

Bridging the weather and climate divide with artificial intelligence

Received: 13 January 2026

Accepted: 10 July 2026

Published online: 18 August 2026

 Check for updates

Gustau Camps-Valls ¹✉, Alberto Carrassi ^{2,3}, Francisco de Melo Viríssimo ⁴, Kevin Debeire ^{5,6}, Miguel-Ángel Fernández-Torres ⁷, Philipp Hess^{8,9}, Helge Heuer ⁵, Nathan Mankovich ¹, David Montero ¹⁰, Esther Rodrigo-Bonet ¹, Federico Serva ¹¹, Vasileios Sitokoustantinou ¹², Alistair White^{8,9}, Kai-Hendrik Cohrs ¹, Jonathan Wider ¹³, Christian Reimers ¹⁴, Nikolaos Ioannis Bountos ^{15,16}, Ioannis Papoutsis ^{15,16,17}, Pierre Gentine ¹⁸ & Veronika Eyring ^{5,19}

Weather and climate shape every aspect of our lives, yet for decades they have been studied as separate worlds—one concerned with what happens tomorrow, the other with what awaits us over the coming decades. This long-standing divide has slowed progress in both fields. Here, we argue that artificial intelligence (AI) is helping to dissolve the boundary between them. By learning directly from vast streams of Earth observations, AI is starting to predict both tomorrow's storm and next century's climate within a single, unified framework. We show how this convergence is already underway, why weather and climate science have much to gain from each other, and what still stands in the way—from trust, transparency, and fairness to the energy cost of computation and the risk of widening global inequalities. Realizing this vision will demand not only better algorithms but a new culture of collaboration across scientific communities, operational centers, fields and disciplines. Done well, a unified, AI-driven Earth science could deliver faster, more reliable, and more actionable predictions—helping societies anticipate and adapt to a rapidly warming planet.

Weather and climate science describe the same Earth system on different horizons—one predicting conditions hours to days ahead, the other projecting changes over decades to centuries—and have long evolved as separate scientific communities, with distinct models, data,

and goals. Over the past decade, artificial intelligence (AI) has rapidly advanced in weather and climate science. Deep learning (DL) architectures—such as graph neural networks, transformers, and diffusion models—have improved forecasting accuracy, efficiency, and

¹Image Processing Laboratory, Universitat de València, València, Spain. ²Department of Physics, University of Bologna, Bologna, Italy. ³Department of Meteorology, University of Reading, Reading, UK. ⁴Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, London, UK. ⁵Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany. ⁶Technical University of Munich, Munich, Germany. ⁷Department of Signal Theory and Communications, Universidad Carlos III de Madrid (UC3M), Madrid, Spain. ⁸Earth System Modelling, School of Engineering and Design, Technical University of Munich, Munich, Germany. ⁹Complexity Science, Potsdam Institute for Climate Impact Research, Potsdam, Germany. ¹⁰Institute for Earth System Science and Remote Sensing, Leipzig University, Leipzig, Germany. ¹¹National Research Council, Rome, Italy. ¹²Artificial Intelligence Group, Wageningen University & Research, Wageningen, The Netherlands. ¹³Helmholtz-Centre for Environmental Research/ScaDS.AI, Leipzig, Germany. ¹⁴Max Planck Institute for Biogeochemistry, Jena, Germany. ¹⁵National Observatory of Athens, Athens, Greece. ¹⁶National Technical University of Athens, Athens, Greece. ¹⁷Archimedes - Athena Research Center, Marousi, Greece. ¹⁸Columbia University, New York, NY, USA. ¹⁹Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany. ✉e-mail: gustau.camps@uv.es

accessibility^{1–3}. They have also enhanced climate model parameterization, downscaling, and emulation^{4,5}, effectively addressing the multi-scale, heterogeneous, and nonlinear nature of Earth's systems^{6–8}.

Weather and climate systems are inherently multi-scale, nonlinear, and chaotic⁹, making them difficult to model and predict. Although closely related, the fields operate under different hypotheses, temporal scales, and dynamics, which lead to divergent challenges, methodologies, and solution strategies. In weather forecasting, for example, AI models often rely on reanalysis data yet struggle to capture extreme events or maintain physical balance^{8,10}. In climate science, the challenge is deeper: weather provides many realizations of short events, whereas climate offers only one observed realization of a long-term trajectory. This limits uncertainty quantification and generalization, especially under climate change¹¹.

Despite these difficulties, AI-enabled progress has been a perspective. Recent breakthroughs—FourCastNet¹², GraphCast¹³, NeuralGCM³, and AIFS¹⁴—show that AI trained solely on reanalysis data (e.g., ERA5¹⁵) can rival physics-based models in short-term atmospheric forecasting. Such models are already deployed or tested operationally, including Pangu-Weather¹⁶, Google's Weather Lab, ECMWF's AIFS¹⁴, DWD's AICON¹⁷, NOAA's AIGFS¹⁸, and Met Norway's Bris¹⁹—a rapid transition from research to real-time forecasting. Climate modeling remains harder, requiring AI to capture responses to long-term, time-dependent forcings such as greenhouse gases and aerosols whose future evolution is uncertain. Emerging applications in climate emulation^{20,21}, process-based learning^{22,23}, and bias correction^{24–26} are beginning to address these, while probabilistic and generative methods efficiently produce synthetic weather and climate states that better capture extremes and changing-climate scenarios^{3,27,28}.

Despite the traditional weather-climate schism²⁹, we observe a convergence in the use of AI tools for weather and climate applications (Fig. 1). This shift is driven by AI: the fraction of research jointly addressing weather and climate exceeds the level expected under independence, with a statistically robust catalysis effect over the last decade. Tracking domain-specific terminology while excluding generic AI keywords, the keyword panels show this synergy rests on shared methodological foundations—emulation, downscaling, and forecasting—that increasingly interlink the two disciplines. Consolidated development approaches, such as the seamless daily-to-seasonal frameworks pioneered by the Met Office and ECMWF^{29,30}, demonstrate that integrating both fields accelerates innovation and improves forecasting skill³¹. Still, building on these examples, a collaborative, high-resolution, cross-disciplinary framework *powered by AI* can drive a truly unified Earth system science represented by a seamless prediction system^{32,33}. Such an integration is already happening in subseasonal-to-seasonal (S2S) forecasting^{34,35}, supported by recent work demonstrating convergence across timescales^{36,37}.

To achieve a unified framework, progress must be evaluated using measurable milestones and multicriteria performance metrics that encompass not only accuracy but also physical consistency, causality, and stability. We highlight three concrete success stories of unifying approaches: a temporal bridge (S2S as the natural middle between both fields), a causal bridge (analyzing extreme weather in different counterfactual climates), and a structural bridge (joint techniques and hybrid process-level modeling where components are interchangeably used for both tasks).

This paper highlights the key distinctions between weather and climate modeling and shows how current and emerging AI methods could close this gap (Table 1), enabling the seamless integration long sought by the community. We also address social and environmental challenges—the effect of ML on greenhouse-gas emissions³⁸, trust and accountability³⁹, and explainability⁴⁰—and propose a forward-looking agenda in which AI is the central catalyst for methodological advances and community convergence.

Distinct challenges in weather and climate

The idea of climate prediction emerged from weather forecasting. Early efforts extended atmospheric model runs and included only very crude representations of other parts of the climate system^{41–43}. Over time, the distinction between the two became clear: weather forecasting and climate prediction (and/or projection) are fundamentally different tasks. They pose distinct challenges in scale, dynamics, and objectives—a divide that persists even when both adopt a common framework such as AI. Table 2 summarizes the key challenges of AI in weather and climate modeling with respect to data, models, and assumptions, as well as the ethical and social challenges arising from the pursuit of a seamless integration.

Forecast, prediction, and projection

The terms *forecast*, *prediction*, and *projection* are often used interchangeably, but they have distinct meanings that are central to the weather-climate distinction. Weather forecasts (we use “forecast” and “prediction” interchangeably for weather) give probabilistic estimates from minutes to weeks, occasionally months, grounded in well-defined probability distributions calibrated against historical data⁴⁴: ideally, an 80% chance of rain means rain occurred in 80% of past cases with similar conditions⁴⁵, though achieving such calibration remains an active research area. This probabilistic framing partly explains the strong performance of AI weather models. Climate prediction and projection differ conceptually. Climate prediction occupies a grey zone between initialized near-term forecasts (interannual to decadal) and long-term estimates; the latter are more rigorously termed *climate projections*—conditional statements about future climate built from ensembles of simulations conditioned on hypothetical scenarios, typically future greenhouse-gas concentrations. Their probabilities are therefore scenario-relative rather than grounded in the verification-based calibration expected of weather forecasts.

The distinction can also be framed by timescale and anthropogenic forcing. Climate prediction spans seasonal-to-decadal horizons, where initial conditions still matter and scenario effects are present but small—motivating data assimilation to initialize predictions (e.g., NorCPM)⁴⁶.

Climate projections, by contrast, address change over decades to centuries, driven primarily by assumed future radiative forcing (greenhouse gases, aerosols, land use). Initial-condition information is largely lost at these timescales, so projections instead explore plausible futures via separate ensembles per emission scenario—four in CMIP5 (the Representative Concentration Pathways, RCPs) and eight in CMIP6 (the Shared Socioeconomic Pathways, SSPs)—rather than a single initial-condition ensemble⁴⁷. These nuanced differences create distinct challenges across data, modeling, assumptions, and operations—challenges AI can address in a unified way.

Data challenges

A shared challenge is not the absence of observations per se, but the uneven availability of independent, long-term, in situ measurements. While station data remain sparse and geographically biased, particularly over oceans and the Southern Hemisphere, satellite observations now dominate global numerical weather prediction skill and have largely mitigated these gaps; their assimilation raised Southern Hemisphere forecast skill to Northern Hemisphere levels⁴⁸, and they remain increasingly central to modern forecasting systems (see, e.g., ECMWF reports on AI-DOP forecasts). Beyond weather prediction, improved exploitation of satellite data also holds promise for regional and seasonal-to-decadal climate prediction, where initial conditions retain predictive power^{49,50}. Observations play a dual role: as inputs to the data assimilation systems that produce reanalyses, and as independent references for evaluating them—both functions being critical, since reanalyses form the backbone of many data-driven models yet can inherit systematic biases from models and assimilation systems.

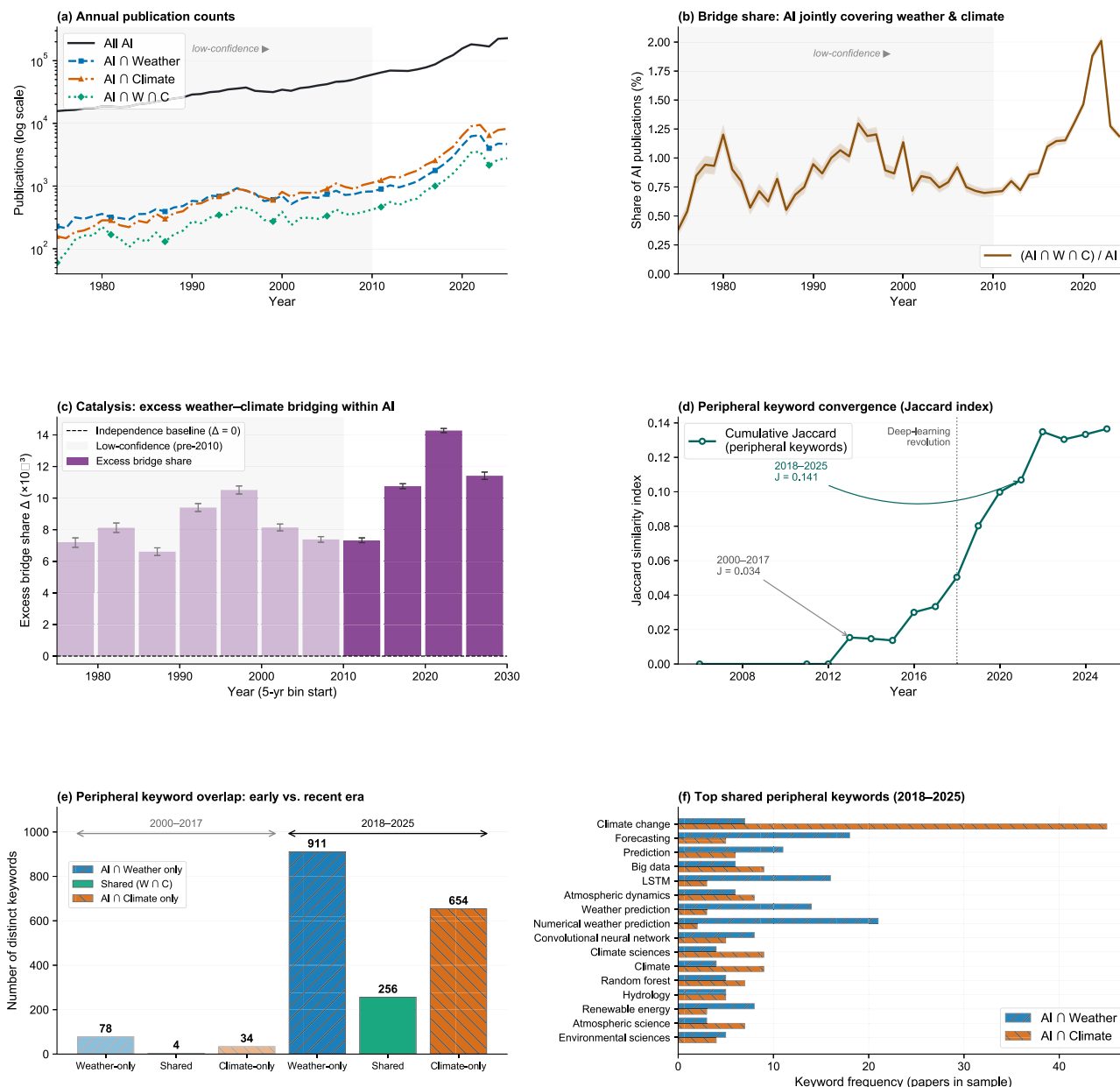


Fig. 1 | Artificial intelligence (AI) as a catalyst for convergence between weather and climate sciences (1975–2025). **a–c** Exponential growth of publications related to AI in weather and climate, with the *excess bridge share* isolating the catalysis effect of AI (showing overlap beyond what is expected under independence). The full 1975–2025 horizon establishes a historical baseline of low weather–climate overlap; early bins (pre-2010, grey shading) denote periods of low publication volume and high uncertainty, while the statistically supported catalysis effect is concentrated in the post-2012 period (non-shaded). **d–f** Peripheral keyword

convergence and overlap. The cumulative Jaccard index (**d**) and the distribution of distinct and shared peripheral keywords (**e**) reveal a noticeable increase in thematic overlap driven by AI adoption, particularly after the deep-learning revolution (post-2017). Ranking the top shared peripheral keywords (**f**) highlights shared methodological foundations such as forecasting, downscaling, and emulation. Generic AI seed terms (e.g., machine learning, deep learning) have been excluded to ensure the analysis reflects domain-specific thematic convergence. See Supplementary Note 1 for methodological details.

While weather relies largely on atmospheric and oceanic observations—though land-surface variables such as soil moisture, surface temperature, and vegetation also matter, particularly at subseasonal-to-decadal timescales—climate science is more demanding, requiring data from across Earth’s interconnected spheres. The challenge is compounded because most measurements integrate processes acting over multiple timescales: observations mix immediate plant responses to warming with longer-term shifts in species composition. Separating these influences is crucial for reliable generalization, yet creates trade-offs, for instance between intra- vs. interannual change and temporal vs. spatial variability⁵¹.

Another related challenge for climate is the uneven availability of observations across scales. For example, while plant-level responses to elevated CO₂ are well understood and supported by abundant data, our understanding of how entire forests respond remains limited due to complex interspecies interactions. This gap makes it challenging to translate rich microscale knowledge into improved AI performance at the macroscale level.

Simulating rare extreme events, a primary goal in both weather and climate, remains difficult due to limited data^{52,53}, despite their outsized impacts and strong feedbacks. In climate prediction, and even more in climate projections, this challenge is amplified by a non-stationary climate. Similarly, other high-frequency, small-scale

Table 1 | Integrated framework highlighting the differences between weather and climate modeling and how artificial intelligence (AI) addresses their respective challenges

	Weather	Climate
Consistency/model properties	Data-driven, accuracy-focused Data assimilation complexity, real-time constraints, atmosphere GCM plus forcings. ML-driven data assimilation improvements, gap filling, hybrid ML-NWP.	Knowledge-driven, stability-focused Sparse long-term data, inconsistencies, paleoclimate reconstructions, multicomponent Earth system model, structural uncertainty. Emulation of climate models, AI-enhanced reanalysis, data fusion for paleoclimate.
Temporal scales	Minutes to weeks Forecast horizon limited by chaos, model bias. Hybrid ML-NWP models, GNNs for process representation.	Years to centuries Long simulations, limited training data, non-stationarity. AI-assisted parameterization, surrogate models for faster projections.
Spatial scales	Regional Bias correction, ensemble validation. ML-based downscaling, ML-based bias correction.	Global Regional downscaling, evaluating long-term projections. ML-based downscaling, ML-based uncertainty estimation.
Prediction target	Trajectory-focused evolution Communicating short-term extreme events, public engagement. Real-time risk messaging, AI-driven impact modeling.	Distribution-focused climate statistics Uncertainty communication over decades, decision-maker engagement. Explainable AI for risk communication, interactive AI-driven decision tools.

Weather focuses on data-driven accuracy, regional and weekly scales, and trajectory forecasts, while climate emphasizes knowledge-driven stability, global and multi-decadal scales, and statistical distributions. AI methods act as a bridge, supporting data assimilation, hybrid modeling, downscaling, emulation, parameterization, and explainable risk communication across both domains.

processes are poorly captured in most data-driven parameterizations. For this reason, traditional numerical weather models still outperform current AI systems in representing record-breaking extremes⁵⁴, despite AI superiority over them in non-extreme conditions.

Modeling challenges

Both weather forecasting and climate prediction now employ coupled Earth system models. In weather forecasting, these models are optimized for high-resolution, short-term atmospheric predictability, whereas climate prediction and projection emphasize long-term interactions, feedbacks, and equilibrium responses of the coupled Earth system⁵⁵. Earth System Model (ESMs) are computationally demanding⁵⁶, difficult to calibrate⁵⁷, and subject to structural uncertainties^{58,59} and modeler biases⁶⁰, especially in components like biogeochemistry⁶¹.

Weather processes operate on shorter time frames and finer spatial scales, where forecasting challenges (e.g., chaotic instabilities) can generally be tackled with better observations⁶², assimilation^{63,64}, and higher-resolution ensemble forecasting⁶⁵, though at a high computational cost. Both weather and climate systems exhibit chaotic behavior. However, instabilities due to deterministic chaos (i.e., the amplification of initial condition errors) grow much faster in weather forecasting, dominating predictability on timescales of days to weeks. On the other hand, climate and Earth system models simulate a variety of mutually interactive long-term processes with high variability and uncertainty in their structure/formulation (e.g., nitrogen fixation⁶⁶; ocean carbon attenuation⁶⁷), with their impact being also dependent on the time horizon under consideration (e.g., decadal, centennial, etc.)⁶⁸.

Weather models emphasize short-term forecasting of atmospheric variables (e.g., temperature, precipitation), whereas climate models target long-term projections and trends of a much broader set (e.g., ocean primary production, carbon sequestration, land cover) under non-periodic anthropogenic GHG forcing. Climate evolution is thus a non-autonomous, non-stationary system⁶⁹—an immense yet appealing challenge for training AI and for any unifying framework.

These differences also manifest spatially. Earth system models typically run at coarse resolutions (0.70°–3.75°, ≈78–416 km)⁴⁷, whereas state-of-the-art global weather models already reach

≈9–10 km (e.g., IFS, ICON, UM)⁷⁰, with km-scale global simulations emerging (e.g., Destination Earth, nextGEMS). Bridging these disparate temporal, spatial, and dynamical regimes remains a central obstacle to seamless integration and a defining challenge for any unified, AI-driven framework.

Assumptions and hypotheses challenges

Predictive modeling across weather and climate rests on different assumptions about which set of variables matter, how they interact, and how their influence evolves across timescales. Variables that play a negligible role in short-term weather prediction can become critical for seasonal forecasts and indispensable for long-term climate projections. For example, ocean salinity has limited impact on day-to-day weather forecasts but increasingly shapes predictability at seasonal scales through ocean-atmosphere coupling and is central to long-term climate evolution^{69,71,72}.

While weather is near-stationary on forecast timescales, the climate system is inherently non-stationary, with evolving patterns and shifting distributions⁶⁹. This requires models that can dynamically adapt to long-term changes and emerging trends. Moreover, while weather forecasting models are much affected by dynamical instabilities, climate models are also susceptible to qualitative changes in behavior due to changes in the forcing. These changes in behavior include the so-called Earth tipping points⁷³, which are also sensitive to the structural uncertainty in those models⁷⁴.

Weather forecasting, being near-stationary and observation-rich, enables objective validation and robust probabilistic prediction^{28,75,76}. In contrast, climate predictions and climate projections depend on long-term extrapolations and, particularly, the latter, on counterfactual scenarios with no clear ground truth⁷⁷. Out-of-distribution (OOD) challenges also differ: in weather, OOD mainly means unseen initial conditions or rare extremes within a largely stationary system (e.g., “gray swan” events)⁷⁸; in climate, it involves entirely new forcing trajectories that shift the data distribution itself, making extrapolation qualitatively harder. The broader literature stresses both the promise and limits of seamless prediction in data-driven contexts: conceptual work highlights the continuity of Earth system processes across timescales and the potential to unify prediction systems, while noting persistent challenges in physical consistency, skill, and tractability⁷⁹.

Table 2 | Challenges ahead on artificial intelligence (AI) for climate

Challenge		Current AI	Needs	Future AI
Data	Sparse and Biased Data	Bias correction, data augmentation (including synthetic observations), interpolation	Uniform and fair spatial generalization across the globe, better integration of heterogeneous, sparse, and noisy data sources, Enhanced satellite and reanalysis data streams	General-purpose Earth system foundation models trained on simulations, generated and imperfect data, few-shot adaptation in data-sparse regions
	Entangled Measurements	Linear disentanglement (PCA/EOFs), latent-space separation (autoencoders)	Improved separation of overlapping processes and drivers in observations	Causal and generative models capable of isolating physical mechanisms across scales
	Utilizing Multi-scale Data	Models trained at single or limited scales, basic multi-scale architectures	Better integration of fine-coarse scale information and consistent cross-scale representations	Scale-aware or scale-independent models (e.g., with scale encodings, hierarchical or multi-resolution FMs) that learn unified representations across scales from multimodal satellite data, reanalysis, and ESM simulations
	Extrapolation and Unknown Regimes	Uncontrolled extrapolation beyond training data, regularization-based stabilization	Principled methods for stable, verifiable behavior outside observed regimes	Physics-integrated, causally invariant, and statistically robust extrapolation
Modeling	Simulating Low-Probability Extreme Events	GANs/VAEs for synthetic extremes, often with tail smoothing or mode collapse	Accurate representation and sampling of rare, heavy-tailed events	Diffusion and tail-aware generative models with regularization against mode collapse, rare-event and surrogate sampling techniques
	Combining Different Modeling Approaches	Offline hybrid models, soft physics constraints, ML-assisted parameterizations, emulators	Stable, hybrid AI-physics models that enforce constraints and train in-the-loop with the physical system	Online-trained, differentiable physics-AI systems coupled with fast emulators for stable, efficient simulation
	Spatio-Temporal Scales and Uncertainty	Deterministic downscaling and ensemble-style approximations, basic Bayesian/ensemble DNNs	Cheap and realistic high-resolution sampling with consistent spatio-temporal structure and uncertainty	Probabilistic multi-scale generative models (e.g., diffusion or flow-based) enabling coherent, distribution-faithful high-resolution ensembles
	Multiple Forecasting Tasks	Short-term neural models optimized only for limited lead times	Models that remain stable and informative at subseasonal-to-seasonal horizons	Unified AI-physics forecasting systems spanning weather to S2S timescales
	Simulating Dynamical Extreme Events	ML emulators with limited stability and weak physical consistency in extreme regimes	Physically grounded, reliable simulation of evolving extremes (e.g., storms, heatwaves, floods)	Hybrid AI-physics models capable of robust extreme-event dynamics and generalization with scarce examples
	Non-Stationary Processes	Static timeseries models	Dynamic adaptation to shifting regimes	Continual and online-learning AI with uncertainty-aware updates
Assumptions and Hypothesis	Structural Uncertainty and Modeler Bias	Hand-designed parameterizations, heuristics, and manual calibration	More objective, data-driven methods that reduce subjective modeling choices	Self-organizing AI-physics models with learnable structures and automated calibration (e.g., through differentiable modeling)
	Inter-annual and Spatial Variability	Data-driven models that struggle to capture large-scale modes and regional variability	Better representation of teleconnections and spatial-temporal variability	Physics-AI integration for coherent global variability across years and regions
	Ethical and Democratization Concerns	Resource-intensive large models	Open, transparent, and accessible AI systems usable across regions and institutions	Efficient, publicly available models with small-footprint variants
Social and Ethical	Fairness in AI Predictions	Bias correction, debiasing techniques	Better ethics frameworks in AI for climate	AI fairness-aware algorithms
	Transparency and Accountability	Explainable AI (XAI)	More accessible AI for researchers, reproducibility, scrutiny and accountability	Interpretable AI, Regulated, auditable AI

Categorization of key challenges in climate prediction and projection into four sections (along rows)—Data, Modeling, Assumptions/Hypothesis, Social and Ethical—highlighting the current AI approaches, existing gaps, and future AI directions. The Data section addresses issues such as sparse and biased data, entangled measurements, and challenges in utilizing multi-scale data. The Modeling section addresses issues in transitioning from weather to climate modeling frameworks, scale and resolution, and challenges in simulating extreme events. The Assumptions/Hypothesis rows are non-stationarity, structural uncertainty/modeler bias, and inter-annual/spatial variability. The Assumptions/Hypothesis section addresses non-stationarity, structural uncertainty and modeler bias, and inter-annual and spatial variability. The Social and Ethical section addresses concerns about the democratization of AI resources, the fairness of AI-driven climate predictions, and the need for transparency and accountability in model development. For each challenge, the table lists (along columns) current AI methods being used, the needs and gaps in these methods, and potential future AI solutions that could improve climate prediction and projection.

More critical perspectives stress fundamental limits of data-driven approaches⁸⁰, arguing that skill alone is insufficient without physical interpretability and robustness under novel conditions⁸¹, and that models calibrated on past data cannot be expected to generalize to fundamentally new regimes such as future climate forcing⁸².

Together, these works support three points: (i) seamless prediction is well-motivated by the continuity of Earth system processes across timescales (a continuum of predictability⁷⁹ and continuous scaling regimes⁸³); (ii) practical implementations improve consistency and efficiency but remain constrained; and (iii) purely data-driven approaches face fundamental limits in extrapolation and uncertainty quantification, especially under climate-change forcing. Recent hybrid and analog-based advances further illustrate the reliance on assumptions about the transferability of predictability. This motivates our emphasis on hybrid, physically informed AI and careful evaluation under OOD conditions^{4,83–85}. OOD nonetheless remains a major barrier across both timescales^{86,87}: both domain shift and extrapolation to unknown extremes hinder reliability, underscoring the difficulty of learning without verifiable reference data.

Social and ethical challenges

Training large-scale AI models requires vast computational and energy resources, raising concerns about environmental impact, sustainability, and research accessibility. While training requires substantial computational resources, inference is relatively inexpensive: once deployed, AI models can generate weather forecasts and climate simulations at a fraction of the traditional computational cost. This efficiency offers a path toward democratizing state-of-the-art modeling, which currently remains accessible mainly to a few well-resourced institutions^{3,88}. Yet “democratization” often means access to run existing models rather than to develop and shape them. Disparities exceed a simple Global North–South divide: many European countries rely on shared infrastructures (e.g., NEMO⁸⁹), while emerging economies such as Brazil and India have built substantial independent capabilities, and the WMO supports AI capacity building and broader participation. Ultimately, AI may both lower and reinforce barriers—reducing entry costs via shared infrastructures yet concentrating power through massive compute and data requirements⁹⁰. In climate science, AI-based models face additional challenges, including structural uncertainty^{58,59}, modeler bias⁶⁰, and unverifiable projections, which limit transparency and accountability⁹¹.

The “black-box” nature of many AI models makes results hard to interpret, and for climate projections and decadal prediction, this opacity can erode trust among policymakers and the public. We must distinguish model explainability, scientific interpretability, and decision-maker trust—the last also depending on robustness, accountability, and clarity under uncertainty^{92,93}. Physics-based models already face trust limits from structural uncertainty, and AI may exacerbate them. Recent mechanistic-interpretability advances move beyond post-hoc explainability⁹⁴ to uncover internal representations in models like AlphaEarth and GraphCast^{95,96}, linking AI behavior to known physical processes; bridging interpretability to actionable trust remains an open interdisciplinary challenge.

Fairness and ethics matter too. Models trained on historical data may reflect existing biases and data gaps, yielding uneven regional performance that risks disproportionately affecting underrepresented or vulnerable communities—for example, underestimating impacts in data-sparse areas such as parts of the Global South with fewer ground stations, thereby misallocating adaptation resources and widening the climate information divide (and AI skill-formation gaps)⁹⁷. It also carries geopolitical weight, as models increasingly influence decisions on bonds, insurance, credit ratings, and investment affecting these countries.

Finally, it is worth noting that the booming climate analytics industry is dominated by a few private players, raising concerns about

transparency and equitable access to climate information^{98,99}. This has prompted calls for open-source, publicly funded climate data platforms to democratize climate knowledge and reduce reliance on proprietary “black-box” solutions¹⁰⁰. Embedding strong ethical frameworks in AI-based climate modeling is also crucial to avoid unchecked modeler biases and ensure trustworthy, justice-oriented climate projections^{101,102}.

AI potential for bridging the weather and climate gap

We survey state-of-the-art AI methodologies applied to weather and climate science, focusing on key components of the data pipeline and modeling strategies⁸, see Table 3. This taxonomic view helps clarify the potential of AI methods for weather and climate science.

Data preprocessing and postprocessing

Different data modalities require tailored processing methods and standards, with specialized tools needed to handle simulation datasets (e.g., ERA5) and improve input quality. Over the years, a few methodologies emerged to fill these gaps. In weather, data augmentation and generative AI can fill gaps and produce realistic synthetic weather patterns¹⁰³ (see Fig. 2). When it comes to regional climate downscaling, techniques such as convolutional neural networks (CNNs), encoder-decoder architectures (for example, U-Net), and diffusion models help downscale input data, thereby enabling finer climate output resolution, which in turn enables better assessment of climate impacts¹⁰⁴. Large language models (LLMs) can serve as post-processing tools for climate or weather model output, helping to convert numerical predictions into natural-language summaries that improve communication with non-experts.

Dimensionality reduction and dynamics

Dimensionality reduction (DR) helps to devise AI algorithms that are more consistent with the data and more accurate, and it can potentially enhance explainability and causal discovery (see Fig. 2, bottom arrows in the illustration panel). DR is used to extract a few essential features from high-dimensional, redundant weather and climate data. Ubiquitous dimensionality-reduction methods, such as empirical orthogonal functions (EOF)¹⁰⁵, are used in both fields. Although EOFs capture dominant variance, they are linear and time-independent, which can limit their applicability to dynamical modeling.

In weather forecasting, DR is usually designed to emphasize initial-value sensitivity. Singular vectors¹⁰⁶ and finite-time Lyapunov exponents¹⁰⁷ are central because they identify the fastest-growing perturbations, revealing flow-dependent directions of instability that dominate short-term forecast error growth. In climate science, DR typically relies on stochastic modeling. The goal is to capture low-frequency variability modes, identify dynamical regimes, and build reduced-order descriptions of the climate system¹⁰⁸. Stochastic climate modeling frameworks^{109,110} gave rise to Linear Inverse Modeling (LIM)¹¹¹, which can represent phenomena such as SST variability as low-dimensional linear systems (Fig. 2). Its nonlinear generalization, Koopman mode decomposition^{112,113}, extends this framework beyond the linear regime.

Modern DR for weather and climate is done using autoencoders¹¹⁴ and can guide causal inference and interpretability^{40,115} (Fig. 2). Recently, foundation models have shown promise in capturing general features and bridging the gap between DR for short-term weather dynamics and long-term climate variability¹¹⁶.

Encoding data structure and knowledge

Domain knowledge embedded in AI is key, and several approaches and frameworks are already showing excellent results in improving accuracy, physical consistency, extrapolation, and speed. A relevant approach is to encode data characteristics into AI architectures. For

Table 3 | Taxonomy of artificial intelligence (AI) and statistical methods used in weather and climate science

	Methodologies	Purpose	Challenges
Data Pre/Post-processing	Simulation and Reanalysis Data	Enhance input data quality, fill missing data, generate realistic weather patterns	Data inconsistencies, observational gaps, limited historical records
	Regional Climate Downscaling	Downscale coarse climate model outputs to high-resolution for impact assessments	Loss of information in downscaling, stochastic uncertainty
	Large Language Models (LLMs)	Assist in natural language explanations of forecasts for non-expert users	Potential biases, lack of robustness in generated summaries
Dimensionality Reduction and Dynamics	Dimensionality Reduction (EOFs, Autoencoders)	Extract key spatial features, remove noise, improve climate variability representation	Difficulty in capturing nonlinear dynamics, limited generalizability
	Dynamical Systems Modeling (SV, LIM, Koopman)	Extract modes of variability, emulate climate dynamics	Limited training data, challenges in representing chaotic climate processes
Encoding Data Structure and Knowledge	Graph Neural Networks	Represent Earth as graph structure. Encode complex spatial and temporal relationships in geospatial data	Computational complexity, need for temporal graph models
	Hybrid Modeling (AI + Physical Models)	Integrate AI with conservation laws and physics-based models for greater accuracy	Balancing physical constraints with AI-driven flexibility
Foundation Models	Earth system FMs (e.g., ClimaX, Prithvi, Aurora), Multi-modal FMs	Create general-purpose, task-agnostic representations; enable efficient transfer learning for diverse downstream tasks with limited data	High pre-training costs; handling heterogeneous data sources; ensuring stability across multiple scales
Generative and Probabilistic AI	Generative Models (Diffusion, GANs, VAEs)	Enable uncertainty quantification and scenario-based climate modeling	High computational costs, stability issues in training
	Probabilistic AI	Provide full probability distributions for forecasts instead of single predictions	Over-smoothing, loss of small-scale variability
	Deterministic Emulators	Enable fast ensemble simulations, improving extreme event forecasting	Approximations may reduce fidelity to real-world physics
XAI and Interpretable AI	Explainable AI (XAI)	Enhance trust and interpretability of AI models for climate applications	Difficulties in defining explainability metrics for AI models
	Causal AI and Interpretable AI	Discover governing equations, parameterizations, and causal relationships in climate data	Causal discovery remains challenging due to limited observational constraints

The table categorizes methodologies across different aspects of the data pipeline, from data preprocessing to forecasting, uncertainty quantification, and explainability, illustrating their role in improving accuracy, efficiency, and interpretability in climate and weather modeling.

example, Graph Neural Networks (GNNs) represent geospatial data as graphs, in which nodes, edges, and messages capture spatial relations, thereby these networks excel at modeling complex relational structures. A prime example is GraphCast¹³, which uses a multi-scale graph mesh and attention-based message passing to predict global weather at 0.25° resolution, outperforming operational deterministic models on over 90% of 1380 targets¹³. Beyond static topologies, dynamic GNNs adapt to evolving interactions over time and scale, while recent graph transformers merge message passing with temporal attention to jointly model short-term weather and longer-term climate dependencies (Fig. 2).

Other AI model architectures may include intrinsic properties of data-generating processes, such as symmetries or invariances. Equivariant architectures explicitly encode symmetry and geometric structure (e.g., rotations or spherical geometry), making them especially well-suited to Earth system modeling, where the domain is inherently spherical, and symmetry is present. For example, the Spherical Fourier Neural Operator (SFNO) uses spherical harmonic transforms to respect rotations and the sphere's geometry, producing more stable long-rollout predictions in atmospheric dynamics^{117,118}. While applications in climate science are still emergent, these models are increasingly used in both weather forecasting and climate modeling¹¹⁹.

The emerging field of hybrid ML, also referred to as physics-aware, physics-guided, or knowledge-based AI, is widely used in Earth system modeling⁷. These approaches combine ML with physical models (e.g., data assimilation or combining outputs from both model types). They can enhance adherence to physical laws in AI systems (e.g., conservation of mass, energy, and balance laws) and strengthen interpretability and extrapolation capabilities⁷ (Fig. 2). Hybrid modeling plays different roles across timescales: in weather, it primarily targets short-term forecast skill and stability, whereas in climate, it must ensure long-term stability, avoid drift, and produce consistent forced responses. At the same time, both domains use hybrid

approaches for tasks such as parameterization, emulation, and bias correction, making this a key area of overlap and convergence. They allow us to obtain robust model runs over long periods, necessary from subseasonal forecasts to climate projections, and to build trust in the predictions, which is crucial for decision-making^{3,120}.

However, hybrid physics-AI approaches face significant limitations, including stability issues and the risk that AI components inherit and amplify biases from the physical models they emulate—a “bias amplification” effect noted in recent work (e.g., NeuralGCM³). While AI components may correct biases within the training regime, they can behave unpredictably outside it. As a mitigation strategy, training approaches based on data assimilation increments leverage observational constraints and can reduce the reinforcement of bias. Additionally, hybrid approaches have shown promising results when combined with transfer learning strategies that allow models to be efficiently adapted as the underlying physical model evolves, without requiring full retraining (e.g., transfer-learning hybrid climate models^{121,122}).

ML and data assimilation are increasingly converging^{123–126}. In hybrid DA–ML systems, DA assimilates sparse, noisy data within ML dynamical models, providing state estimates and uncertainty quantification^{127–129}, while designs such as variational autoencoders embedded in DA workflows handle nonlinear, non-Gaussian dynamics^{130,131}. ML can also fully emulate the DA process^{132–136}, with end-to-end systems learning the full DA–forecasting cycle directly from data^{137,138} (Fig. 2). Extending this to climate prediction requires coupled DA, propagating observational influence from all components (atmosphere, ocean, land, cryosphere) across the multi-scale system¹³⁹.

Foundation models

Traditional DL models need tailored datasets for each task, which is infeasible given the wide array of weather and climate problems.

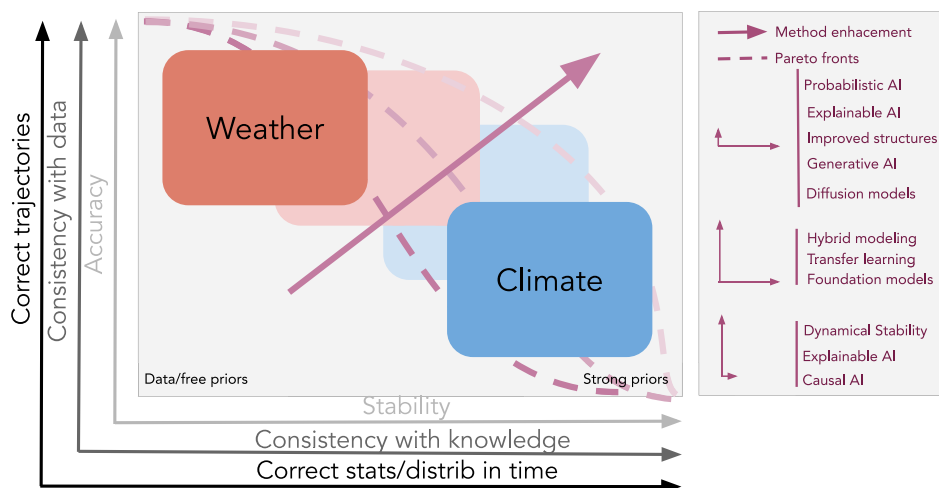


Fig. 2 | Artificial intelligence (AI) bridging weather and climate across key modeling dimensions. Idealized, conceptual, representation of the trade-off space between model *stability* and *accuracy* (or, equivalently, consistency with domain knowledge vs. data fidelity). AI methods progressively shift the Pareto front (dashed lines) toward higher performance and stability. Arrows highlight areas of advancement driven by explainable AI, improved architectures, generative and

probabilistic approaches, diffusion models, transfer learning, hybrid modeling, and foundation models—collectively pushing toward unified, physically consistent AI frameworks for weather and climate prediction. Colored dashed lines denote successive Pareto fronts (lighter to darker with increasing AI capability), and colored arrows mark the corresponding methodological advances.

Foundation models (FMs) instead leverage large, unlabeled, heterogeneous datasets to learn general-purpose representations adaptable to any downstream task¹⁴⁰. FMs lower barriers to scientific use by providing a unified approach across the Earth system sciences. These representations form a reusable backbone that can be fine-tuned with limited additional data, enabling efficient transfer to downstream applications¹⁴¹. This paradigm is particularly fitting to the weather and climate domains, given the freely available observational and reanalysis data sources.

FMs remain an emerging direction rather than a demonstrated capability for unification across Earth system sciences. Key open questions concern physical consistency and OOD behavior, since evidence for robust generalization across model generations (e.g., CMIP5 to CMIP6) or under changing forcing is still lacking. Recent approaches either enforce constraints in the architecture (e.g., ACE2¹⁴², with mass and moisture conservation) or learn physically consistent behavior implicitly from large datasets (e.g., Aurora¹⁴³, Aardvark¹³⁸), while multimodal and time-series FMs¹⁴⁴ target non-stationarity and extrapolation; concrete benchmarks (S2S skill, conservation-error thresholds, robustness to perturbed forcings) will be vital for validation. The domain's multimodality and complex variable interactions enable diverse pretraining paradigms, including autoregressive training, information reconstruction, and instance discrimination. Earth observation has seen rapid adoption of such models¹⁴⁵, a domain that shares fundamental characteristics and challenges with weather and climate science: multi-modality, multiple scales, and spatio-temporal variability. Recent Earth observation models have demonstrated a paradigm shift towards task and sensor-agnostic FMs that flexibly handle varying sensor combinations tailored to the needs of the downstream task without sacrificing efficiency^{146–148}. Similar solutions are needed for weather and climate models to accommodate diverse downstream applications across heterogeneous training datasets, given the wide range of atmospheric variables, data sources (reanalysis, observations, and simulations), and spatio-temporal resolutions.

Complementary to generalist feature extractors, other approaches have attempted to provide general-purpose, location-specific representations by aligning location embeddings with rich satellite image embeddings¹⁴⁹. Extending this to account for temporal variations and diverse weather and climate information holds great promise

for a more holistic representation. Targeted AI systems are an important complement to foundation models. For example, Google's [Weather Lab for tropical cyclones illustrates](#) how focused applications can deliver actionable improvements. However, such systems carry risks, including potential overfitting to specific tasks, limited transferability, and dependence on proprietary data and infrastructure. The increasing role of private actors may also introduce constraints on data access, geopolitical pressures, and coordination (e.g., Reuters satellite imaging delay; WMO cooperation challenges).

Similarly, FMs have been widely adopted in climate and weather studies¹⁵⁰, with models trained on decades of climate simulations (e.g., *Climax*¹¹⁶, pre-trained on CMIP6), reanalysis data (e.g., *Prithvi WxC*¹⁵¹, pre-trained on MERRA-2), or varied climate datasets (e.g., *Aurora*¹⁴³, pre-trained on ERA5, and others). Recent FMs bridge scales and modalities such as *WeatherGenerator*, which includes data ranging from satellite imagery and reanalysis to ground and ocean measurements, and *AlphaEarth Foundations*¹⁵², which fuses multi-source satellite imagery and text embeddings along with reanalysis and geophysical data to provide unified annual embeddings.

These examples show how a single pre-training supports diverse downstream tasks—weather forecasting, climate modeling, downscaling, land-cover mapping, and extreme-event detection—progressively pushing the Pareto front toward better stability and accuracy (Fig. 2). By synthesizing vast datasets with minimal task-specific training, FMs lower barriers for scientific and operational use, potentially accelerating climate and weather insights.

Generative AI

Generative models (e.g., diffusion models, flow-based and flow-matching models, variational autoencoders (VAEs), generative adversarial networks (GANs) enable sampling from complex, high-dimensional distributions and are increasingly applied to weather and climate problems. Recent applications demonstrate their ability to generate physically realistic high-resolution climate fields, perform statistical downscaling, augment or gap-fill incomplete datasets, emulate expensive numerical simulations, and sample diverse weather scenarios, including extreme events. These approaches represent a shift along the Pareto front toward improved stability and correct data distribution (Fig. 2), though

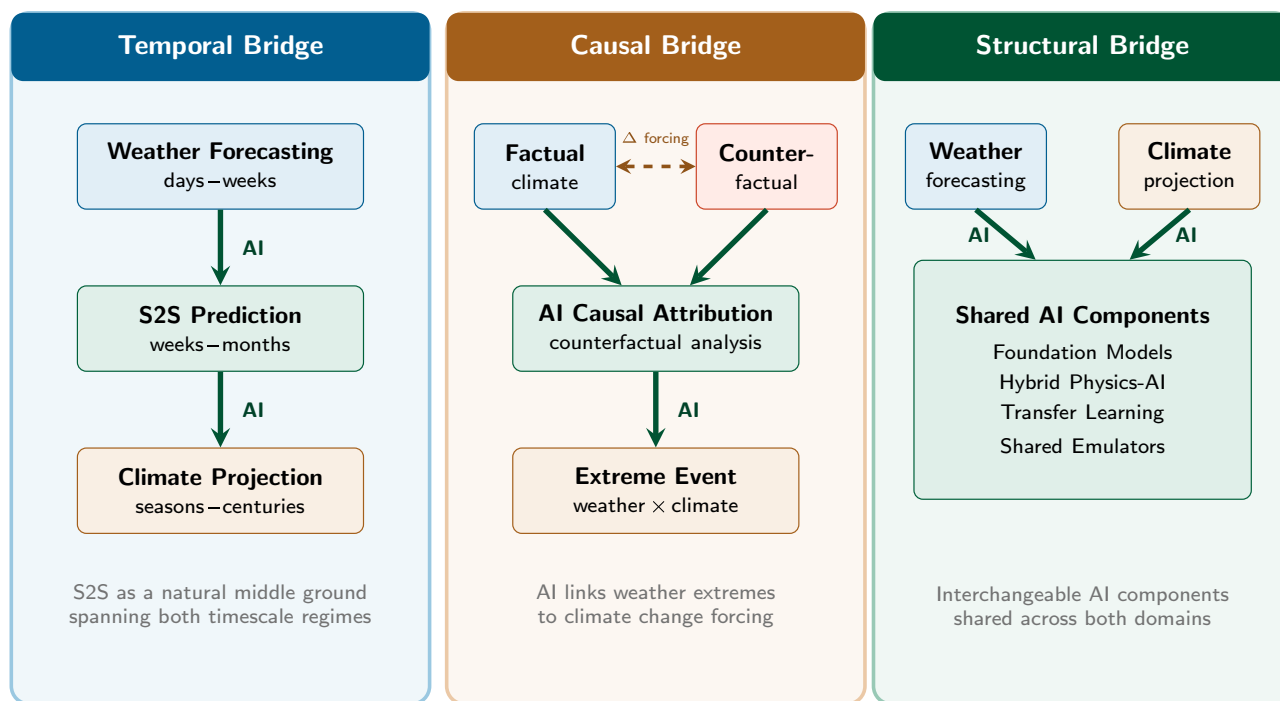


Fig. 3 | Three artificial intelligence (AI)-mediated bridges between weather and climate science. (Left) The *temporal bridge*: subseasonal-to-seasonal (S2S) prediction sits naturally between weather forecasting (days–weeks) and climate projection (seasons–centuries), enabling AI models to span the full timescale continuum^{34–37}. (Center) The *causal bridge*: AI causal attribution methods compare factual and counterfactual climate states to quantify the contribution of climate

change to individual extreme weather events, linking weather-scale dynamics to climate forcing^{116,161}. (Right) The *structural bridge*: foundation models, hybrid physics-AI, transfer learning and shared emulators serve as interchangeable components shared across weather forecasting and climate projection, creating a common technical infrastructure^{125,147}.

they typically prioritize statistical realism over strict physical consistency.

GANs, flow-matching and diffusion-based approaches have shown particular promise for spatial refinement, refining coarse-resolution inputs to kilometer-scale fields while preserving fine-scale variability and physical realism (e.g., *CorrDiff*⁴⁵³, *TAUDiff*⁴⁵⁴, *GenFocal*⁴⁵⁵). Latent diffusion methods perform downscaling in compressed representation spaces to improve computational efficiency⁴⁵⁶, achieving competitive results at a reduced cost.

For global weather forecasting, diffusion models trained on reanalysis data produce probabilistic ensemble forecasts with improved skill compared to operational systems, particularly for extreme events and tropical cyclones (e.g., *GenCast*²⁸), while functional generative networks achieve faster inference with competitive or superior performance (e.g., *WeatherNext* 2⁵⁷).

Modern generative methods offer fast inference, making large ensembles feasible and providing diverse, physically consistent realizations that are useful for quantifying uncertainty and characterizing extremes. Accurately representing tail behavior is crucial for extreme-event assessment and attribution studies, where probabilistic models offer clear advantages⁸. Complementary probabilistic ML approaches, such as Bayesian neural networks, deep ensembles, and conformal prediction, further support uncertainty quantification by providing calibrated distributional forecasts²⁸, collectively enhancing decision-making under uncertainty and risk assessment for high-impact weather and climate events.

Explainable and causal AI

Explainable and causal AI techniques are critical for bridging the gap between opaque data-driven models and physically interpretable methods in weather and climate science. These approaches enable transparency, ensure physical consistency, and support hybrid modeling that spans weather to climate prediction^{7,115}.

Explainable AI (XAI) uses attribution methods, saliency maps, and physics-informed feature analysis to reveal the mechanisms behind model predictions. This transparency enhances trust and enables identification of key physical drivers behind weather or climate phenomena¹¹⁵ (Fig. 2). Relatedly, Interpretable AI (IAI) aims to build glass-box models or extract governing equations or simplified laws from data, advancing parameterization and enabling operational use of ML-derived insights in simulation and forecasting systems¹⁵⁸.

Explaining AI models does not imply understanding processes and their relations, either mechanistically or causally. Causal AI integrates causal reasoning into hybrid Earth system models, enabling the identification of cause-and-effect relationships rather than mere correlations (Fig. 2). This capability improves subgrid-scale parameterization, supports model evaluation by aligning causal mechanisms in observations and simulations, and strengthens generalization across regime changes and scales from weather to climate^{115,159,160}.

Three AI-mediated bridges

The methods surveyed in the preceding sections collectively enable three concrete, complementary bridges between weather and climate science (Fig. 3): temporal, causal and structural bridges.

The *temporal bridge* is realized through subseasonal-to-seasonal (S2S) prediction, which occupies the intermediate timescale regime between weather forecasting (days to weeks) and climate projection (seasons to centuries). AI models have recently demonstrated skillful predictions at S2S lead times—a regime historically underserved by both communities^{34–37}. By learning representations that remain coherent across medium-range atmospheric dynamics and slowly evolving boundary-forced variability, these models serve as a natural conduit between the two domains. Progress in the *temporal bridge* directly benefits both communities: improved S2S skill informs climate model initialization and provides a testbed for evaluating whether AI

representations remain physically consistent beyond the medium range.

The *causal bridge* connects individual extreme weather events to the long-term evolution of the climate system. AI causal inference methods^{115,159} enable efficient comparison of large ensembles of factual and counterfactual climate simulations, identifying the causal pathway from radiative forcing to observable extremes¹⁶⁰. This bridge is bidirectional: causal AI applied to weather-scale events can test and constrain climate model projections of extremes, while climate-scale counterfactual analyses sharpen attribution statements at the event level, directly linking weather dynamics to climate change.

The *structural bridge* emerges from the growing convergence of AI architectures and model components used in both fields. Foundation models, hybrid physics-AI systems, and transfer learning strategies are increasingly applied to weather forecasting and climate projection alike^{3,121,143}. The interchangeability of these components (e.g., ML-based parameterizations developed for climate models being reused in weather systems, or weather-scale data assimilation techniques informing climate initialization) creates a shared technical layer that accelerates progress in both domains simultaneously. This structural convergence is also reflected in the bibliometric analysis (Fig. 1), where shared methodological vocabulary has grown substantially since the deep-learning revolution.

Outlook: open problems and future directions

AI as a catalyst for seamless weather-climate modeling

Building on the section Distinct challenges in weather and climate, we highlight key research directions at the interface of AI, weather, and climate (Fig. 2). Weather and climate models have traditionally occupied different regions of this trade-off space, shaped by distinct objectives, methods, and metrics. Recent AI approaches are relaxing these trade-offs, enabling models that better balance data fidelity, physical constraints, and dynamical behavior, and opening a shared modeling regime spanning timescales. AI is thus not just another tool but a catalyst for tighter integration across scales, model components, and communities.

Recent progress in hybrid modeling, foundation models, and probabilistic learning shows AI can cross long-standing weather-climate boundaries. Efforts such as ECMWF's *WeatherGenerator* show how AI can augment weather prediction, while ML-based parameterizations have improved climate models in convection, cloud microphysics, land-atmosphere coupling, and sea-ice-radiation interactions^{3,161–165}. Yet a coherent framework unifying the two remains lacking.

Towards a unifying methodology

We identify two directions central to this integration. First, end-to-end hybrid physics-AI systems offer a principled way to learn from observations while preserving physical consistency, supporting stable forecasts and long-term projections. Second, multi-scale, scale-aware foundation models can learn shared representations that transfer across tasks, from weather prediction and seasonal forecasting to downscaling and climate projection. Together, these advances point toward seamless Earth system AI models that span nowcasting, medium-range forecasting, subseasonal-to-seasonal prediction, and decadal-to-centennial timescales.

Realizing this vision requires key capabilities and specific actors to drive change. First, modeling centers and researchers should develop differentiable Earth system components over the next 3–5 years; these are essential for end-to-end training, sensitivity analysis, and uncertainty propagation¹⁶⁶. Concretely, modeling centers such as ECMWF, NCAR, and national meteorological services should prioritize releasing modular, open-source, differentiable components compatible with existing ESM frameworks, with CMIP-compatible interfaces to enable community reuse. These components must be modular and

composable, allowing parameterizations and emulators to be exchanged, optimized, and audited, while making structural uncertainty explicit. This shift calls for automated calibration and uncertainty quantification in place of ad hoc tuning. Finally, closer integration of data assimilation and machine learning is needed to address observational constraints and uncertainties arising from initial conditions and model structure in a unified way across coupled components and timescales.

Evaluation, robustness and trust across timescales

As a second action, international consortia such as CMIP, with operational centers including ECMWF and WMO, must jointly establish robust validation benchmarks for AI-enabled ESMs within a few years: joint weather-climate AI benchmark suites with standardized metrics, CMIP-compatible AI training splits under open-access mandates, and intercomparison experiments such as AIMIP¹⁶⁷ that hold AI and physics-based models to the same standards. This includes physics-aware diagnostics, such as conservation checks, as well as stability and behavior stress tests across regimes to avoid brittle extrapolation. For seamless prediction, benchmarks must span timescales and resolutions, covering short-term forecasts, non-stationary behavior, and responses to changing forcing. Concrete examples of “robust evaluation metrics” go beyond traditional grid-point error measures such as RMSE. They include (i) physical consistency, enforcing conservation of energy, mass, and moisture; (ii) probabilistic skill via proper scoring rules for distributional forecasts^{168,169}; (iii) spatio-temporal coherence of forecast trajectories; (iv) model-evidence strategies¹⁷⁰; and (v) causal or process-level consistency^{171–174}. Due to the lack of ground truth for climate projections, evaluation practices must differ, relying on surrogate validation strategies like perfect-model experiments, hindcasts, and paleoclimate analogs. Climate model evaluation (e.g., CMIP-style intercomparisons¹⁷⁵) relies on a combination of historical skill¹⁷⁶, process-based diagnostics testing causal dynamics^{177,178}, and inter-model consistency¹⁷⁹ rather than strict out-of-sample verification. Joint evaluation is critical to ensure that progress along the Pareto front does not come at the expense of performance in other dimensions (Fig. 2).

AI approaches reach their full potential when combined with complementary methodological advances. Tail-aware sampling and rare-event emulation, supported by probabilistic and generative frameworks, are thus essential for efficient representation of extremes and risk. Interpretable machine-learning methods can mitigate black-box behavior, support physical insight, and improve trust in extrapolation. Finally, causal surrogate models enable policy-relevant scenario analysis by supporting counterfactual “what if” questions for decision-making.

Towards social and ethical AI-Earth systems

Expanding the shared weather-climate regime in Fig. 2 depends on boundary conditions shaping the sustainability, equity, and trustworthiness of AI advances. A third action: funding agencies and the global research community must prioritize sustainability and democratization, with funding calls—starting immediately—that demand energy-efficient models, scalable training, and reduced-complexity variants enabling participation beyond a few well-resourced institutions. Funding agencies should explicitly require energy and compute reporting in grant applications, and support dedicated compute allocations for under-resourced institutions in the Global South. Open-access mandates for reanalysis-derived AI training corpora—modeled on initiatives such as ERA5's Copernicus CLIMATE Data Store—would lower entry barriers without compromising data quality. Neglecting these constraints risks concentrating climate modeling capacity and undermining reproducibility.

Open and shared datasets, pre-training corpora, and benchmarks are equally essential. Open access to reanalysis and observational data,

model code, parameters, and evaluation protocols underpins transparency and cumulative progress. This must be accompanied by strong standards for documentation, skill reporting, and reproducibility, which are critical for operational use and policy-relevant applications.

Bridging communities

Realizing this vision ultimately requires seamless collaboration among domain scientists (weather, climate, Earth systems), AI researchers (mathematics, computer science, engineering), and institutions (funding agencies, operational centers, and public and private organizations). Progress is encouraging, but it must proceed in all directions, fostering an environment where expertise is genuinely shared and mutually reinforcing.

This calls for a fourth line of action related to changes in how communities interact, including more fluid and informal modes of collaboration that enable cross-fertilization and serendipity^{180,181}. The scientific community and institutional organizers must immediately implement practical steps, including regular interdisciplinary workshops, joint weather-climate-AI conferences, and shared benchmark efforts and competitions^{182,183}. Initiatives such as the Earth Virtualization Engines (EVE) illustrate how shared platforms and digital twins can align modeling groups, AI developers, and stakeholders around common components, data standards, and evaluation pipelines, accelerating the transfer of advances from research to operations.

Strategic commitments to AI integration at major centers, such as ECMWF, show how operational settings can serve as testbeds for methodological advances under real-world constraints. The private sector and foundation model developers also play a key role by providing scalable compute, large pre-training corpora, and shared model evaluations, while adhering to open standards and equitable access. AI is a critical catalyst, but it cannot address these challenges in isolation. Stronger links among meteorologists, climate and Earth scientists, AI researchers, engineers, mathematicians, funders, and policymakers are essential. Operational centers should open testbed programs for external AI researchers; funding agencies should mandate interdisciplinary co-supervision of doctoral students; and the community should establish recurring joint weather-climate-AI summer schools targeting early-career researchers, building on models such as the ECMWF training courses, the ESA-NASA ML4Earth Workshop Series and the NSF NCAR's AI strategy¹⁸⁴.

Through sustained and structured collaboration—across institutions, career stages, and disciplines—AI can help advance modeling, improve forecasts, and support more informed decisions in a changing climate. The steps outlined here are not merely aspirational: several, such as benchmark coordination through CMIP, open training data releases, and interdisciplinary summer schools, can begin immediately with existing infrastructure and community engagement.

References

- de Burgh-Day, C. O. & Leeuwenburg, T. Machine learning for numerical weather and climate modelling: a review. *Geosci. Model Dev.* **16**, 6433–6477 (2023).
- Conti, S. Artificial intelligence for weather forecasting. *Nat. Rev. Electr. Eng.* **1**, 8 (2024).
- Kochkov, D. et al. Neural general circulation models for weather and climate. *Nature* **632**, 1060–1066 (2024).
- Eyring, V., Gentile, P., Camps-Valls, G., Lawrence, D. M. & Reichstein, M. AI-empowered next-generation multiscale climate modelling for mitigation and adaptation. *Nat. Geosci.* **17**, 963–971 (2024).
- Eyring, V. et al. Pushing the frontiers in climate modelling and analysis with machine learning. *Nat. Clim. Change* **14**, 916–928 (2024).
- Camps-Valls, G., Tuia, D., Zhu, X. X. & Reichstein, M. *Deep learning for the Earth Sciences: A comprehensive approach to remote sensing, climate science and geosciences* (John Wiley & Sons, 2021).
- Reichstein, M. et al. Deep learning and process understanding for data-driven Earth system science. *Nature* **566**, 195–204 (2019).
- Camps-Valls, G. et al. Artificial intelligence for modeling and understanding extreme weather and climate events. *Nat. Commun.* **16**, 1919 (2025).
- Ghil, M. & Lucarini, V. The physics of climate variability and climate change. *Rev. Mod. Phys.* **92**, 035002 (2020).
- Bonavita, M. On some limitations of current machine learning weather prediction models. *Geophys. Res. Lett.* **51**, e2023GL107377 (2024).
- Schmidt, G. AI Caramba. <https://www.realclimate.org/index.php/archives/2024/12/ai-caramba/> (2024).
- Pathak, J. et al. FourCastNet: a global data-driven high-resolution weather model using adaptive Fourier neural operators. Preprint at <https://doi.org/10.48550/arXiv.2202.11214> (2022).
- Lam, R. et al. Learning skillful medium-range global weather forecasting. *Science* **382**, 1416–1421 (2023).
- McNally, A. et al. Data driven weather forecasts trained and initialised directly from observations. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2407.15586> (2024).
- Hersbach, H. et al. The era5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
- Bi, K. et al. Accurate medium-range global weather forecasting with 3D neural networks. *Nature* **619**, 533–538 (2023).
- Deutscher Wetterdienst (DWD). *AICON: Artificial Intelligence for the Continuous Numerical Weather Prediction* (Deutscher Wetterdienst (DWD), 2025). https://www.dwd.de/EN/Home/home_node.html.
- National Oceanic and Atmospheric Administration (NOAA). *NOAA Artificial Intelligence Global Forecast System (AIGFS)* (NOAA, 2025) <https://www.noaa.gov/news-release/noaa-deploys-new-generation-of-ai-driven-global-weather-models>.
- Nipen, T. N. et al. Regional data-driven weather modeling with a global stretched-grid. *Artif. Intell. Earth Syst.* **5**, 250001 (2026).
- Dheeshjith, A. V. et al. Transfer learning for emulating ocean climate variability across CO₂ forcing. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2405.18585> (2024).
- Bouabid, S. et al. FaIRGP: a Bayesian energy balance model for surface temperature emulation. *J. Adv. Model. Earth Syst.* **16**, e2023MS003926 (2024).
- Schneider, T., Leung, L. & Wills, R. Optimizing climate models with process knowledge, resolution, and artificial intelligence. *Atmos. Chem. Phys.* **24**, 7041–7060 (2024).
- Wilkinson, R. et al. Process-based machine learning observationally constrains climate model uncertainty in future regional temperature projections. *J. Geophys. Res. Atmos.* **130**, e2025JH000698 (2025).
- Pasula, A. & Subramani, D. N. Global climate model bias correction using deep learning. *Mach. Learn. Earth* **1**, 015001 (2025).
- Nivron, O., Wischik, D. J., Vrac, M., Shuckburgh, E. & Archibald, A. T. A temporal stochastic bias correction using a machine learning attention model. *Environ. Data Sci.* **3**, e36 (2024).
- Gao, W. et al. Deconfounding multi-cause latent confounders: a factor-model approach to climate model bias correction. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2408.12063> (2024).
- Li, L., Carver, R., Lopez-Gomez, I., Sha, F. & Anderson, J. Generative emulation of weather forecast ensembles with diffusion models. *Sci. Adv.* **10**, eadk4489 (2024).
- Price, I. et al. Probabilistic weather forecasting with machine learning. *Nature* **637**, 84–90 (2025).

29. Randall, D. A. & Emanuel, K. The weather-climate schism. *Bull. Am. Meteorol. Soc.* **105**, E300 – E305 (2024).
30. Benjamin, S. G. et al. 100 years of progress in forecasting and NWP applications. *Meteorol. Monogr.* **59**, 13.1–13.67 (2018).
31. Vitart, F. & Brown, A. S2S forecasting: towards seamless prediction. *WMO Bull.* **68**, 70–74 (2019).
32. Ruti, P. M. et al. Advancing research for seamless earth system prediction. *Bull. Am. Meteorol. Soc.* **101**, E23 – E35 (2020).
33. Palmer, T. N., Doblas-Reyes, F. J., Weisheimer, A. & Rodwell, M. J. Toward seamless prediction: calibration of climate change projections using seasonal forecasts. *Bull. Am. Meteorol. Soc.* **89**, 459–470 (2008).
34. Li, G. et al. TianQuan-S2S: a subseasonal-to-seasonal global weather model via incorporate climatology state. In *Int. Conf. Learn. Representations* (2025).
35. Chen, L. et al. A machine learning model that outperforms conventional global subseasonal forecast models. *Nat. Commun.* **15**, 6425 (2024).
36. Karlbauer, M. et al. Advancing parsimonious deep learning weather prediction using the healpix mesh. *J. Adv. Model. Earth Syst.* **16**, e2023MS004021 (2024).
37. Kent, C. et al. Skilful global seasonal predictions from a machine learning weather model trained on reanalysis data. *npj Clim. Atmos. Sci.* **8**, 314 (2025).
38. Kaack, L. H. et al. Aligning artificial intelligence with climate change mitigation. *Nat. Clim. Change* **12**, 518–527 (2022).
39. McGovern, A. et al. The value of convergence research for developing trustworthy AI for weather, climate, and ocean hazards. *npj Nat. Hazards* **1**, 13 (2024).
40. Bommer, P. L., Hedström, A., Bareeva, D. & Höhne, M. M.-C. Finding the right XAI method - a guide for the evaluation and ranking of explainable AI methods in climate science. *Artif. Intell. Earth Syst.* **3**, e230074 (2024).
41. Phillips, N. A. The general circulation of the atmosphere: a numerical experiment. *Q. J. R. Meteorol. Soc.* **82**, 123–164 (1956).
42. Edwards, P. N. History of climate modeling. *WIREs Clim. Change* **2**, 128–139 (2011).
43. Geophysical Fluid Dynamics Laboratory. Brief history of global atmospheric modeling at GFDL <https://www.gfdl.noaa.gov/brief-history-of-global-atmospheric-modeling-at-gfdl/> (2025) accessed 1 Aug 2026.
44. Mylne, K. Probability forecasts—part 1: ensembles and probabilistic forecasts. *Weather* **81**, 82–88 (2026).
45. Spiegelhalter, D. *The Art of Statistics: Learning from Data* (Penguin UK, 2019).
46. Bethke, I. et al. NorCPM1 and its contribution to CMIP6 DCP. *Geosci. Model Dev.* **14**, 7073–7116 (2021).
47. Soares, P. M. M. et al. High-resolution downscaling of cmip6 earth system and global climate models using deep learning for iberia. *Geosci. Model Dev.* **17**, 229–259 (2024).
48. Bauer, P., Thorpe, A. & Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55 (2015).
49. Hawkins, E. & Sutton, R. The potential to narrow uncertainty in regional climate predictions. *Bull. Am. Meteorol. Soc.* **90**, 1095–1108 (2009).
50. Lehner, F. et al. Partitioning climate projection uncertainty with multiple large ensembles and cmip5/6. *Earth Syst. Dyn.* **11**, 491–508 (2020).
51. Nelson, J. A. et al. X-base: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, fluxcom-x. *Biogeosciences* **21**, 5079–5115 (2024).
52. Ragone, F., Wouters, J. & Bouchet, F. Computation of extreme heat waves in climate models using a large deviation algorithm. *Proc. Natl. Acad. Sci. USA* **115**, 24–29 (2018).
53. Coles, S. *An Introduction to Statistical Modeling of Extreme Values* <https://doi.org/10.1007/978-1-4471-3675-0> (Springer-Verlag, 2001).
54. Zhang, Z., Fischer, E., Zscheischler, J. & Engelke, S. Numerical models outperform AI weather forecasts of record-breaking extremes <https://arxiv.org/abs/2508.15724> (2025).
55. Flato, G. M. Earth system models: an overview. *WIREs Clim. Change* **2**, 783–800 (2011).
56. de Melo Virissimo, F. & Stainforth, D. A. Micro- and macroparametric uncertainty in climate change prediction: a large ensemble perspective. *Bull. Am. Meteorol. Soc.* **106**, E1319 – E1341 (2025).
57. Hourdin, F. et al. The art and science of climate model tuning. *Bull. Am. Meteorol. Soc.* **98**, 589–602 (2017).
58. Zheng, Z., Zhao, L. & Oleson, K. Large model structural uncertainty in global projections of urban heat waves. *Nat. Commun.* **12**, 3736 (2021).
59. Holland, M. M. et al. New model ensemble reveals how forcing uncertainty and model structure alter climate simulated across cmip generations of the community earth system model. *Geosci. Model Dev.* **17**, 1585–1602 (2024).
60. Pulkkinen, K. et al. The value of values in climate science. *Nat. Clim. Change* **12**, 4–6 (2022).
61. Martin, A. P. et al. When to add a new process to a model - and when not: a marine biogeochemical perspective. *Ecol. Model.* **498**, 110870 (2024).
62. Roebber, P. J., Schultz, D. M., Colle, B. A. & Stensrud, D. J. Toward improved prediction: high-resolution and ensemble modeling systems in operations. *Weather Forecast.* **19**, 936–949 (2004).
63. Simmons, A. J. & Hollingsworth, A. Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol. Soc.* **128**, 647–677 (2002).
64. Carrasi, A., Bocquet, M., Bertino, L. & Evensen, G. Data assimilation in the geosciences: an overview of methods, issues, and perspectives. *Wiley Interdiscip. Rev. Clim. Change* **9**, e535 (2018).
65. Mass, C. F., Ovens, D., Westrick, K. & Colle, B. A. Does increasing horizontal resolution produce more skillful forecasts?: the results of two years of real-time numerical weather prediction over the Pacific Northwest. *Bull. Am. Meteorol. Soc.* **83**, 407–430 (2002).
66. Davies-Barnard, T., Zaehle, S. & Friedlingstein, P. Assessment of the impacts of biological nitrogen fixation structural uncertainty in CMIP6 earth system models. *Biogeosciences* **19**, 3491–3503 (2022).
67. de Melo Virissimo, F., Martin, A. P. & Henson, S. A. Influence of seasonal variability in flux attenuation on global organic carbon fluxes and nutrient distributions. *Glob. Biogeochem. Cycles* **36**, e2021GB007101 (2022).
68. Huntzinger, D. N. et al. Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Sci. Rep.* **7**, 4765 (2017).
69. de Melo Virissimo, F., Stainforth, D. A. & Bröcker, J. The evolution of a non-autonomous chaotic system under non-periodic forcing: a climate change example. *Chaos* **34**, 013136 (2024).
70. Rasp, S. et al. WeatherBench 2: a benchmark for the next generation of data-driven global weather models. *J. Adv. Model. Earth Syst.* **16**, e2023MS004019 (2024).
71. Lee, E. & Hong, S.-Y. Impact of the sea surface salinity on simulated precipitation in a global numerical weather prediction model. *J. Geophys. Res. Atmos.* **124**, 719–730 (2019).
72. Deser, C. et al. Effects of macro vs. micro initialization and ocean initial-condition memory on the evolution of ensemble spread in the CESM2 large ensemble. *Clim. Dyn.* **63**, 62 (2025).
73. Lenton, T. M. et al. Tipping elements in the earth's climate system. *Proc. Natl. Acad. Sci. USA* **105**, 1786–1793 (2008).

74. Lucarini, V. & Bódai, T. Global stability properties of the climate: melancholia states, invariant measures, and phase transitions. *Nonlinearity* **33**, R59 (2020).
75. Buizza, R. The value of probabilistic prediction. *Atmos. Sci. Lett.* **9**, 36–42 (2008).
76. Bouallègue, Z. B. et al. The rise of data-driven weather forecasting: a first statistical assessment of machine learning-based weather forecasts in an operational-like context. *Bull. Am. Meteorol. Soc.* **105**, E864 – E883 (2024).
77. Stainforth, D. A. *Predicting Our Climate Future: What We Know, What We Don't Know, And What We Can't Know* 1st edn (Oxford University Press, 2023).
78. Sun, Y. Q. et al. Can AI weather models predict out-of-distribution gray swan tropical cyclones? *Proc. Natl. Acad. Sci. USA*. **122**, e2420914122 (2025).
79. Hoskins, B. The potential for skill across the range of the seamless weather-climate prediction problem: a stimulus for our Science. *Q. J. R. Meteorol. Soc.* **139**, 573–584 (2013).
80. Knüsel, B. et al. Applying machine learning to climatology: a review. *Wiley Interdiscip. Rev. Clim. Change* **11**, e632 (2020).
81. Acosta Navarro, J. C. et al. Seamless seasonal to multi-annual predictions of temperature and standardized precipitation index by constraining transient climate model simulations. *Earth Syst. Dyn.* **16**, 1723–1737 (2025).
82. Thompson, E. *Escape from Model Land: How Mathematical Models Can Lead Us Astray and What We Can Do About It* (Basic Books, 2022).
83. Lovejoy, S. Review article: Scaling, dynamical regimes, and stratification. How long does weather last? How big is a cloud? *Non-linear Process. Geophys.* **30**, 311–374 (2023).
84. Chen, M. et al. Enhancing reproducibility in hybrid Earth system models. *Nat. Mach. Intell.* <https://doi.org/10.1038/s42256-026-01299-5> (2026).
85. Kashinath, K. et al. Physics-informed machine learning: case studies for weather and climate modelling. *Phil. Trans. R. Soc. A* **379**, 20200093 (2021).
86. Van Katwyk, P., Fox-Kemper, B., Hewitt, H. T. & Bergen, K. J. Rewiring climate modeling with machine learning emulators. *Commun. Earth Environ.* **7**, 107 (2026).
87. Landsberg, J. B. & Barnes, E. A. Forecasting the future with yesterday's climate: temperature bias in AI weather and climate models. *Geophys. Res. Lett.* **53**, e2025GL119740 (2026).
88. Nayak, M. S., Armour, K. C. & Battisti, D. S. Evaluating physical climate model emulators for global warming projections. *Geophys. Res. Lett.* **52**, e2025GL118591 (2025).
89. NEMO Consortium. *Nucleus for European Modelling of the Ocean* (NEMO, 2025) <https://nemo-ocean.eu/>.
90. Reichstein, M. et al. Early warning of complex climate risk with integrated artificial intelligence. *Nat. Commun.* **16**, 2564 (2025).
91. Betz, G. Are climate models credible worlds? prospects and limitations of possibilistic climate prediction. *Eur. J. Philos. Sci.* **5**, 191–215 (2015).
92. Hrst Essenfelder, A., Toreti, A. & Seguini, L. Expert-driven explainable artificial intelligence models can detect multiple climate hazards relevant for agriculture. *Commun. Earth Environ.* **6** <https://doi.org/10.1038/s43247-024-01987-3> (2025).
93. Dramsch, J. S. et al. Explainability can foster trust in artificial intelligence in geoscience. *Nat. Geosci.* **18**, 112–114 (2025).
94. O'Loughlin, R. J., Li, D., Neale, R. & O'Brien, T. A. Moving beyond post hoc explainable artificial intelligence: a perspective paper on lessons learned from dynamical climate modeling. *Geosci. Model Dev.* **18**, 787–802 (2025).
95. MacMillan, T. & Ouellette, N. T. Towards mechanistic understanding in a data-driven weather model: internal activations reveal interpretable physical features. Preprint at arXiv <https://doi.org/10.48550/arXiv.2512.24440> (2025).
96. Benavides-Martinez, I. F. et al. What on earth is AlphaEarth? Hierarchical structure and functional interpretability for global land cover. Preprint at arXiv <https://doi.org/10.48550/arXiv.2603.16911> (2026).
97. McGovern, A. et al. Identifying and categorizing bias in ai/ml for Earth sciences. *Bull. Am. Meteorol. Soc.* **105**, E567–E583 (2024).
98. UNEP Finance Initiative. The climate data challenge: the critical role of open-source and neutral data platforms. Tech. Rep. UNEP Finance Initiative. <https://www.unepfi.org/wordpress/wp-content/uploads/2024/05/Dataland-Final-Report-The-Climate-Data-Challenge-1.pdf> (2024).
99. Nguyen, T., Jewik, J., Bansal, H., Sharma, P. & Grover, A. Climate-learn: benchmarking machine learning for weather and climate modeling. In *Thirty-Seventh Conference on Neural Information Processing Systems (NeurIPS 2023)* (2023).
100. Hamman, J. et al. The pangeo platform: a community-driven open-source big data environment. *ESSOAR* <https://doi.org/10.1002/essoar.10501751.1>. (2020).
101. McGovern, A., Ebert-Uphoff, I., Gagne, D. J. & Bostrom, A. Why we need to focus on developing ethical, responsible, and trustworthy artificial intelligence approaches for environmental science. *Environ. Data Sci.* **1**, e6 (2022).
102. Wynsberghe, A. V. Sustainable AI: AI for sustainability and the sustainability of AI. *AI Ethics* **1**, 213–218 (2021).
103. Edel, L., Xie, J., Korosov, A., Brajard, J. & Bertino, L. Reconstruction of Arctic sea ice thickness (1992–2010) based on a hybrid machine learning and data assimilation approach. *Cryosphere* **19**, 731–752 (2025).
104. Rampal, N. et al. Enhancing regional climate downscaling through advances in machine learning. *Artif. Intell. Earth Syst.* **3**, 230066 (2024).
105. Hannachi, A., Jolliffe, I. T., Stephenson, D. B. et al. Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.* **27**, 1119–1152 (2007).
106. Diaconescu, E. P. & Laprise, R. Singular vectors in atmospheric sciences: A review. *Earth Sci. Rev.* **113**, 161–175 (2012).
107. Vannitsem, S. Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics. *Chaos* **27**, 032101 (2017).
108. Hummel, F., Ashwin, P. & Kuehn, C. Reduction methods in climate dynamics—a brief review. *Phys. D Nonlinear Phenom.* **448**, 133678 (2023).
109. Hasselmann, K. Stochastic climate models part I. theory. *Tellus* **28**, 473–485 (1976).
110. Frankignoul, C. & Hasselmann, K. Stochastic climate models, part ii application to sea-surface temperature anomalies and thermocline variability. *Tellus* **29**, 289–305 (1977).
111. Penland, C. & Magorian, T. Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Clim.* **6**, 1067–1076 (1993).
112. Navarra, A., Tribbia, J. & Klus, S. Estimation of Koopman transfer operators for the equatorial Pacific SST. *J. Atmos. Sci.* **78**, 1227–1244 (2021).
113. Lorenzo-Sánchez, P., Newman, M., Albers, J., Subramanian, A. C. & Navarra, A. Koopman theory for enhanced tropical Pacific SST forecasting. *Artif. Intell. Earth Syst.* **4**, e240088 (2025).
114. Tibau, X.-A., Reimers, C., Requena-Mesa, C. & Runge, J. Spatio-temporal autoencoders in weather and climate research. In *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences* 186–203 (John Wiley & Sons, 2021).

115. Runge, J., Gerhardus, A., Varando, G., Eyring, V. & Camps-Valls, G. Causal inference for time series. *Nat. Rev. Earth Environ.* **10**, 487–505 (2023).
116. Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K. & Grover, A. Climax: a foundation model for weather and climate. In *Proc. 40th Int. Conf. Machine Learn.* Vol. 202, 25904–25938 (2023).
117. Mahesh, A. et al. Huge ensembles – part 1: design of ensemble weather forecasts using spherical Fourier neural operators. *Geosci. Model Dev.* **18**, 5575–5603 (2025).
118. Mahesh, A. et al. Huge ensembles – part 2: properties of a huge ensemble of hindcasts generated with spherical Fourier neural operators. *Geosci. Model Dev.* **18**, 5605–5633 (2025).
119. Garrett, R. C., Harris, T., Wang, Z. & Li, B. Validating climate models with spherical convolutional Wasserstein distance. In *Proc. 38th Conference on Neural Information Processing Systems (NeurIPS 2024)*, vol. 37 <https://doi.org/10.5555/3737916.3739802> (2024).
120. Finn, T. S. et al. Generative AI models capture realistic sea-ice evolution from days to decades. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2508.14984> (2025).
121. Immorlano, F. et al. Transferring climate change physical knowledge. *Proc. Natl. Acad. Sci. USA* **122**, e2413503122 (2025).
122. De Cillis, G. et al. Hybrid physics–data-driven modeling for sea ice thermodynamics and transfer learning. *Q. J. R. Meteorol. Soc.* **e70200** <https://doi.org/10.1002/qj.70200> (2026).
123. Bonavita, M. et al. Machine learning for Earth system observation and prediction. *Bull. Am. Meteorol. Soc.* **102**, E710–E716 (2021).
124. Buizza, C. et al. Data learning: integrating data assimilation and machine learning. *J. Comput. Sci.* **58**, 101525 (2022).
125. Cheng, S. et al. Machine learning with data assimilation and uncertainty quantification for dynamical systems: a review. *IEEE/CAA J. Autom. Sin.* **10**, 1361–1387 (2023).
126. Arcucci, R. et al. The convergence of machine learning and data assimilation in earth system science. *npj Artif. Intell.* **2**, 48 (2026).
127. Brajard, J., Carrassi, A., Bocquet, M. & Bertino, L. Combining data assimilation and machine learning to infer unresolved scale parametrization. *Philos. Trans. R. Soc. A* **379**, 20200086 (2021).
128. Lafon, N., Fablet, R. & Naveau, P. Uncertainty quantification when learning dynamical models and solvers with variational methods. *J. Adv. Model. Earth Syst.* **15**, e2022MS003446 (2023).
129. Durand, C. et al. Four-dimensional variational data assimilation with a sea-ice thickness emulator. *EGU sphere* **2025**, 1–36 (2025).
130. Si, P. & Chen, P. Latent-EnSF: a latent ensemble score filter for high-dimensional data assimilation with sparse observation data. In *Int. Conf. Learn. Representations* (2024).
131. Pasmans, I., Chen, Y., Sebastian Finn, T., Bocquet, M. & Carrassi, A. Ensemble Kalman filter in latent space using a variational auto-encoder pair. *Q. J. R. Meteorol. Soc.* **152**, e70070 (2025).
132. Cintra, R. S. & de Campos Velho, H. F. in *Advanced Applications for Artificial Neural Networks* (ed. El-Shahat, A.) Ch. 14 (IntechOpen, 2018).
133. Keller, J. D. & Potthast, R. AI-based data assimilation: learning the functional of analysis estimation. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2406.00390> (2024).
134. Boudier, P., Fillion, A., Gratton, S., Gürol, S. & Zhang, S. Data assimilation networks. *J. Adv. Model. Earth Syst.* **15**, e2022MS003353 (2023).
135. Bach, E., Baptista, R., Luk, E. & Stuart, A. Learning optimal filters using variational inference. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2406.18066> (2024).
136. Bocquet, M. et al. Accurate deep learning-based filtering for chaotic dynamics by identifying instabilities without an ensemble. *Chaos* **34**, 091104 (2024).
137. Alexe, M. et al. GraphDOP: towards skilful data-driven medium-range weather forecasts learnt and initialised directly from observations. *arXiv* <https://doi.org/10.48550/arXiv.2412.15687> (2024).
138. Allen, A. et al. End-to-end data-driven weather prediction. *Nature* **641**, 1172–1179 (2025).
139. Garcia-Oliva, L., Carrassi, A. & Counillon, F. Exploring the influence of spatio-temporal scale differences in coupled data assimilation. *Nonlinear Process. Geophys.* **32**, 439–456 (2025).
140. Bommasani, R. et al. On the opportunities and risks of foundation models. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2108.07258> (2021).
141. Chen, S., Long, G., Jiang, J., Liu, D. & Zhang, C. Foundation models for weather and climate data understanding: a comprehensive survey. Preprint at *arXiv* <https://arxiv.org/abs/2312.03014> (2024).
142. Watt-Meyer, O. et al. ACE2: accurately learning subseasonal to decadal atmospheric variability and forced responses. *npj Clim. Atmos. Sci.* **8**, 205 (2025).
143. Bodnar, C. et al. A foundation model for the Earth system. *Nature* **641**, 1180–1187 (2025).
144. Zhu, X. X. et al. On the foundations of Earth foundation models. *Commun. Earth Environ.* **7** <https://doi.org/10.1038/s43247-025-03127-x> (2026).
145. Xiao, A. et al. Foundation models for remote sensing and Earth observation: a survey. *IEEE Geosci. Remote Sens. Mag.* **13**, 297–324 (2025).
146. Wang, Y. et al. Towards a unified Copernicus foundation model for Earth vision. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)* 9888–9899 (IEEE, 2025).
147. Bountos, N. I., Ouaknine, A., Papoutsis, I. & Rolnick, D. Fomo: multi-modal, multi-scale and multi-task remote sensing foundation models for forest monitoring. In *Proc. AAAI Conference on Artificial Intelligence*, Vol. 39, 27858–27868 (AAAI, 2025).
148. Tseng, G. et al. Galileo: learning Global & local features of many remote sensing modalities. In *Proc. 42nd Int. Conf. Machine Learn.* Vol. 267, 60280–60300 (PMLR, 2025).
149. Klemmer, K., Rolf, E., Robinson, C., Mackey, L. & Rußwurm, M. Satclip: Global, general-purpose location embeddings with satellite imagery. In *Proc. AAAI Conference on Artificial Intelligence*, Vol. 39, 4347–4355 (AAAI, 2025).
150. Shi, J. et al. Deep learning and foundation models for weather prediction: a survey. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2501.06907> (2025).
151. Schmude, J. et al. Prithvi wxC: foundation model for weather and climate. Preprint at *arXiv* <https://arxiv.org/abs/2409.13598> (2024).
152. Brown, C. F. et al. Alphaearth foundations: an embedding field model for accurate and efficient global mapping from sparse label data. Preprint at *arXiv* <https://arxiv.org/abs/2507.22291> (2025).
153. Mardani, M. et al. Residual corrective diffusion modeling for km-scale atmospheric downscaling. *Commun. Earth Environ.* **6**, 124 (2025).
154. Sundar, R., Hu, Y., Parashar, N., Blanchard, A. & Dodov, B. TAUDiff: highly efficient kilometer-scale downscaling using generative diffusion models. https://casml.cc/wp-content/uploads/2024/12/Rahul_Paper_TAUDiff.pdf (2025).
155. Wan, Z. Y. et al. Regional climate risk assessment from climate models using probabilistic machine learning. Preprint at *arXiv* <http://arxiv.org/abs/2412.08079> (2025).
156. Tomasi, E., Franch, G. & Cristoforetti, M. Can AI be enabled to perform dynamical downscaling? A latent diffusion model to mimic kilometer-scale COSMO5.0_cfm9 simulations. *Geosci. Model Dev.* **18**, 2051–2078 (2025).
157. Alet, F. et al. Skillful joint probabilistic weather forecasting from marginals. Preprint at *arXiv* <https://arxiv.org/abs/2506.10772> (2025).

158. Toms, B. A., Barnes, E. A. & Ebert-Uphoff, I. Physically interpretable neural networks for the geosciences: applications to earth system variability. *Geophys. Res. Lett.* **46**, 13135–13144 (2019).
159. Camps-Valls, G. et al. Discovering causal relations and equations from data. *Phys. Rep.* **1044**, 1–68 (2023).
160. Runge, J. et al. Inferring causation from time series in Earth system sciences. *Nat. Commun.* **10**, 2553 (2019).
161. Grundner, A., Beucler, T., Gentine, P. & Eyring, V. Data-driven equation discovery of a cloud cover parameterization. *J. Adv. Model. Earth Syst.* **16**, e2023MS003763 (2024).
162. Grundner, A. et al. Deep learning based cloud cover parameterization for ICON. *J. Adv. Model. Earth Syst.* **14**, e2021MS002959 (2022).
163. Driscoll, S. et al. Parameter sensitivity analysis of a sea ice melt pond parameterisation and its emulation using neural networks. *J. Comput. Sci.* **79**, 102231 (2024).
164. Hu, Z. et al. Stable machine-learning parameterization of subgrid processes in a comprehensive atmospheric model learned from embedded convection-permitting simulations. *J. Adv. Model. Earth Syst.* **17**, e2024MS004618 (2025).
165. Heuer, H. et al. Beyond the training data: confidence-guided mixing of parameterizations in a hybrid AI-climate model. *J. Adv. Model. Earth Syst.* **18**, e2025MS005544 (2026).
166. Shen, C. et al. Differentiable modelling to unify machine learning and physical models for geosciences. *Nat. Rev. Earth Environ.* **4**, 552–567 (2023).
167. Henn, B. et al. AIMIP phase 1: systematic evaluations of AI weather and climate models. *arXiv* <https://doi.org/10.48550/arXiv.2605.06944> (2026).
168. Ben Bouallègue, Z. What is a realistic forecast? Assessing data-driven weather forecasts, a journey from verification to falsification. *arXiv* <https://doi.org/10.48550/arXiv.2602.00622> (2026).
169. Gneiting, T. et al. Probabilistic measures afford fair comparisons of AIWP and NWP model output. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2506.03744> (2025).
170. Metref, S. et al. Estimating model evidence using ensemble-based data assimilation with localization - the model selection problem. *Q. J. R. Meteorol. Soc.* **145**, 1571–1588 (2019).
171. Zhang, B. & Merlis, T. M. The equilibrium response of atmospheric machine-learning models to uniform sea surface temperature warming. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2510.02415> (2025).
172. Wang, X. et al. Condensnet: enabling stable long-term climate simulations via hybrid deep learning models with adaptive physical constraints. *npj Clim. Atmos. Sci.* **9**, 7 (2026).
173. Sha, Y., Schreck, J. S., Chapman, W. & Gagne, D. J. Improving AI weather prediction models using Global mass and energy conservation schemes. *J. Adv. Model. Earth Syst.* **17**, e2025MS005138 (2025).
174. Hickman, S. et al. Causal climate emulation with Bayesian filtering. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2506.09891> (2025).
175. Katzenberger, A. et al. Developing guidelines for working with multi-model ensembles in CMIP. *Earth Syst. Dyn.* **17**, 495–532 (2026).
176. Simpson, I. R. et al. Confronting Earth system model trends with observations. *Sci. Adv.* **11**, eadt8035 (2025).
177. Neelin, J. D. et al. Process-oriented diagnostics: principles, practice, community development, and common standards. *Bull. Am. Meteorol. Soc.* **104**, E1452–E1468 (2023).
178. Heinze, C. et al. ESD reviews: climate feedbacks in the Earth system and prospects for their evaluation. *Earth Syst. Dyn.* **10**, 379–452 (2019).
179. Merrifield, A. L., Brunner, L., Lorenz, R., Humphrey, V. & Knutti, R. Climate model selection by independence, performance, and spread (ClimSIPS v1.0.1) for regional applications. *Geosci. Model Dev.* **16**, 4715–4747 (2023).
180. Chen, M. et al. Collaboration between artificial intelligence and earth science communities for mutual benefit. *Nat. Geosci.* **17**, 949–952 (2024).
181. Camps-Valls, G. Serendipity's role in advancing geoscience. *Nat. Geosci.* **18**, 684–685 (2025).
182. Lin, J. et al. LEAP – atmospheric physics using AI (ClimSim). <https://kaggle.com/competitions/leap-atmospheric-physics-ai-climsim> (2024).
183. Gruca, A. et al. Weather4cast: NeurIPS 2025 challenge. <https://proceedings.mlr.press/v220/gruca23a.html> (2025).
184. Richter, J. H. et al. Advancing Earth system science with and for the community: shaping NSF NCAR's AI strategy. *Bull. Am. Meteorol. Soc.* <https://doi.org/10.1175/BAMS-D-26-0202.1> (2026).

Author contributions

This piece originated from discussions in an ELLIS workshop titled “AI for Learning Weather and Climate” held in València, Spain, in 2024 as part of the ELLIS Research Program on “Machine Learning for Earth and Climate Sciences,” where more than 30 experts at the intersection of AI, weather and climate sciences gathered for a week. G.C.-V. designed and structured the study with the help of A.C., F.d.M.V. and K.-H.C. G.C.-V. wrote the first version of the manuscript with the help of all coauthors. G.C.-V. supervised the study. G.C.-V., A.C., F.d.M.V., K.-H.C., F.S. and A.W. wrote the introduction section. F.d.M.V., K.D., P.H., N.M., D.M., F.S., H.H., M.-Á.F.-T., J.W., C.R., and E.R.-B. worked on the challenges section. I.P., N.I.B., K.-H.C., M.-Á.F.-T., and V.S. worked on the “AI potential for bridging the weather and climate gap” section. A.C., F.d.M.V., M.-Á.F.-T., D.M., V.S., H.H., E.R.-B., P.G., and V.E. worked on the “Three AI-mediated bridges” section. A.C., F.d.M.V., K.-H.C., I.P., and H.H. worked on the outlook concepts. All authors reviewed the manuscript.

Funding

G.C.-V., M.-Á.F.-T., H.H., K.D., P.G. and V.E. acknowledge support from the European Research Council (ERC) under the ERC Synergy Grant USMILE (grant agreement No. 855187). F.d.M.V. acknowledges partial support from the U.K. Natural Environment Research Council through the Optimising the Design of Ensembles to Support Science and Society project (ODESSS, grant agreement NE/V011790/1). A.C. has been funded by the project MELTED—*Machine Learning for Arctic ice prediction*, from the Italian Ministry of Universities and Research (MUR) through the 2022 PRIN (Progetti di Rilevante Interesse Nazionale) call, and by the Scale-Aware Sea Ice Project (SASIP, Grant Number: G-24-67790) of Schmidt Sciences, a philanthropic initiative that seeks to improve societal outcomes by advancing emerging science and technologies. E.R.-B., K.-H.C., and G.C.-V. thank the Horizon project ELIAS (grant agreement 101120237). G.C.-V., H.H., and V.E. acknowledge the Horizon project AI4PEX (grant agreement 101137682). V.S. acknowledges the project “WUR FM” (project number KB-57-001002). G.C.-V., M.-Á.F.-T., N.I.B., and I.P. acknowledge the Horizon Europe project MeDiTwin (grant agreement 101159723) and the Horizon Europe project ThinkingEarth (grant agreement 101130544). M.-Á.F.-T. thanks the Universidad Carlos III de Madrid, Spain, for the support provided through the Funding for Research Activities Program (project 2025/00803/001) and the Regional Government of Madrid through project TEC-2024/COM-322. P.H. acknowledges funding by the Volkswagen Foundation. D.M. acknowledges support from the European Space Agency—ESA (cube.ng—No Gap Earth System Data Cubes, CCN1 of ARCEME—Adaptation and Resilience to Climate Extremes and Multi-hazard Events, ESA Contract No. 4000144482/24/I-KE). V.E. was additionally supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Gottfried Wilhelm Leibniz Prize (reference No. EY 22/2-1).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-75787-y>.

Correspondence and requests for materials should be addressed to Gustau Camps-Valls.

Peer review information *Nature Communications* thanks Florence Rabier, Eulalie Boucher and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026