core stratospheric aerosol/chemistry sampling	Occasional free-troposphere aerosol / cloud context
Boundary-layer / cloud aircraft	Aerosol-cloud closure, low clouds, mixed phase	Support aircraft for plume descent / background where useful	Core cloud-base, in-cloud, CCN, drizzle, turbulence sampling
Remote-sensing context aircraft	Cloud geometry, lidar/radar, polar/convection context	Plume mapping and independent context	Cloud-field mapping, plume-cloud geometry
Rapid-response aircraft access	Natural perturbations and rare events	Volcanic/stratospheric opportunities	Shipping, smoke, cloud-regime opportunities

Shared instrument modules

Module	Core measurements	Use / boundary
Radiation / polarimetry	Broadband SW/LW, spectral radiance, polarization	All three; GCM calibration backbone
Lidar / remote profiling	Layer height, extinction/backscatter, cloud structure	All three; altitude and eye-safety requirements vary
Aerosol microphysics / optics	Size, number, composition, scattering, absorption, hygroscopicity	All three; SAI adds stratospheric chemistry emphasis
Cloud microphysics / CCN	CCN, droplet/ice spectra, LWC/IWC, drizzle	GCM and MCB core; SAI mostly supporting context
Meteorology / turbulence / timing	T/RH, winds, turbulence, GPS/INS, synchronized clocks	Common to all campaigns and essential for attribution
Data / mission systems	Common formats, quick-look, calibration history, provenance	Joint infrastructure; campaign-specific analysis remains separate

Avoiding duplicate spending

- Maintain one portfolio asset register and one calibration history for shared payloads.
- Charge each initiative only for its incremental integration, unique instruments, operating time, consumables, and release-specific hardware.
- A new general-purpose aircraft should not be a first-line investment; use existing research-aircraft operators unless a documented capability gap survives the shared-asset review.

Existing Infrastructure to Leverage

The following research programs are potential partners for collaboration.

DOE Atmospheric Radiation Measurement (ARM) user facility
https://www.arm.gov/

Coupled Model Intercomparison Project (CMIP)
https://wcrp-cmip.org/

Cloud Feedback Model Intercomparison Project (CFMIP)
https://cfmip.github.io/

Geoengineering Model Intercomparison Project (GeoMIP)
https://climate.envsci.rutgers.edu/GeoMIP/

Argo global ocean observing program
https://argo.ucsd.edu/

OSNAP overturning observations
https://www.o-snap.org/

RAPID AMOC observing project
https://rapid.ac.uk/

Important caution

This document is a research-program concept and uses order-of-magnitude planning estimates. It is not an engineering design, deployment authorization, environmental assessment, legal opinion or substitute for independent scientific review. Any controlled atmospheric perturbation would require its own scientific justification, governance process, environmental analysis, permits and public consultation.