Authors' Response to the Referees

EGUsphere

Manuscript ID: EGUSPHERE-2026-3813

Sunlit and Unlit: Kernel Ridge Regression for Dynamic Albedo Modeling and Prediction

Gavin K. Blair, Mohamad H. Kazma, Ahmad F. Taha, Ralf Bennartz, Sankaran Mahadevan

Response submitted: November 17, 2026

We thank the anonymous referees for their detailed comments on our manuscript. Below we provide a point-by-point response to every comment raised. We note the following:

- Referee comments are shown in blue boxes; authors' responses are shown in white boxes beneath them.
- Text highlighted in blue denotes new or substantially revised passages of the manuscript.
- The major changes made to the paper are included in the following textbox.

Summary of Changes

1. We revise the manuscript's title, abstract, case studies, and conclusions to emphasize the methodological contributions, predictive performance, and climate-relevant evaluation of the proposed framework.
2. We revise the introduction to situate EBMs within the broader literature on reduced-complexity climate modeling.
3. Most importantly, we introduce an **insolation-banded kernel ridge regression (KRR)** formulation that incorporates seasonal and geographical information. **It is shown to improve albedo predictive accuracy, and substantially reduces computational cost relative to the original KRR approaches.**
4. We replace the previous coupled kernel case study with an analysis of empirical top-of-atmosphere (TOA) shortwave radiative sensitivities. This study furthers the climate relevance of the paper by evaluating whether accurate state predictions are accompanied by agreement with radiative sensitivities as evidenced in ERA5, providing an additional criterion for assessing the coupled behavior of identified parameterizations.
5. We clarify mathematical notation for dimensional consistency and edit any typos in our initial submission.
6. In short, the revised paper now focuses on deriving and computing an albedo model with good prediction qualities compared to those available in the literature.
7. We revise the manuscript figures to maintain a consistent aesthetic.

In what follows, we respond to the individual comments by Referee 1 beginning on Page 3. We then respond to Referee 2 beginning on Page 9.

Response to Referee 1

General Comments (GC) by Referee 1

Comment GC.1

EBMs involve only a small portion of the physics contained in GCMs, so calling them simplified GCMs is missing the point. I think what the authors mean to refer to is Reduced Complexity Models, which have quite a bit of associated research and a model intercomparison project devoted to them (RCMIP).

The referee is right and we agree with the referee's assessment. The presumption of a stark EBM/GCM divide was not the correct framing for our literature review. We have therefore revised our introduction to situate EBMs within the broader literature on reduced-complexity climate modeling, emphasizing their physical interpretability and their utility for investigating climate response and feedback mechanisms.

Specifically, we have expanded the literature review to discuss representative approaches to reduced-complexity climate modeling. We additionally cite the Reduced Complexity Model Intercomparison Project (RCMIP) Phase 1 to provide a broader context for these approaches. The relevant revised passage reads as follows.

Reduced-complexity climate models (RCMs) and climate emulators support this objective through simplified physical representations, statistical approximations, or combinations of the two (Nicholls et al., 2020). These approaches address different tasks.

This brief overview then transitions into our discussion of simple theoretical models such as box climate models and the lessons which they can provide:

Among physical RCMs, energy balance models (EBMs) provide a physically interpretable framework for studying climate response. EBMs describe temperature through a balance among absorbed solar radiation, outgoing longwave radiation, and energy storage, with spatially resolved formulations additionally representing heat redistribution (North, 1975). Their representations range from a small number of interacting thermal compartments to finer temperature fields.

Comment GC.2

I know what uses RCMs can have for understanding basic climate response. But EBMs are so limited and so abstracted from GCMs that the specific purpose for using them needs to be stated, as well as a proposed way of parameterizing all of the processes that you're not explicitly representing.

We thank the referee for this observation and agree that the original manuscript did not sufficiently motivate the use of an EBM or clarify the treatment of unresolved physical processes. We have revised the introduction to better establish the scientific objectives of the proposed framework and clarify the role of its effective parameterizations. The paper has a much clearer objective now that goes beyond EBMs per se—its objective is to obtain a model that predicts albedo substantially better than current methods.

We emphasize that the purpose of the present work is not to reproduce the full physical complexity of comprehensive climate models, but to investigate whether the evolution of effective TOA albedo can be identified and coupled with a reduced-complexity model. The compact structure of EBMs provides a useful setting for isolating the influence of an identified albedo parameterization on temperature predictions and evaluating its associated feedback behavior. We have revised the introduction to emphasize the identification of dynamical surrogate models and their application to climate-relevant analyses:

The reduced physical structure of EBMs provides a useful framework for investigating the development and coupling of empirical parameterizations for otherwise unresolved processes. This perspective thus enables the identification of dynamical surrogate models for evaluating climate-relevant feedback and robustness under varying environmental conditions. A central challenge in developing such parameterizations lies in the representation of unresolved surface and radiative processes, particularly albedo.

We recognize that this representation does not independently resolve the physical mechanisms responsible for variations in albedo. In particular, the model does *not* explicitly simulate the evolution of snow cover, sea-ice extent, or melt ponds. We have retained this distinction in our discussion of model limitations:

Furthermore, the present formulation does not explicitly resolve the evolution of surface properties such as snow cover, sea ice extent, or melt ponds, which influence both albedo and effective heat capacity (Hunke et al., 2013; Krebs-Kanzow et al., 2021).

Comment GC.3

Geoffroy et al. (2015)'s land/sea contrast model is a good example of doing this. Without such physical grounding, you're left with a paper whose point is, essentially determining what top-of-atmosphere albedo "should" be for given climate data, even though we are able to both measure and model it quite well.

We agree that the original manuscript did not sufficiently establish the scientific purpose of identifying albedo dynamics. We have therefore revised the manuscript to better motivate the proposed parameterization and evaluate its behavior within a coupled climate model.

First, we introduce a new insolation-banded kernel ridge regression (KRR) formulation that incorporates additional information about seasonality and geographical location. Unlike the original formulations, which assigned predictors according to individual latitudes or selected regions, the new approach groups observations according to incoming solar radiation. The inclusion of the insolation trend $\dot{s}$ and sine of the latitude μ as predictor inputs further distinguishes between different seasonal and geographical conditions. In particular, the former distinguishes increasing from decreasing solar forcing, while the banded structure allows observations from different locations experiencing comparable insolation conditions to inform the same predictor. The training regimes for individual kernel predictors are exemplified in Fig. 1. This formulation is shown to be a much better predictor than the original version of the method, while also being computationally faster. The mathematical formulation of this method is included in the revised **Section**: **Insolation-Banded Kernel Regression** of the manuscript. We invite the referee to look at the updated Algorithm 1 as well which has been revised.

We re-simulated the entire case studies given this new approach. Here, in this document, we provide the referee with a representative subset of these simulations in Tab. 1 below.

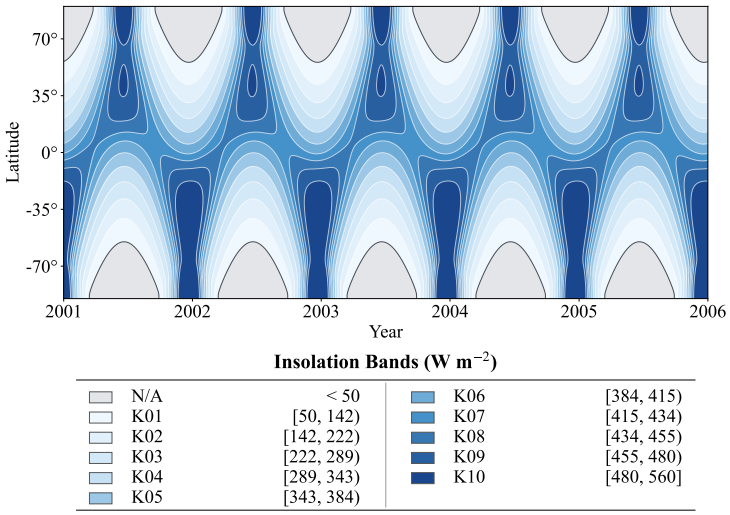

Figure 1: Training regimes for the insolation-banded KRR model identified by (φ_i, t_k) across all latitudes from January 1, 2001 to December 31, 2005. All tabulated intervals denote the selected insolation quantiles for $N_s = 10$ models. Unsupported regimes identify all $s_i^k < 50\,\mathrm{W/m^2}$ in gray.

Table 1: Computational efficiency and predictive accuracy of selected albedo parameterizations.

Model	$\overline{\text{RMSE}}$ (°C)	Albedo MAE	Time (s)
Constant	5.67	0.0385	17.5
RBF-72 Informed	**5.38**	0.0333	135.2
RBF-721 Naive	5.82	0.0405	1081.0
RBF-10 Banded	5.65	**0.0277**	**35.4**

Second, we introduce a new section, **Coupled Climate Sensitivity and Radiative Response** to investigate whether accurate historical predictions are accompanied by faithful representations of observed responses in climate. One way we attempt to display this is by calculating a diagnostic feedback parameter, $\lambda_\alpha^{\text{TOA}}$ in $\text{W m}^{-2}\,^{\circ}\text{C}^{-1}$, simplified from the form proposed in Roe et al. (2015). This parameter is obtained from the coupled EBM–KRR simulations against corresponding estimates derived from ERA5 over the 2010–2011 evaluation period. We kindly invite the referee to read the additional section in the paper, showing the value of this test.

The results indicate that the insolation-banded formulation most closely reproduces the ERA5 feedback profiles among the investigated KRR approaches. **In particular, it achieves the lowest RMSE** in both the Arctic and Antarctic regions. We show the results in both Fig. 2 and Tab. 2 below. It also shows that our previous KRR models struggled to recreate this observed feedback, demonstrating that accurate reconstruction of historical albedo alone is insufficient to establish the effectiveness of an identified parameterization. We emphasize that this diagnostic does *not* constitute a complete representation of ice-albedo feedback, nor does agreement with ERA5 establish that the underlying physical mechanisms have been recovered. It instead provides an additional criterion for evaluating the behavior of empirical albedo parameterizations beyond simple predictive accuracy.

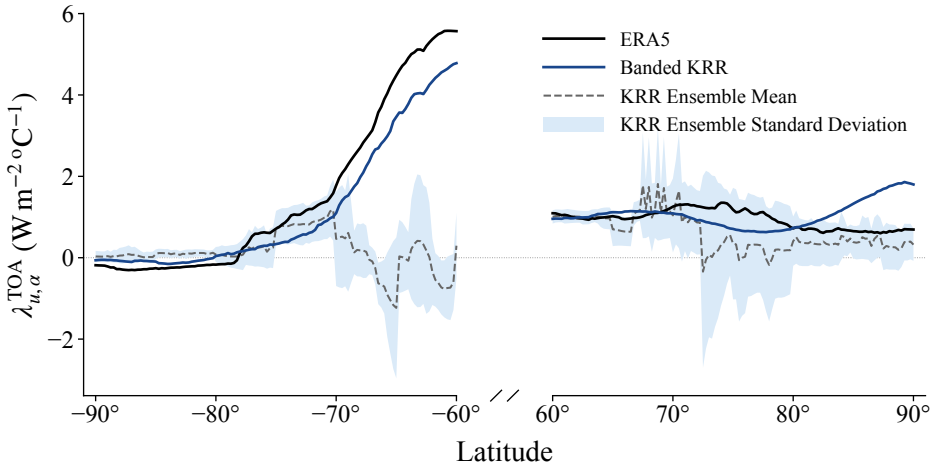

Figure 2: Annual mean latitudinal feedback profile of $\lambda_\alpha^{\text{TOA}}(\varphi_i)$ across ERA5, the insolation-banded model, and the ensemble of selected KRR predictors.

Table 2: Regionally averaged $\lambda_\alpha^{\mathrm{TOA}}$ and the area-weighted RMSE between the modeled and ERA5 feedback profiles. All values are in W m^{-2} °C^{-1}.

		Arctic		**Antarctic**	
Case-N	Model	$\lambda_\alpha^{\mathrm{TOA}}$	Error	$\lambda_\alpha^{\mathrm{TOA}}$	Error
RBF-36	Naïve	0.59	0.79	0.28	3.21
RBF-36	Informed	1.13	0.32	−0.45	4.09
RBF-72	Naïve	−0.19	1.87	−0.02	3.80
RBF-72	Informed	1.31	0.48	0.08	3.48
RBF-721	Naïve	0.96	0.95	0.65	3.30
RBF-10	Banded	**0.98**	**0.30**	**2.00**	**0.74**
ERA5	—	**1.03**	—	**2.61**	—

Comment GC.4

So essentially the premise the authors claim to be pursuing isn't all that interesting or useful. Your conclusion, that there are limits to albedo model identification and prediction, is a strawman argument – I don't know of anyone who would think that pure energy balance models should do a great job of reproducing real climate data beyond a very coarse aggregate. I will point out that the introduction does a decent job of reviewing work on EBMs. That in itself could be an interesting contribution.

We agree with the referee's assessment that the original conclusion placed undue emphasis on the limitations of albedo prediction in EBMs without properly establishing the contributions of the proposed framework. We have therefore substantially revised the abstract, Introduction, and conclusions to emphasize the development and evaluation of empirical albedo dynamics within a reduced-complexity climate model.
The revised manuscript incorporates the following principal contributions:

1. **Dynamic Albedo Representation:** We demonstrate that the evolution of effective TOA albedo can be empirically identified using a limited set of variables available within an EBM of the style in North (1975). The identified parameterizations are coupled with the physical temperature model and evaluated through daily simulations.
2. **Effective Prediction:** We introduce an insolation-banded KRR formulation that incorporates seasonal and geographical information into the identification procedure. This approach improves albedo predictive accuracy and substantially reduces computational cost relative to several earlier KRR formulations. Our analysis additionally identifies low-insolation conditions and seasonal transitions as important sources of prediction difficulty, motivating the revised formulation.
3. **Coupled Climate Response:** We extend our analysis beyond conventional assessments of prediction error, examining whether our models reproduce temperature-albedo sensitivities observed in ERA5. The new analysis conveys that agreement with historical albedo and temperature data does not necessarily imply comparable agreement with observed responses in ERA5. This motivates the use of complementary climate-relevant diagnostics when evaluating empirically identified parameterizations.

Given the paper's reframing from limitations to a proposition of satisfactory parameterizations, we changed the title to best reflect these changes to *Sunlit and Unlit: Kernel Ridge Regression for Dynamic Albedo Modeling and Prediction*.

Comment GC.5

I think ultimately this is the wrong journal for this sort of work. GMD is primarily focused on developing tools for new geoscience discovery, and I don't think this paper provides or enables new climate or Earth system science in any convincing capacity. Nor do I think that, given the framework the authors have set up, it can do that. This is an interesting formulation for using simple, physically-based models and constraining them with real-world data for, essentially, optimization. That strikes me as a better fit for a SIAM journal or something like that. And I suspect that if the authors submitted to a SIAM journal, they would have a more friendly review process that is focused on their methods, rather than their contributions to climate science.

We thank the referee for voicing their concern on journal fit, but we disagree that this is the wrong journal for this work. We recognize that the initial manuscript placed great emphasis on the mathematical identification of albedo dynamics without sufficiently demonstrating how the resulting parameterizations could contribute to reduced-complexity climate modeling. In response, the revisions described in GC.1–GC.4 were undertaken to address this concern.

The revised manuscript maintains the mathematical foundation while expanding the applicability of our initial proposal for model identification as denoted in GC.4. In addition, it introduces an insolation-banded formulation that improves computational efficiency and albedo predictive performance, while the new coupled feedback analysis evaluates whether the identified dynamics reproduce selected empirical relationships between temperature and the TOA shortwave radiative response. These experiments provide a climate-relevant assessment of the parameterizations beyond their ability to reconstruct historical albedo.

Several recent studies published in Geoscientific Model Development have demonstrated the continued development and application of energy balance models (EBMs) as computationally efficient frameworks for investigating climate dynamics. For example, Ziegler and Rehfeld (2021) introduced TransEBM, a transient two-dimensional EBM calibrated against reanalysis data, while Gunning et al. (2025) developed ZEMBA, a zonally averaged energy and moisture balance model incorporating parameterizations of albedo, sea ice, and atmospheric heat transport. The former provides many of the parameterizations and methods of model construction, while the latter is cited in our manuscript for its implementation of a particularly thorough physical parameterization of albedo. We hope the referee agrees with us on this, in light of the changes that we made and the reframing of the paper's scope and contributions.

Comment GC.6

On a minor note, there are lots of places where you want to use `\citep` instead of `\cite`.

We thank the referee for pointing this out. This is fixed in our updated submission.

Response to Referee 2

General Comments (GC) by Referee 2

Comment GC.1

I think there is an inconsistency in the temperature convention. In eq. (4), the outgoing longwave radiation is parameterised as

$$A(t) + B(u - 273.15),$$

where u is expressed in Kelvin. However, eq. (9) contains the term

$$A(t) + Bu.$$

I suspect that, switching between the two formulations, the authors are implicitly redefining the coefficients $A(t)$. This could be ambiguous for some readers and make not clear which choice is adopted in the numerical implementation.

We thank the referee for identifying this inconsistency. Equation (9) is edited in our updated manuscript as

$$\dot{\mathbf{u}}(t) = \mathrm{diag}(\mathbf{c}(t))^{-1}\big[\mathbf{s}(t) \odot (\mathbf{1} - \boldsymbol{\alpha}(t)) - A(t)\mathbf{1} - B(\mathbf{u}(t) - 273.15\mathbf{1}) + \mathbf{L}(t)\mathbf{u}(t)\big].$$

The definition of $\mathbf{g}$ is updated accordingly,

$$\mathbf{g}^k = \mathbf{s}^k \odot \left(\mathbf{1} - \boldsymbol{\alpha}^k\right) - A^k\mathbf{1} - B(\mathbf{u}^k - 273.15\mathbf{1}).$$

Comment GC.2

I have a question regarding the Crank-Nicolson scheme in eq. (14). For me, the Crank-Nicolson method for an ODE of the form $\dot{u}(t) = f(t, u(t))$ is given by

$$u^{k+1} = u^k + \frac{\Delta t}{2}\left(f(t_k, u^k) + f(t_{k+1}, u^{k+1})\right).$$

In the notations of the authors, writing $F^k := \omega^k + g^k$, the ODE for the temperature u is

$$\dot{u}(t) = \mathrm{diag}(c(t))^{-1} F(t, u(t)).$$

Thus, the Crank-Nicolson discretisation gives

$$u^{k+1} = u^k + \frac{\Delta t}{2}\left[\mathrm{diag}(c^k)^{-1} F^k + \mathrm{diag}(c^{k+1})^{-1} F^{k+1}\right].$$

Instead, eq. (14) uses

$$u^{k+1} = u^k + \frac{\Delta t}{2}\,\mathrm{diag}(c^k)^{-1}(F^k + F^{k+1}),$$

which effectively freezes the heat capacity over each time step. Since c depends on time (see Eq. (35)), I am not sure that this modification preserves the usual second-order accuracy of the Crank-Nicolson method. I have no objection to the discretisation in eq. (14) itself, but I would appreciate a justification or reference for the claimed second-order accuracy (line 226) in the present setting.

We thank the referee for raising this point. We provide a brief analysis below, divided between two parts. We first investigate how much and where the heat capacities change according to our thresholds and we then review whether our Crank-Nicolson scheme is second-order as we claimed in our initial draft. In summary, our current Crank-Nicolson model differs little from the referee's proposed scheme for the selected simulation horizon. Our claim of second-order accuracy is edited in the current manuscript to reflect first-order temporal error in our current assumption.

Our motivation for the proposed numerical integration scheme arose from a limitation in how heat capacities were evaluated. Whereas insolation has a self-contained model dependent only on the time of year and selected orbital parameterizations, our paper did not have such a convenient estimator for a parameter such as $\mathbf{c}(t)$. In this way, we assumed $\mathbf{c}(t) = \mathbf{c}(t_k)\ \forall t \in [t_k, t_{k+1})$ regardless of the temporal resolution using the values in Ziegler and Rehfeld (2021). Additionally, we assumed that for constant land and ocean compositions (*lsm* in ERA5) and constant snow cover (*sd*) as of 00:00 on January 1, 2001, this would reserve changes in heat capacity to only latitudes with varying sea ice concentrations (*siconc*). We show this change in expected heat capacity in Fig. 3. This plot shows that the heat capacities are subject to change at select subpolar latitudes and that their change is, in some instances, significant. We perform an independent EBM run utilizing the Crank-Nicolson formulation proposed by the referee, but the discrepancy between the two runs is small in our daily projections with $\overline{\mathrm{RMSE}}$ changing by only thousandths of a degree Celsius for the proposed simulation horizon.

To investigate the error polynomial corresponding to our scheme, we first perform a Crank Nicolson simulation of the form proposed by the referee using $\Delta t = 7.5$ minutes. We then evaluate the RMSE between the Crank-Nicolson models operating at various Δt and this reference simulation. The results may be viewed in Fig. 5 where additional error may be modeled between the reference model u_{ref} and the coarser model $u_{\Delta t}$ as

$$E(u_{\Delta t}) = \sqrt{\frac{1}{MN}\sum_{i=1}^{N}\sum_{k=1}^{M-1}\left(u_{\Delta t}(\varphi_i, t_k) - u_{\mathrm{ref}}(\varphi_i, t_k)\right)}.$$

Our current scheme is in line with first-order error as Δt increases, but we note that our current scheme is closer to the reference at $\Delta t = 1$ day in comparison to the proposed alternative as in Fig. 4. Our manuscript is edited to reflect this change in the derivation of eq. (14):

This form makes the assumption that heat capacity is constant throughout the interval, decoupling the problem of temperature prediction from that of underlying surface state prediction. Thus, Crank-Nicolson's second-order accuracy in time is true if and only if $\mathbf{c}^k = \mathbf{c}^{k+1}$. The failure of this condition renders the scheme first-order accurate in time.

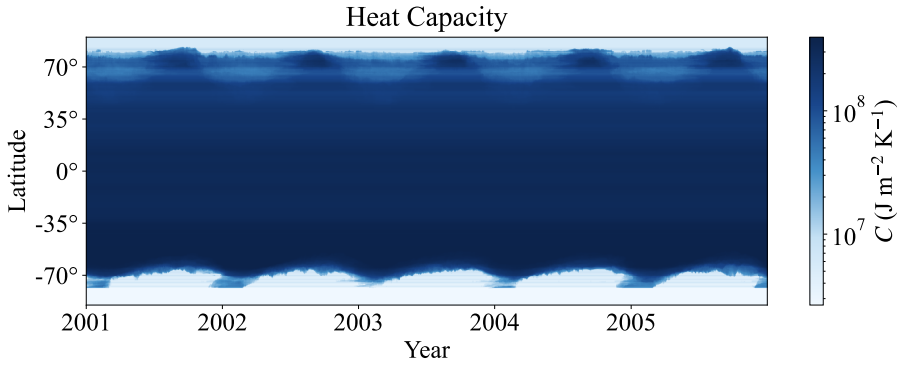

Figure 3: Projected heat capacity (J m^{-2} K^{-1}) at each latitude evaluated from January 1, 2001 through December 31, 2005.

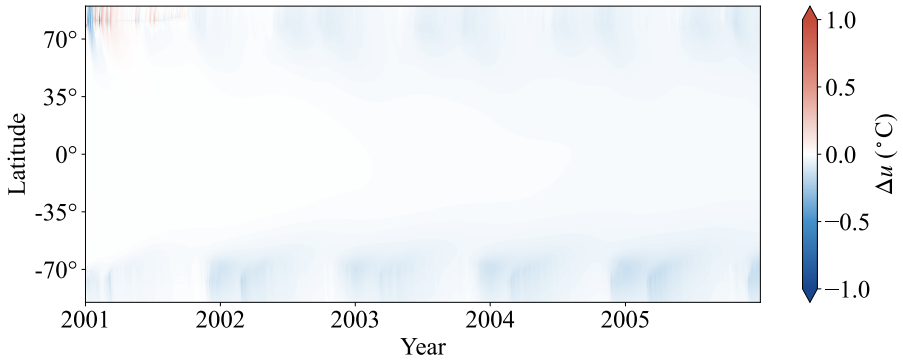

Figure 4: Projected temperature difference between proposed numerical integration scheme and current integration scheme over January 1, 2001 to December 31, 2005 simulation horizon.

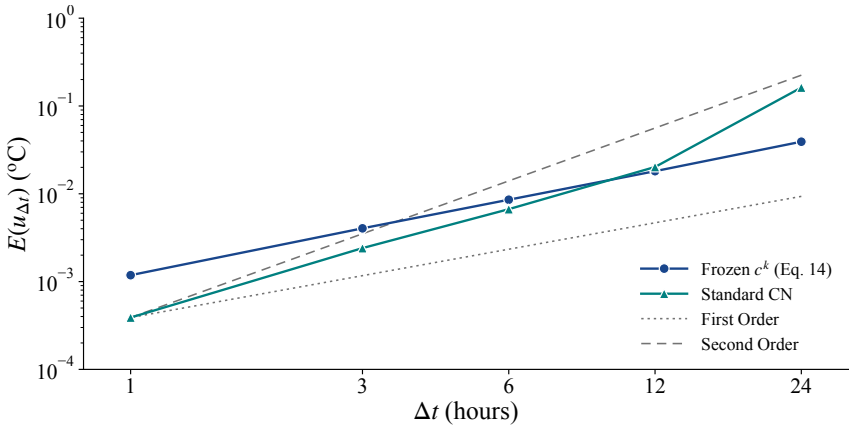

Figure 5: Additional Error contributed by selected Δt between numerical integration schemes. All errors taken in reference to standard Crank-Nicolson operating at $\Delta t = 7.5$ minutes.

Comment GC.3

In Section 2.2.2, the authors assume that $\Delta\mu = \mu_{i+1} - \mu_i$ is constant (lines 238–240). However, immediately afterwards (lines 241–244), they state that temperature and radiative data are evaluated on a uniformly spaced latitude grid, with constant $\Delta\varphi = \varphi_{i+1} - \varphi_i$. Since $\mu = \sin(\varphi)$, these two assumptions cannot hold simultaneously on the same grid. Can you please clarify this point?

The referee is indeed justified in raising this point. Our model currently assumes that all absorbed shortwave radiation, outgoing longwave radiation, and temperature change vectors operate on a constant $\Delta\varphi$ assumption. The only distinction, however, arises in the formulation of $\boldsymbol{\omega} = \mathbf{L}\mathbf{u}$, which operates on the assumption of constant $\Delta\mu$. We investigate an alternative form of $\boldsymbol{\omega} = \mathbf{L}_\varphi\mathbf{u}$, where the temperature values are defined on a uniformly spaced latitude grid ordered from north to south, with $\varphi_1 = \pi/2$, $\varphi_N = -\pi/2$, and $\Delta\varphi = \varphi_i - \varphi_{i+1} > 0$. The interior cell boundaries may then be written as

$$\varphi_{i+1/2} = \frac{\varphi_i + \varphi_{i+1}}{2}, \qquad i = 1, \ldots, N-1,$$

with exterior boundaries $\varphi_{1/2} = \pi/2$ and $\varphi_{N+1/2} = -\pi/2$. The spherical area weight of cell i is

$$w_i = \sin\varphi_{i-1/2} - \sin\varphi_{i+1/2}, \qquad i = 1, \ldots, N.$$

The entries of $\boldsymbol{\omega}^k$ are then given by

$$\omega_i^k = \begin{cases} \dfrac{d_1^k}{w_1}(u_2^k - u_1^k), & i = 1, \\ -\dfrac{d_{N-1}^k}{w_N}(u_N^k - u_{N-1}^k), & i = N, \\ \dfrac{d_i^k(u_{i+1}^k - u_i^k) - d_{i-1}^k(u_i^k - u_{i-1}^k)}{w_i}, & 2 \le i \le N-1, \end{cases}$$

where the interface exchange coefficients are

$$d_i^k = \frac{D_{i+1/2}^k \cos\varphi_{i+1/2}}{\Delta\varphi}, \qquad i = 1, \ldots, N-1.$$

Zero flux is imposed at both polar boundaries. Consequently, the discretization satisfies

$$\mathbf{L}_\varphi^k\mathbf{1} = \mathbf{0}, \qquad \mathbf{w}^\mathsf{T}\mathbf{L}_\varphi^k = \mathbf{0},$$

where $\mathbf{w} = [w_1, \ldots, w_N]^\mathsf{T}$. While indeed this model remains consistent with its spacing assumption, it is found to produce larger predictive errors relative to that of the current EBM utilized in the manuscript. From 2001–2005, our current EBM produces a mean RMSE of 5.19°C, while the tuned alternative produces a mean RMSE of 6.01 °C. We compare our current model's output and that of the proposed alternative in Figs. 6 and 8. In comparison to our current implementation, this alternative consistently underestimates temperatures in the Southern Hemisphere, while producing largely similar error modes to those of our existing EBM. **Given that the focus of this study is one of parameterizing albedo and evaluating its predictive capabilities subject to partial state prediction by the EBM, we opt for the less erroneous of the two predictors in the main text.**

We have also considered the alternative of evaluating all terms of the 1D EBM on a constant $\Delta\mu$ assumption; however, this spacing is less preferable for two reasons. First, the weighting of uniform $\Delta\mu$ samples more sparsely towards the polar latitudes, where our albedo prediction methods are of particular concern. Second, as we state in the manuscript, the selected ERA5 data operates on the assumption of constant $\Delta\varphi$, requiring interpolation on the reference data.

Our current revision of the manuscript makes this discrepancy clear, where we add **Appendix D** titled **Discretized Modeling of Meridional Diffusion**. In this appendix, we include the derivation of the proposed approximation, the error mode as shown in Fig. 6(a), and its corresponding sensitivity curves and parameterizations in Fig. 7 and Tab. 3.

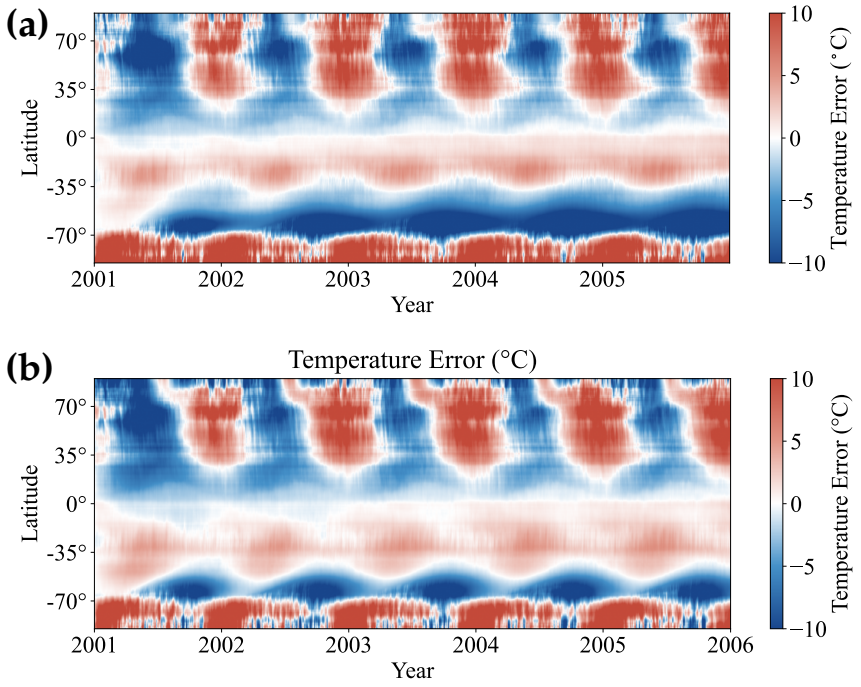

Figure 6: Tuned EBM performances over the 2001–2005 evaluation period. Model forms differ in their implementation of $\boldsymbol{\omega}$ operating on assumptions of (a) constant $\Delta\varphi$ and (b) constant $\Delta\mu$.

Table 3: Comparison of parameterizations between the original EBM assuming the form of $\boldsymbol{\omega}$ as used in the main text and the proposed EBM assuming $\boldsymbol{\omega} = \mathbf{L}_\varphi \mathbf{u}$. Both models are tuned using a daily time step over 2001–2005. Effective regional diffusion constants are expressed as $\eta'_n = D_0 \eta_n$.

Parameter	Units	Original EBM	Proposed EBM
A_{ref}	W/m^2	180.34	199.35
B	$\text{W}/(\text{m}^2\,^{\text{o}}\text{C})$	1.62	2.20
w_c	–	1.71	2.27
η'_1	$\text{W}/(\text{m}^2\,\text{K})$	0.311	0.1947
η'_2	$\text{W}/(\text{m}^2\,\text{K})$	0.506	0.6785
η'_3	$\text{W}/(\text{m}^2\,\text{K})$	1.830	1.7700
η'_4	$\text{W}/(\text{m}^2\,\text{K})$	1.324	1.7700
η'_5	$\text{W}/(\text{m}^2\,\text{K})$	0.659	0.8201
η'_6	$\text{W}/(\text{m}^2\,\text{K})$	0.866	0.4956
$\overline{\text{RMSE}}$	$^{\text{o}}\text{C}$	5.19	6.01

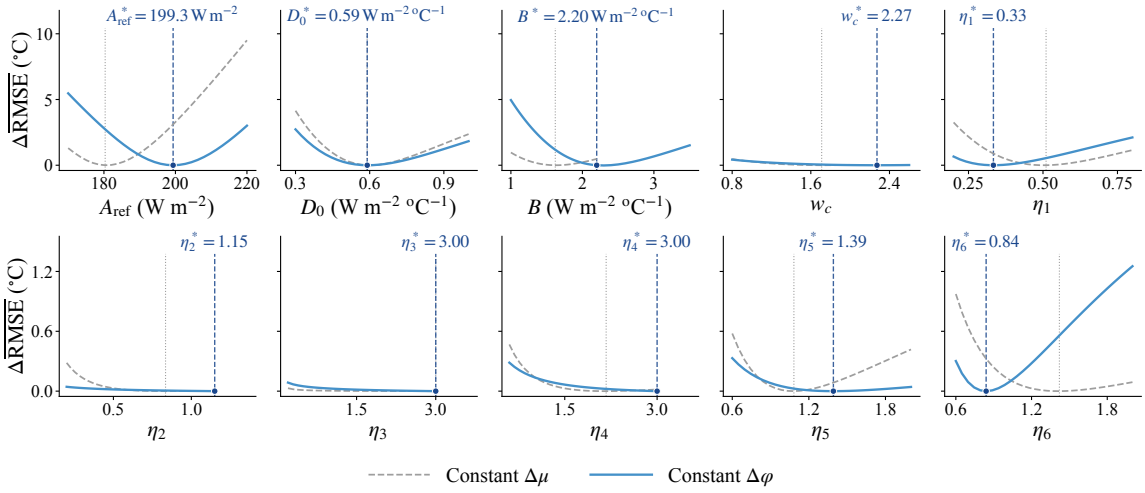

Figure 7: Sensitivity curves for EBMs utilizing assumptions of $\boldsymbol{\omega}$ operating on constant $\Delta\mu$ in gray and $\Delta\varphi$ in blue.

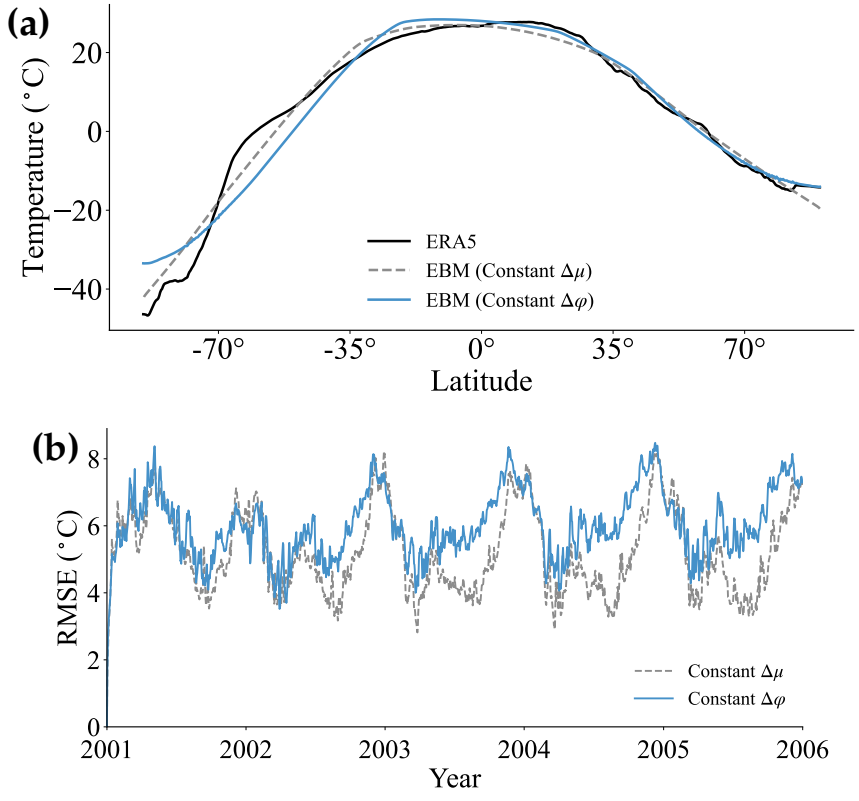

Figure 8: Performance of the 1D EBM utilizing $\boldsymbol{\omega} = \mathbf{L}_\varphi\mathbf{u}$ over the 2001–2005 evaluation period. Panel (a) compares the latitudinal temperature prediction profiles averaged over the same period. The model using $\boldsymbol{\omega} = \mathbf{L}\mathbf{u}$ as in the manuscript is shown in gray, whereas the model using $\boldsymbol{\omega} = \mathbf{L}_\varphi\mathbf{u}$ is shown in blue. Panel (b) reports the temperature RMSE over the same evaluation period.

Comment GC.4

I think that only $D_0\eta_n$ can be identified from temperature observations. Indeed, the diffusion coefficient D is defined in eq. (13) as

$$D(\mu, t_k) = D_0 \left(\sum_{n=1}^{N_d} \eta_n \chi_n(\mu) \right) \left(1 + \frac{\mu}{2} \cos \nu^k \right).$$

I think that, at the moment, D_0 and η_n are not uniquely identifiable. Indeed, for any $a > 0$, the transformation

$$D_0 \mapsto aD_0, \qquad \eta_n \mapsto \frac{\eta_n}{a}, \quad n = 1, \ldots, N_d,$$

keeps $D(\mu, t_k)$ unchanged, and shows that it is not possible to identify D_0 and η_n uniquely. I think this is possible, provided that some additional conditions (constraints or normalisation) are added.

We agree with the referee's assessment. As proposed, this formulation uniquely identifies only the individual regional parameterizations. In the manuscript, we employ a sequential one-at-a-time (OAT) approach, where we first hold $D_0 = 0.6$ and evaluate all of the regional multipliers η_n for the selected bands. We first tune the nine parameters (all except D_0) for their optimal values. Then given the proportions of each η_n, we begin to vary D_0, rescaling the diffusivity globally. However, we recognize that this procedure does not, by itself, eliminate the ambiguity identified by the referee. Our revisions denote that D_0 and its corresponding curve in Fig. 4 are intended to display a sensitivity analysis of temperature by the global scaling of diffusive strength. We make this distinction clear in our manuscript:

Herein, discussions of D_0 and η_n are intended to distinguish the global scaling of diffusive strength, and the individual regional strengths, respectively. We define $\eta'_n = D_0\eta_n$ to be the effective regional diffusion coefficient. In total, this problem entails the selection of nine parameters.

We then clarify our tuning procedure:

For parameters corresponding to the model's heat transport, D_0 and its corresponding regional multipliers η_n, we directly compare our results to those of Ziegler and Rehfeld (2021), evaluating the proposed diffusion coefficient in (14) by latitude and time of year. We initially hold $D_0 = 0.6$ and tune all other parameters. Once complete, we vary D_0 while holding the regional multipliers fixed to investigate the model's sensitivity to changes in global diffusion strength, preserving the relative regional distribution of diffusivity.

We then edit Tab. 1 in the manuscript to follow Tab. 4 below.

Table 4: Optimal OAT parameterizations and 2001–2005 mean RMSE across specified Δt. Effective regional diffusion coefficients $\eta'_n = D_0\eta_n$ represent the diffusion strength in the n-th region.

Parameter	**Units**	**1 hour**	**1 day**	**3 days**	**6 days**
A_{ref}	W/m^2	175.17	180.34	180.34	180.34
B	$\text{W/(m}^2\,^\circ\text{C)}$	1.66	1.62	1.62	1.62
w_c	–	2.00	1.71	1.71	1.67
η'_1	$\text{W/(m}^2\,\text{K)}$	0.288	0.311	0.301	0.301
η'_2	$\text{W/(m}^2\,\text{K)}$	0.237	0.506	0.490	0.490
η'_3	$\text{W/(m}^2\,\text{K)}$	1.920	1.830	1.770	1.770
η'_4	$\text{W/(m}^2\,\text{K)}$	1.920	1.324	1.280	1.280
η'_5	$\text{W/(m}^2\,\text{K)}$	1.030	0.659	0.637	0.637
η'_6	$\text{W/(m}^2\,\text{K)}$	1.280	0.866	0.838	0.838
$\overline{\text{RMSE}}$	$^\circ$C	7.23	5.19	5.19	5.24

Comment GC.5

Related to the previous point, how are the coefficients η_n and D_0 determined in Table 1 and Figure 4?

We answer this in GC.4.

Comment GC.6

Typo in eq. (24). The last component of the vector on the RHS contains $\hat{y}_1$, but I think it should be $\hat{y}_N$.

This entry is changed.

Comment GC.7

Possible incosistency in the construction of the regional KRR predictors in Section 2.3.4 and Algorithm 1. Let $R_j = \{i : a(i) = j\}$ denote the set of latitudes belonging to the j-th region. Algorithm 1 constructs the regional input vectors by averaging the individual inputs, so that

$$\bar{\mathbf{v}}_j^k = \frac{1}{|R_j|} \sum_{i \in R_j} \mathbf{v}_{\mathrm{inf},i}^k. \tag{1}$$

However, in eq. (23) it is written that

$$\mathbf{v}_{\mathrm{inf},1}^k, \mathbf{v}_{\mathrm{inf},N}^k \in \mathbb{R}^5, \qquad \mathbf{v}_{\mathrm{inf},i}^k \in \mathbb{R}^7, \quad 2 \le i \le N-1.$$

Thus, if a region contains both a polar grid point and an interior grid point, the average in (1) is not well defined, since the vectors have different dimensions. This seems relevant, for instance, when $N_r = 72$ and $N = 721$, so that each region contains approximately ten latitudes.
Moreover, the subsequent kernel evaluation in eqs. (18)–(21) requires the local input vector to have the same dimension as the corresponding regional training vectors.
Could the authors clarify how the polar grid points are handled in the regional averaging and kernel evaluation?

We thank the referee for catching this inconsistency. Our previous draft represented the boundary input vectors $\mathbf{v}_{\mathrm{inf},1}^k, \mathbf{v}_{\mathrm{inf},N}^k \in \mathbb{R}^5$ to depict the distinct entries corresponding to those predictions. Our manuscript is updated to reflect the inputs as they are implemented in our code:

For latitude φ_i, the informed input vector is defined as

$$\mathbf{v}_{\mathrm{inf},i}^k = \begin{cases} [u_i^k, u_i^k, u_{i+1}^k, \alpha_i^k, \alpha_i^k, \alpha_{i+1}^k, s_i^k]^\top, & i = 1, \\ [u_{i-1}^k, u_i^k, u_i^k, \alpha_{i-1}^k, \alpha_i^k, \alpha_i^k, s_i^k]^\top, & i = N, \\ [u_{i-1}^k, u_i^k, u_{i+1}^k, \alpha_{i-1}^k, \alpha_i^k, \alpha_{i+1}^k, s_i^k]^\top, & \text{else.} \end{cases}$$

Thus, $\mathbf{v}_{\mathrm{inf},i}^k \in \mathbb{R}^7$ for all latitudes. The repetition of u_i^k, α_i^k in both $\mathbf{v}_{\mathrm{inf},1}^k$ and $\mathbf{v}_{\mathrm{inf},N}^k$ is used to maintain a consistent dimension between all interior and boundary vectors $\mathbf{v}_{\mathrm{inf},i}^k$.

Whereas interior latitudes have the benefit of neighboring features in $\varphi_{i\pm1}$, the boundaries in the problem are limited to only two of the three. We implement repetition of the polar latitude's albedo and temperature features to maintain consistent dimensions and keep an interpretable polar boundary condition during training. This maintains the dimension of the kernel vector during regional averaging. Evaluation of the kernel follows from the updated dimension.

Comment GC.8

I have a question regarding the definition of the albedo error in Section 3.1, immediately after eq. (36). The MAE sums over all N latitudes. However, Section 2.4.3 explains that observed albedo is undefined during polar night and that α_{PN} is only a numerical placeholder.
Are these undefined observations excluded when computing the reported errors? For instance, if I_k denotes the set of latitudes with valid albedo observations at time t_k, is the intended definition

$$\mathrm{MAE}(t_k) = \frac{1}{|I_k|} \sum_{i \in I_k} \left|\alpha^k_{\mathrm{true},i} - \alpha^k_i\right| ?$$

Since the number of valid albedo observations varies with time, could the authors clarify how these missing observations are treated in the computation of MAE, and whether the normalisation $\frac{1}{|I_k|}$ accounts for the number of valid observations?

We agree with the referee's assessment and proposed notation. Our calculation of mean absolute error (MAE) excludes all observations of polar night, excluding any and all application of α_{PN} from the averaging. This change is reflected in our manuscript:

Albedo error metrics are given by mean absolute error (MAE) in the same manner:

$$\mathrm{MAE}(t_k) = \frac{1}{|\mathcal{I}_k|} \sum_{i \in \mathcal{I}_k} \left|\alpha^k_{\mathrm{true},i} - \alpha^k_i\right|, \qquad \overline{\mathrm{MAE}} = \sum_{k=1}^{M} \frac{\mathrm{MAE}(t_k)}{M},$$

where $\mathcal{I}_k$ denotes the set of latitudinal observations with defined values of albedo at time t_k, and $|\mathcal{I}_k|$ denotes its corresponding cardinality.

Works Cited

References

Olivier Geoffroy, David Saint-Martin, and Aurore Voldoire. Land-sea warming contrast: the role of the horizontal energy transport. *Climate Dynamics*, 45(11–12):3493–3511, 2015. doi: 10.1007/s00382-015-2552-y.

D. F. J. Gunning, K. H. Nisancioglu, E. Capron, and R. S. W. van de Wal. Zemba v1.0: an energy and moisture balance climate model to investigate quaternary climate. *Geoscientific Model Development*, 18(9):2479–2508, 2025. doi: 10.5194/gmd-18-2479-2025. URL https://gmd.copernicus.org/articles/18/2479/2025/.

Elizabeth C. Hunke, David A. Hebert, and Olivier Lecomte. Level-ice melt ponds in the Los Alamos sea ice model, CICE. *Ocean Modelling*, 71:26–42, 2013. doi: 10.1016/j.ocemod.2012.11.008.

Uta Krebs-Kanzow, Paul Gierz, Christian B. Rodehacke, Shan Xu, Hu Yang, and Gerrit Lohmann. The diurnal energy balance model (debm): a convenient surface mass balance solution for ice sheets in earth system modeling. *The Cryosphere*, 15:2295–2313, 2021. doi: 10.5194/tc-15-2295-2021.

Z. R. J. Nicholls, M. Meinshausen, J. Lewis, R. Gieseke, D. Dommenget, K. Dorheim, C.-S. Fan, J. S. Fuglestvedt, T. Gasser, U. Golüke, P. Goodwin, C. Hartin, A. P. Hope, E. Kriegler, N. J. Leach, D. Marchegiani, L. A. McBride, Y. Quilcaille, J. Rogelj, R. J. Salawitch, B. H. Samset, M. Sandstad, A. N. Shiklomanov, R. B. Skeie, C. J. Smith, S. Smith, K. Tanaka, J. Tsutsui, and Z. Xie. Reduced complexity model intercomparison project phase 1: introduction and evaluation of global-mean temperature response. *Geoscientific Model Development*, 13(11):5175–5190, 2020. doi: 10.5194/gmd-13-5175-2020. URL https://gmd.copernicus.org/articles/13/5175/2020/.

Gerald R. North. Theory of energy-balance climate models. *Journal of the Atmospheric Sciences*, 32(11):2033–2043, 1975. doi: 10.1175/1520-0469(1975)032⟨2033:TOEBCM⟩2.0.CO;2.

Gerard H. Roe, Nicole Feldl, Kyle C. Armour, Yen-Ting Hwang, and Dargan M. W. Frierson. The remote impacts of climate feedbacks on regional climate predictability. *Nature Geoscience*, 8(2):135–139, 2015. doi: 10.1038/ngeo2346. URL https://doi.org/10.1038/ngeo2346.

Elisa Ziegler and Kira Rehfeld. Transebm v. 1.0: description, tuning, and validation of a transient model of the earth's energy balance in two dimensions. *Geoscientific Model Development*, 14:2843–2866, 2021. doi: 10.5194/gmd-14-2843-2021. URL https://gmd.copernicus.org/articles/14/2843/2021/.