Sensitivity Analysis for Climate Science with Generative Flow Models

Alex Dobra[*]
Dept. of Engineering
University of Oxford

Jakiw Pidstrigach
Dept. of Statistics
University of Oxford

Tim Reichelt
Dept. of Physics
University of Oxford

Paolo Fraccaro
IBM

Johannes Jakubik
IBM

Anne Jones
IBM

Christian Schroeder de Witt
Dept. of Engineering
University of Oxford

Philip Torr
Dept. of Engineering
University of Oxford

Philip Stier
Dept. of Physics
University of Oxford

Abstract

Sensitivity analysis is a cornerstone of climate science, essential for understanding phenomena ranging from storm intensity to long-term climate feedbacks. However, computing these sensitivities using traditional physical models is often prohibitively expensive in terms of both computation and development time. While modern AI-based generative models are orders of magnitude faster to evaluate, computing sensitivities with them remains a significant bottleneck. This work addresses this challenge by applying the adjoint state method for calculating gradients in generative flow models. We apply this method to the cBottle generative model, trained on ERA5 and ICON data, to perform sensitivity analysis of any atmospheric variable with respect to sea surface temperatures. We quantitatively validate the computed sensitivities against the model's own outputs. Our results provide initial evidence that this approach can produce reliable gradients, reducing the computational cost of sensitivity analysis from weeks on a supercomputer with a physical model to hours on a GPU, thereby simplifying a critical workflow in climate science. The code can be found at `https://github.com/Kwartz18/cbottle_adjoint_sensitivity`.

1 Introduction

Sensitivity analysis is widely applied throughout climate science, e.g. for investigating the drivers of storms [1], the importance of observations [2] or understanding climate feedbacks [3, 4]. However, conducting sensitivity analysis with classical physical simulation models is prohibitively expensive, relying either on deriving and implementing adjoints by hand [5, 6, 7] or running finite difference approximations [3]. For example, a standard approach for estimating the response of the atmosphere to spatial changes in temperature fields relies on Green's function method which requires executing an expensive General Circulation Model (GCM) for thousands of model years which can take weeks, even on a supercomputer [8, 9, 10, 11, 12, 3, 4].

Recently, there has been an explosion of development of large AI models for weather and climate. While expensive to train, they are significantly faster and differentiable by default, warranting an

[*]Corresponding author: alex.dobra@robots.ox.ac.uk

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2025.

investigation into their potential to simplify the sensitivity analysis workflow. Modern AI models are developed in frameworks like PyTorch [13] and JAX [14], which are equipped with automatic differentiation (AD) engines [15]. Consequently, obtaining the gradient of any scalar quantity derived from the model's output with respect to any model input only requires a single function call to the AD engine, which has motivated their use for initial condition sensitivity studies [1, 16, 17].

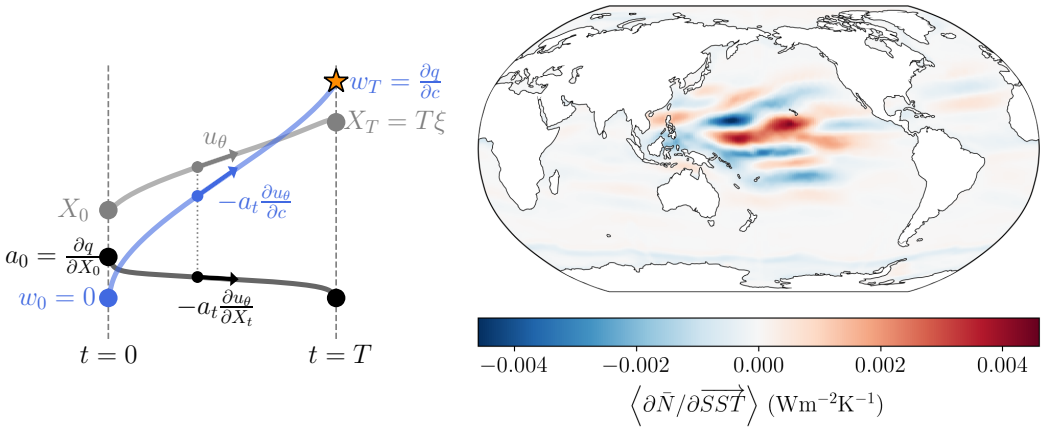

Figure 1: **Left.** An illustration of obtaining the sensitivity $\frac{\partial q}{\partial c}$ with the adjoint method. The velocity network $u_\theta(X_t, t, c)$ turns the clean vector X_0 into a scaled noise latent X_T. As u_θ is steered by conditioning c throughout the reverse sampling process, we need a variable w_t which accumulates the gradient from all noise levels. The gradient accumulator depends on the adjoint $a_t = \frac{\partial q}{\partial X_0}\frac{\partial X_0}{\partial X_t}$, which expresses the sensitivity of q with respect to a partially noised vector X_t. Integrating their respective ODEs together yields $\frac{\partial q}{\partial c} = w_T$. **Right.** Net global radiation flux sensitivity to SST obtained from cBottle, averaged over historical AMIP SST 1971-2020.

Notably, flow and diffusion models [18, 19, 20, 21] have gained popularity in weather and climate due to their powerful generative properties and their ability to model uncertainty. Diffusion models have demonstrated state-of-the-art results in weather forecasting [22] and are able to reproduce atmospheric states [23]. The method presented in this paper provides a way to extract gradients from these types of models, reducing the cost of sensitivity analysis from days to weeks on a supercomputer to minutes or hours on a GPU. However, it is not as clear if these models can produce informative gradients even if they provide a good fit to the data. To begin the work of evaluating their quality, this paper introduces a method of extracting gradients from flow models and checking their consistency with model predictions through finite differences.

2 Background and notation

Climate in a Bottle. The experiments in this work are done with the simplest Climate in a Bottle diffusion model [23] (henceforth called cBottle). Given the day of year τ, second of day ζ, sea surface temperature forcing c, and a high dimensional random vector ξ, the cBottle model generates 45 atmospheric variables (4 3D-variables at 8 pressure levels and 13 2D-variables) at a ground resolution of ~ 100 km. See App. C for details. cBottle has been trained on both ERA5 reanalysis data [24] and outputs from the ICON climate model [25], providing a unique test-bed for gradient extraction methods. While the empirical results are computed with cBottle, the method presented here is not specific to cBottle and applies to generative flow models in general.

Flow models. Diffusion models [18, 19, 20], and the more recent flow matching [21, 26, 27] are all special cases of flow models, a family of generative models which learn a time-varying velocity field $u_\theta(X_t, t)$ that transports an easy-to-sample probability distribution p_{init} to a data probability distribution p_{data}. In what follows, we will use the parameterization of the Elucidated Diffusion Model (EDM) [28] with deterministic sampling.

Conditioning. Flow models can be guided by input variables c to sample from the conditional distribution $p_{data}(\cdot|c)$ [29, 30, 31]. Therefore, the procedure for generation implies sampling ξ from

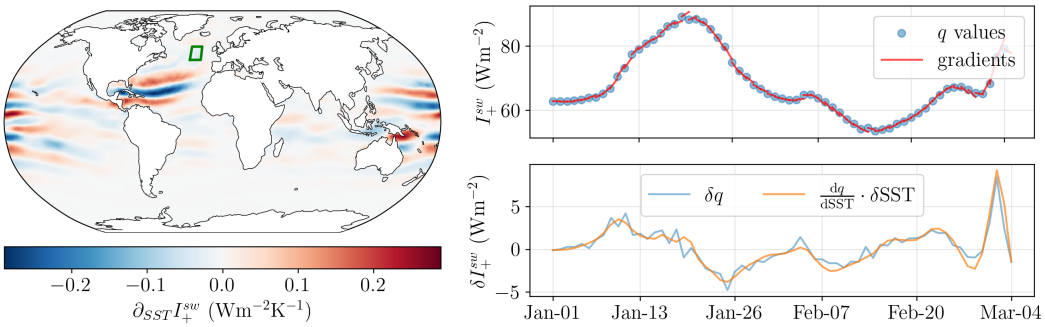

Figure 2: **Left.** A sensitivity map $\frac{\partial q}{\partial \text{SST}}$ on February 1st 2020, at 17:30 UTC. Here, q is the average outgoing shortwave radiation in the green patch I_+^{sw}. **Right.** The top plot shows how q changes over the course of two months, and the extracted gradients. The bottom plot shows the true finite difference δq and the linearized difference $\frac{\text{d}q}{\text{dSST}} \cdot \delta\text{SST}$, see equation 3. The RMSE (equation 2) between the two is 0.76 Wm^{-2}.

the standard Gaussian, picking a conditioning variable c and solving the ODE

$$\text{d}X_t = u_\theta(X_t, t, c)\text{d}t, \text{ with } X_T = T\xi \text{ and } \xi \sim \mathcal{N}(0, \text{Id})$$

backwards in time, from T to 0, obtaining the generated sample $X_0 \sim p_{data}(\cdot|c)$.

3 Method

We are interested in a scalar quantity derived from a generated sample $q(X_0)$ and how this quantity varies with conditioning $\frac{\partial q}{\partial c}$. For example, in the context of cBottle, q can be net outgoing radiation globally, precipitation on the coast of Peru or wind shear in the tropical Pacific, and c is global sea surface temperatures. As flow models have very deep computational graphs due to the recurrent calls to the model, it is not feasible to simply use AD to obtain this derivative. We first write down a continuous counterpart of the chain-rule

$$\frac{\partial q}{\partial c} = \frac{\partial q}{\partial X_0}\frac{\partial X_0}{\partial c} = \frac{\partial q}{\partial X_0}\int_T^0 \frac{\partial X_0}{\partial X_t}\frac{\partial u_\theta}{\partial c}(X_t, t, c)\text{d}t.$$

The key step is to use the adjoint state $a_t = \frac{\partial q}{\partial X_0}\frac{\partial X_0}{\partial X_t}$, which satisfies the ODE

$$\frac{\text{d}}{\text{d}t}a_t = -a_t\frac{\partial u_\theta}{\partial X_t}.$$

Letting $w_{T-t} = \int_T^t a_s \frac{\partial u_\theta}{\partial c}\text{d}s$ we see the desired result is $\frac{\partial q}{\partial c} = a_0\frac{\partial X_0}{\partial c} = w_T$. Solving the system of ODEs

$$\frac{\text{d}}{\text{d}t}\begin{bmatrix} X_t \\ a_t \\ w_t \end{bmatrix} = \begin{bmatrix} u_\theta \\ -a_t\frac{\partial u_\theta}{\partial X_t} \\ -a_t\frac{\partial u_\theta}{\partial c} \end{bmatrix}, \text{ with } \begin{bmatrix} X_0 \\ a_0 \\ w_0 \end{bmatrix} = \begin{bmatrix} X_0 \\ \frac{\partial q}{\partial X_0} \\ 0 \end{bmatrix}, \tag{1}$$

from 0 to T yields $\frac{\partial q}{\partial c} = w_T$. The process is depicted in Figure 1, and proofs can be found in [32, 33, 34].

4 Results with Climate in a Bottle

We conduct our experiments with cBottle, a diffusion model of the EDM type [28]. It can be easily reparameterized as a flow model $\text{d}X_t = u_\theta(X_t, t, c, \tau, \zeta)\text{d}t$, with $X_T = T\xi$ and $\xi \sim \mathcal{N}(0, \text{Id})$ via the probability flow ODE [20]. The conditioning variables are c – sea surface temperature (SST) forcing, τ – day of the year and ζ – UTC second of day. We generate samples from scratch given the conditioning and a randomly drawn noise sample, and then we apply the adjoint sensitivity algorithm

with the same conditioning. For the gradient self-consistency checks, we keep the noise samples ξ constant.

Gradient self-consistency check. The approximate nature of the flow model mapping from c to X_0 means the reliability of its gradients is not immediately clear. While the model's outputs are typically validated for consistency against real-world data or physical models [23], the trustworthiness of the gradients is commonly not assessed. We bridge this gap by proposing a self-consistency check that validates the model's gradients against its own outputs, thereby using the validation of the predictive samples to establish confidence in the model gradients. Figure 2 shows the gradient self-consistency check procedure for a selected quantity q chosen to be the average outgoing shortwave radiation I_+^{sw} in a small area in the north Atlantic. The gradients obtained through the adjoint sensitivity method seem consistent with cBottle's outputs. To quantify this, we define a simple RMSE metric over K samples, following [3]:

$$\text{RMSE} = \sqrt{\frac{1}{K}\sum_k \left(\delta q - \frac{\mathrm{d}q}{\mathrm{d}c}\cdot\delta c\right)^2}, \tag{2}$$

where the total linearized difference is obtained by the chain rule:

$$\frac{\mathrm{d}q}{\mathrm{d}c}\cdot\delta c = \frac{\partial q}{\partial c}\cdot\delta c + \frac{\partial q}{\partial \tau}\delta\tau + \frac{\partial q}{\partial \zeta}\delta\zeta. \tag{3}$$

The derivatives with respect to the other conditioning variables are simply obtained by setting up gradient accumulators for each one, see Appendix A for details.

The sensitivity of global net radiative flux with respect to SST patterns. While the gradient self-consistency check confirms the correctness of the algorithm, it does not guarantee physical meaning. To test this, we follow [3] (henceforth called GFMIP) and calculate the sensitivity of average global net radiative flux with respect to SST $\frac{\partial \bar{N}}{\partial \mathrm{SST}}$, averaged over $K = 2536$ samples of SST in the AMIP dataset from 1971-2020. For this run, the noise latent ξ was sampled from the standard Gaussian. The timestep was set to $169\,\mathrm{h}$, which is 1 week and 1 hour, to uniformly sample the two time conditioning variables in cBottle, τ and ζ. The resulting sensitivity map is presented in the right panel of Figure 1. The RMSE for the entire period is $0.465\,\mathrm{Wm}^{-2}$, and if we take the RMSE of the yearly averaged differences like in GFMIP, we get $0.07\,\mathrm{Wm}^{-2}$, which is close to their value, $0.23\,\mathrm{Wm}^{-2}$. GFMIP do not have high spatial frequencies like ours, but this is to be expected because the patch perturbations used are much larger than the $\sim 100\,\mathrm{km}$ pixels in cBottle and they smooth out the pattern. There are clear negative feedbacks around the Maritime Continent, which might be explained by an increase in convective clouds and thus reduced outgoing long-wave radiation. A faint ring-shaped negative signal around the Antarctic ice-sheet can also be observed, which can be explained by a decrease in global albedo caused by ice-sheet melting. However, the sensitivity in the Pacific is 1-2 orders of magnitude larger than in GFMIP, even as the signals in the other oceans are around the same order of magnitude.

5 Discussion and future work

Limitations. In the case of the sensitivity of global net radiation with respect to SST patterns, the large departure from previous results warrants further investigation. It is possible that cBottle's outputs are very strongly conditioned on the day of year τ, which implies that simply averaging $\frac{\partial q}{\partial c}$ is not the full picture and the procedure described in Appendix A must be followed to add the contribution of the $\frac{\partial q}{\partial \tau}$ gradients.

It must also be said that there is no guarantee cBottle produces physical atmospheres for out-of-distribution SST values (for example $+2\,\mathrm{K}$). In fact, we noticed tiling artifacts appearing for gradients with respect to climatology SST when trying to recreate GFMIP's results.

Future work. A promising avenue for greatly extending the applications of this method is model guidance. To condition the model to generate desired weather states from the exact posterior (for example, show a realistic hurricane in the Atlantic), we can train a separate model [29, 30, 31] that uses a guiding variable $y = G(X_0)$, where G is an observation operator. Gradients with respect to y can then be pulled through the small guiding model. This method would allow calculations of

gradients of anything with respect to anything, including variables not included in the original model, for example CO_2 concentrations, aerosol optical depths or ocean salinity.

Acknowledgements

AD is part of the Intelligent Earth CDT supported by funding from the UK Research and Innovation Council (UKRI) grant number EP/Y030907/1. JP acknowledges support from EPSRC (grant number EP/Y018273/1). TR is supported by the EU's Horizon Europe program under grant agreement number 101131841 and also received funding from UK Research and Innovation (UKRI).

AD thanks Milan Kloewer for the insightful discussions on the physical meaning of extracted sensitivities, Lilli Freischem for an introduction to regridding and Peter Manshausen for feedback on early results.

References

[1] Jorge Baño-Medina, Ankur Sengupta, James D. Doyle, et al. Are ai weather models learning atmospheric physics? a sensitivity analysis of cyclone xynthia. *npj Climate and Atmospheric Science*, 8(92), 2025.

[2] Ryan D Torn and Gregory J Hakim. Ensemble-based sensitivity analysis. *Monthly Weather Review*, 136(2):663–677, 2008.

[3] Jonah Bloch-Johnson, Maria AA Rugenstein, Marc J Alessi, Cristian Proistosescu, Ming Zhao, Bosong Zhang, Andrew IL Williams, Jonathan M Gregory, Jason Cole, Yue Dong, et al. The green's function model intercomparison project (gfmip) protocol. *Journal of Advances in Modeling Earth Systems*, 16(2):e2023MS003700, 2024.

[4] Senne Van Loon, Maria Rugenstein, and Elizabeth A Barnes. Reanalysis-based global radiative response to sea surface temperature patterns: Evaluating the ai2 climate emulator. *Geophysical Research Letters*, 52(14):e2025GL115432, 2025.

[5] Ronald M Errico. What is an adjoint model? *Bulletin of the American Meteorological Society*, 78(11):2577–2592, 1997.

[6] Yannick Trémolet. Model-error estimation in 4d-var. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 133(626):1267–1280, 2007.

[7] James D Doyle, Clark Amerault, Carolyn A Reynolds, and P Alex Reinecke. Initial condition sensitivity and predictability of a severe extratropical cyclone using a moist adjoint. *Monthly Weather Review*, 142(1):320–342, 2014.

[8] Grant Branstator. Analysis of general circulation model sea-surface temperature anomaly simulations using a linear model. part i: Forced solutions. *Journal of Atmospheric Sciences*, 42(21):2225–2241, 1985.

[9] Mark Holzer and Timothy M Hall. Transit-time and tracer-age distributions in geophysical flows. *Journal of the atmospheric sciences*, 57(21):3539–3558, 2000.

[10] Joseph J Barsugli and Prashant D Sardeshmukh. Global atmospheric sensitivity to tropical sst anomalies throughout the indo-pacific basin. *Journal of Climate*, 15(23):3427–3442, 2002.

[11] Chen Zhou, Mark D Zelinka, and Stephen A Klein. Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a green's function approach. *Journal of Advances in Modeling Earth Systems*, 9(5):2174–2189, 2017.

[12] Yue Dong, Cristian Proistosescu, Kyle C Armour, and David S Battisti. Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a green's function approach: The preeminence of the western pacific. *Journal of Climate*, 32(17):5471–5491, 2019.

[13] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[14] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.

[15] Atilim Gunes Baydin, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18(153):1–43, 2018.

[16] P Trent Vonich and Gregory J Hakim. Predictability limit of the 2021 pacific northwest heatwave from deep-learning sensitivity analysis. *Geophysical Research Letters*, 51(19):e2024GL110651, 2024.

[17] Kinya Toride, Matthew Newman, Andrew Hoell, Antonietta Capotondi, Jakob Schlör, and Dillon J Amaya. Using deep learning to identify initial error sensitivity for interpretable enso forecasts. *Artificial Intelligence for the Earth Systems*, 4(2):e240045, 2025.

[18] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, pages 2256–2265, 2015.

[19] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, 2020.

[20] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

[21] Yaron Lipman, Ricky T. Q. Chen, Haggai Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

[22] Ilan Price, Alvaro Sanchez-Gonzalez, Ferran Alet, Tom R. Andersson, Andrew El-Kadi, Dominic Masters, Timo Ewalds, Jacklynn Stott, Shakir Mohamed, Peter Battaglia, Remi Lam, and Matthew Willson. Probabilistic weather forecasting with machine learning. *Nature*, 637(8044):84–90, 2025.

[23] Noah D. Brenowitz, Tao Ge, Akshay Subramaniam, Peter Manshausen, Aayush Gupta, David M. Hall, Morteza Mardani, Arash Vahdat, Karthik Kashinath, and Michael S. Pritchard. Climate in a bottle: Towards a generative foundation model for the kilometer-scale global atmosphere, 2025.

[24] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly journal of the royal meteorological society*, 146(730):1999–2049, 2020.

[25] Cathy Hohenegger, Peter Korn, Leonidas Linardakis, René Redler, Reiner Schnur, Panagiotis Adamidis, Jiawei Bao, Swantje Bastin, Milad Behravesh, Martin Bergemann, et al. Icon-sapphire: simulating the components of the earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, 16(2):779–811, 2023.

[26] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow, 2022.

[27] Michael S. Albergo and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. In *Advances in Neural Information Processing Systems*, 2023.

[28] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Advances in Neural Information Processing Systems*, 2022.

[29] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. In *ICCV*, pages 3813–3824. IEEE, 2023.

[30] Alexander Denker, Francisco Vargas, Shreyas Padhy, Kieran Didi, Simon V Mathis, Riccardo Barbano, Vincent Dutordoir, Emile Mathieu, Urszula Julia Komorowska, and Pietro Lio. DEFT: Efficient fine-tuning of diffusion models by learning the generalised h-transform. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

[31] Jakiw Pidstrigach, Elizabeth Louise Baker, Carles Domingo-Enrich, George Deligiannidis, and Nikolas Nüsken. Conditioning diffusions using malliavin calculus. In *Forty-second International Conference on Machine Learning*, 2025.

[32] Lev Semenovich Pontryagin, E. F. Mishchenko, V. G. Boltyanskii, and R. V. Gamkrelidze. *The Mathematical Theory of Optimal Processes*. Interscience Publishers, New York, 1962.

[33] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 6572–6583, Red Hook, NY, USA, 2018. Curran Associates Inc.

[34] Ilya V. Schurov. Adjoint state method, backpropagation and neural odes, August 5 2022. Accessed on 2025-08-19.

A Getting total derivatives with the adjoint sensitivity method

In the case in which the inputs to the flow model are dependent on each other, as is the case with cBottle's forcing SSTs c and day of year τ, we might be interested in total derivatives, in addition to the partial derivatives obtainable through the adjoint sensitivity method. To reiterate, obtaining partial derivatives (from a velocity-reparameterized EDM) with the adjoint sensitivity method implies solving the system of ODEs

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} X_t \\ a_t \\ w_t \\ v_t \end{bmatrix} = \begin{bmatrix} u_\theta \\ -a_t \frac{\partial u_\theta}{\partial X_t} \\ -a_t \frac{\partial u_\theta}{\partial c} \\ -a_t \frac{\partial u_\theta}{\partial \tau} \end{bmatrix}, \text{ with } \begin{bmatrix} X_0 \\ a_0 \\ w_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} X_0 \\ \frac{\partial q}{\partial X_0} \\ 0 \\ 0 \end{bmatrix}$$

from $t = 0$ to $t = T$, where $w_T = \frac{\partial q}{\partial c}$ and $v_T = \frac{\partial q}{\partial \tau}$. In practice, calculating the additional derivative $\frac{\partial u_\theta}{\partial \tau}$ adds negligible computational cost, as the backpropagation process through the computational graph happens the same way and with the same starting tangent $-a_t$. Any input of the model can be added to the ODE this way. The total derivative is then

$$\frac{\mathrm{d}q}{\mathrm{d}c} = \frac{\partial q}{\partial c} + \frac{\partial q}{\partial \tau}\frac{\partial \tau}{\partial c} = w_T + v_T \frac{\partial \tau}{\partial c}.$$

To approximate $\frac{\partial \tau}{\partial c}$, one could train a regressor and get this gradient with an AD call.

For calculating the RMSE on historical data this is not needed, as we have both δc and $\delta \tau$:

$$\frac{\mathrm{d}q}{\mathrm{d}c} \cdot \delta c = \frac{\partial q}{\partial c} \cdot \delta c + \frac{\partial q}{\partial \tau}\frac{\partial \tau}{\partial c} \cdot \delta c \approx w_T \cdot \delta c + v_T \delta \tau.$$

B More sensitivity maps

See Figure 3 for monthly sensitivity maps, each one averaged over $K/12 \approx 211$ samples.

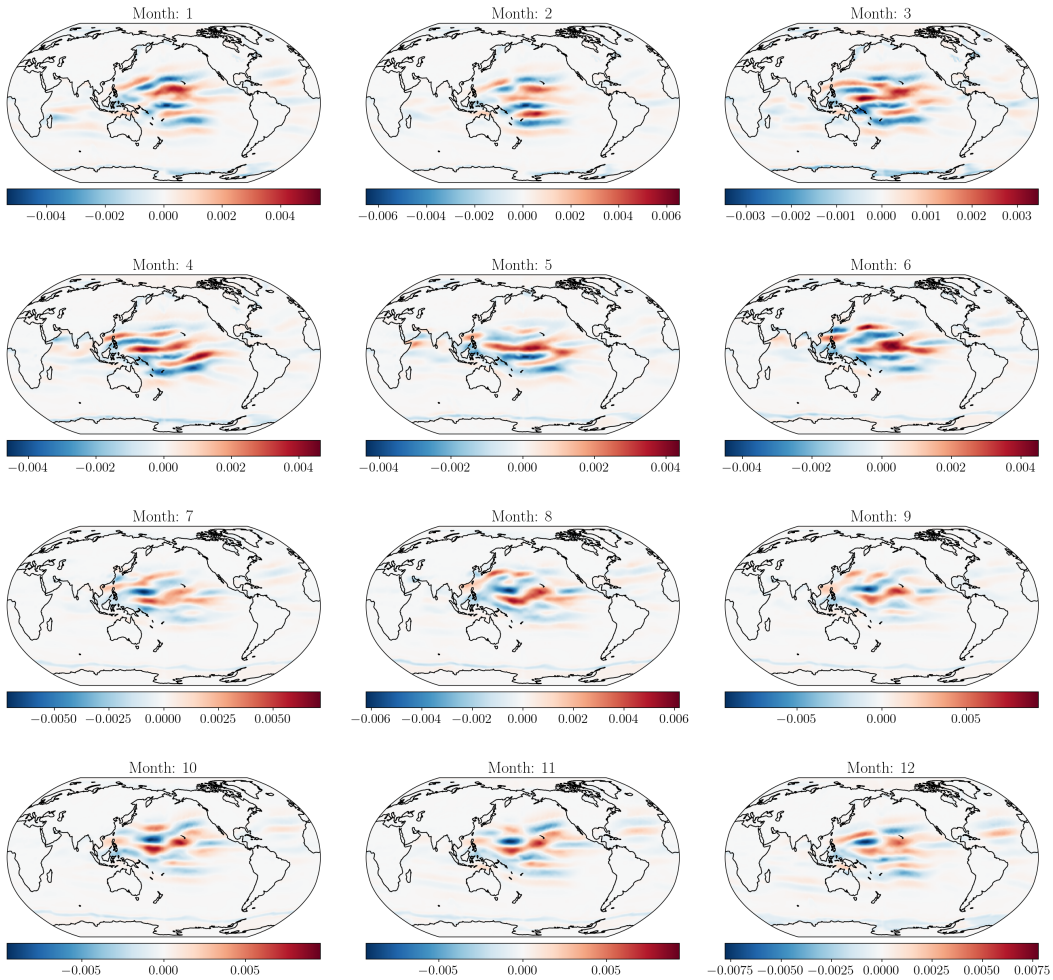

Figure 3: Monthly averaged sensitivity maps $\frac{\partial \bar{N}}{\partial \text{SST}}$ in the period 1971-2020.

C cBottle details

C.1 Generated variables

See Table 1.

C.2 The nature of SST conditioning data

Climate in a Bottle is using one of its conditioning variables, the day of year τ, to obtain the sea surface temperature forcing $c(\tau)$. This procedure is mentioned but not detailed in their paper, and so we will describe here. They use the AMIP dataset `input4MIPs.CMIP6Plus.CMIP.PCMDI.PCMDI-AMIP-1-1-9.ocean.mon.tosbcs.g` which can be downloaded with a script from here. This dataset contains the estimated global mid-month, monthly mean SST from January 1870 to December 2022, all dated on the 16th of the month (except February, which has the 15th). If the model is given a day of year τ, $c(\tau)$ is linearly interpolated from the nearest two data points:

$$c(\tau) = c(\tau_i) + \frac{\tau - \tau_i}{\tau_{i+1} - \tau_i} \left(c(\tau_{i+1}) - c(\tau_i) \right),$$

where τ_i is the mid-month day-of-year of month i and τ satisfies $\tau_i \leq \tau < \tau_{i+1} < \tau + 31$.

For simplicity, we have maintained that τ is the day of the year, but it would be more precise to say it is the day of the year plus the day fraction: $\tau = 1.5$ means January 1st 12:00 UTC.

Short name	Description	Units
Profile variables @ 1000, 850, 700, 500, 300, 200, 50, 10 hPa		
`T`	air temperature (profile)	K
`U`	zonal wind (profile)	$\mathrm{m\,s^{-1}}$
`V`	meridional wind (profile)	$\mathrm{m\,s^{-1}}$
`Z`	geopotential height	m
2D variables		
`clivi`	column integrated cloud ice	$\mathrm{kg\,m^{-2}}$
`cllvi`	column integrated cloud water	$\mathrm{kg\,m^{-2}}$
`rlut`	TOA outgoing longwave radiation	$\mathrm{W\,m^{-2}}$
`rsut`	TOA outgoing shortwave radiation	$\mathrm{W\,m^{-2}}$
`rsds`	surface downwelling shortwave radiation	$\mathrm{W\,m^{-2}}$
`tcwv`	column integrated water vapour	$\mathrm{kg\,m^{-2}}$
`pr`	precipitation flux	$\mathrm{kg\,m^{-2}\,s^{-1}}$
`pres_msl`	mean sea level pressure	Pa
`sic`	sea ice concentration (fractional)	–
`sst`	sea surface temperature	K
`tas`	2m air temperature	K
`uas`	zonal wind at 10m	$\mathrm{m\,s^{-1}}$
`vas`	meridional wind at 10m	$\mathrm{m\,s^{-1}}$

Table 1: Variables grouped into profile, column, and surface categories with their ICON names, descriptions, and units.