Physics-Informed WAIS Uncertainty Framework: Correcting Systematic Biases in Antarctic Ice Sheet Projections for Sea-Level Rise Forecasting

Supplementary Material

February 16, 2026

Abstract

The IPCC AR6 Antarctic Ice Sheet (AIS) projections likely understate the true range of West Antarctic Ice Sheet (WAIS) uncertainty due to three identified systematic biases: (1) the near-universal use of Glen's flow law with stress exponent $n = 3$ despite evidence supporting $n = 4$; (2) the omission of irreducible stochastic amplification from marine ice-sheet instability (MISI); and (3) the absence of key physical processes from most contributing models. We develop four complementary approaches to construct a physics-informed σ_{ice} for use in hierarchical sea-level rise forecasting. The approaches converge on $\sigma_{\text{ice}} \approx$ 320–491 mm at 2100, compared to ∼150 mm (medium confidence) and ∼290 mm (low confidence) from IPCC AR6. The approach combining restructured scenario weighting, quantile-dependent rheology correction, and stochastic amplification during marine ice-sheet instability—yields $\sigma_{\text{ice}} \approx$ 491 mm from WAIS alone. This uncertainty is largely decoupled from atmospheric CO2 concentrations, meaning these contributions uncertainties add to the largely "direct" thermodynamic SLR contributions and uncertainties from thermosteric effects, Greenland, and glaciers not in the ice sheets. This should provide a pretty clean separation between the predictable (everything but WAIS) and (currently) unpredictable (WAIS) contributions **and** a separation in quantified uncertainties between those uncertainties that are dominated by societal choices (future emissions and other 'geoengineering' schemes) – which is everything but WAIS – and those uncertainties that are dominated by internal processes and feedbacks (WAIS). This is one way for us to quantify and articulate the uniqueness and value of our mission in the broader set of climate mitigation, adaptation, and intervention.

1 Motivation

Our hierarchical SLR forecasting framework decomposes total projection uncertainty into three additive components:

$$\sigma^2_{\text{total}}(t) = \sigma^2_{\text{constrained}}(t) + \sigma^2_{\text{scenario}}(t) + \sigma^2_{\text{ice}}(t) \tag{1}$$

where $\sigma_{\text{constrained}}$ represents calibrated DOLS parameter uncertainty, σ_{scenario} represents scenario (SSP) spread, and σ_{ice} represents deep uncertainty from Antarctic ice sheet dynamics.

In the original decomposition using IPCC AR6 data, σ_{ice} is estimated directly from the AIS component 5–95% range in the confidence-level NetCDF files. Even at low confidence (which includes structured expert judgment on poorly constrained processes), σ_{ice} accounts for only ∼7% of total variance at 2100—a result that appears inconsistent with the depth of uncertainty articulated in the recent ice-sheet literature.

Three lines of evidence suggest the IPCC AIS projections systematically understate uncertainty:

1. **Rheological bias**: Nearly all contributing ice-sheet models use $n = 3$ in Glen's flow law, despite laboratory and field evidence favoring $n = 4$ for Antarctic conditions. This is a one-directional systematic bias that grows nonlinearly with forcing magnitude.

2. **Stochastic amplification**: Marine ice-sheet instability exponentially amplifies ensemble spread, generating irreducible uncertainty from internal climate variability that is not captured in deterministic model intercomparisons.

3. **Missing processes**: Grounding-zone tidal pumping, subglacial hydrology coupling, cold-to-warm ocean cavity transitions, and marine ice-cliff instability are absent from most contributing models.

We address these through four complementary approaches, each grounded in specific published results.

2 Approach 1: Rheology Correction ($n = 3 \to n = 4$)

2.1 Physical basis

The standard Glen–Nye flow law relates deviatoric stress τ to strain rate $\dot{\varepsilon}$:

$$\dot{\varepsilon} = A\,\tau^n \tag{2}$$

where A is a temperature-dependent rate factor and n is the stress exponent. The value $n = 3$ has been used since Glen (1955), but multiple lines of evidence support $n = 4$ for Antarctic conditions:

- Laboratory experiments at Antarctic temperatures and stress levels (Goldsby and Kohlstedt 2001)
- Remotely sensed observations of Antarctic ice shelves yield $n = 4.1 \pm 0.4$ in fast-flowing extensional regimes (Millstein et al. 2022)
- The inversion procedure using $n = 3$ compensates through a viscosity multiplier ϕ that masks the rheological error

2.2 Quantitative impact

Two recent studies provide direct estimates of the $n = 3 \to n = 4$ bias:

- **Martin et al. (in review)**: Using coupled Amundsen Sea Embayment (ASE) simulations, $n = 3$ underestimates SLR contribution by 21% under moderate ocean forcing and 35% under extreme forcing. The bias grows nonlinearly because the inversion-calibrated viscosity multiplier becomes increasingly inadequate as forcing departs from calibration conditions.
- **Getraer and Morlighem (2025)**: Pan-Antarctic simulations show $32 \pm 14\%$ more ice loss by 2100 with $n = 4$ vs. $n = 3$, growing to $\sim 70 \pm 15\%$ by 2300. Crucially, by 2200–2250, uncertainty from n exceeds uncertainty from the climate forcing scenario.

2.3 Implementation

We apply quantile-dependent correction factors $f(q,t)$ to the IPCC AIS distribution:

$$Q^{\text{corrected}}(q,t) = f(q,t) \times Q^{\text{IPCC}}(q,t) \tag{3}$$

The correction factor is parameterized as:

$$f(q,t) = 1 + \Delta f(q) \cdot g(t) \tag{4}$$

where:

$$\Delta f(q) = 0.05 + 0.35\,q \qquad \text{(quantile-dependent amplitude)} \tag{5}$$

$$g(t) = \left(\frac{t - t_{\text{ref}}}{2100 - t_{\text{ref}}}\right)^{\gamma} \qquad \text{(temporal ramp, } \gamma = 1.5\text{)} \tag{6}$$

This gives:

$$f(Q_{5\%}, 2100) \approx 1.07 \qquad \text{(minimal correction at low end)} \tag{7}$$

$$f(Q_{50\%}, 2100) \approx 1.22 \qquad \text{(Martin et al. moderate forcing)} \tag{8}$$

$$f(Q_{95\%}, 2100) \approx 1.38 \qquad \text{(Getraer \& Morlighem: } 32 \pm 14\%\text{)} \tag{9}$$

The correction is asymmetric by design: the $n = 3$ bias is larger for faster retreat (upper quantiles) because stronger forcing amplifies the difference between $n = 3$ and $n = 4$ flow dynamics.

The temporal ramp uses $\gamma = 1.5 > 1$ to reflect the nonlinear growth of the bias with cumulative ice loss. Near present (2020), the correction is negligible because ice flow is still close to the calibration state. By 2100, the full correction applies.

For years beyond 2100, the correction continues to grow (following the $\sim 70\%$ by 2300 result from Getraer and Morlighem 2025), reaching $f \approx 1.8 \times \Delta f$ by 2300.

3 Approach 2: Stochastic Amplification

3.1 Physical basis

Robel et al. (2019) demonstrated that marine ice-sheet instability (MISI) does not merely increase the mean SLR projection—it fundamentally changes the character of the uncertainty distribution through three mechanisms:

1. **Exponential amplification**: An ensemble that begins with modest spread (from parameter or initial-condition uncertainty) develops exponentially growing spread once MISI is triggered, because the instability amplifies initial differences.

2. **Positive skew**: The distribution of SLR contributions becomes positively skewed (heavy right tail in SLR space), allocating more probability mass to large contributions than a Gaussian would predict.

3. **Irreducible floor**: Even with perfectly known mean forcing, internal climate variability on multidecadal timescales generates substantial irreducible uncertainty. For Thwaites Glacier alone: $\sim$20 cm uncertainty during fast retreat (conservative, known mean forcing), $\sim$40 cm with multidecadal variability, and $\sim$60 cm when the mean forcing itself is uncertain.

3.2 Implementation

We model the stochastic amplification as an additive, independent uncertainty component:

$$\sigma_{\text{ice}}^{2,\text{total}}(t) = \sigma_{\text{ice}}^{2,\text{A1}}(t) + \sigma_{\text{internal,WAIS}}^{2}(t) \tag{10}$$

where $\sigma_{\text{ice}}^{2,\text{A1}}$ is the rheology-corrected IPCC uncertainty from Approach 1, and $\sigma_{\text{internal,WAIS}}$ represents the irreducible stochastic component.

The temporal evolution of σ_{internal} follows an exponential ramp-up after an estimated MISI onset:

$$\sigma_{\text{internal}}(t) = \sigma_{\max}\left(1 - e^{-(t-t_{\text{onset}})/\tau_{\text{amp}}}\right) \quad \text{for } t > t_{\text{onset}} \tag{11}$$

Table 1: Parameter sets for stochastic amplification model

Estimate	t_{onset} (yr)	τ_{amp} (yr)	$\sigma_{\max}$ (mm)	Skewness α	$\sigma(2100)$ (mm)
Conservative	2040	50	150	2.0	$\sim$105
Central	2030	40	200	3.0	$\sim$165
Aggressive	2025	30	250	4.0	$\sim$229

The skewness is represented through a skew-normal distribution with shape parameter α, normalized to unit variance. This ensures that the upper-tail quantiles are larger than the lower-tail quantiles for a given σ.

Parameter choices are motivated by:

- t_{onset}: The WAIS temperature threshold is estimated at > 1.5°C above pre-industrial, which could be reached by 2025–2040 depending on the scenario (Fricker et al. 2025).
- τ_{amp}: The e-folding time for MISI amplification is $\sim$30–50 years based on the marine ice-sheet instability timescale for Thwaites Glacier (Robel et al. 2019).
- $\sigma_{\max}$: Scaled up from the Thwaites-only results ($\sim$20–40 cm) to full WAIS ($\times$1.5–2, accounting for Pine Island Glacier and other basins), giving 150–250 mm.

The addition in quadrature (eq. (10)) is justified because the stochastic component arises from internal climate variability, which is independent of both the model structural uncertainty captured in the IPCC ensemble and the systematic rheological bias corrected in Approach 1.

4 Approach 3: Process-Informed Scenario Weighting

4.1 Rationale

Approaches 1 and 2 adjust the IPCC distribution parametrically. Approach 3 takes a fundamentally different path: we construct a **discrete mixture model** that explicitly represents physically distinct WAIS dynamical regimes. This approach:

- Makes the assessed probabilities of different outcomes transparent and debatable
- Naturally produces a multimodal, heavy-tailed distribution
- Can incorporate processes (MICI, GZ amplifiers) not captured in Approaches 1–2
- Serves as an independent cross-check on the parametric approaches

4.2 Scenario definitions

Table 2: Process-informed WAIS scenario definitions at 2100

Scenario	Description	P	SLR range (mm)	Basis
S1: Status quo	Current acceleration continues	0.10	30–80	IMBIE trend
S2: MISI (n=3)	GL retreat, standard rheology	0.20	150–400	IPCC low conf.
S3: MISI (n=4)	GL retreat, corrected rheology	0.35	200–550	S2 × Approach 1
S4: MISI + amplifiers	+ GZ tidal pumping, subglacial hydro.	0.25	400–1000	Fricker et al.
S5: MISI + MICI	+ marine ice-cliff instability	0.10	600–2000+	DeConto & Pollard

Each scenario's SLR range defines the 5th–95th percentiles of a log-normal distribution (ensuring positive support and right skew). The mixture distribution is:

$$p(\mathrm{SLR}) = \sum_{k=1}^{5} w_k \cdot p_k(\mathrm{SLR}) \tag{12}$$

4.3 Probability assessment

The scenario weights are assessed as follows:

- **S1** (P = 0.10): The "no instability" scenario. WAIS continues losing mass at accelerating rates consistent with current observations, but without triggering widespread MISI. This receives low weight because observational evidence for ongoing grounding-line retreat at Thwaites and Pine Island Glacier, intrusion of warm Circumpolar Deep Water onto the continental shelf, and theoretical understanding of MISI all favor instability being underway or imminent. The weight is nonzero to hedge against the possibility that the deceleration observed in IMBIE data around 2010 reflects a real dynamical pause rather than variability.

- **S2** (P = 0.20): MISI is triggered for Thwaites and/or Pine Island Glacier, but with the standard $n = 3$ rheology used in IPCC contributing models. This is consistent with the median IPCC low-confidence AIS projection. It receives lower weight than S3 because the evidence for $n = 3$ is weak: it originates from early laboratory experiments at stress and temperature conditions unrepresentative of Antarctic ice streams, and the inversion procedure compensates for the rheological error through a viscosity multiplier ϕ that masks the bias (Martin et al. in review).

- **S3** (P = 0.35): Same dynamical scenario as S2 but with corrected rheology ($n = 4$). This is the highest-weighted scenario because it combines the most likely dynamical regime (MISI triggered) with the best-supported rheology. Laboratory experiments at Antarctic-relevant conditions (Goldsby and Kohlstedt 2001), remotely sensed inversions of Antarctic ice shelves (Millstein et al. 2022), and two independent modeling studies (Martin et al. in review, Getraer and Morlighem 2025) all support $n \geq 4$.

- **S4** (P = 0.25): MISI with additional process amplifiers identified by Fricker et al. (2025): grounding-zone tidal pumping (which can double melt rates), subglacial hydrology coupling (∼30% amplification), and cold-to-warm ocean cavity transitions. This receives substantial weight because each of these processes is physically documented and observed, even though their combined quantitative impact remains poorly constrained. The fact that *none* of these processes are represented in the models contributing to IPCC AR6 means the current projections systematically omit known amplifying feedbacks.

- **S5 ($P = 0.10$)**: The extreme scenario including marine ice-cliff instability (MICI), which could produce rapid, self-sustaining retreat of tall ice cliffs exposed after ice-shelf loss. While MICI remains uncertain (Edwards et al. 2019), it cannot be ruled out. We assign 10% weight—double the previous 5%—because recent observations of accelerating ice-shelf thinning and collapse (Conger, Larsen B) demonstrate that the preconditions for MICI (exposure of tall ice cliffs) are plausible within the 21st century.

4.4 Temporal scaling

Each scenario's temporal evolution is scaled from the 2100 value using a physically motivated growth curve:

- S1 (status quo): $g(t) = \Delta t^2$ (quadratic accumulation, consistent with constant acceleration)
- S2–S3 (MISI): Delayed onset (∼2030) followed by exponential ramp-up, $g(t) = (1 - e^{-3\Delta t_{\mathrm{eff}}}) \cdot \Delta t_{\mathrm{eff}}$
- S4–S5 (MISI + amplifiers/MICI): Earlier onset (∼2025), faster ramp-up

5 Approach 4: Rheology-Corrected Scenario Mixture + Stochastic Amplification

Approach 4 is the most complete framework, combining all three preceding approaches. The key insight motivating this combination is that the $n = 3 \to n = 4$ rheology correction (Approach 1) affects *all* ice flow dynamics—not just MISI onset—and should therefore be applied to every scenario, not only parameterized as a separate scenario branch.

5.1 Restructured scenarios

We merge the Approach 3 scenarios S2 (MISI, $n = 3$) and S3 (MISI, $n = 4$) into a single MISI scenario, since the $n = 3/n = 4$ distinction is now handled by the systematic rheology correction rather than by separate scenarios with heuristically adjusted ranges. This yields four scenarios:

Table 3: Restructured scenario definitions for Approach 4 (SLR ranges are $n = 3$ baselines)

Scenario	Description	P	SLR range (mm, n=3)	Basis
S1: Status quo	Current acceleration continues	0.10	30–80	IMBIE tre
S2: MISI	GL retreat, base dynamics	0.55	150–400	IPCC low co
S3: MISI + amplifiers	+ GZ tidal pumping, subglacial hydro.	0.25	400–1000	Fricker et a
S4: MISI + MICI	+ marine ice-cliff instability	0.10	600–2000+	DeConto & Po

The merged S2 weight ($P = 0.55$) equals the sum of the old S2 and S3 weights ($0.20 + 0.35$), reflecting that MISI (regardless of the assumed rheology) is the most likely dynamical regime. The SLR ranges are $n = 3$ *baselines*—the rheology correction inflates them.

5.2 Combined pipeline

For each Monte Carlo sample k in scenario s at year t:

$$X_{k,s}^{\mathrm{A4}}(t) = f(q_k, t) \times X_{k,s}^{\mathrm{base}}(t) + \epsilon_k(t) \tag{13}$$

where:

- $X_{k,s}^{\mathrm{base}}(t)$ is drawn from the scenario's log-normal distribution and temporally scaled (§4)
- $f(q_k, t)$ is the Approach 1 rheology correction (eq. (3)), applied using the empirical quantile q_k of sample k within its scenario batch (Hazen plotting position). This correction is applied to *every* sample in *every* scenario, including S1
- $\epsilon_k(t) \sim \mathrm{SkewNormal}(0, \sigma_{\mathrm{internal}}(t), \alpha)$ is the Approach 2 stochastic amplification, applied only to MISI scenarios ($s \in \{S2, S3, S4\}$). S1 has no MISI, so $\epsilon_k = 0$ for status-quo samples

This is the most complete approach, incorporating: (1) the discrete scenario structure with assessed probabilities, (2) the non-Gaussian shape from log-normal within-scenario distributions, (3) the quantile-dependent $n = 3 \to n = 4$ rheology correction applied to all ice dynamics, and (4) the irreducible stochastic amplification during MISI. Approach 4 serves as our recommended estimate of the full physics-informed σ_{ice}.

6 Results and Comparison

6.1 σ_{ice} at 2100

All four approaches independently suggest that the true AIS uncertainty is substantially larger than the IPCC estimates:

Table 4: Comparison of σ_{ice} at 2100 (mm) across approaches

Source	σ_{ice} (mm)	Ratio to IPCC Med.
IPCC Medium Confidence	150	1.0×
IPCC Low Confidence	290	1.9×
A1: Rheology Correction	393	2.6×
A2: A1 + Stochastic Amplification	426	2.8×
A3: Scenario Mixture	320	2.1×
A4: Rheology + Scenarios + Stochastic	491	3.3×

Approaches 2–4 yield $\sigma_{\mathrm{ice}} \approx$ 320–491 mm from complementary methodologies, spanning a range ∼1.1–1.7× the IPCC low-confidence estimate and ∼2.1–3.3× the medium-confidence estimate. The parametric approach (A2) and the hybrid approach (A4) both include the rheology correction, while the structural approach (A3) does not. Approach 4 ($\sigma_{\mathrm{ice}} = 491$ mm) is the most complete, incorporating assessed scenario probabilities, non-Gaussian within-scenario distributions, the quantile-dependent rheology correction applied to all ice dynamics, and irreducible stochastic amplification during MISI.

6.2 Impact on variance decomposition

With physics-informed σ_{ice}, the ice-sheet fraction of total variance at 2100 increases substantially:

- **IPCC Medium Confidence**: $f_{\text{ice}} \approx 2\%$ at 2100
- **IPCC Low Confidence**: $f_{\text{ice}} \approx 7\%$ at 2100
- **A2 (Rheology + Stochastic)**: $f_{\text{ice}} \approx 14\%$ at 2100
- **A3 (Scenario Mixture)**: $f_{\text{ice}} \approx 8\%$ at 2100
- **A4 (Recommended)**: $f_{\text{ice}} \approx 18\%$ at 2100

While scenario uncertainty (σ_{scenario}) still dominates at 2100 due to the large spread in temperature pathways, the physics-informed corrections reveal that ice-sheet uncertainty is a substantially more important contributor than the raw IPCC numbers suggest—increasing f_{ice} from 2% (medium confidence) to 8–18% depending on the approach—and becomes dominant at earlier time horizons (∼2040–2060).

6.3 Distribution shape

The Approach 3 mixture distribution at 2100 exhibits positive skewness (∼2.0) and heavy tails (excess kurtosis ∼6.2), reflecting the substantial combined weight (35%) on high-impact scenarios S4 and S5. Its 95th percentile (∼1048 mm) substantially exceeds the IPCC low-confidence 95th percentile, and its median (∼342 mm) exceeds the IPCC low-confidence median.

Approach 4 substantially increases the spread ($\sigma_{\text{ice}} = 491$ vs. 320 mm) through the combined effect of the rheology correction and stochastic amplification, and shifts the median upward (∼574 mm). The skewness (∼1.75) and kurtosis (∼5.7) remain high, reflecting both the discrete scenario structure and the asymmetric rheology correction (which inflates the upper tail more than the lower tail). The 95th percentile increases to ∼1589 mm, and the 5th percentile shifts modestly from 49 to 59 mm because even the status-quo scenario receives the rheology correction.

This non-Gaussian shape has practical implications: a Gaussian approximation to σ_{ice} underestimates the probability of outcomes in the upper tail.

7 Emergent Thermodynamic Sensitivity: DOLS vs. IPCC Process Models

A key assumption of the hierarchical framework (eq. (1)) is that the DOLS rate–temperature relationship, calibrated on ∼120 years of observations, captures the thermodynamic response (steric expansion + glaciers + Greenland) that governs $\sigma_{\text{constrained}}$. If IPCC AR6 process models produce a fundamentally different thermodynamic sensitivity, either the DOLS calibration or the process models contain systematic error. We test this by fitting the rate–temperature relationship directly to the IPCC AR6 component projections.

7.1 Method

For each SSP (1-1.9 through 5-8.5), we construct the *thermodynamic* sea level signal from IPCC AR6 FACTS output as the sum of ocean dynamics (thermosteric), glacier, and Greenland components—the IPCC analogue of the observational DOLS target, which uses GMSL minus terrestrial water storage. Antarctica is excluded because it is not a significant thermodynamic contributor within the calibration record and is dominated by marine ice-sheet dynamics (treated separately as σ_{ice}).

We compute finite-difference rates from the decadal IPCC projections (2020–2100; 8 rate estimates from 9 time points) and fit linear and quadratic models to the rate–temperature relationship using `np.polyfit`, comparing them via AIC, BIC, F-test, and the p-value on the quadratic coefficient. DOLS (`calibrate_dols`, $n_{\text{lags}} = 0$) is run as a consistency check.

7.2 Results

All five SSPs unanimously prefer a linear rate–temperature relationship (4/4 model selection criteria for SSP1-2.6 through SSP5-8.5; 2/4 for SSP1-1.9 where the temperature range is too narrow for discrimination). The linear sensitivity α_0 from the higher-forcing SSPs is $\sim$1.9–2.5 mm/yr/°C, compared with the observational DOLS value of $\sim$5.3 mm/yr/°C—lower by roughly a factor of two. No SSP produces a statistically significant quadratic term ($p > 0.2$ for all).

A statistical power analysis confirms that the absence of quadratic detection is not merely a sample-size artifact for the higher-forcing SSPs. For SSP2-4.5, SSP3-7.0, and SSP5-8.5, the standard error on the quadratic coefficient is $\sim$0.1–0.9 mm/yr/°C^2, and the minimum detectable effect at 80% power is $\sim$0.4–2.5 mm/yr/°C^2—well below the observational value of $\sim$5.1 mm/yr/°C^2. The observational quadratic falls *outside* the IPCC 95% confidence intervals for these SSPs. The IPCC thermodynamic component is genuinely more linear than what observations suggest.

7.3 Consistency with the literature

This discrepancy is consistent with a growing body of literature documenting that IPCC process-based projections underestimate the observed thermodynamic sensitivity:

- Grinsted and Christensen (2021) defined the "Transient Sea Level Sensitivity" (TSLS) and showed that IPCC projections fall below what the observational record would predict, proposing the TSLS as an emergent constraint.
- Jevrejeva et al. (2021) found that the CMIP6 ensemble-mean thermosteric rate over 1940–2005 ($\sim$0.2 mm/yr) is less than half the observed rate ($\sim$0.5 mm/yr)—a systematic offset of $\sim$10 cm per century per degree of warming.
- Grinsted et al. (2022) extended the TSLS framework to individual components in CMIP6, finding that glacier sensitivity is strongly time-dependent ($\sim$2.8 mm/yr/K before 2050 vs. $\sim$0.7 mm/yr/K after) and that total GMSL projections may underestimate the observational relationship.
- Rahmstorf et al. (2012) compared IPCC projections to satellite-era observations and found that observed sea level rise exceeded the IPCC best estimate by $\sim$60%, despite temperature projections tracking observations well.
- Törnqvist et al. (2025) showed that the accurate net IPCC projection over 30 years was "somewhat fortuitous," arising from compensating errors: overestimated thermal expansion offset by underestimated ice-sheet contributions.

A physical mechanism for this discrepancy has been proposed by Kuhlbrodt and Gregory (2012): CMIP-class models may have ocean stratification that is too weak, causing them to absorb heat too efficiently into the deep ocean rather than manifesting it as near-surface thermal expansion. This would suppress the transient thermosteric sensitivity while preserving the equilibrium sensitivity—consistent with our finding that IPCC process models produce a lower α_0 than observations over centennial timescales.

7.4 Implications for the framework

These findings have two implications for the hierarchical framework. First, they validate the DOLS approach: by calibrating directly on observations, the DOLS model captures the realized thermodynamic sensitivity rather than the model-dependent sensitivity embedded in IPCC process-model output. The systematic underestimation by IPCC process models means that using their thermodynamic projections directly would understate $\sigma_{\text{constrained}}$.

Second, the discrepancy reinforces the importance of separating the ice-sheet term. The IPCC's accurate *total* GMSL projection (despite component-level biases) arises partly from compensating errors between thermodynamic and cryospheric components (Törnqvist et al. 2025). The hierarchical framework avoids this by calibrating the thermodynamic response independently (via DOLS) and treating ice-sheet uncertainty separately (via σ_{ice}).

Multi-centennial projections from Turner et al. (2023), which extend the IPCC AR6 component-resolved framework to 2500 under SSP1-2.6 and SSP2-4.5, could provide the longer thermal lever arm needed to test whether the observational quadratic sensitivity eventually emerges in process models over multi-century timescales where the temperature signal is much larger.

8 Multi-Dataset Robustness of the DOLS Calibration

A potential criticism of the DOLS framework is that the calibrated sensitivities α_0 and $d\alpha/dT$ depend on the particular choice of GMSL reconstruction and GMST product. To address this, we run DOLS on every combination of available GMSL and GMST datasets, producing a coefficient matrix that quantifies the sensitivity of the results to observational data choice.

8.1 Datasets

Seven GMSL records are used:

1. **Frederikse (total)**: 1900–2018, full budget reconstruction with component decomposition (Frederikse et al. 2020);
2. **Frederikse thermodynamic**: GMSL minus terrestrial water storage (TWS), isolating the temperature-driven signal;
3. **Dangendorf (total)**: 1900–2021, an independent tide-gauge-based reconstruction (Dangendorf et al. 2019);
4. **Dangendorf sterodynamic**: the ocean thermal expansion plus dynamic sea-level component only—not equivalent to GMSL$-$TWS since it excludes all land-ice contributions;
5. **IPCC observed (total)**: 1950–2020, the AR6 observational composite;
6. **IPCC observed thermodynamic**: IPCC GMSL minus Frederikse TWS (interpolated to IPCC years over the 1950–2018 overlap);

7. **Horwath**: 1993–2016 monthly budget (Horwath et al. 2022), annualized—excluded from ensemble statistics due to its short record (24 yr produces unstable quadratic estimates with $n_{\text{lags}} = 2$).

Four GMST products are used: Berkeley Earth (Rohde and Hausfather 2020), GISTEMP, HadCRUT5, and NOAA GlobalTemp. All are annualized over their common period starting at 1950 (the beginning of the IPCC observed GMSL record).

8.2 Results

The $6 \times 4 = 24$ valid DOLS fits (excluding Horwath) are summarized in Table 5. Two key patterns emerge:

1. **GMST choice has modest impact**: For a given GMSL dataset, α_0 varies by $\lesssim$1 mm/yr/°C across GMST products, and $d\alpha/dT$ by $\lesssim$1 mm/yr/°C^2. The four GMST products are sufficiently similar over the post-1950 period that they do not materially affect the DOLS calibration.

2. **GMSL choice dominates the spread**: The largest differences arise between total GMSL and thermodynamic (GMSL−TWS) variants. Total GMSL datasets yield higher α_0 (∼2–4 mm/yr/°C) because they include ice-sheet and TWS contributions that add to the background acceleration. Thermodynamic variants yield lower α_0 (∼0–1.3 mm/yr/°C) but higher $d\alpha/dT$ (∼2–3.3 mm/yr/°C^2), consistent with the quadratic acceleration arising primarily from the temperature-sensitive components.

Dangendorf sterodynamic is a notable outlier: it shows $d\alpha/dT \approx 0$ across all GMST products. This is expected because the sterodynamic component isolates ocean thermal expansion and dynamics, excluding glacier and ice-sheet mass loss—the components that drive the observed acceleration. This result confirms that the quadratic term in DOLS captures the accelerating cryospheric contribution to sea level, not a nonlinearity in ocean thermal expansion alone.

Table 5: Ensemble DOLS coefficients from the multi-dataset robustness matrix (order=2, $n_{\text{lags}} = 2$, start≥1950). Horwath excluded ($n = 20$, unstable estimates); Dangendorf sterodynamic excluded from thermodynamic ensemble (see text).

Ensemble	α_0 (mm/yr/°C)	$d\alpha/dT$ (mm/yr/°C^2)
Thermodynamic only (8 pairs)	0.49 ± 0.58	2.85 ± 0.38
All datasets (24 pairs)	1.73 ± 1.21	1.83 ± 1.09
By GMSL dataset (mean across 4 GMST products):		
Frederikse (total)	3.33 ± 0.44	1.09 ± 0.60
Frederikse thermo	0.76 ± 0.32	2.65 ± 0.42
Dangendorf (total)	2.78 ± 0.41	1.99 ± 0.57
Dangendorf sterodynamic	1.12 ± 0.10	0.01 ± 0.15
IPCC observed (total)	2.06 ± 0.40	2.22 ± 0.53
IPCC obs thermo	-0.04 ± 0.30	3.05 ± 0.25

8.3 Implications for the framework

The robustness matrix demonstrates that the positive quadratic sensitivity ($d\alpha/dT > 0$) is a robust feature of observational sea-level records—it is present across all GMSL datasets except the Dangendorf sterodynamic component (which by construction excludes the accelerating cryospheric signal). Dangendorf sterodynamic is excluded from the thermodynamic ensemble because it isolates only ocean thermal expansion and dynamics, whereas the other "thermodynamic" variants (GMSL$-$TWS) retain glacier and ice-sheet contributions that respond to temperature. With this exclusion, the thermodynamic ensemble tightens substantially: $d\alpha/dT = 2.85 \pm 0.38$ mm/yr/°C^2 (8 pairs), compared with 1.90 ± 1.38 when Dangendorf sterodynamic was included (12 pairs). This provides a robust multi-dataset estimate that is consistent with, though somewhat lower than, the single-dataset WLS calibration (Frederikse thermo × Berkeley: $d\alpha/dT \approx 5.1$ mm/yr/°C^2).

The spread across datasets also provides a natural estimate of structural uncertainty in the DOLS calibration—an uncertainty component that is not captured by the within-model HAC standard errors. We recommend reporting coefficients as the WLS point estimate (which optimally weights the highest-quality data) with the multi-dataset spread as a supplementary robustness check.

8.4 Sensitivity to start date

The results in Table 5 use a common start year of 1950, matching the beginning of the IPCC observational GMSL record. Several datasets extend substantially earlier: Frederikse to 1900 and Dangendorf to 1900. Extending the calibration window to include the early 20th century reveals a striking sensitivity of the DOLS coefficients to the choice of start date (Table 6).

Table 6: DOLS coefficients (mean $\pm$ 1σ across 4 GMST products) for 1950-start vs. native-start calibrations. IPCC datasets are unchanged (native start is 1950).

	α_0 (mm/yr/°C)		$d\alpha/dT$ (mm/yr/°C^2)	
GMSL dataset	1950-start	Native start	1950-start	Native start
Frederikse (total)	3.33 ± 0.44	0.45 ± 0.17	1.09 ± 0.60	3.27 ± 0.21
Frederikse thermo	0.76 ± 0.32	1.01 ± 0.19	2.65 ± 0.42	0.19 ± 0.21
Dangendorf (total)	2.78 ± 0.41	-0.30 ± 0.14	1.99 ± 0.57	5.18 ± 0.17
Dangendorf sterodynamic	1.12 ± 0.10	1.87 ± 0.08	0.01 ± 0.15	-0.53 ± 0.10
Thermo ensemble	0.70 ± 0.57	0.92 ± 0.80	1.90 ± 1.38	0.90 ± 1.55
All-dataset ensemble	1.73 ± 1.21	0.82 ± 0.93	1.83 ± 1.09	2.23 ± 1.94

Two patterns are noteworthy. First, α_0 and $d\alpha/dT$ trade off against one another: extending the record to 1900 substantially changes the partition between a constant linear sensitivity (α_0) and an accelerating sensitivity ($d\alpha/dT$), while the overall rate–temperature relationship remains well-fit. For Frederikse thermodynamic, $d\alpha/dT$ drops from 2.65 to 0.19 mm/yr/°C^2 when the early century is included, while for Dangendorf total it *rises* from 1.98 to 5.18. The quadratic and linear coefficients are absorbing each other's signal in a manner that depends on which epoch dominates the fit.

Second, the across-dataset spread increases substantially with native start dates: the thermodynamic ensemble $d\alpha/dT$ coefficient of variation rises from 0.73 to 1.72, and the all-dataset ensemble from 0.60 to 0.87. The additional half-century of early 20th-century data—where observational uncertainty is larger and potential non-stationarity in the rate–temperature relationship is strongest—amplifies inter-dataset disagreement.

This epoch dependence has important implications. It suggests that the DOLS rate–temperature relationship may itself be *time-varying*: the effective α_0 and $d\alpha/dT$ evolve as the climate system traverses different regimes of ocean heat uptake, glacier response, and ice-sheet dynamics. If so, a single static quadratic calibrated over the entire record conflates different dynamical epochs, and the specific coefficient values depend on which epoch receives the most statistical weight.

This motivates exploring the sliding-window DOLS analysis described in §10—the "Dynamic" in DOLS—where a kernel-weighted window is moved through the record to estimate $\alpha_0(t)$ and $d\alpha/dT(t)$ as functions of time. Such an analysis would directly reveal whether the temperature sensitivity is stationary or evolving, and would help distinguish genuine nonstationarity from the aliasing of different observational biases across epochs.

8.5 Sliding-window DOLS: time-varying sensitivity

The epoch dependence revealed in Table 6 motivates a direct test: rather than choosing a single start date, we slide a kernel-weighted window through the observational record and estimate $\alpha_0(t)$ and $d\alpha/dT(t)$ at each center year. This "Dynamic" OLS approach uses a tri-cube kernel with bandwidth h (the half-width of the compact support), producing kernel-weighted WLS fits at each center time t_0:

$$w_i = K\left(\frac{t_i - t_0}{h}\right) \times \frac{1}{\sigma_i^2} \tag{14}$$

where K is the tri-cube kernel and σ_i is the GMSL measurement uncertainty (WLS weights).

8.5.1 Coefficient evolution

Running multi-bandwidth sliding-window DOLS (h = 30, 40, 50, 60 yr) on Frederikse thermodynamic × Berkeley Earth reveals a smooth evolution of both coefficients. At the narrowest bandwidth ($h = 30$ yr), $d\alpha/dT$ increases from ∼0–1 mm/yr/°C^2 for windows centered on the early-to-mid 20th century to ∼3–5 mm/yr/°C^2 for late-century windows. Simultaneously, α_0 decreases, confirming the tradeoff identified in the start-date analysis: the quadratic and linear coefficients exchange signal in an epoch-dependent manner.

Different GMSL datasets agree on the qualitative structure of this evolution—$d\alpha/dT$ increasing and α_0 decreasing toward the present—though they differ in absolute magnitudes. This cross-dataset agreement suggests the temporal structure is physical rather than an artifact of any single reconstruction.

8.5.2 SAOD reconsideration

The static DOLS calibration found that stratospheric aerosol optical depth (SAOD) is not statistically significant (γ_{saod} t-statistic = 0.26; §2a). However, the sliding-window analysis reveals a more nuanced picture.

When Mauna Loa Observatory (MLO) transmission-derived SAOD is included in the sliding-window DOLS, γ_{saod} achieves statistical significance ($|t| > 2$) in approximately 50% of center years. Significance is concentrated in windows centered before ∼1980, where the fit window encompasses the major 1963 Agung and 1982 El Chichón eruptions. By contrast, GloSSAC satellite-derived SAOD (available only from 1979) is significant in only 2–5% of windows, because its shorter record excludes the pre-satellite volcanic events that drive the MLO signal.

This finding does not contradict the static-DOLS conclusion that SAOD can be omitted from the baseline model—which uses the full record where the volcanic signal averages out—but it does indicate that epoch-specific calibrations (e.g., for hindcast validation over specific time periods) may benefit from including volcanic forcing. The fact that α_0 and $d\alpha/dT$ shift only modestly when SAOD is added (<0.3 pooled SE on average) confirms that volcanic forcing does not materially alias into the temperature sensitivity coefficients, even at the local level.

9 Caveats and Limitations

1. **Scenario probabilities are subjective**: The weights in Approach 3 reflect our assessment of current literature and are intended to be transparent and debatable, not definitive. Sensitivity to weight choices should be explored.
2. **Rheology correction is extrapolated**: The Martin et al. and Getraer & Morlighem results are for specific model configurations. Applying them as multiplicative corrections to the IPCC ensemble implicitly assumes the bias is representative.
3. **Independence assumption**: Combining rheology correction and stochastic amplification in quadrature assumes independence. Some correlation may exist if MISI dynamics depend on rheology.
4. **East Antarctic Ice Sheet**: Our framework focuses on WAIS. The EAIS is treated as having lower deep uncertainty, consistent with IPCC AR6, but recent evidence of increased EAIS mass loss may warrant similar analysis.
5. **Temporal scaling**: The growth curves for Approach 3 are idealized. Real WAIS dynamics will be more complex, with potential for abrupt transitions.

10 Recommendations and Future Work

For use in the hierarchical SLR forecasting framework, we recommend:

1. **Approach 4 as the primary σ_{ice}**: It is the most complete approach, combining restructured 4-scenario weighting, quantile-dependent rheology correction for all ice dynamics (Approach 1), and stochastic amplification during MISI (Approach 2). It produces $\sigma_{\text{ice}} \approx 491$ mm at 2100, raising the ice-sheet fraction from 7% to 18%. The rheology correction is applied to every sample in every scenario—reflecting that the $n = 3 \to n = 4$ bias affects all ice flow, not only MISI dynamics—while stochastic perturbations are applied only to MISI scenarios (S2–S4).
2. **Approach 2 as a parametric comparison**: The purely parametric correction ($\sigma_{\text{ice}} \approx 426$ mm, $f_{\text{ice}} \approx 14\%$) does not depend on subjective scenario weights and provides a useful comparison. That A4 ($\sigma_{\text{ice}} = 491$ mm) exceeds A2 ($\sigma_{\text{ice}} = 426$ mm) reflects the additional structural uncertainty captured by the scenario mixture that is absent from the parametric approach.
3. **Report results at four levels**: IPCC low confidence (baseline), Approach 3 (scenario structure without rheology correction), Approach 2 (parametric rheology + stochastic), and Approach 4 (recommended: full framework), allowing readers to assess sensitivity to each component of the uncertainty treatment.

4. **Sliding-window DOLS ("Dynamic OLS")**: The epoch sensitivity revealed in Table 6 motivated the sliding-window DOLS analysis described in §8.5. This analysis confirms that the rate–temperature relationship evolves over time: $d\alpha/dT$ increases toward the present while α_0 decreases, and the SAOD volcanic term becomes significant in $\sim$50% of epoch-specific windows. Future work should investigate whether this evolution reflects genuine physical nonstationarity (e.g., changing ocean heat uptake efficiency, emerging ice-sheet acceleration) or observational inhomogeneity in the early record.

References

DeConto, R. M. and Pollard, D. (2016). Contribution of Antarctica to past and future sea-level rise. *Nature*, 531(7596):591–597.

Edwards, T. L., Brandon, M. A., Durand, G., et al. (2019). Revisiting Antarctic ice loss due to marine ice-cliff instability. *Nature*, 566(7742):58–64.

Fricker, H. A., Becker, M. K., Dow, C. F., et al. (2025). Antarctica in 2025: Drivers of deep uncertainty in projected ice loss. *Science*, 387(6736):758–765.

Getraer, R. D. and Morlighem, M. (2025). Increasing the Glen–Nye power-law exponent accelerates ice-loss projections. *Geophysical Research Letters*, 52.

Glen, J. W. (1955). The creep of polycrystalline ice. *Proceedings of the Royal Society of London A*, 228(1175):519–538.

Goldsby, D. L. and Kohlstedt, D. L. (2001). Superplastic deformation of ice: Experimental observations. *Journal of Geophysical Research*, 106(B6):11017–11030.

Martin, D. F., Kachuck, S. B., Trevers, K., Millstein, J. D., Cornford, S. L., and Minchew, B. M. (in review). Impact of the stress exponent on ice sheet simulations. *AGU Advances*.

Millstein, J. D., Minchew, B. M., and Pegler, S. S. (2022). Ice viscosity is more sensitive to stress than commonly assumed. *Communications Earth & Environment*, 3:57.

Robel, A. A., Seroussi, H., and Roe, G. H. (2019). Marine ice sheet instability amplifies and skews uncertainty in projections of future sea-level rise. *Proceedings of the National Academy of Sciences*, 116(30):14887–14892.

Grinsted, A. and Christensen, J. H. (2021). The transient sensitivity of sea level rise. *Ocean Science*, 17:181–186.

Grinsted, A., Bamber, J. L., Bingham, R. J., et al. (2022). The Transient Sea Level Response to External Forcing in CMIP6 Models. *Earth's Future*, 10:e2022EF002696.

Jevrejeva, S., Palanisamy, H., and Jackson, L. P. (2021). Global mean thermosteric sea level projections by 2100 in CMIP6 climate models. *Environmental Research Letters*, 16:014028.

Kuhlbrodt, T. and Gregory, J. M. (2012). Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophysical Research Letters*, 39:L18608.

Rahmstorf, S., Foster, G., and Cazenave, A. (2012). Comparing climate projections to observations up to 2011. *Environmental Research Letters*, 7:044035.

Törnqvist, T. E., Conrad, C. P., Dangendorf, S., and Hamlington, B. D. (2025). Evaluating IPCC Projections of Global Sea-Level Change From the Pre-Satellite Era. *Earth's Future*, 13:e2025EF006533.

Turner, F. E., Malagon Santos, V., Edwards, T. L., Slangen, A. B. A., Nicholls, R. J., Le Cozannet, G., O'Neill, J., and Adhikari, M. (2023). Illustrative Multi-Centennial Projections of Global Mean Sea-Level Rise and Their Application. *Earth's Future*, 11(12):e2023EF003550.

Frederikse, T., Landerer, F., Caron, L., et al. (2020). The causes of sea-level rise since 1900. *Nature*, 584:393–397.

Dangendorf, S., Hay, C., Calafat, F. M., et al. (2019). Persistent acceleration in global sea-level rise since the 1960s. *Nature Climate Change*, 9:705–710.

Horwath, M., Gutknecht, B. D., Cazenave, A., et al. (2022). Global sea-level budget and ocean-mass budget, with a focus on advanced data products and uncertainty characterisation. *Earth System Science Data*, 14:411–447.

Rohde, R. A. and Hausfather, Z. (2020). The Berkeley Earth Land/Ocean Temperature Record. *Earth System Science Data*, 12:3469–3479.