https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Emergence of a non-linear global surface air temperature response to rapidly increasing effective radiative forcing

Craig R. Smith
Independent scientist, Naples, Florida, USA
Correspondence to: Craig R. Smith (smithcr239@gmail.com)

Abstract

A time series of global surface air temperature anomalies (GSAT) and effective radiative forcing (ERF) obtained from IPCC Sixth Assessment Report (AR6) for 1977 to 2025 provides evidence that GSAT is a non-linear, power-law function of ERF when ERF is rapidly increasing. With ERF > 0 W m^{-2} (n = 48), the GSAT–ERF relationship is described by a power-law, GSAT = 0.41 + 0.11 × $ERF^{1.90}$ (R^2 = 0.896; likely range = $ERF^{1.74–2.24}$). A linear relationship was rejected ($F(1,45) = 15.41$, $p < 0.0003$). GSAT, $ERF^{1.90}$, and the Niño 3.4 sea surface temperature anomaly index (ONI) were used in an autoregressive distributed lag model [ARDL(1,0,0)] yielding $GSAT_t = 0.32 + 0.23\ GSAT_{t-1} + 0.08\ ERF^{1.90} + 0.11\ ONI$. This model is equivalent to a two-layer energy-balance model with the deep-ocean temperature held constant. The ARDL model explains 96 % of GSAT variability from 1977 to 2025 (R^2 = 0.959). Residual diagnostics indicate no evidence of residual autocorrelation or heteroskedasticity and bounds testing (F = 40.51) supports cointegration. The model was validated by expanding-window recursive out-of-sample prediction from 1994 to 2025 (Theil's U = 0.53). Earth's climate system has entered a new regime in which the radiative response to the energy imbalance at the top of the atmosphere is progressively decreasing. ARDL projections for GSAT in 2041–2060 with ONI = 0 exceed the AR6 best estimate by 0.6 °C. If ERF continues to rise 0.7 W m^{-2} $decade^{-1}$, GSAT by mid-century will likely be 2.8 ± 0.3 °C.

1. Introduction

Since the late 19th century (1850-1900), Earth has retained over 500 ZJ of energy and the global surface air temperature (GSAT) has increased about 1.5 °C (Rohde and Hausfather, 2020; von Schuckmann et al., 2023). Using combined satellite radiation and ocean heat content measurements, Loeb et al. showed that Earth's energy imbalance doubled from 0.5 ± 0.2 W m^{-2} over 2000–2010 to 1.0 ± 0.2 W m^{-2} over 2013–2022 (Loeb et al., 2021). Earth is currently retaining more than 20 ZJ (~1.3 W m^{-2}) of energy per year (von Schuckmann et al., 2023; Cheng et al., 2022; Forster et al., 2026). Understanding how this rapidly increasing EEI will change the climate system in the next 2-3 decades is essential for strategic planning and policy development.

According to climate system theory, any energy imbalance at the top of the atmosphere (TOA) can be decomposed into 1) radiative forcing due to factors external to the climate system, together with the

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

adjustments they induce that are independent of surface temperature, 2) adaptive climate feedback, and 3) radiative variability unrelated to forcing or climate feedback. In the standard global mean energy-balance framework, the relationship between these factors is expressed as a linear equation:

$$\Delta N = \Delta F - \lambda\Delta T + V,$$

where ΔN is a change in EEI, ΔF is a change in effective radiative forcing (ERF), λ is the climate feedback parameter, ΔT is a change in GSAT, and V is radiative variability unrelated to forcing or climate feedback (Gregory et al., 2004). λ is not constant: the strength of the individual feedbacks depends both on the background climate state and on the spatial pattern of surface warming. Following the convention of Gregory et al. (2004), λ is positive, although λ can also be defined as a negative parameter in which case $\Delta N = \Delta F + \lambda\Delta T + V$.

Paleoclimate analyses over the past ~784,000 years show that climate sensitivity ($\Delta T/\Delta F$) is strongly state dependent, with significantly larger values attained during warm phases than during glacial periods (von der Heydt et al., 2014; Caballero and Huber, 2013). Recently, Meyssignac et al. (2023) found empirical evidence that λ varied within the range −1.0 to −3.2 W m^{-2} K^{-1} (using the alternative sign convention for λ) from 1970-2005 due to sea-surface temperature pattern changes associated with the Pacific Decadal Oscillation and concluded that climate models underestimate the pattern effect at decadal time scales.

A central feature of energy-balance climate models is the assumption that the relationship between ΔT and ERF is approximately linear, ERF = $(\lambda + \kappa)\ \Delta T$ (Forster et al., 2021). This assumption rests on the near-linearity of ΔN with ΔT. Gregory et al. (2004) show that in a coupled atmosphere–ocean general circulation model (AOGCM) ordinary least-squares regression of N on ΔT yields an estimate of λ and equilibrium climate sensitivity (ECS). Over multi-decadal time scales, this method assumes ocean heat uptake is proportional to ΔT. Two-layer energy-balance models embed the same linear GSAT–forcing structure in a linear dynamical system (Held et al., 2010; Geoffroy et al., 2013a, b). In this formulation the global climate response is governed by a small set of constant parameters, and the dependence of GSAT on N is linear, with ocean heat uptake and feedbacks introducing time dependence but not fundamental nonlinearity in the EEI–GSAT relationship.

Over much of the last 150 years, a first-order Taylor series expansion of net radiative flux from the Stefan-Boltzmann law is well motivated. Empirical observation and simulations of twentieth-century climate change indicate a linear relationship, ERF = $(\lambda + \kappa)\Delta T$ (Gregory and Andrews, 2016). Across sixteen coupled atmosphere–ocean general circulation models forced with CO_2 increasing at 1 % per year, the correlation between decadal-mean F and decadal-mean ΔT exceeded 0.98 in almost every case, indicating that linearity is a very good assumption (Gregory and Forster, 2008). Hébert et al. (2021) construct an observationally constrained projection of global temperature to 2100 and note that their model assumes a

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

linear stationary relationship between forcing and temperature and thereby neglects nonlinear interactions that could arise as the system continues to warm. The linear assumption has been verified for a doubling of CO_2, corresponding to a 3.93 ± 0.47 W m^{-2} increase in ERF, and the observed twentieth-century change of about 2 W m^{-2} lies well within that range (Forster et al., 2021). When a CO_2 quadrupling has been evaluated, corresponding to an increase of approximately 6.9 W m^{-2}, nonlinearity emerges.

Recent observations suggest the climate system may have entered a new regime. Forster et al. (2026) showed that anthropogenic ERF, EEI, and the rate of human-induced warming have all increased sharply in the last decade, with warming rates that are unprecedented in the instrumental record. Hansen et al. (2023, 2025) showed that global warming has accelerated since about 2010, with the post-2010 trend (~ 0.3 – 0.4 °C/decade) more than 50 % higher than the 1970–2010 rate of 0.18 °C decade^{-1}. Foster and Rahmstorf (2026) found that the increase in GSAT since 2015 (~ 0.35 °C decade^{-1}) substantially exceeds the 1970 to 2015 rate (~ 0.2 °C decade^{-1}) even after removing ENSO, volcanic, and solar variability. Minière et al. (2023) reported warming of the Earth's oceans has accelerated by 0.15 ± 0.05 W m^{-2} per decade since 1960 while warming of the land, cryosphere, and atmosphere has accelerated by 0.013 ± 0.003 W m^{-2} per decade. Pan et al. (2026) found that upper-2000 m ocean heat content increased by about 23 ± 8 ZJ in 2025 relative to 2024 and that the 0 – 2000 m warming rate rose from 0.14 ± 0.03 W m^{-2} decade^{-1} over 1960–2025 to 0.32 ± 0.14 W m^{-2} decade^{-1} over 2005–2025 (Pan et al., 2026). Van Loon et al. (2025) found a 0.5 to 0.9 W m^{-2} increase in observation-based effective radiative forcing since 2001, with 2023–2024 warmth consistent with the increase in radiative forcing detected directly in the satellite radiation record over the same period (Kramer et al., 2021). These converging lines of evidence suggest the climate system may have entered a regime in which linear approximations of the global warming response to rapidly increasing radiative forcing should be re-evaluated.

GSAT, ERF, and the Niño 3.4 sea surface temperature anomaly index (ONI) are continuous variables ordered in a time series. Time series often have statistical properties, including autocorrelation, non-stationarity, and endogeneity, that require use of specialized analytical methods. Other properties such as multicollinearity and cointegration must also be considered in order to obtain consistent, unbiased effect estimates. Econometric models can be used to emulate two-layer energy balance models (Pretis, 2020). The strength of econometric modeling lies in its empirical, observation-based approach that can detect relationships between dependent and independent variables without requiring extensive knowledge of underlying physical mechanisms. Pretis (2020) demonstrated the mathematical equivalence between a two-component energy-balance model of global mean temperature, ocean heat content, and radiative forcing and a restricted cointegrated vector autoregression (VAR)/autoregressive distributed lag (ARDL) system. He showed that climate physics can be represented in a discrete-time econometric framework with only a handful of physically interpretable parameters. Using observations of surface temperature, ocean heat content, and forcing, Pretis modeled this simple system and found cointegrating relationships

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

consistent with the underlying energy-balance equation. These results show that parsimonious econometric climate models, when grounded in energy-balance theory and estimated with appropriate time-series methods, can legitimately capture key aspects of the global climate response and provide robust scientific insight. ARDL models excel at capturing short-term and long-term relationships within a single framework, incorporating lagged effects and autoregressive components that reflect climate system inertia (Pesaran et al., 2001). ARDL models are particularly valuable for small sample sizes, provide unbiased coefficient estimates, and can accommodate structural breaks (Hamilton, 1989). Statistical models of this kind have been used to generate probabilistic global temperature projections directly from the observational record. Bennedsen et al. (2023) developed an energy-balance model on multiple historical data sources and produced projections comparable to those of process-based, mechanistic climate models.

The current study was designed to determine if the observed GSAT-ERF relationship is nonlinear over the last five decades, and if a simple econometric model could be constructed from observational data that would accurately and reliably reproduce observed global warming on a decadal time scale. The methodological approach combined non-linear modeling of GSAT as a function of ERF, and use of an ARDL model for time series analysis to predict GSAT from measurements of thermal inertia ($GSAT_{t-1}$), ERF and ONI. The results show that GSAT is a non-linear function of ERF when ERF is rapidly increasing, and an ARDL model of GSAT using ERF^v, $GSAT_{t-1}$, and ONI as independent predictor variables can accurately reproduce observed GSAT over the last 48 years. The model makes specific falsifiable predictions about future changes in GSAT through 2050, conditioned on future ERF and ONI estimates. The model could prove to be useful for government policy development and strategic business planning.

2. Two-compartment Energy-Balance Model

A two-compartment energy-balance model describes a mixed upper layer of low heat capacity, consisting of Earth's surface and upper ocean layer, exchanging heat with a deep-ocean layer of high heat capacity:

$$C_m \frac{dT_m}{dt} = -\lambda T_m - \kappa(T_m - T_d) + \mathrm{ERF} + V, \qquad C_d \frac{dT_d}{dt} = \gamma\,(T_m - T_d)$$

where T_m and T_d are the mixed-layer and deep-ocean temperature anomalies, C_m and C_d the corresponding effective heat capacities, λ the climate feedback parameter, γ the interlayer heat-transfer coefficient, and $\kappa = \varepsilon\gamma$ its efficacy(ε)-weighted counterpart, λ, γ and κ all in W m^{-2} K^{-1}, ERF is effective radiative forcing and V is internal climate variability (Held et al., 2010; Geoffroy et al., 2013a, b). Effective radiative forcing (ERF) is the top-of-atmosphere energy balance change following a perturbation, inclusive of adjustments in the stratosphere, the troposphere, and in surface albedo but excluding any radiative response to a change in global surface temperature (Forster et al., 2021). ERF is not measured directly. For the well-mixed greenhouse gases, it is constructed as observed concentration × radiative efficiency from radiative transfer models × (1 + a GCM-derived adjustment factor). The aerosol, volcanic, solar and land-use terms are

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

constructed from emissions inventories, stratospheric aerosol optical depth, total solar irradiance and satellite albedo retrievals respectively, each scaled by model-derived efficiencies (Forster et al., 2021).

Pretis (2020) showed that this system is mathematically equivalent to a cointegrated system in T_m, T_d and ERF, mappable to a cointegrated vector autoregression and estimable in discrete time. The equivalence yields two cointegrating relations,

$$\mathrm{ERF} - \lambda T_m \qquad T_m - T_d$$

the first the top-of-atmosphere radiative imbalance and the second the interlayer heat transfer. Pretis enters forcing linearly, so his cointegrating relation is linear in ERF. The specification estimated here is the discrete-time counterpart of the first relation, obtained under the following assumptions with T_m = GSAT:

A1. The deep-ocean temperature is constant over the estimation window, $T_{d,t} \approx T_D$,

A2. ERF is weakly exogenous (Engle et al., 1983),

A3. Internal variability, V in W m^{-2}, is independent of forcing, and enters the surface equation additively, orthogonal to the ERF regressor,

A4. Forcing acts on the surface through the normalized forcing $\left(\frac{\mathrm{ERF}_t}{F_{\mathrm{ref}}}\right)^{\nu}$, with ν the forcing exponent imposed from the power law non-linear regression. The constant $F_{\mathrm{ref}} = 1$ W m^{-2} is a dimensional normalization, not a physical parameter. It renders the regressor dimensionless so that the exponent can be applied. The ERF term in the ARDL GSAT equation becomes $F_{\mathrm{ref}}\left(\frac{\mathrm{ERF}_t}{F_{\mathrm{ref}}}\right)^{\nu}$.

A5. The surface response is linear in the normalized non-linear forcing of A4, so all departure from a linear ERF–GSAT relation is carried by the exponent ν rather than by variation in the coefficient acting on that forcing.

A backward-Euler step of $\Delta t = 1$ yr applied to the surface equation, collecting terms in $T_t \equiv T_{m,t}$, yields an ARDL equation of order (1,0,0):

$$\mathrm{GSAT}_t = \alpha + \rho\,\mathrm{GSAT}_{t-1} + \beta\left(\frac{\mathrm{ERF}_t}{F_{\mathrm{ref}}}\right)^{\nu} + \delta\,\mathrm{ONI}_t + \epsilon_t$$

where ONI_t represents internal climate variability, and ϵ_t is the regression error. The energy-balance parameters map onto the estimated coefficients as

$$\beta = \frac{\rho\Delta t}{C_m} F_{ref} \qquad \alpha = \frac{\rho\Delta t}{C_m}\kappa T_D$$

Where ρ is the estimated GSAT lag coefficient and Δt is the discretization interval of the backward-Euler step, equal to one year at the annual resolution of the data.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

The ARDL model can be re-parameterized by subtracting $GSAT_{t-1}$ from both sides of the equation to compute the long-run change in $GSAT_t$ that follows from changes in the joint effect of $GSAT_{t-1}$ and ERF^{ν}. This re-parameterization defines $\mu = \frac{\alpha}{(1-\rho)}$ and $\theta = \frac{\beta}{(1-\rho)}$, where (1– ρ) is the rate of $GSAT_t$ adjustment per year, μ is the intercept and θ is the long run multiplier.

$$\mathrm{GSAT}_t - \mathrm{GSAT}_{t-1} = \Delta\mathrm{GSAT}_t = -(1-\rho)\left[\mathrm{GSAT}_{t-1} - \mu - \theta\left(\frac{\mathrm{ERF}_t}{F_{\mathrm{ref}}}\right)^{\nu}\right] + \delta\,\mathrm{ONI}_t + \epsilon_t$$

$\Delta\mathrm{GSAT}_t$ is the long-run change in GSAT that occurs due to the joint effect of $GSAT_{t-1}$ and ERF^{ν}. The long-run multiplier and the intercept can be further specified:

$$\theta = \frac{F_{\mathrm{ref}}}{\Omega} \qquad \mu = \frac{\kappa T_D}{\Omega}$$

where Ω is the internal scale coefficient, in W m^{-2} K^{-1} held constant by A5. Both C_m and Δt cancel in θ, so the long-run coefficient is a property of the climate system rather than of the sampling interval.

$GSAT(ERF_t)$ is the long-run level of GSAT supported by a forcing held permanently at ERF_t, with ONI at zero. If $\Delta GSAT_t = 0$ with forcing held at ERF_t:

$$0 = -(1-\rho)\left[\mathrm{GSAT}(\mathrm{ERF}_t) - \mu - \theta\left(\frac{\mathrm{ERF}_t}{F_{\mathrm{ref}}}\right)^{\nu}\right]$$

Since $1 - \rho \neq 0$:

$$\mathrm{GSAT}(\mathrm{ERF}_t) = \mu + \theta\left(\frac{\mathrm{ERF}_t}{F_{\mathrm{ref}}}\right)^{\nu}$$

$GSAT(ERF_t)$ is the level toward which the error-correction term drives GSAT at rate $1 - \rho$. It is determined by the forcing level alone. The difference between $GSAT_{t-1}$ and $GSAT(ERF_t)$ is the adjustment not yet realized at that forcing level (i.e., the error correction term). Because A1 holds the deep-ocean temperature at T_D, $GSAT(ERF_t)$ is a long-run level of the surface compartment conditional on the deep-ocean state, not a long-run state of the coupled system, and it therefore does not correspond to the equilibrium warming that defines ECS.

Climate sensitivity is the change in surface temperature in response to a change in the atmospheric CO_2 concentration or other radiative forcing.

$$\frac{d\mathrm{GSAT}}{d\mathrm{ERF}} = \frac{\nu\theta}{F_{ref}}\left(\frac{\mathrm{ERF}}{F_{ref}}\right)^{\nu-1} = \frac{\nu\beta}{(1-\rho)F_{ref}}\left(\frac{\mathrm{ERF}}{F_{ref}}\right)^{\nu-1}, \qquad \text{in K per W m}^{-2}$$

Because $\lambda + \kappa = \frac{d\mathrm{ERF}}{d\mathrm{GSAT}}$, the reciprocal of climate sensitivity defines the aggregate rate of heat dissipation per unit of surface warming with the deep-ocean temperature held constant. The reciprocal is a transient quantity and does not describe equilibrium conditions.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

3. Methods

A dataset consisting of GSAT (annual_averages.csv) and effective radiative forcing (ERF_best.csv) from 1871–2025 was obtained from the Indicators of Global Climate Change update to the IPCC AR6 Working Group I assessment (Forster et al., 2026; Smith et al., 2026). The GSAT series is a blend of the principal observational surface temperature analyses, including GISTEMP (Lenssen et al., 2019; GISTEMP Team, 2025) and the Berkeley Earth analysis (Rohde and Hausfather, 2020). Total ERF for each year was calculated by summing the AR6 forcing estimates for solar, volcanic, aerosols, contrails, land use, black carbon on snow, atmospheric greenhouse gas concentrations, ozone, and stratospheric water vapor for that year. The greenhouse-gas component derives from the measured concentration record maintained by NOAA (Butler and Montzka, 2025). GSAT and ERF uncertainty estimates were obtained from the same data source. Monthly Niño 3.4 sea surface temperature anomalies were obtained from the NOAA Physical Sciences Laboratory (NOAA PSL, 2026).

For the period from 1974-2025, the relationship between GSAT and ERF was investigated using either a linear or power law specification:

1) Linear: GSAT = $\alpha + \beta \times$ ERF
2) Power Law: GSAT = $\alpha + \beta \times \mathrm{ERF}^{\nu}$

where α represents the baseline temperature, β the scaling coefficient, and ν is an ERF exponent. Years with ERF <0 (n=1) were excluded from the analysis because negative ERF values with a fractional exponent produce complex numbers, invalidating standard nonlinear regression methods. Linear parameters were estimated using ordinary least squares. Power law parameters were estimated using nonlinear least squares with multiple starting values to ensure global convergence. Model selection employed F-statistics for overall fit comparing linear versus power law specifications complemented by Bayesian information criteria and bootstrap confidence intervals (n = 1000). Statistical significance was assessed using t-tests, and model goodness-of-fit was evaluated using R^2, root mean square error (RMSE), and F-statistics.

The Akaike information criterion (AIC) and Bayesian information criterion (BIC) of seven power-law models with start years from 1974 to 1980 were compared to determine which starting year for the power-law analysis provided the lowest AIC and BIC. Other functional forms of the GSAT–ERF relationship were also evaluated. The power law was compared to linear, linear-log, and quadratic models fit to the 1977–2025 GSAT–ERF sample (n = 48) by least squares. Model comparison used the BIC. Following the Kass and Raftery (1995) convention, ΔBIC < 2 indicates no meaningful preference between models, $2 \leq \Delta\mathrm{BIC} < 6$ positive evidence against, $6 \leq \Delta\mathrm{BIC} < 10$ strong evidence, and $\Delta\mathrm{BIC} \geq 10$ decisive evidence.

Because ERF is derived rather than directly observed, the first-stage exponent is subject to attenuation bias from measurement error in the regressor. A simulation–extrapolation (SIMEX) correction was therefore

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

applied (Cook and Stefanski 1994; Carroll et al., 2006). Measurement error was modelled as additive and Gaussian with year-specific standard deviation implied by the AR6 5–95 % range (treated as the 95 % interval) for total ERF, averaging 0.435 W m^{-2} over 1977–2025, or 23 % of mean total forcing. The power law was re-estimated by nonlinear least squares at each point of a variance-inflation grid $\lambda \in \{0, 0.25, \ldots, 2.00\}$ with 1000 replicates per point, and the grid means extrapolated to $\lambda = -1$. The linear extrapolant is reported. Standard errors use the SIMEX variance estimator of Carroll et al (2006). Because a fractional power is undefined at non-positive values, replicates containing a perturbed forcing value at or below 0.02 W m^{-2} were rejected and redrawn. The full 1000 replicates were obtained for $\lambda \leq 1.25$, falling to 494 at $\lambda = 2.00$.

To examine both short-term adjustments and long-term equilibrium relationships between GSAT and ERF while accommodating non-stationary time series, an ARDL model was used. The ARDL model specification was:

$$GSAT_t = \alpha + \Sigma_{i=1}^{p} \beta_i \, GSAT_{t-i} + \Sigma_{j=0}^{q} \, ERF^{\,\nu_j}_{\;t-j} + \delta \, ONI_t + \varepsilon_t$$

where p and q represent optimal lag orders determined using information criteria with maximum lag length set to 4 years. The ONI_t term controls for natural climate variability associated with ENSO cycles. Lag-order selection over $p, q \in \{0,\ldots,4\}$ led to a preferred ARDL(1,0,0) specification for the main analysis. 1992 is excluded because total ERF is negative in that year following the Mt. Pinatubo eruption, rendering the power transformation undefined. The exponent ν is estimated once, in the power-law regression of GSAT on ERF, and is thereafter held fixed. It is not re-estimated within the ARDL. The uncorrected first-stage ν estimate is carried forward into the ARDL, so the results are conditional on that value, and the ARDL standard errors reported below do not propagate first-stage uncertainty in ν. The SIMEX-corrected exponent is used to bound the sensitivity of the second stage to first-stage attenuation.

Candidate lag orders were further restricted to those consistent with the energy-balance interpretation of the specification. Eliminating the interior temperatures of a linear energy-balance model with n reservoirs yields an n^{th} order difference equation carrying forcing terms to order $n - 1$, so that an n-reservoir model corresponds to an ARDL with n lags of GSAT and $n - 1$ lags of transformed forcing. Specifications of this form were estimated for n = 1, 2 and 3 and compared by AIC and BIC. Lag structures outside this family were also fitted, and the implied characteristic roots of the GSAT lag polynomial were examined to determine whether they were consistent with a system of coupled heat reservoirs, for which all relaxation modes are real.

Although the bounds procedure accommodates a mixture of I(0) and I(1) regressors, it is invalid if any series is integrated of order two. Integration orders were therefore established before estimation. Augmented Dickey–Fuller (ADF) tests with lag length selected by AIC and Kwiatkowski-Phillips-

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Schmidt-Shin (KPSS) tests of the null of level stationarity were applied to GSAT, ERF, the transformed regressor ERF^{v} and ONI, in levels and in first differences. Because 1992 is excluded, the transformed regressor could be tested only over the contiguous 1993–2025 span, and that test therefore carries lower power than the others.

The ARDL specification was tested for a level relationship between GSAT and ERF, with ONI treated as an exogenous stationary regressor, using the Pesaran–Shin–Smith bounds test.[27] The test was applied in the unrestricted-intercept, no-trend specification (Case III) with a single long-run forcing regressor ($k = 1$). Both the F-statistic on the joint exclusion of the lagged levels and the t-statistic on the lagged level of GSAT were evaluated against the reported asymptotic bounds. Where a level relationship was supported, the corresponding error-correction representation was estimated to separate short-run dynamics from long-run adjustment. For cointegrated cases, the corresponding error-correction representation was estimated to separate short-term dynamics from long-term adjustment.

Model validation included comprehensive diagnostic testing: residual autocorrelation (Breusch–Godfrey Lagrange Multiplier test at one and two lags), heteroskedasticity (Breusch-Pagan and White tests), and normality (Jarque-Bera test). Recursive coefficient estimates were examined to identify any time-varying relationships within each regime period. Optimal lag structures were determined through systematic testing of lag combinations from (0,0) to (4,4), with final selection based on information criteria consensus and residual diagnostic performance. Coefficient standard errors are reported both conventionally and in heteroskedasticity- and autocorrelation-consistent form with three lags. The standard error of the long-run coefficient, formed as the ratio of the forcing coefficient to one minus the sum of the autoregressive coefficients, was obtained by the delta method. Weak exogeneity of forcing with respect to the long-run parameter was assessed by estimating the marginal equation for the transformed forcing regressor including the lagged error-correction term, in the specification corresponding to Model C of Pretis, in which forcing is permitted to adjust to the cointegrating relation. This was supplemented by an augmented-regression test for contemporaneous endogeneity and by a Granger causality test from GSAT to forcing. As an estimator-robustness check on the long-run coefficient, dynamic OLS estimates including one and two leads and lags of the differenced regressor were also computed.

To evaluate predictive skill and address the potential for overfitting in a model with four parameters estimated from 48 yearly observations, an expanding-window recursive validation was performed. This approach is a standard method for evaluating the forecasting performance of econometric time-series models because it respects temporal ordering, avoids data leakage, and produces a distribution of out-of-sample errors across multiple forecast origins rather than a single split-sample estimate. The initial training window spanned 1977–1993, providing 16 observations, which is sufficient to estimate four parameters while preserving the high-ERF years (post-2000) for out-of-sample testing. At each recursive step, the ARDL(1,0,0) model was re-estimated on all available data through year t, and a one-step-ahead prediction was generated for year t+1 using the observed $GSAT_t$, observed ERF_{t+1}, and the contemporaneous ONI_{t+1}:

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

$$\widehat{\mathrm{GSAT}}_{t+1} = \hat{\alpha}_t + \hat{\beta}_{1,t}\,\mathrm{GSAT}_t + \hat{\beta}_{2,t}\,\mathrm{ERF}^{\nu}_{t+1} + \hat{\beta}_{3,t}\,\mathrm{ONI}_{t+1},$$

where $\hat{\alpha}_t$, $\hat{\beta}_{1,t}$, $\hat{\beta}_{2,t}$, and $\hat{\beta}_{3,t}$ are OLS estimates based only on GSAT data up to year t. The power-law exponent ν was held fixed at its value from the independent power-law regression. The procedure was repeated for each year from 1994 through 2025, yielding 32 consecutive out-of-sample predictions. Predictive accuracy was assessed using out-of-sample Root Mean Square Error (OOS RMSE), Mean Absolute Error (MAE), and mean bias (mean signed residual). Prediction intervals for each forecast were constructed as:

$$\widehat{\mathrm{GSAT}}_{t+1} \pm z_{0.975} \cdot \hat{\sigma}_t \cdot \sqrt{1 + \boldsymbol{x}^{\mathrm{T}}_{t+1}\,(\boldsymbol{X}^{\mathrm{T}}_t \boldsymbol{X}_t)^{-1}\,\boldsymbol{x}_{t+1}}$$

where $\hat{\sigma}^2_t$ is the residual mean square from the training-sample fit, $\boldsymbol{x}_{t+1}$ is the predictor vector for the forecast year, and $\boldsymbol{X}_t$ is the training design matrix. This formulation accounts for both residual variance and coefficient estimation uncertainty.

GSAT projections were generated by recursive simulation of the estimated ARDL(1,0,0), initialized at the observed 2025 GSAT anomaly of 1.391 °C and run annually through 2050. Five ERF trajectories were prescribed, each rising linearly from 3.4 W m^{-2} in 2026 to a 2050 endpoint of 4.0, 4.5, 5.0, 5.15 or 5.5 W m^{-2}; the 5.15 W m^{-2} path extends the observed 2015–2025 ERF trend of approximately +0.73 W m^{-2} decade^{-1}. The Niño 3.4 Index was set to zero throughout the projection period, so the projections describe ENSO-neutral conditions. Because the exponent ν is estimated in a first stage and then held fixed when the ARDL is estimated, intervals conditional on ν = 1.90 would understate projection uncertainty. Uncertainty was therefore propagated by Monte Carlo simulation with 20,000 replicates (random seed 20260808). Each replicate proceeds in four steps. First, ν is drawn from N(1.8902, 0.2627^2), its first-stage sampling distribution, truncated to the interval [0.5, 4.0]; the truncation is more than five standard deviations from the mean in either direction and is non-binding in practice. Second, the transformed regressor ERF$^{\nu}$ is reconstructed at the drawn exponent and the ARDL is re-estimated by ordinary least squares over 1977–2025 (n = 48), so that the intercept, the autoregressive coefficient and the forcing and ONI coefficients each adjust to that value of ν rather than being held at their ν = 1.90 estimates. Third, a coefficient vector is drawn from a multivariate normal distribution with mean equal to the re-estimated coefficients and covariance equal to their estimated covariance matrix. Fourth, the model is iterated forward from 2026 to 2050, adding in each year an independent innovation drawn from N(0, $\hat{\sigma}^2$), where $\hat{\sigma}$ is the residual standard error of the re-estimated model. Because the lagged dependent variable carries each innovation forward, residual uncertainty accumulates over the horizon rather than entering as a fixed band. A single draw of ν and of the coefficient vector is applied to all five trajectories within a replicate, so that scenario ordering is preserved draw by draw and differences between trajectories reflect forcing alone. The reported values are empirical percentiles of the 20,000 simulated 2050 outcomes: the median, the 17th and 83rd percentiles for the 66 % intervals shaded in Figure 6. The procedure therefore propagates first-stage exponent uncertainty, second-stage coefficient uncertainty conditional on the exponent, and compounding residual uncertainty. It

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

does not propagate uncertainty in the ERF trajectories themselves, which are prescribed rather than forecast, and the intervals are conditional on the estimated functional form.

The Finite-Amplitude Impulse-Response (FaIR 2.1) simple (reduced complexity) climate model was used as a benchmark. FaIR output was taken as published, using the 841-member reweighted constrained posterior of fair-calibrate v1.4.1, run with FaIR 2.1.3 on historical and harmonized SSP2-4.5 emissions (Smith et al., 2024). Annual values were read from the archived timebound temperatures, taking calendar year Y as the timebound at $Y + 1$, with anomalies relative to 1850–1900 formed as in the calibration. Correct extraction was confirmed against two published diagnostics. The observational series used here agrees with the Indicators of Global Climate Change reference underlying the calibration (Forster et al., 2026) to a correlation of 0.9995 and a mean difference of 0.009 °C. Emissions in the published run are observation-based to 2022 and follow SSP2-4.5 thereafter, and forcing pathways differ negligibly by 2025. The ARDL model was evaluated as a free-running simulation. From the observed 1976 value the model was iterated forward using only the observed forcing and the Niño 3.4 Index, with 1992 excluded. Accuracy over the 48 years was compared using the root-mean-square error of each model's median trajectory, the distribution of member-wise errors across the 841 FaIR members, and Diebold–Mariano tests on the squared-error loss differential with heteroskedasticity- and autocorrelation-consistent standard errors at three lags.

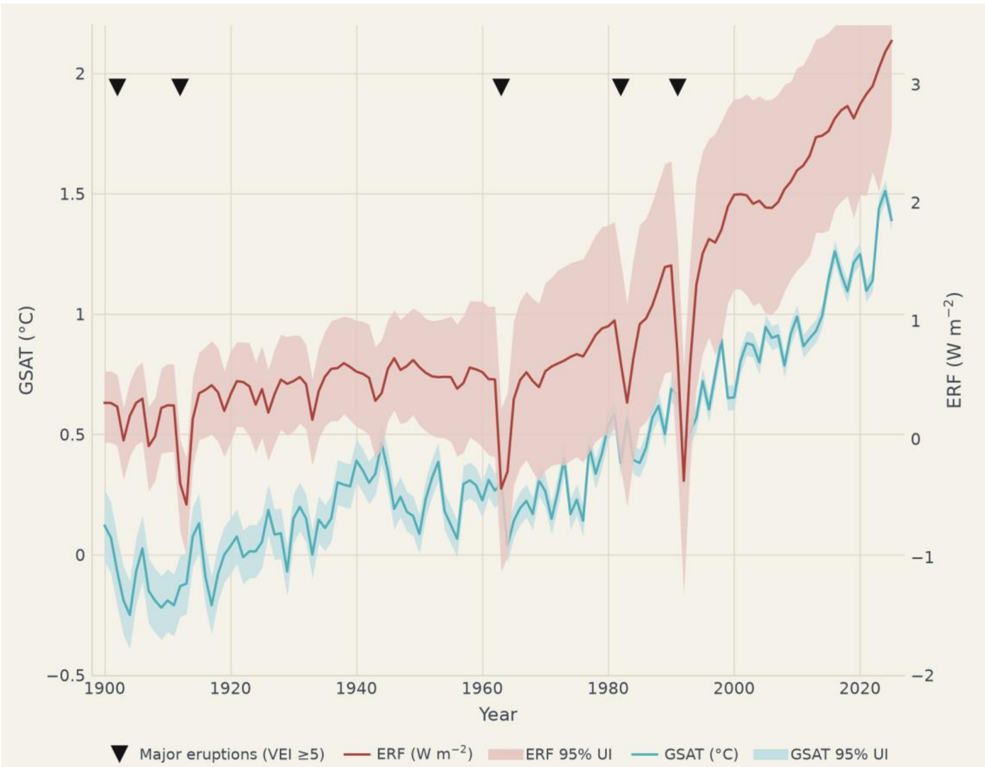
All analyses were performed using Python 3.9 with scipy, statsmodels, and pandas libraries. Nonlinear optimization employed scipy.optimize.curve_fit with convergence tolerance of 1×10^{-8}. The SIMEX simulation step used a fixed random seed (12345) with 1000 replicates per grid point, so that reported values are exactly reproducible. ARDL estimation used statsmodels tsa.ardl with maximum lag order determined by sample size constraints. Statistical significance was assessed at $\alpha = 0.05$ unless otherwise noted.

4. Results

The relationship between GSAT and ERF from 1900–2025 is shown in Figure 1. There are 7 years with negative ERF, 6 of them before 1977, of which all 6 fall within three years of a major volcanic eruption. From 1977–2025 only one year has a negative ERF (1992) following the Mt. Pinatubo eruption in 1991. Annual GSAT anomalies over 1977–2025 are estimated to within ± 0.04–0.06 °C (95 % UI), while annual ERF values carry substantially larger relative uncertainties (5-95 % Range half-width: ± 0.72–0.95 W m^{-2}), dominated by aerosol forcing uncertainty. The 50-year warming trend, defined as the linear trend in annual GSAT over 1976–2025, is 0.203 °C per decade (95% CI: 0.176, 2.31). Pearson correlation analysis revealed a strong positive association between GSAT and ERF from 1977–2025, excluding 1992 ($r = 0.93$, $p < 0.001$). Augmented Dickey–Fuller tests confirm that both series are non-stationary in levels over 1977–2025.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

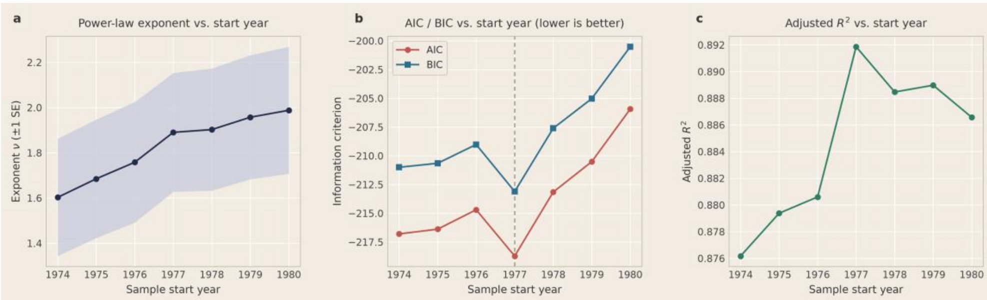
Figure 1: GSAT and ERF from 1900-2025. GSAT in °C is shown in blue with the 95 % UI shaded. ERF in W m^{-2} is shown in red with the 95 % UI shaded. Major volcanic eruptions (VEI ≥ 5) are shown as black triangles.



A sensitivity analysis was done to determine the starting year for a non-linear analysis. Years from 1974-1980, representing the beginning of the modern instrumental record, were evaluated. Figure 2 shows the starting year with the highest R^2 and the lowest AIC and BIC was 1977.

Figure 2: Comparison of starting years for the power law analysis from 1974 to 1980. The year with the lowest AIC/BIC and the highest R^2 is 1977. Note the power-law exponent increases each year and reaches a value of 1.99 (95 % UI: 1.44, 2.54) in 1980.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026



A power law analysis was conducted for 1977-2025, excluding the year with ERF < 0 (1992), to examine potential nonlinear dynamics in the GSAT-ERF relationship. The power law model is:

$$GSAT = \alpha + \beta \text{ x } ERF^{\nu}$$

The 1977-2025 period (n=48) encompasses an ERF range of 0.784 to 3.366 W m^{-2}, representing a 4-fold variation in forcing magnitude that provides sufficient statistical power for detecting nonlinear relationships. The estimation sample begins in 1977 for two physical reasons. Global anthropogenic Sulphur dioxide emissions peaked in the early-to-mid 1970s and declined thereafter (Smith, S. J. et al., 2011), so from the mid-1970s onward the greenhouse and aerosol contributions to total ERF ceased to act in opposition and the aggregate forcing behaves as a coherent trend. The 1976/77 North Pacific regime shift, an abrupt transition of the Pacific Decadal Oscillation from its negative to its positive phase in mid-1976 (Hartmann and Wendler, 2005; Miller et al., 1994), falls immediately before this start date, so the sample lies wholly within a single decadal regime and the ENSO covariate describes a relationship of consistent amplitude throughout.

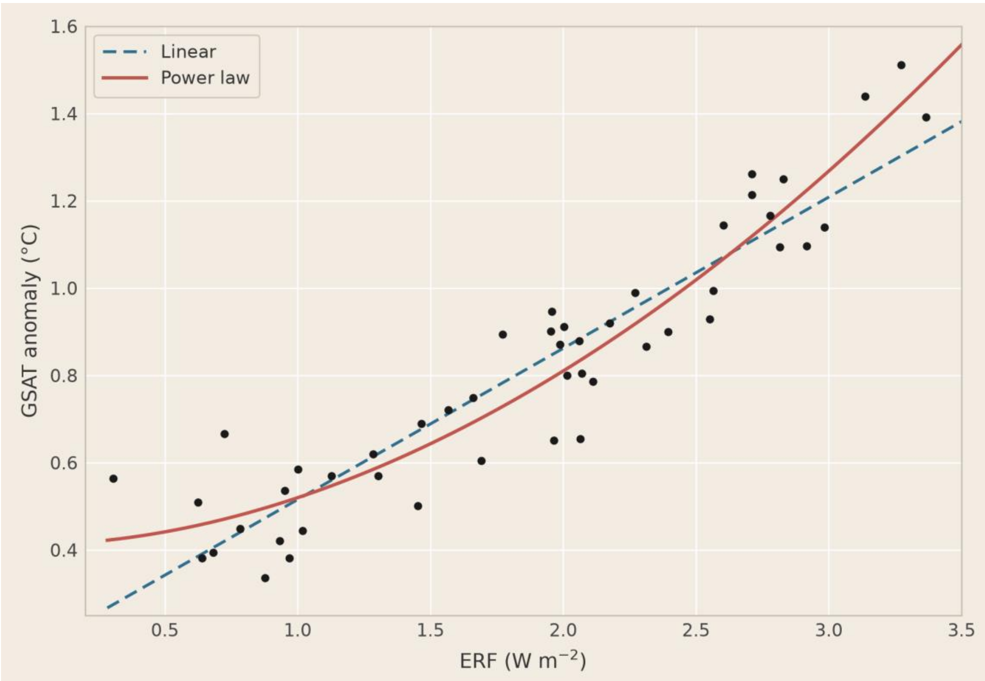
Power law specification yielded an exponent of $\nu = 1.890 \pm 0.263$. To check that this estimate is robust to systematic error in the forcing series, ERF uncertainty was propagated using a Monte Carlo simulation (MC). Over 2000 draws the MC returned $\nu = 1.90 \pm 0.34$ (95% CI [1.28, 2.55]), excluding a linear response. The power law model is:

$$GSAT = 0.41 + 0.11 \times ERF^{1.90}$$

$R^2 = 0.896$ and provided a highly significant improvement over the linear alternative $R^2 = 0.861$ ($F(1,45) = 15.41$, $p < 0.0003$). (Figure 3) The 95 % confidence interval of the estimated exponent is [1.23, 2.57], which is above the critical threshold of $\nu = 1.0$ ($t(45) = 3.389$, $p < 0.006$).

Figure 3: Comparison of linear and power law models for 1977-2025 (ERF > 0). Black circles represent observed data points (n=48), the dashed blue line shows the linear model fit ($R^2 = 0.861$), and the solid red line shows the power law model fit ($R^2 = 0.896$).

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026



Inference for ν was assessed using two complementary procedures. A residual bootstrap (N = 5000) treating ERF as fixed yielded a 95 % CI of $\nu \in [1.42, 2.42]$ and rejected $\nu \leq 1$ at $p < 0.0002$. Correcting for measurement error in ERF via linear SIMEX increases the estimated power-law exponent from 1.89 (95 % UI: 1.38–2.40) to 2.09 (95 % UI: 1.76–2.43), indicating that the ERF–GSAT non-linear response is stronger than fits treating ERF as error-free.

To assess the stability of the power-law relationship, sensitivity analyses were conducted, including systematic data exclusions and leave-one-out jackknife resampling. Excluding the record-warm years 2023 and 2024 reduced the exponent by 8.6 % to 1.73 ± 0.29, but it remained significantly greater than unity (95 % CI [1.17, 2.29]; $t(43) = 2.54$, one-sided $p = 0.007$). This is a demanding test, since those two years supply two of the three highest-forcing observations in the sample. Jackknife analysis revealed distributed leverage: only 4.2 % of years (2 of 48) exerted individual influence exceeding 5 % on the exponent estimate and mean absolute year-by-year influence was 1.6 %, indicating that no single observation dominates the fitted curvature. The most influential observations were 1983 and 2025, which act in opposite directions (−7.9 % and +5.7 %). 1983 is the minimum-forcing year in the sample, with ERF depressed to 0.305 W m^{-2} by the El Chichón eruption while the 1982–83 El Niño elevated GSAT, making it a high-leverage point at the left end of the forcing range. Residuals showed no systematic pattern with forcing level, with the largest deviations occurring at low-to-middle forcing (1991, 2000, 1998, 1978)

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

rather than at the high-forcing endpoints, where 2023, 2024 and 2025 all deviate by less than 0.10 °C. The power-law form is therefore not an artifact of recent extreme observations.

The power law model is compared with other admissible functional forms in Table 1. A functional form was admissible if it was defined over the sampled forcing range, entered GSAT only up to an additive constant, and admitted an interpretation within a two-layer energy-balance framework. Among eligible forms, the linear and linear-log specifications are decisively rejected (ΔBIC = +10.3 and +47.1). The linear term of the quadratic form is indistinguishable from zero ($b = +0.0192$, SE 0.0857, $t = 0.22$), reducing the quadratic fit to $T = a + c \cdot ERF^2$, a power law with $\nu = 2$.

Table 1: Alternative admissible functional forms for the GSAT–ERF relationship. The power law was fit by nonlinear least squares; the remaining forms were fit by ordinary least squares. ΔBIC is relative to the power law.

Model	Number of parameters	R^2	RMSE	BIC	ΔBIC
Power Law	3	0.896	0.096	−73.02	–
Quadratic	3	0.896	0.096	−72.88	0.14
Linear	2	0.861	0.112	−62.75	10.27
Linear-log	2	0.700	0.164	−25.90	47.12

For the period from 1977–2025, four ARDL specifications were tested using the empirically derived $ERF^{1.90}$ functional form: ARDL(0,0,0), ARDL(0,1,0), ARDL(1,0,0), and ARDL(1,1,0). The lag term for GSAT in 1977 was provided by 1976 GSAT. The analysis incorporated ONI as a control variable to account for natural climate variability. Selection was based on the lowest BIC value. The selected model is ARDL(1,0,0):

$$GSAT_t = \alpha + \rho \times GSAT_{t-1} + \beta \times ERF^{1.90} + \delta \times ONI_t$$

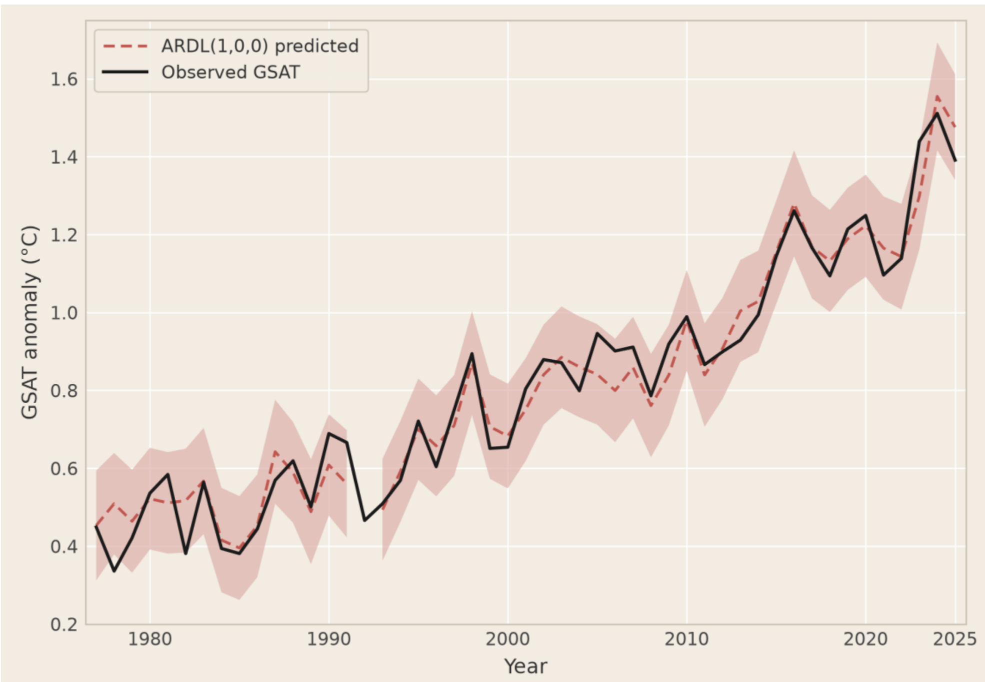
where α represents the intercept, ρ captures temperature persistence, β governs the power law relationship with ERF, and δ quantifies the ONI effect. The estimated coefficients with 95 % confidence intervals are: $\alpha = 0.319$ (SE = 0.036, $p < 0.0001$), $\rho = 0.235$ (SE = 0.086, $p = 0.009$), $\beta = 0.082$ (SE = 0.010, $p < 0.0001$), and $\delta = 0.108$ (SE = 0.014, $p < 0.0001$). The final model is:

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

$$GSAT_t = 0.32 + 0.235\ GSAT_{t-1} + 0.082\ ERF^{1.90} + 0.108\ ONI_t$$

The ARDL(1,0,0) model $R^2 = 0.959$ (adjusted $R^2 = 0.957$) with RMSE = 0.063 °C, AIC = –125.32, BIC = –117.83. The Observed vs predicted GSAT from the ARDL model is shown in Figure 4.

Figure 4: Observed vs predicted GSAT from the ARDL model. The observed GSAT is the solid black line and the GSAT predicted by the ARDL model is the dashed red line with the UI shaded.



β and ρ are strongly collinear (r = –0.94, VIF = 8.27) and are not separately identified with useful precision. No interpretation is placed on either coefficient in isolation. θ is more precisely determined than either ρ or β; $\theta = 0.1071$ K per unit normalized ERF, with a standard error of 0.0045 (95 % UI: 0.0982 to 0.1161). The long-run intercept is $\mu = 0.4164 \pm 0.0210$ K. The coefficient of variation for θ is 4.2 %, against 11.9 % for β and 36.6 % for ρ. It is θ rather than β or ρ individually on which the ARDL ΔGSAT projections depend. The ENSO coefficient (δ = 0.108) indicates that a one-unit change in the ONI index corresponds to a 0.11 K change in GSAT within the same year, consistent with established ENSO-temperature relationships (Johnson and Kosaka, 2016).

θ is the fixed point of the model's adjustment path rather than a value attained at a stated horizon. For a sustained unit change in normalized forcing $(\mathrm{ERF}/F_{\mathrm{ref}})^{\nu}$, the cumulative response after τ years is $\theta\ (1 - \rho^{\tau+1})$, so the share of θ realised is $1 - \rho^{\tau+1}$. At $\hat{\rho} = 0.2345$ this gives 76.6 % in the year of the change,

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

94.5 % after one year, 98.7 % after two and 99.7 % after three; the mean lag, $\rho/(1-\rho)$, is 0.31 years. GSAT adjustment is therefore effectively complete within about three years. This is the response of the mixed layer alone: under A1 the deep-ocean temperature is held fixed, so the centennial adjustment associated with deep-ocean heat uptake lies outside the estimated specification by construction. The parameter θ carries no equilibrium interpretation.

A comprehensive battery of diagnostic tests was conducted to verify the validity of the ARDL(1,0,0) model's statistical inference. Residual autocorrelation was assessed using both the Ljung-Box Q statistic (lags 1–8, all $p \geq 0.33$) and the Breusch-Godfrey LM test at lag 4 (LM = 4.43, p = 0.35); neither test rejected the null of no serial correlation. The Breusch-Pagan test found no evidence of heteroskedasticity (LM = 0.86, p = 0.834). The Jarque-Bera test found no significant departure from residual normality (JB = 1.065, p = 0.587). Multicollinearity diagnostics showed elevated but acceptable variance inflation factors (VIF) of 8.27 for both GSAT(t−1) and $ERF^{1.90}$. Augmented Dickey–Fuller tests over 1977–2025 fail to reject a unit root in either GSAT or ERF in levels at every lag order from 0 to 8, while both reject decisively in first differences (−5.63 and −6.62, $p < 0.001$), confirming that neither series is I(2) and that the bounds-testing framework is applicable. ONI is stationary in levels over the same period (ADF = −6.83, $p < 0.001$). We therefore tested for a long-run equilibrium relationship between GSAT and $ERF^{1.90}$ using the Pesaran, Shin and Smith (Pesaran et al., 2001) bounds procedure, estimated over 1977–2025 excluding 1992 (n = 48). The bounds F-statistic (F = 40.51, $p < 0.0001$) and the accompanying t-statistic on the lagged level (t = −8.98, $p < 0.0001$) both exceed the 1 % upper bounds under the Pesaran et al. (2001) asymptotic critical values (k = 1, Case III; upper bound at 1 % = 7.84) and under the Narayan (2005) small-sample critical values. A parametric-in-residuals bootstrap of the bounds test under the imposed null of no cointegration (2000 replicates) gave a finite-sample p-value of 0.0005 (bootstrap 99 % critical value = 6.40). As corroborating evidence, an Engle–Granger residual-based test of the bivariate regression also rejected the null of no cointegration (ADF = −6.79, 1 % critical value = −4.15, $p < 0.0001$). Taken together, these tests provide strong and consistent evidence that GSAT and $ERF^{1.90}$ share a genuine long-run equilibrium relationship.

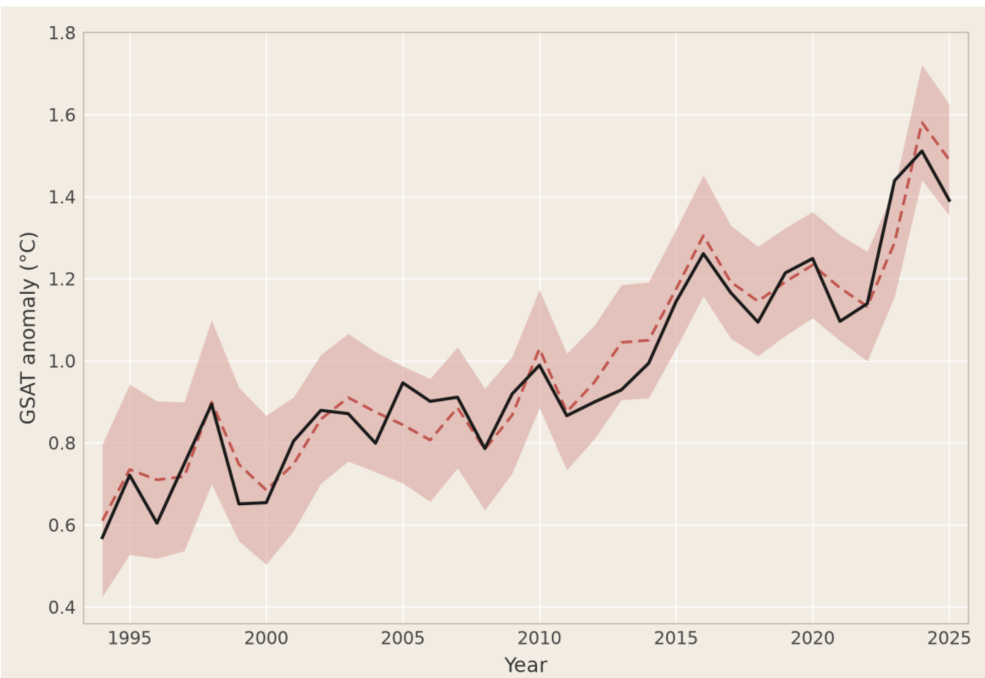
To assess whether the ARDL(1,0,0) model generalizes beyond its estimation sample and to directly address the concern that a 4-parameter model fitted to 48 annual observations may overfit the data, a formal expanding-window recursive validation was conducted. The validation procedure began with an initial training window spanning 1977 to 1993, excluding 1992 (n = 16). The model was fitted on this period and used to generate a single out-of-sample prediction for 1994. The training window was then extended by one year to include 1994, the model was re-estimated on this expanded sample, and a prediction was generated for 1995. This procedure was repeated sequentially through 2025, yielding 32 consecutive out-of-sample predictions covering the period 1994 to 2025. At each step, the model was re-estimated using only data available at that time, with no knowledge of future observations. The power law exponent $\nu = 1.90$ was

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

held fixed, consistent with its derivation from the independent power law analysis. All four coefficients (α, ρ, β, δ) were re-estimated at each expansion step.

The results of the recursive validation are shown in Figure 5. The out-of-sample RMSE (0.064 °C for 32 recursive forecasts spanning 1994–2025) is nearly identical to the in-sample RMSE (0.063 °C) and represents a 47 % reduction relative to a naive persistence benchmark (Theil's U = 0.53). The near-equivalence of in-sample and expanding-window out-of-sample RMSE indicates that the ARDL(1,0,0) specification is not overfit and that its parameters are stable as the training window is progressively extended. The mean absolute error was 0.052 °C and the mean bias was −0.016 °C, indicating no systematic directional error across three decades of prediction. All but one out-of-sample prediction fell within the 95 % prediction interval, the exception being the record-warm year 2023. Retrospective comparison of projections against subsequent observations is the established test of projection skill (Hausfather et al., 2020).

Figure 5. Expanding-window recursive validation of the ARDL(1,0,0) model. Observed GSAT (solid black), one-step-ahead out-of-sample prediction (dashed red), and 95 % prediction interval (shaded).

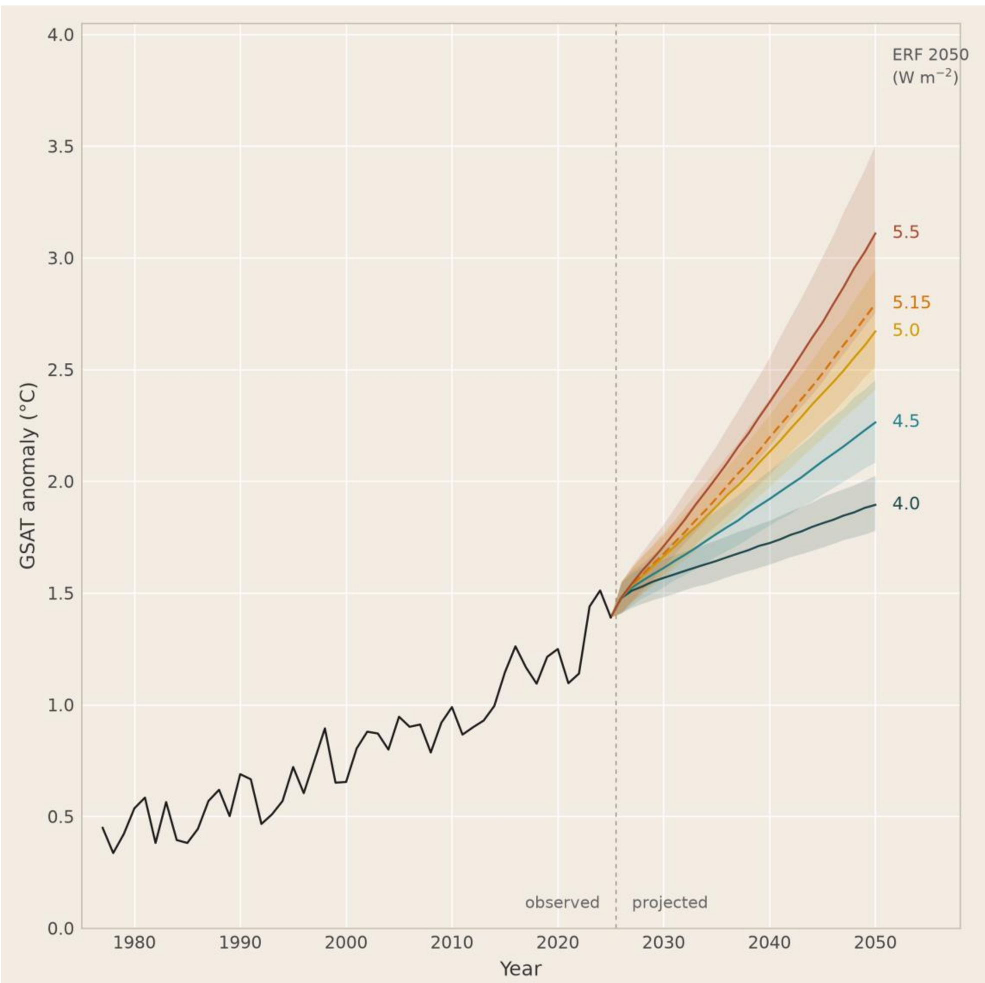


The ARDL(1,0,0) model can be used to project GSAT anomalies through 2050. These projections are conditional on four linear ERF trajectories rising from 3.4 W m^{-2} in 2026 to 4.0, 4.5, 5.0 and 5.5 W m^{-2} at 2050, plus a fifth trajectory reaching 5.15 W m^{-2} obtained by extending the observed 2015–2025 ERF trend

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

of approximately 0.7 W m^{-2} decade^{-1}. The 25-year projection horizon relative to the 49-year calendar span of the estimation window (1977–2025) yields a horizon-to-window ratio of 0.51, within the conventional envelope for semi-empirical climate projections (Rahmstorf 2007). Figure 6 shows the annual projections with 66 % uncertainty intervals through 2050, assuming ONI = 0, for each of the five ERF trajectories.

Figure 6: GSAT projections to 2050. Black: observed GSAT, 1977 to 2025 (1850 to 1900 baseline). Colored lines: median ARDL(1,0,0) projections under five linear ERF trajectories rising from 3.4 W m^{-2} in 2026 to the 2050 endpoints labelled at right, the dashed path (5.15 W m^{-2}) extending the observed 2015 to 2025 ERF trend. Shading: 66 % uncertainty intervals. ONI = 0 throughout. The vertical dashed line separates observed from projected.

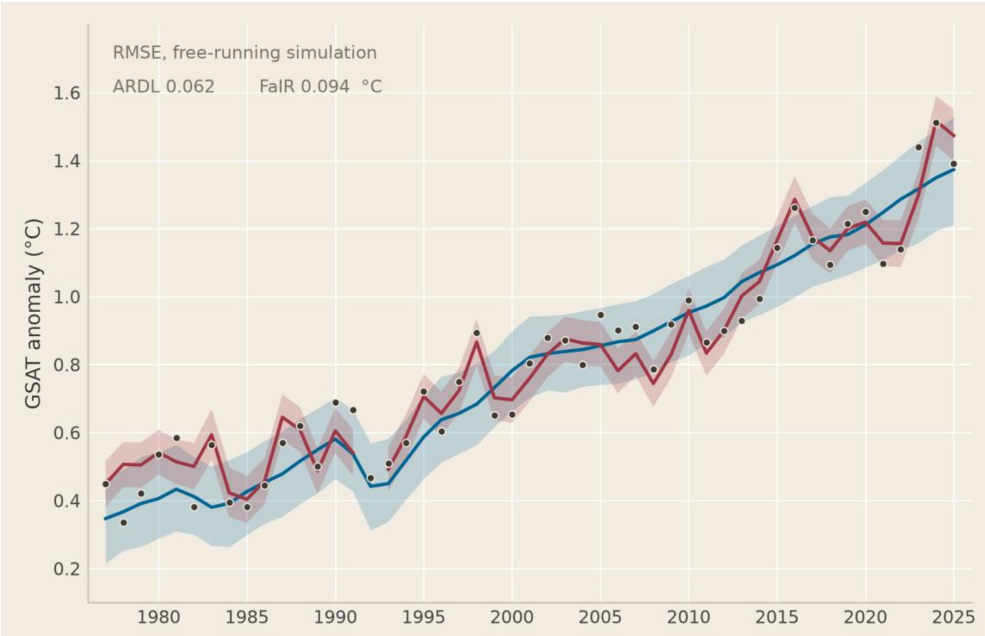


https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

The Finite Amplitude Impulse Response model (FaIR) and the Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC) are standard SCMs. In the IPCC Sixth Assessment Report both were calibrated to the assessed physical indicators and were found to reproduce the assessment to within ± 5 % for central estimates and ± 10 % for ranges across more than 80 % of evaluated metrics. Neither model was identified as superior. The FaIR and ARDL simulations of annual GSAT from 1977 to 2025 are compared in Figure 7. Root-mean-square error of the median path over the 48 estimation years is 0.062 °C for the ARDL model and 0.094 °C for FaIR. Across the 841 individual members, the ARDL simulation is more accurate than every member. A two-sided Diebold–Mariano test on the squared-error loss differential gives $t = -3.15$ and $p = 0.003$ against the median trajectory, and $t = -3.25$ and $p = 0.002$ against the single best-fitting member. The difference between the two models lies in year-to-year tracking. With the Niño 3.4 Index set to zero, the ARDL error rises to 0.092 °C, which is indistinguishable from FaIR at $p = 0.68$.

Figure 7: Observed annual GSAT (black points) and free-running annual GSAT simulations from the ARDL(1,0,0) (red) and FaIR (blue) for 1977 to 2025. Lines are medians; shaded regions are 66 % uncertainty intervals (17th–83rd percentiles). The ARDL simulation is initialized at the observed 1976 value and uses no observed temperatures thereafter, and its range is across 20,000 replicates drawing the exponent, coefficients and error terms from their estimated distributions, conditional on the observed forcing. The FaIR range is across the 841 published members of the fair-calibrate v1.4.1 constrained posterior, which include forcing uncertainty. No alignment to the observed record was applied. The UI widths are therefore not comparable. The ARDL simulation is undefined in 1992 because ERF is negative. Root-mean-square errors for the two models over the 48 years are shown in gray.

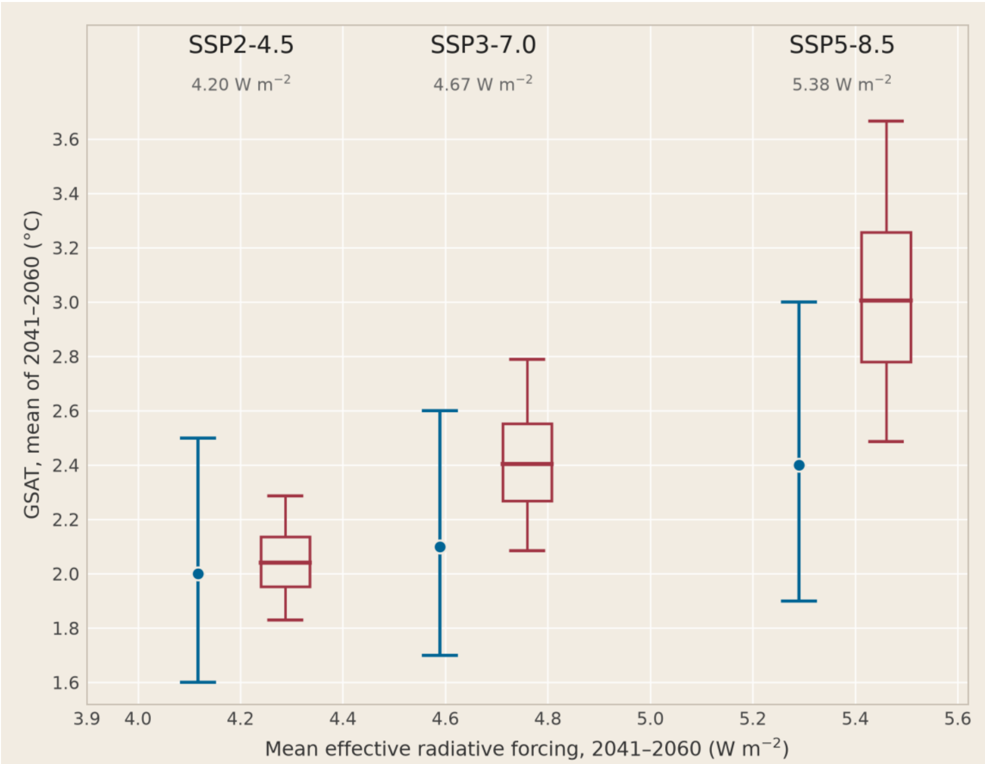
https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026



The IPCC AR6 assessment of global warming over 2041 to 2060 is conditioned on five Shared Socioeconomic Pathways, each generating a trajectory of effective radiative forcing from projected greenhouse gas and aerosol emissions. Figure 8 compares the ARDL projection of mean GSAT over 2041 to 2060 with the assessed estimates for three of these pathways, each placed at its mean forcing over the same window. The two agree closely at SSP2-4.5, where mean forcing is 4.20 W m^{-2}. They separate as forcing rises, to 2.40 against 2.1 °C at SSP3-7.0 and 3.01 against 2.4 °C at SSP5-8.5. For ERF of 5.5 W m^{-2} in 2050, the AR6 best estimate is 2.40 °C [1.90, 3.00 °C] and the ARDL model best estimate is 3.1 °C [2.44, 3.96 °C].

Figure 8. GSAT change averaged over 2041 to 2060, under three SSP scenarios. Red boxes show the ARDL projection: the horizontal line is the median, the box spans the interquartile range and the whiskers span the 5th to 95th percentiles of 20,000 Monte Carlo replicates with ONI set to zero. Blue symbols show the best estimate and very likely range from AR6 WG1 Chapter 4, Table 4.5. Each scenario is positioned at its mean effective radiative forcing over 2041 to 2060, computed from the anchors published in AR6 WG1 Annex III, Table AIII.5 (Lee et al., 2021; Dentener et al., 2021).

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026



5. Discussion

The current study has three main findings. First, since 1977 the GSAT response to increasing ERF is non-linear. The long-run relationship is most likely of the form GSAT = $\mu + \theta \cdot ERF^{\nu}$, with ν likely in the range of 1.74 – 2.24. GSAT increases more per unit ERF as ERF increases, consistent with weakening negative climate feedbacks in a warming climate (Andrews et al. 2015; Armour 2017; Friedrich et al. 2016). Second, an ARDL model of GSAT that combines thermal inertia ($GSAT_{t-1}$), ERF^{ν} and ONI reproduces annual GSAT from 1977 to 2025 with high fidelity ($R^2 = 0.96$, RMSE = 0.063 °C). This model corresponds to a two-layer energy-balance model with the deep-ocean temperature held constant. Lastly, if ERF continues to rise at 0.7 W m^{-2} per decade and exceeds 5 W m^{-2} by 2050, GSAT in 2050 will likely be about 0.5 °C higher than IPCC AR6 projections and reach 2.8°C. The last time GSAT exceeded 2.5°C was during the mid-Pliocene Warm Period 3.3 million years ago.

The GSAT-ERF analysis uses information obtained from authoritative, curated data sources. Prior to 1979 there is considerable uncertainty in both the GSAT and ERF estimates. Starting in 1980, the accuracy and reliability of GSAT and ERF data improved due to the availability of direct observation from satellites, ocean buoys, and measurement of atmospheric greenhouse gas concentrations. The uncertainty in ERF estimates represents measurement error that is independent of GSAT measurement error, since the two

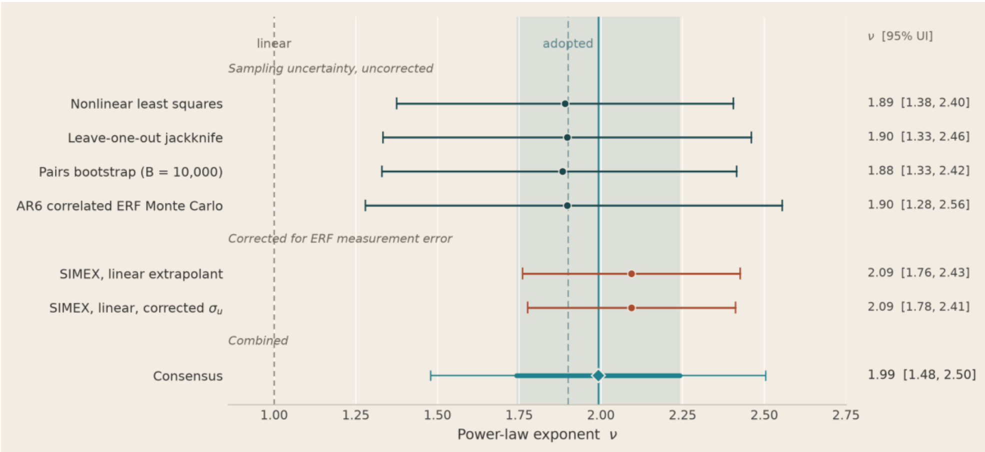
https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

metrics are derived from different observational and computational systems. Forster et al. (2026) reported that 2025 mean Total Anthropogenic ERF was 3.10 W m^{-2} with a range of 2.35 to 3.83 W m^{-2}. Atmospheric aerosols dominate the uncertainty in ERF, accounting for roughly three-quarters of the total assessed variance, while CO_2 and tropospheric ozone each contribute about 9%. The current measurement error in ERF leads to estimating ERF regression exponents that are biased toward one. Detecting associations in the presence of large random uncertainty strengthens the evidence for underlying physical relationships, rather than weakening them (Fuller, 1987; Carroll et al., 2006).

The finding that the GSAT–ERF relationship is non-linear over the 1977 to 2025 forcing range was subjected to a series of residual diagnostic tests, none of which suggested the finding is a statistical artifact. The result is robust to influential observations, with no single year materially altering the estimate. The point estimates from most adjustments lie between 1.88 and 2.09, with a midpoint of 1.99. A linear relationship was rejected ($F(1,45) = 15.41$, $p < 0.0003$). Taking the clustered estimates as the primary evidence and allowing for their sampling uncertainty, a likely range for ν is 1.74 to 2.24 (Figure 9). The power-law fit returns $\nu = 1.89$. A Monte Carlo simulation over the assessed ERF uncertainty adjusts the ERF estimate for the fixed component of the associated uncertainty and returns $\nu = 1.90$, the value adopted for the ARDL analysis (Figure 9). Simulation-extrapolation was used to adjust for ERF uncertainty bias in the estimate of ν. A linear extrapolant yielded $\nu = 2.09$, indicating that ERF uncertainty biases the unadjusted estimate of ν low. The SIMEX correction was not used for the ARDL analysis because it assumes ERF uncertainty is independent across years, whereas the assessed ERF uncertainty is dominated by aerosol forcing and is persistent.

Figure 9: Estimates of the power-law exponent ν in $GSAT = \mu + \theta \cdot ERF^{\nu}$ over 1977 to 2025 ($n = 48$), shown as central estimates with 95 % uncertainty intervals: four uncorrected estimates differing in how sampling uncertainty is obtained (1.88–1.90), and two linear SIMEX estimates corrected for measurement error in ERF (2.09). Dashed lines mark the linear response $\nu = 1$ and the value adopted for the ARDL model $\nu = 1.90$. The solid blue line is the consensus estimate ($\nu = 1.99$) and the shaded band is the 66% consensus range, 1.74–2.24.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026



Alternative functional forms were also evaluated. Linear and Linear-log forms can be confidently excluded, with ΔBIC of 10.27 and 47.12 relative to the power-law form. Over the 1977 to 2025 observation period, the power-law and quadratic forms cannot be distinguished. The quadratic form is simply the second-order Taylor series expansion of the top-of-atmosphere energy budget. Bloch-Johnson et al. (2021) used a quadratic Taylor expansion of the TOA energy budget across fourteen coupled general circulation models and showed that equilibrium climate sensitivity increases with CO_2 concentration in thirteen of the fourteen, with positive feedback temperature dependence explaining 69 % of the average increase. The present study provides the first empirical support for this prediction. The power-law exponent, with a consensus range of 1.74–2.24, is the observational signature of the feedback temperature dependence that Bloch-Johnson et al. (2021) identified in GCMs. The consistency between the empirical power law structure and the theoretical quadratic prediction suggests the nonlinearity identified in this study reflects a genuine physical property of the climate system rather than a statistical artifact. The non-linear results also align with paleoclimate evidence suggesting that past climate state transitions involved a non-linear, accelerating GSAT response to changes in radiative forcing (Caballero and Huber, 2013; von der Heydt et al., 2014).

Knutti and Rugenstein (2015) showed that the assumption of constant radiative feedback, the foundational premise of the classical $\Delta T = -F\lambda^{-1}$ framework, breaks down under conditions of large forcing changes, evolving sea surface temperature (SST) patterns, and state-dependent feedback interactions. They showed that the climate sensitivity parameter varies systematically with forcing magnitude and climate state, and that climate sensitivity estimates based on the observed warming that assume a constant λ would be biased low. The present study provides time-series-based confirmation of these theoretical predictions using the instrumental GSAT-ERF record. The power law exponent is the corollary of the accelerating feedback temperature dependence that Knutti and Rugenstein (2015) identified as a source of the climate state and forcing dependence of λ.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

An ARDL model was used to evaluate GSAT_t as a function of $\mathrm{ERF}_t^{1.90}$, GSAT_{t-1}, and ONI_t. GSAT_{t-1} captures thermal inertia in the climate system. The resulting model:

$$\mathrm{GSAT}_t = 0.32 + 0.23\,\mathrm{GSAT}_{t-1} + 0.082\,\mathrm{ERF}_t^{1.90} + 0.11\,\mathrm{ONI}_t$$

explained 96 % of the variability in GSAT from 1977 to 2025 (n = 48). The Pesaran–Shin–Smith bounds test and Engle–Granger residual test both reject the null hypothesis of no cointegration at $p < 0.001$, establishing a genuine long-run relationship between GSAT_t and $\mathrm{ERF}^{1.90}$. Expanding-window recursive validation (EWR) over 1994 to 2025 produced 32 out-of-sample forecasts whose RMSE (0.064 °C) matches the in-sample fit and reduces error by 47 % relative to a persistence benchmark (Theil's U = 0.53). The EWR validation provides confirmation that the ARDL(1,0,0) model with a nonlinear $\mathrm{ERF}^{1.90}$ forcing term is robust and representative of post-1977 GSAT dynamics. Despite its parsimony, the model accurately reproduces observed GSAT, suggesting that low-dimensional, observation-driven approaches are a viable and informative part of the climate model hierarchy. The ARDL model is parsimonious, accurate, and robust.

The ARDL specification is the discrete-time counterpart of the surface equation of a two-layer energy-balance model, obtained under A1 to A5. Its two forcing-related regressors, GSAT_{t-1} and the normalized ERF, are strongly collinear (r = 0.94, VIF = 8.27), as expected, and jointly capture the net effect of feedbacks in the climate system, including cloud effects, albedo changes and pattern effects (Andrews et al., 2018; Ceppi and Gregory, 2017), together with heat uptake by the deep ocean. Combined with the ERF exponent ν, the specification separates two physically distinct aspects of the forcing–temperature response: the state dependence of the response, carried by ν, and the magnitude of the net feedback at a given forcing level, carried jointly by β and ρ through the long-run multiplier $\theta = \beta\,(1-\rho)^{-1}$.

The forcing required to sustain a degree of warming falls as forcing accumulates, which is a statement that the system becomes progressively less effective at dissipating additional forcing. Equivalently, and more directly, the marginal response $\frac{\mathrm{dGSAT}}{\mathrm{dERF}}$ rises: 0.204 K per W m^{-2} at ERF = F_{ref}, 0.607 K per W m^{-2} at 2025 ERF and 0.890 K per W m^{-2} at ERF = 5.15 W m^{-2}. The factor by which it rises between any two forcing levels is $\left(\frac{\mathrm{ERF}_2}{\mathrm{ERF}_1}\right)^{\nu-1}$ and depends on the exponent alone, not on θ.

The two terms in the composite $\lambda + \kappa$ are both sinks of mixed-layer energy but are individually distinct. The radiative term λT_m captures energy lost to space and scales with the surface anomaly itself, whereas $\kappa(T_m - T_D)$ captures energy transferred downward into the deep ocean, which remains within the Earth system, and scales with the vertical temperature gradient. Under A1 the deep-ocean temperature is held fixed, so over the estimation period, λ and κ act jointly as restoring terms on the mixed layer and only their sum is identified. The reciprocal of the marginal response, $\frac{\mathrm{dERF}}{\mathrm{dGSAT}}$, is consequently a transient quantity rather than an equilibrium one, and this is the reason θ is not an equilibrium sensitivity. At equilibrium the

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

interlayer flux vanishes as T_D approaches T_m, leaving λ alone to balance the forcing. The $\frac{\mathrm{dERF}}{\mathrm{dGSAT}}$ estimates capture the transient response over the observed period and carry no implication for equilibrium behavior. The reciprocal $\frac{\mathrm{dERF}}{\mathrm{dGSAT}}$ is the empirical counterpart of the climate-feedback-plus-ocean-heat-uptake parameter: 1.65 W m^{-2} K^{-1} at 2025 ERF. Across sixteen CMIP5 models calibrated to the same two-layer structure at $\varepsilon = 1$, the climate feedback parameter has a multi-model mean of 1.13 W m^{-2} K^{-1} (standard deviation 0.31) and the interlayer ocean heat-exchange coefficient a mean of 0.73 W m^{-2} K^{-1} (standard deviation 0.18, range 0.5 to 1.2). Their sum is 1.86 W m^{-2} K^{-1} and spans 1.20 to 2.86 across models (Geoffroy et al., 2013).

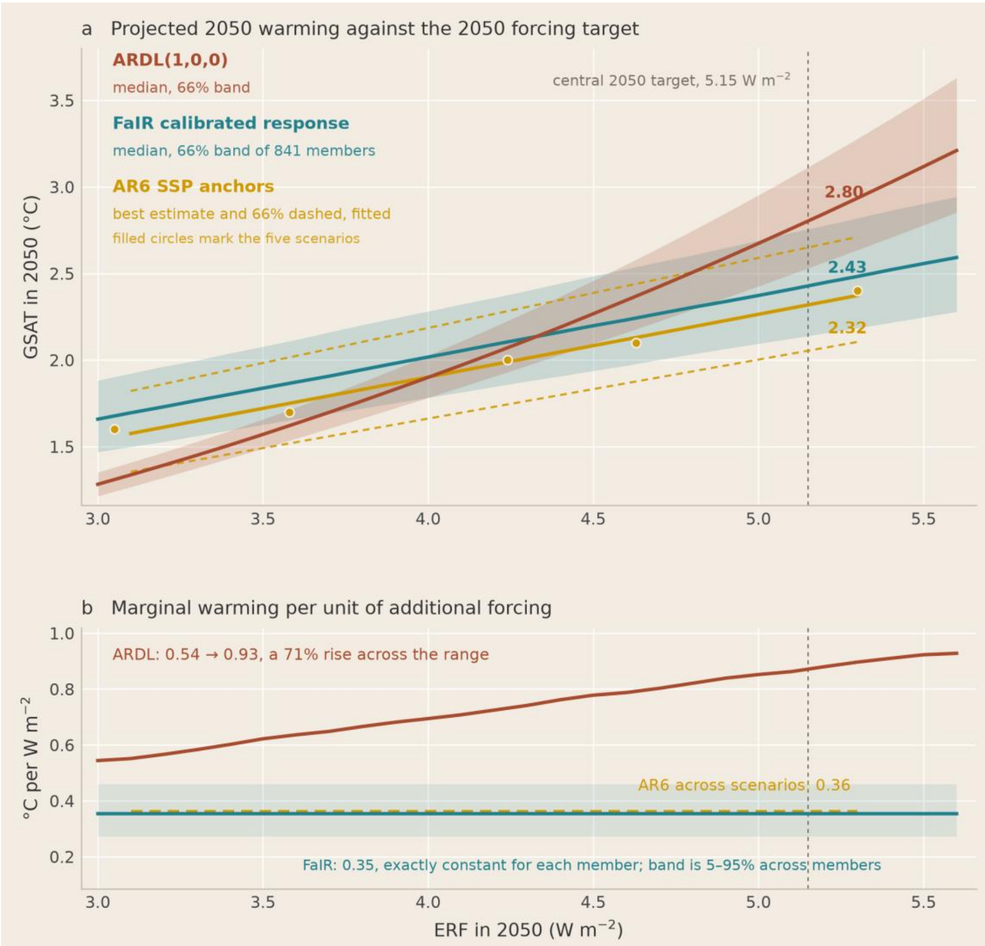
The model confirms that El Niño-Southern Oscillation (ENSO) variability, represented by ONI, contributes 0.108 °C (95 % UI: 0.080, 0.136 °C) per unit increase to short-term global temperature fluctuations. This effect size is consistent with previous observational estimates and demonstrates the importance of controlling for natural climate variability when assessing forced warming trends (Johnson and Kosaka, 2016). Inclusion of upper-ocean heat content (OHC) over 0 – 2000 m as an additional predictor in the ARDL model was explored, using the gridded objective analysis of Cheng et al. (2024). OHC and GSAT_{t-1} were strongly collinear and highly correlated over 1978 to 2025 ($R^2 = 0.878$). In models with both variables, OHC coefficients were unstable (VIF = 26.2) and did not improve in-sample fit once the added parameter is accounted for (BIC = +2.93) or out-of-sample forecast skill relative to specifications using GSAT_{t-1} alone. Given this redundancy and the risk of multicollinearity, GSAT_{t-1} was retained as a parsimonious proxy for thermal inertia and net feedback response. The ARDL coefficients are interpreted as capturing the combined effects of ocean heat uptake and other state-dependent feedbacks on GSAT.

The ARDL model projections of GSAT at progressively increasing ERF levels are compared to the projections from FaIR and the AR6 assessed estimates. At ERF levels of 4.0 – 4.3 the three models agree within < 0.1 °C. At the central 2050 target of 5.15 W m^{-2} the ARDL median is 2.80 °C, against 2.43 °C for the calibrated-constrained FaIR ensemble and 2.32 °C for the AR6 assessed best estimate. Panel (b) identifies the source. Marginal warming per unit of additional forcing is constant in both comparison responses, 0.35 °C per W m^{-2} for every FaIR member and 0.36 °C per W m^{-2} across the AR6 scenarios, whereas in the ARDL specification it rises from 0.54 to 0.93 °C per W m^{-2} across the same forcing range, a 71% increase. FaIR lies 0.11 °C above the AR6 line at every forcing level while remaining parallel to it, so the two differ in level and not in structure. The disagreement is therefore not between the ARDL specification and any one comparison, but between a response whose marginal sensitivity is fixed and one in which the marginal sensitivity depends on the forcing level already reached.

Figure 10. Comparison of 2050 GSAT estimates for ERF ranging from 3.0-5.5 W m^{-2}. Panel (a) shows GSAT in 2050 as a function of the prescribed 2050 ERF target for the ARDL(1,0,0) specification with ν = 1.90 (median and 66 % band, 20,000 draws), the calibrated-constrained FaIR ensemble (median and 66 % band, 841 members), and the AR6 assessed best estimate with its 66 % range, fitted across the five SSP

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

anchors. Panel (b) shows the corresponding marginal warming per unit of additional forcing. FaIR lies 0.11 °C above the AR6 line at every forcing level while remaining parallel to it.



The findings of this study concern the transient response of GSAT to ERF during a period of rapidly rising forcing. Prior use of a linear approximation of the GSAT ERF relationship in climate models was well motivated and is fully supported by the 20th century observational record. The argument advanced here is that forcing has rapidly increased over the last 48 years, and that this magnitude and rate of change has fundamentally changed the GSAT–ERF relationship. No estimate of equilibrium climate sensitivity, the transient climate response, or effective climate sensitivity is offered, and no result bears on the equilibrium or multi-century response to any given forcing level. Those quantities are assessed from multiple independent lines of evidence (Sherwood et al., 2020), and the historical record alone constrains them weakly, because the effective sensitivity inferred from the historical period differs systematically from the

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

equilibrium value (Gregory et al., 2020; Tokarska et al., 2020). The specification is estimated over forcing from 0.31 to 3.37 W m^{-2}, rising at 0.7 W m^{-2} per decade for the past 20 years, and is applied over a horizon of two to three decades. The ARDL specification is applicable at ERF > 0.3 Wm^{-2} and predicted ΔGSAT is expected to develop within three years. No physical mechanism for the observed change in the GSAT ERF relationship can be inferred from this analysis.

Reducing uncertainty in aerosol effective radiative forcing would be the single most valuable improvement to this class of climate models. Aerosols account for roughly three-quarters of the assessed variance in total ERF, which both dominates the width of the 2050 projections and biases the estimated exponent downward. In the current model, correcting for measurement error in ERF raises ν from 1.89 to 2.09. A narrower ERF record would also make it feasible to estimate the exponent jointly in the ARDL regression rather than impose it, which the collinearity between lagged GSAT and ERF currently prevents. The most important open question is whether climate state dependence survives when the deep-ocean temperature is allowed to evolve. Testing this question requires estimating the full two-equation cointegrated system rather than adding ocean heat content to a single equation. The mechanism for the non-linear relationship between GSAT and ERF also remains open. Observational data separating cloud feedback, the sea surface temperature pattern effect and sea-ice albedo would establish which of them carries the dependence on ERF level. Finally, applying this specification to output from ESMs and from simple two-layer energy balance models would show whether non-linearity emerges there as well, and fitting the same specification to land, ocean and hemispheric series would show where the nonlinearity resides.

Code and data availability

All data and code required to reproduce the results, tables and figures in this paper are archived at https://doi.org/10.5281/zenodo.22069204 (Smith, 2026). The archive contains the analysis input data, the Python scripts for every estimation and figure, the cached Monte Carlo ensembles required to regenerate Figures 7, 8 and 10, and the tabulated output. Code is released under the MIT License and data and figures under CC-BY-4.0. Global surface air temperature and effective radiative forcing series are from the Indicators of Global Climate Change 2025 dataset, version v2026.06.02 (https://doi.org/10.5281/zenodo.20499280). The FaIR comparison uses the 841-member constrained posterior of fair-calibrate v1.4.1 (https://doi.org/10.5194/gmd-17-8569-2024).

Author contributions

Craig R. Smith (CRS) is solely responsible for the design of this study, assembling the database, directing the analysis of the data, and the content of the manuscript.

Competing interests

CRS declares that he has no conflict of interest.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Acknowledgements

CRS acknowledges the use of Perplexity AI (https://www.perplexity.ai). Perplexity assisted with retrieving publicly available data and literature, writing the Python code used for the statistical and econometric analyses and for figure production, assembling and formatting the reference list, and editing the manuscript. All cited references were independently obtained and read by the author. All interpretation of results and all conclusions are the author's own.

Financial support

This research received no external funding.

References

Andrews, T., Gregory, J. M., and Webb, M. J.: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models, J. Climate, 28, 1630–1648, https://doi.org/10.1175/JCLI-D-14-00545.1, 2015.

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., Webb, M. J., Armour, K. C., Forster, P. M., and Titchner, H.: Accounting for changing temperature patterns increases historical estimates of climate sensitivity, Geophys. Res. Lett., 45, 8490–8499, https://doi.org/10.1029/2018GL078887, 2018.

Armour, K. C.: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks, Nat. Clim. Change, 7, 331–335, https://doi.org/10.1038/nclimate3278, 2017.

Bennedsen, M., Hillebrand, E., and Zhou Lykke, J.: Global temperature projections from a statistical energy balance model using multiple sources of historical data, J. Climate, 36, 6817–6838, https://doi.org/10.1175/JCLI-D-22-0460.1, 2023.

Bloch-Johnson, J., Rugenstein, M., Stolpe, M. B., Rohrschneider, T., Zheng, Y., and Gregory, J. M.: Climate sensitivity increases under higher CO_2 levels due to feedback temperature dependence, Geophys. Res. Lett., 48, e2020GL089074, https://doi.org/10.1029/2020GL089074, 2021.

Butler, J. H. and Montzka, S. A.: The NOAA Annual Greenhouse Gas Index (AGGI), NOAA Global Monitoring Laboratory, Boulder, CO, USA [data set], https://gml.noaa.gov/aggi/aggi.html (last access: 21 August 2026), 2025.

Caballero, R. and Huber, M.: State-dependent climate sensitivity in past warm climates and its implications for future climate projections, Proc. Natl. Acad. Sci. USA, 110, 14162–14167, https://doi.org/10.1073/pnas.1303365110, 2013.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M.: Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn., Monographs on Statistics and Applied Probability, Vol. 105, Chapman & Hall/CRC, Boca Raton, FL, USA, 488 pp., https://doi.org/10.1201/9781420010138, 2006.

Ceppi, P. and Gregory, J. M.: Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget, Proc. Natl. Acad. Sci. USA, 114, 13126–13131, https://doi.org/10.1073/pnas.1714308114, 2017.

Cheng, L., Abraham, J., Trenberth, K. E., Fasullo, J., Boyer, T., Mann, M. E., Zhu, J., Wang, F., Locarnini, R., Li, Y., Zhang, B., Tan, Z., Yu, F., Wan, L., Chen, X., Song, X., Liu, Y., Reseghetti, F., Simoncelli, S., Gouretski, V., Chen, G., Mishonov, A., and Reagan, J.: Another record: ocean warming continues through 2021 despite La Niña conditions, Adv. Atmos. Sci., 39, 373–385, https://doi.org/10.1007/s00376-022-1461-3, 2022.

Cheng, L., Pan, Y., Tan, Z., Zheng, H., Zhu, Y., Wei, W., Du, J., Yuan, H., Li, G., Ye, H., Gouretski, V., Li, Y., Trenberth, K. E., Abraham, J., Jin, Y., Reseghetti, F., Lin, X., Zhang, B., Chen, G., Mann, M. E., and Zhu, J.: IAPv4 ocean temperature and ocean heat content gridded dataset, Earth Syst. Sci. Data, 16, 3517–3546, https://doi.org/10.5194/essd-16-3517-2024, 2024.

Cook, J. R. and Stefanski, L. A.: Simulation-extrapolation estimation in parametric measurement error models, J. Am. Stat. Assoc., 89, 1314–1328, https://doi.org/10.1080/01621459.1994.10476871, 1994.

Dentener, F. J., Hall, B., and Smith, C. (Eds.): Annex III: Tables of historical and projected well-mixed greenhouse gas mixing ratios and effective radiative forcing of all climate forcers, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2139–2152, https://doi.org/10.1017/9781009157896.017, 2021.

Engle, R. F., Hendry, D. F., and Richard, J.-F.: Exogeneity, Econometrica, 51, 277–304, https://doi.org/10.2307/1911990, 1983.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's energy budget, climate feedbacks, and climate sensitivity, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, UK and New York, NY, USA, 923–1054, https://doi.org/10.1017/9781009157896.009, 2021.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Forster, P. M., Walsh, T., Smith, C., Lamb, W. F., Lamboll, R., Cassou, C., Hauser, M., Hausfather, Z., Lee, J.-Y., Palmer, M. D., von Schuckmann, K., Slangen, A. B. A., Szopa, S., Trewin, B., Yun, J., Gillett, N. P., Jenkins, S., Matthews, H. D., Raghavan, K., Ribes, A., Rogelj, J., Rosen, D., Zhang, X., Allen, M., Andrew, R. M., Atkinson, C., Betts, R. A., Bombelli, A., Burgess, S. N., Cheng, L., Claxton, H. E., Friedlingstein, P., Frölicher, T. L., Domingues, C. M., Gasser, T., Gregory, C. H., Hoesly, R. M., Huppmann, D., Ishii, M., Kadow, C., Karwat, A., Kennedy, J., Killick, R. E., Kovilakam, M. V. M., Krummel, P. B., Lan, X., Lamarque, J.-F., Liné, A., Martín-Míguez, B., Monselesan, D. P., Morice, C., Mühle, J., Mussak, P., Peters, G. P., Pirani, A., Pongratz, J., Rigby, M., Rohde, R., Savita, A., Seneviratne, S. I., Smith, S. J., Taha, G., Tassone, C., Thorne, P., Wells, C., Western, L. M., van der Werf, G. R., Wijffels, S. E., Zecchetto, M., Zhong, J., Zhang, X.-Y., Masson-Delmotte, V., and Zhai, P.: Indicators of Global Climate Change 2025: annual update of key indicators of the state of the climate system and human influence, Earth Syst. Sci. Data, 18, 3889–3933, https://doi.org/10.5194/essd-18-3889-2026, 2026.

Foster, G. and Rahmstorf, S.: Global warming has accelerated significantly, Geophys. Res. Lett., 53, e2025GL118804, https://doi.org/10.1029/2025GL118804, 2026.

Friedrich, T., Timmermann, A., Tigchelaar, M., Elison Timm, O., and Ganopolski, A.: Nonlinear climate sensitivity and its implications for future greenhouse warming, Sci. Adv., 2, e1501923, https://doi.org/10.1126/sciadv.1501923, 2016.

Fuller, W. A.: Measurement Error Models, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, USA, 440 pp., https://doi.org/10.1002/9780470316665, 1987.

Geoffroy, O., Saint-Martin, D., Olivié, D. J. L., Voldoire, A., Bellon, G., and Tytéca, S.: Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments, J. Climate, 26, 1841–1857, https://doi.org/10.1175/JCLI-D-12-00195.1, 2013a.

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D. J. L., and Tytéca, S.: Transient climate response in a two-layer energy-balance model. Part II: Representation of the efficacy of deep-ocean heat uptake and validation for CMIP5 AOGCMs, J. Climate, 26, 1859–1876, https://doi.org/10.1175/JCLI-D-12-00196.1, 2013b.

GISTEMP Team: GISS Surface Temperature Analysis (GISTEMP), version 4, NASA Goddard Institute for Space Studies [data set], https://data.giss.nasa.gov/gistemp/ (last access: 21 August 2026), 2025.

Gregory, J. M. and Andrews, T.: Variation in climate sensitivity and feedback parameters during the historical period, Geophys. Res. Lett., 43, 3911–3920, https://doi.org/10.1002/2016GL068406, 2016.

Gregory, J. M., Andrews, T., Ceppi, P., Mauritsen, T., and Webb, M. J.: How accurately can the climate sensitivity to CO_2 be estimated from historical climate change?, Clim. Dynam., 54, 129–157, https://doi.org/10.1007/s00382-019-04991-y, 2020.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Gregory, J. M. and Forster, P. M.: Transient climate response estimated from radiative forcing and observed temperature change, J. Geophys. Res., 113, D23105, https://doi.org/10.1029/2008JD010405, 2008.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophys. Res. Lett., 31, L03205, https://doi.org/10.1029/2003GL018747, 2004.

Hamilton, J. D.: A new approach to the economic analysis of nonstationary time series and the business cycle, Econometrica, 57, 357–384, https://doi.org/10.2307/1912559, 1989.

Hansen, J. E., Kharecha, P., Sato, M., Tselioudis, G., Kelly, J., Bauer, S. E., Ruedy, R., Jeong, E., Jin, Q., Rignot, E., Velicogna, I., Schoeberl, M. R., von Schuckmann, K., Amponsem, J., Cao, J., Keskinen, A., Li, J., and Pokela, A.: Global warming has accelerated: are the United Nations and the public well-informed?, Environment: Science and Policy for Sustainable Development, 67, 6–44, https://doi.org/10.1080/00139157.2025.2434494, 2025.

Hansen, J. E., Sato, M., Simons, L., Nazarenko, L. S., Sangha, I., Kharecha, P., Zachos, J. C., von Schuckmann, K., Loeb, N. G., Osman, M. B., Jin, Q., Tselioudis, G., Jeong, E., Lacis, A., Ruedy, R., Russell, G., Cao, J., and Li, J.: Global warming in the pipeline, Oxford Open Clim. Change, 3, kgad008, https://doi.org/10.1093/oxfclm/kgad008, 2023.

Hartmann, B. and Wendler, G.: The significance of the 1976 Pacific climate shift in the climatology of Alaska, J. Climate, 18, 4824–4839, https://doi.org/10.1175/JCLI3532.1, 2005.

Hausfather, Z., Drake, H. F., Abbott, T., and Schmidt, G. A.: Evaluating the performance of past climate model projections, Geophys. Res. Lett., 47, e2019GL085378, https://doi.org/10.1029/2019GL085378, 2020.

Hébert, R., Lovejoy, S., and Tremblay, B.: An observation-based scaling model for climate sensitivity estimates and global projections to 2100, Clim. Dynam., 56, 1105–1129, https://doi.org/10.1007/s00382-020-05521-x, 2021.

Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., and Vallis, G. K.: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing, J. Climate, 23, 2418–2427, https://doi.org/10.1175/2009JCLI3466.1, 2010.

Johnson, N. C. and Kosaka, Y.: The impact of eastern equatorial Pacific convection on the diversity of boreal winter El Niño teleconnection patterns, Clim. Dynam., 47, 3737–3765, https://doi.org/10.1007/s00382-016-3039-1, 2016.

Kass, R. E. and Raftery, A. E.: Bayes factors, J. Am. Stat. Assoc., 90, 773–795, https://doi.org/10.1080/01621459.1995.10476572, 1995.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Knutti, R. and Rugenstein, M. A. A.: Feedbacks, climate sensitivity and the limits of linear models, Philos. Trans. R. Soc. A, 373, 20150146, https://doi.org/10.1098/rsta.2015.0146, 2015.

Kramer, R. J., He, H., Soden, B. J., Oreopoulos, L., Myhre, G., Forster, P. M., and Smith, C. J.: Observational evidence of increasing global radiative forcing, Geophys. Res. Lett., 48, e2020GL091585, https://doi.org/10.1029/2020GL091585, 2021.

Lee, J.-Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J. P., Engelbrecht, F., Fischer, E., Fyfe, J. C., Jones, C., Maycock, A., Mutemi, J., Ndiaye, O., Panickal, S., and Zhou, T.: Future global climate: scenario-based projections and near-term information, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, UK and New York, NY, USA, 553–672, https://doi.org/10.1017/9781009157896.006, 2021.

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R., and Zyss, D.: Improvements in the GISTEMP uncertainty model, J. Geophys. Res.-Atmos., 124, 6307–6326, https://doi.org/10.1029/2018JD029522, 2019.

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., Mishonov, A. V., Reagan, J. R., Seidov, D., Yarosh, E. S., and Zweng, M. M.: World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010, Geophys. Res. Lett., 39, L10603, https://doi.org/10.1029/2012GL051106, 2012.

Loeb, N. G., Johnson, G. C., Thorsen, T. J., Lyman, J. M., Rose, F. G., and Kato, S.: Satellite and ocean data reveal marked increase in Earth's heating rate, Geophys. Res. Lett., 48, e2021GL093047, https://doi.org/10.1029/2021GL093047, 2021.

Meyssignac, B., Chenal, J., Loeb, N., Guillaume-Castel, R., and Ribes, A.: Time-variations of the climate feedback parameter λ are associated with the Pacific Decadal Oscillation, Commun. Earth Environ., 4, 241, https://doi.org/10.1038/s43247-023-00887-2, 2023.

Miller, A., Cayan, D., Barnett, T., Graham, N., and Oberhuber, J.: The 1976–77 climate shift of the Pacific Ocean, Oceanography, 7, 21–26, https://doi.org/10.5670/oceanog.1994.11, 1994.

Minière, A., von Schuckmann, K., Sallée, J.-B., and Vogt, L.: Robust acceleration of Earth system heating observed over the past six decades, Sci. Rep., 13, 22975, https://doi.org/10.1038/s41598-023-49353-1, 2023.

Narayan, P. K.: The saving and investment nexus for China: evidence from cointegration tests, Appl. Econ., 37, 1979–1990, https://doi.org/10.1080/00036840500278103, 2005.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

NOAA PSL: Niño 3.4 monthly sea surface temperature anomaly index, 1870 to present (nino34.long.anom.data), NOAA Physical Sciences Laboratory, Boulder, CO, USA [data set], https://psl.noaa.gov/data/timeseries/month/data/nino34.long.anom.data (last access: 21 August 2026), 2026.

Pan, Y., Cheng, L., Abraham, J., Trenberth, K. E., Reagan, J., Du, J., Wang, Z., Storto, A., von Schuckmann, K., Zhu, Y., Mann, M. E., Zhu, J., Wang, F., Yu, F., Locarnini, R., Fasullo, J., Huang, B., Graham, G., Yin, X., Gouretski, V., Zheng, F., Li, Y., Zhang, B., Wan, L., Chen, X., Wang, D., Feng, L., Song, X., Liu, Y., Reseghetti, F., Simoncelli, S., Chen, G., Zhang, R., Mishonov, A., Wei, W., Tan, Z., Li, G., Cao, L., Chen, L., Yuan, H., Lyu, K., Sulaiman, A., Mayer, M., Wang, H., Ma, Z., Bao, S., Yan, H., Liu, Z., Yang, C., Liu, X., Hausfather, Z., Gues, F., Song, X., Zhang, M., and Chen, L.: Ocean heat content sets another record in 2025, Adv. Atmos. Sci., https://doi.org/10.1007/s00376-026-5876-0, 2026.

Pesaran, M. H., Shin, Y., and Smith, R. J.: Bounds testing approaches to the analysis of level relationships, J. Appl. Econom., 16, 289–326, https://doi.org/10.1002/jae.616, 2001.

Pretis, F.: Econometric modelling of climate systems: the equivalence of energy balance models and cointegrated vector autoregressions, J. Econometrics, 214, 256–273, https://doi.org/10.1016/j.jeconom.2019.05.013, 2020.

Rahmstorf, S.: A semi-empirical approach to projecting future sea-level rise, Science, 315, 368–370, https://doi.org/10.1126/science.1135456, 2007.

Rohde, R. A. and Hausfather, Z.: The Berkeley Earth Land/Ocean Temperature Record, Earth Syst. Sci. Data, 12, 3469–3479, https://doi.org/10.5194/essd-12-3469-2020, 2020.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An assessment of Earth's climate sensitivity using multiple lines of evidence, Rev. Geophys., 58, e2019RG000678, https://doi.org/10.1029/2019RG000678, 2020.

Smith, C., Cummins, D. P., Fredriksen, H.-B., Nicholls, Z., Meinshausen, M., Allen, M., Jenkins, S., Leach, N., Mathison, C., and Partanen, A.-I.: fair-calibrate v1.4.1: calibration, constraining, and validation of the FaIR simple climate model for reliable future climate projections, Geosci. Model Dev., 17, 8569–8592, https://doi.org/10.5194/gmd-17-8569-2024, 2024.

Smith, C., Walsh, T., Gillett, N., Hauser, M., Krummel, P., Lamb, W., Lamboll, R., Mühle, J., Palmer, M., Ribes, A., Schumacher, D., Seneviratne, S., Slangen, A., Trewin, B., von Schuckmann, K., and Forster, P.: Indicators of Global Climate Change 2025, version v2026.06.02, Zenodo [data set], https://doi.org/10.5281/zenodo.20499280, 2026.

https://doi.org/10.5194/egusphere-2026-5141
Preprint. Discussion started: 31 August 2026

Smith, C. R.: Replication data and code for “Emergence of a non-linear global surface air temperature response to rapidly increasing effective radiative forcing”, Zenodo [code], https://doi.org/10.5281/zenodo.22069204, 2026.

Smith, S. J., van Aardenne, J., Klimont, Z., Andres, R. J., Volke, A., and Delgado Arias, S.: Anthropogenic sulfur dioxide emissions: 1850–2005, Atmos. Chem. Phys., 11, 1101–1116, https://doi.org/10.5194/acp-11-1101-2011, 2011.

Tokarska, K. B., Hegerl, G. C., Schurer, A. P., Forster, P. M., and Marvel, K.: Observational constraints on the effective climate sensitivity from the historical period, Environ. Res. Lett., 15, 034043, https://doi.org/10.1088/1748-9326/ab738f, 2020.

Van Loon, S., Rugenstein, M., and Barnes, E. A.: Observation-based estimate of Earth’s effective radiative forcing, Proc. Natl. Acad. Sci. USA, 122, e2425445122, https://doi.org/10.1073/pnas.2425445122, 2025.

von der Heydt, A. S., Köhler, P., van de Wal, R. S. W., and Dijkstra, H. A.: On the state dependency of fast feedback processes in (paleo)climate, Geophys. Res. Lett., 41, 6484–6492, https://doi.org/10.1002/2014GL061121, 2014.

von Schuckmann, K., Minière, A., Gues, F., Cuesta-Valero, F. J., Kirchengast, G., Adusumilli, S., Straneo, F., Ablain, M., Allan, R. P., Barker, P. M., Beltrami, H., Blazquez, A., Boyer, T., Cheng, L., Church, J., Desbruyeres, D., Dolman, H., Domingues, C. M., García-García, A., Giglio, D., Gilson, J. E., Gorfer, M., Haimberger, L., Hakuba, M. Z., Hendricks, S., Hosoda, S., Johnson, G. C., Killick, R., King, B., Kolodziejczyk, N., Korosov, A., Krinner, G., Kuusela, M., Landerer, F. W., Langer, M., Lavergne, T., Lawrence, I., Li, Y., Lyman, J., Marti, F., Marzeion, B., Mayer, M., MacDougall, A. H., McDougall, T., Monselesan, D. P., Nitzbon, J., Otosaka, I., Peng, J., Purkey, S., Roemmich, D., Sato, K., Sato, K., Savita, A., Schweiger, A., Shepherd, A., Seneviratne, S. I., Simons, L., Slater, D. A., Slater, T., Steiner, A. K., Suga, T., Szekely, T., Thiery, W., Timmermans, M.-L., Vanderkelen, I., Wjiffels, S. E., Wu, T., and Zemp, M.: Heat stored in the Earth system 1960–2020: where does the energy go?, Earth Syst. Sci. Data, 15, 1675–1709, https://doi.org/10.5194/essd-15-1675-2023, 2023.