The convergence of machine learning and data assimilation in Earth system science

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Arcucci, R., Healy, S., Dance, S. ORCID: https://orcid.org/0000-0003-1690-3338, Lei, L., Bach, E. ORCID: https://orcid.org/0000-0002-9725-0203, Weaver, A. T., Miyoshi, T., Eugenia Dillon, M., Draper, C., Schneider, R., Lang, S., Dueben, P., Bormann, N., Lean, P., Geer, A., Bonavita, M., Jan van Leeuwen, P., Cheng, S., Bocquet, M., Zagar, N., Fraga de Campos Velho, H., Jose Ruiz, J., Bauer, P., Ahmed Boukabara, S., Carrassi, A. ORCID: https://orcid.org/0000-0003-0722-5600, Treadon, R., Collard, A., Kleist, D., Gholoubi, A., Wang, X., Samrat, N., Ralton, G., Moore, A. M., Lamer, K. and Caltabiano, N. (2026) The convergence of machine learning and data assimilation in Earth system science. npj Artificial Intelligence, 2. 48. ISSN 3005-1460 doi: 10.1038/s44387-026-00107-0 Available at https://centaur.reading.ac.uk/129498/

It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1038/s44387-026-00107-0

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

npj | artificial intelligence

Perspective

https://doi.org/10.1038/s44387-026-00107-0

The convergence of machine learning and data assimilation in Earth system science

Check for updates

Rossella Arcucci[1] ✉, Sean Healy[2], Sarah Dance[3], Lili Lei[4], Eviatar Bach[3], Anthony T. Weaver[5], Takemasa Miyoshi[6], Maria Eugenia Dillon[7], Clara Draper[8], Rochelle Schneider[9], Simon Lang[2], Peter Dueben[2], Niels Bormann[2], Peter Lean[2], Alan Geer[2], Massimo Bonavita[2], Peter Jan van Leeuwen[10], Sibo Cheng[11], Marc Bocquet[11], Nedjeljka Zagar[12], Haroldo Fraga de Campos Velho[13], Juan Jose Ruiz[14], Peter Bauer[15], Sid Ahmed Boukabara[16], Alberto Carrassi[17], Russ Treadon[8], Andrew Collard[8], Daryl Kleist[8], Azadeh Gholoubi[8], Xuguang Wang[18], Nahidul Samrat[19], Gemma Ralton[1], Andrew M. Moore[20], Katia Lamer[21] & Nico Caltabiano[22]

Data assimilation (DA) combines observations with numerical models to estimate evolving Earth system states for forecasting and monitoring. Machine learning (ML) enables surrogate modeling, pattern recognition and Bayesian inference. These fields are converging: ML accelerates DA, while DA provides uncertainty quantification and physical constraints. Hybrid DA-ML systems are promising, yet challenges persist in generalization, consistency and reproducibility. These approaches are increasingly integrated, shaping next-generation prediction systems and observing networks.

There has been remarkable progress in the development of numerical weather prediction (NWP) forecast models based on machine learning (ML) techniques[1–7]. It has been demonstrated that forecasts produced with these models can now outperform state-of-the-art, physically-based forecast models for many key parameters when given the same initial conditions[8,9]. However, this success is, at least in part, still reliant on "traditional", physically based data assimilation (DA) techniques. Firstly, traditional DA techniques provide the initial conditions required to produce the forecast. In addition, the forecast models are trained on multiple decades of atmospheric, ocean and wave state information from global reanalysis systems that also use standard DA techniques.

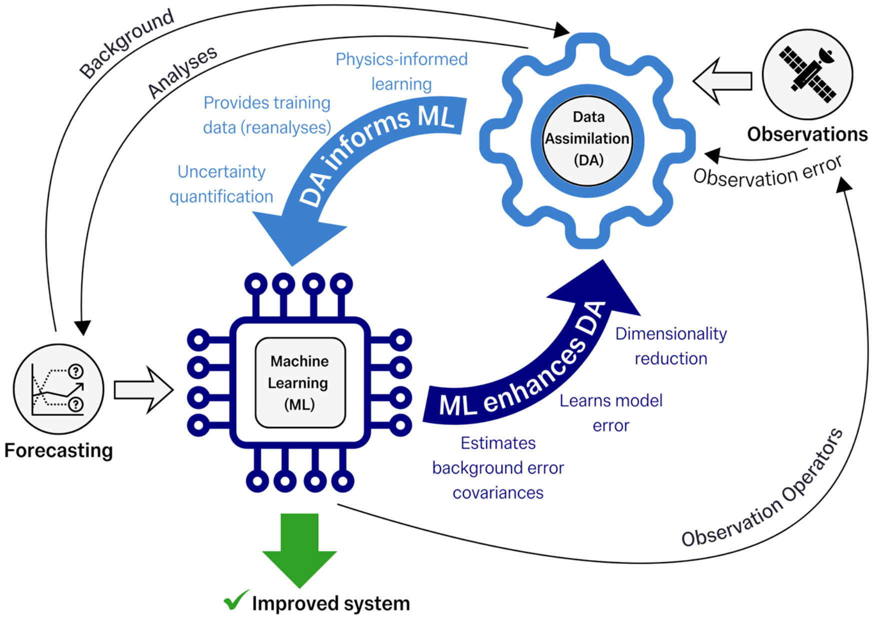
The application of ML techniques to DA is an active area of research. Figure 1 highlights the expanding research landscape at the intersection of ML and DA, which forms the focus of this paper.

DA has long been a cornerstone of NWP and Earth system modeling. DA systems aim to characterize the posterior probability distribution function (pdf) of the system state and to provide the statistically optimal estimate of the initial conditions required for forecasting. This is achieved by combining a prior short-range forecast with a heterogeneous set of noisy and often indirect observations collected over a short assimilation window (typically 6 or 12 h in global atmospheric DA applications). Traditional DA approaches[10], such as variational methods (e.g., 3D Variational Data Assimilation (3D-Var), 4D Variational Data Assimilation (4D-Var)) and ensemble-based schemes (e.g., ensemble Kalman filters (EnKF)), have evolved significantly over the past decades and are now integral to operational forecasting centers around the world. Some fundamental components of these traditional DA systems include good estimates of both the short-range forecast uncertainty and the observation uncertainty, to ensure that these sources of information are combined and weighted in an optimal way. DA systems also rely on a forecast model to interpolate to the time and location of the observation and the development of fast but accurate "observation operators" that transform from geophysical state space of interest represented by the forecast model (temperature, humidity, wind, ...) to observation space (e.g, satellite radiances, bending angles, ...). Incremental improvements in these components have contributed to sustained progress in NWP forecast quality over many decades[11].

Simultaneously, the field of ML has undergone rapid and transformative growth. From foundational developments in deep learning to recent advances in generative modeling and differentiable programming, ML methods have demonstrated the ability to approximate complex, nonlinear

[1]Imperial College London, London, UK. [2]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK. [3]University of Reading, Reading, UK. [4]Nanjing University, Nanjing, China. [5]CERFACS, Toulouse, France. [6]RIKEN, Kobe, Japan. [7]Servicio Meteorológico Nacional, Buenos Aires, Argentina. [8]National Oceanic and Atmospheric Administration (NOAA), Washington D.C., USA. [9]European Space Agency (ESA), Frascati, Italy. [10]Colorado State University, Fort Collins, USA. [11]CEREA Institut Polytechnique de Paris, Paris, France. [12]University of Hamburg, Hamburg, Germany. [13]National Institute for Space Research (INPE), Belem, Brazil. [14]University of Buenos Aires, Buenos Aires, Argentina. [15]Max-Planck-Institute for Meteorology, Hamburg, Germany. [16]NASA Earth Science Division (ESD), Washington D.C., USA. [17]University of Bologna, Bologna, Italy. [18]University of Oklahoma, Norman, USA. [19]Bureau of Meteorology, Sydney, Australia. [20]University of California Santa Cruz, Santa Cruz, USA. [21]Brookhaven National Laboratory, Upton, USA. [22]World Meteorological Organization, Geneva, Switzerland. ✉e-mail: r.arcucci@imperial.ac.uk

Fig. 1 | Circular feedback loop between DA and ML.



functions, extract patterns from massive datasets, and accelerate computational workflows. These capabilities have begun to permeate environmental science and are increasingly viewed as promising complements, or even alternatives, to traditional DA techniques.

In recent years, the intersection between ML and DA has given rise to a new class of hybrid systems that leverage the strengths of both paradigms. ML models are being used to emulate observation operators, learn model errors, improve background error covariance estimation, and even approximate the DA process itself through end-to-end training frameworks. Conversely, DA principles are feeding back into ML by informing uncertainty quantification, sequential learning, and physics-constrained optimization[12]. The convergence of these fields is reshaping the way observational data and physical models are integrated, with profound implications for forecasting, reanalysis, and Earth system understanding.

This paper presents a synthesis of perspectives from 35 experts from 22 institutions on the current and future impact of ML on DA, as well as the reciprocal influence of DA on ML. Through a series of structured questions posed to researchers and practitioners in both fields, we explore the emerging trends, operational challenges, methodological innovations, and philosophical considerations that are guiding this convergence. The perspectives summarized here were collected through a structured questionnaire distributed to experts contributing to the World Weather Research Programme (WWRP) Working Group on Data Assimilation and Observing Systems (DAOS) and beyond, and subsequently synthesized by the authors. Key topics include the rise of hybrid DA-ML systems, the viability of fully data-driven forecasting approaches, the role of DA in training and constraining ML models, and the broader scientific and institutional implications of this interdisciplinary evolution.

Our goal is to provide a clear and comprehensive account of where the field stands today, what developments can be expected in the near term, and how the DA and ML communities can collaborate to build robust, interpretable, and efficient systems for Earth system prediction and analysis. By capturing a range of expert insights, we aim to contribute to a more informed and integrated research agenda that leverages the best of both DA and ML.

Statement of the problem

The core challenge in DA and ML-based prediction is to estimate the evolving state of the Earth system as accurately as possible, given imperfect information. The inputs to the problem consist of: (i) a prior estimate of the system state, often referred to as the *background state*, which is frequently generated by integrating a dynamical model forward in time but may also come from previous analyses, climatological estimates, or other prior information, and (ii) a diverse set of *observations*, which may include a diverse set of both in situ and satellite measurements. The output is an *analysis state*, an optimally combined estimate that can serve as the initial condition for the next forecast cycle (see Fig. 1).

Formally, the system is described in a high-dimensional time-dependent *state space*, denoted $\mathbf{x}(t)$, which evolves under the dynamics of a numerical model. Observations, $\mathbf{y}(t)$, belong to a distinct *observation space*, and are related to the state by an often nonlinear, indirect *observation operator*$\mathcal{H}$. This separation is fundamental: satellites, for example, do not measure temperature or humidity directly, instead measuring quantities like radiances, which must be mapped back into the geophysical state space within the DA system[13,14].

The observations alone are not sufficient to constrain the estimate of the initial conditions, that problem would be mathematically under-determined, and the short-range forecast information is required to stabilize the solution. To resolve this, DA systems employ the background state and its associated error statistics as a form of statistical *regularization*, ensuring that the resulting analysis is statistically optimal and dynamically consistent. This renders the problem inherently *probabilistic*: uncertainties from the forecast and the observations must be quantified and optimally weighted in the assimilation process. Several structural difficulties arise:

- Indirect and incomplete observations: many measurements provide noisy, partial or transformed views of the true state.
- Under-determined problem: the available observations are insufficient to uniquely determine the system state, allowing multiple states to be consistent with the same observational data.
- High dimensionality: operational state vectors may contain on the order of 10^9 variables.
- Time constraints: assimilation cycles must be completed within strict deadlines to deliver forecasts on schedule.
- Model errors: NWP imperfections lead to systematic errors in the system.

ML offers new avenues for addressing several of the open challenges in DA (see Fig. 1). By learning efficient surrogates for expensive physical operators, ML can accelerate the mapping between state space and observation space, particularly in cases involving nonlinear radiative transfer or other complex forward models. ML models can also help characterize and correct systematic model errors, thereby improving the quality of the background state used to regularize the under-determined problem. In high-dimensional systems, ML-driven dimensionality reduction or latent-space representations provide a pathway to more efficient assimilation, while adaptive localization and flow-dependent covariance modeling may be enhanced through data-driven techniques. Finally, the probabilistic nature of the problem aligns naturally with emerging ML methods for uncertainty quantification, generative modeling, and probabilistic inference, offering tools to better represent forecast and observation errors. In this way, ML has the potential to complement the statistical and physical rigor of DA with flexible, data-adaptive solutions, helping to overcome limitations imposed by incomplete observations, and strict operational time constraints.

Recent advances in machine learning for data assimilation

ML is already having an impact in operational NWP systems within the traditional DA frameworks. Recent examples include the development of a new observation operator that enables the assimilation of surface sensitive microwave radiance observations over sea ice[15,16], automated data monitoring and anomaly detection[17], and the correction of forecast model errors in the stratosphere in a new global reanalysis (ERA6)[18]. However, there are many other emerging applications that are not yet operational. For example, at Météo-France, ML tools are being developed for convective-scale ensemble expansion, high-resolution statistical downscaling, and toward fully data-driven kilometer-scale forecasting systems[19]. In the United States, the National Oceanic and Atmospheric Administration (NOAA) has explored AI-driven data fusion and assimilation to manage large observational volumes more efficiently[20]. Meanwhile, Google DeepMind's GenCast[5] and the European Centre for Medium-Range Forecasts (ECMWF) Artificial Intelligence Forecasting System Continuous Ranked Probability Score (AIFS-CRPS)[21] both demonstrate probabilistic ensemble forecasts that outperform traditional medium-range systems while running in only minutes, suggesting near-term operational adoption. More broadly, European services have launched the open-source *Anemoi* framework[22] to integrate ML into operational workflows[23], while systems such as GraphCast[8] and ECMWF's AIFS[9] show that hybrid and ML-native forecasting is moving from research into daily forecasting practice.

A prominent area of development has been the integration of ML into traditional DA systems, forming hybrid methodologies. Hybrid ML-DA approaches typically integrate ML with established DA frameworks, including variational methods (e.g., 3D-Var and 4D-Var) and ensemble-based methods such as the Ensemble Kalman Filter[24]. Given the algorithmic differences between these approaches, ML has been incorporated in distinct ways. First, ML can be used to improve DA components external to the assimilation algorithm itself, such as observation operators. Second, in variational DA systems, ML has been explored as a differentiable surrogate for operators required by the optimization procedure, particularly the tangent-linear and adjoint models, which are costly to develop and maintain. Third, in ensemble-based DA systems, ML has primarily been applied to enhance ensemble methodologies, including ensemble generation, covariance localization, and adaptive inflation. As noted above, observation operators for satellite radiances have been approximated using ML models, offering greater speed and flexibility in complex observational settings[15,16,25]. In variational DA, differentiable ML surrogates[26–28] have been proposed to replace manually developed tangent-linear and adjoint models, reducing technical overhead and enabling rapid iteration[29]. However, despite promising work on transfer learning and the development of adjoint models for ML-based forecast systems, perturbations in these models often fail to propagate forward in time in a physically consistent manner. This limitation poses a major challenge for their integration into 4D-Var systems and more generally raises concerns for DA applications that rely on physically meaningful error propagation[30–32].

Further enhancements include the use of ML for adaptive localization and inflation in ensemble systems, improving covariance modeling in high-dimensional spaces[33]. These hybrid systems offer the promise of reduced computational cost, improved flexibility, and accelerated model development, particularly for high-resolution or convective-scale (≤3 km) applications. Alternatively, ML models have been trained to emulate the analysis step[34–38]. In parallel with hybrid systems, a more radical line of research has focused on developing fully ML-driven DA methods. These approaches aim to bypass the traditional DA cycle by training models to learn direct mappings from observational inputs to analysis or forecast states, often using reanalysis datasets as ground truth. Systems such as Aardvark[4] and AI-Var[39] exemplify this strategy. In contrast, Artificial Intelligence-Direct Observation Prediction (AI-DOP)[40] follows a similar end-to-end learning philosophy but does not rely on reanalysis data, instead learning directly from observation-observation pairs. Such models are typically built using supervised learning architectures, generative frameworks like diffusion networks, or autoencoder-based designs that emulate the DA process in a data-driven manner[41,42]. Their ability to ingest image-like satellite data and to handle complex, heterogeneous observation networks makes them particularly attractive for cases where traditional DA methods struggle. Although these end-to-end systems remain largely at the research stage, initial results suggest they can generate realistic analyses at reduced computational cost and with enhanced architectural flexibility.

Another promising line of research involves latent-space DA, where ML methods are used to project high-dimensional state vectors into reduced representations to facilitate more efficient assimilation[42,43]. Additionally, ML has been employed for uncertainty quantification[44], with Bayesian and ensemble-based methods used to improve the probabilistic characterization of the analysis and forecast[45].

Several challenges have been identified as barriers to wider deployment of ML techniques. Many ML models exhibit strong performance in specific, curated test scenarios but fail to generalize across broader contexts. Purely data-driven methods also struggle to enforce physical constraints, which are critical for ensuring physically consistent analyses and forecasts. This issue is related, but not identical, to the challenge of physically consistent error propagation discussed above: while physical constraints concern the preservation of governing balances or conservation laws, error propagation refers to whether perturbations evolve in a dynamically realistic way within the model[31,32]. Moreover, training robust ML systems typically requires access to large, high-quality labeled datasets, such as those produced by reanalyses, which may not always be available or suitable for all use cases. Another challenge concerns the physical realism of ML-based forecasts: models trained to minimize average error metrics often produce overly smooth fields, which can blur small-scale structures and lead to a systematic underrepresentation of extremes, particularly at longer forecast lead times. Interpretability is another key issue; black-box ML models can be difficult to diagnose and are often viewed with caution in high-stakes, safety-critical forecasting environments. Experts have also voiced concern about the overselling of ML in public forums, stressing the importance of rigorous validation and the integration of ML developments with established physical understanding.

The literature reflects these evolving dynamics. Foundational and recent contributions have played a critical role in shaping the field. Among the most influential works are those focused on hybrid DA-ML methodologies[18,46–50], surrogate modeling and latent-space assimilation[28,42,43], and operationally oriented studies[15,25]. Several review articles[45,51,52] and books[53,54] have also synthesized key developments and challenges in this domain. More recently, the development of fully ML-native DA systems, such as Aardvark[4] AI-DOP[40], AI-Var and FengWu-4DVar[55], illustrates the field's ongoing exploration of alternative paradigms for assimilating observational data.

Looking forward, the consensus within the community is that ML will play an increasingly prominent role in the design and implementation of DA systems. While the transition from research to operational deployment will take time, future systems are likely to adopt hybrid architectures that capitalize on both the flexibility and efficiency of ML and the robustness of physics-based modeling. There is also growing interest in differentiable DA frameworks, data-driven parameter estimation, and the use of DA to support model development and structural correction. However, the rapid rise of ML poses a challenge to sustaining traditional DA expertise. The shift in funding and research focus toward ML-centered projects risks marginalizing the foundational methods upon which modern weather and climate forecasting are built. Ensuring a balanced approach to innovation—one that leverages the strengths of both paradigms while maintaining rigorous scientific standards—will be essential for the long-term success of DA research and its operational relevance.

Anticipated impacts of machine learning on data assimilation over the next five years

Looking ahead, ML is expected to have an increasingly profound impact on the evolution of DA methodologies. While current operational systems remain largely rooted in physics-based ensemble and variational methods, the coming five years are poised to see the emergence of hybrid and fully data-driven paradigms with growing practical relevance. Developments are anticipated across several dimensions, including the unification of DA and forecasting in end-to-end frameworks, the integration of ML to accelerate traditional DA components, and the use of ML to better leverage complex or underutilized observational datasets.

One key direction involves the rise of fully data-driven DA and forecasting systems, wherein neural networks or other ML architectures learn both the system dynamics and the assimilation process directly from data. These models have shown promise in bypassing the need for explicit physical models, potentially offering advantages in speed, simplicity, and data flexibility[4,31]. Nevertheless, there is consensus that these systems still face significant limitations, particularly in generalization beyond training regimes, interpretability, and uncertainty quantification[56]. It is also unclear how these approaches address the inherently under-determined nature of the problem when training directly on observations, as the available observations are typically insufficient to fully constrain the atmospheric state compared with traditional NWP analyses. In practice, the success of several ML-based forecasting systems relies on training with the extensive analysis datasets previously generated by conventional DA systems, rather than on observations alone. As such, while purely ML-based DA-forecasting systems may become viable in niche applications or as complementary tools, they are unlikely to displace physics-based systems in high-stakes operational environments within the next five years.

Instead, hybrid DA-ML systems are anticipated to become the dominant architecture in operational and research settings. These systems preserve the strengths of physically based methods while using ML to augment or replace specific components. ML can be used to learn model errors, emulate observation operators, optimize background error covariances, and enhance the flexibility of assimilation schemes[57–59]. There is growing interest in "physics-aware" machine learning architectures, in which neural networks are trained to respect physical constraints or embedded within differentiable DA frameworks. In these approaches, the learning process is guided by governing equations, conservation laws, or model-consistent loss functions, rather than purely statistical error minimization[12,37]. When integrated within DA frameworks such as 4D-Var or ensemble Kalman filters, the ML components remain constrained by the dynamical model and observational likelihood, which helps maintain physically consistent error propagation during the assimilation cycle. Recent studies have demonstrated that neural networks can be incorporated within variational DA systems to learn model errors or reduced representations while preserving the underlying dynamical constraints[50]. These hybrid, physics-aware architectures offer a promising compromise between physical realism, computational efficiency, and adaptability to new data types.

Moreover, ML will increasingly support the exploitation of complex observing systems. ML-based observation operators, especially for nonlinear satellite radiance data or data from cloudy scenes, will enhance the assimilation of previously underutilized observations[15,16,60]. Similarly, ML can assist in quality control, bias correction, and uncertainty characterization, areas that are becoming more critical as observational datasets grow in volume and diversity[17,61]. ML's ability to learn from raw observational data may be particularly useful for preprocessing data from challenging platforms such as active/passive microwave sensors, lidar, hyperspectral instruments optical detectors, or systems like the Geostationary Lightning Mapper (GLM)[62].

ML will also play a vital role in accelerating components of the DA workflow. Emulators trained on expensive simulations, such as those for radiative transfer or atmospheric dynamics, can replace computational bottlenecks with efficient surrogates[58]. The use of generative models for sampling uncertainty or reconstructing missing variables has the potential to improve ensemble-based DA schemes while reducing computational costs. Additionally, techniques such as latent-space assimilation, enabled by ML-based dimensionality reduction, may offer new paths to scalability in high-resolution or coupled Earth system models[37].

Another important development is the evolving interdependence between ML and DA. High-quality training datasets for ML-based forecasting systems often depend on reanalyses and outputs generated from physics-based DA. Conversely, DA systems are beginning to incorporate transfer learning and online updating strategies to adapt to changes in observational networks and environmental conditions. In this context, transfer learning refers to initializing ML components with models trained on large historical datasets, such as reanalyses, and subsequently adapting them to new observational configurations or regional domains using limited additional data. A related distinction concerns offline versus online training. In offline approaches, ML components are trained in advance and then deployed with fixed parameters, whereas online learning strategies update model parameters sequentially as new observations are assimilated within the DA cycle. Online approaches allow ML components to adapt to nonstationary conditions and evolving observing systems, although ensuring stability and robustness remains an important challenge for operational deployment[48,63].

Missing observations highlight a key contrast between DA and ML. DA has well-established methods for handling under-determined problems through the use of prior information, whereas in ML the a priori is less explicit, often embedded as inductive biases within network architectures that guide learning in ways that are not always transparent[64]. Related research in learning dynamical systems has also pointed to approaches that alternate between expectation-maximization and DA steps, further underscoring the complementarity of the two paradigms.

ML may reshape the assimilation of station-based versus gridded data by reducing representativeness errors and harmonizing scales between sparse in situ records and dense satellite products. Long station time series provide valuable training for bias correction and consistency, while gridded data offer spatial coverage and redundancy. Any improvements will likely depend on ML's flexibility in handling heterogeneous inputs partly stems from advances in multimodal representation learning. In these architectures, different data sources (e.g., satellite imagery, LiDAR point clouds, or in situ measurements) are typically processed by separate feature-extraction modules before being combined in a shared latent representation, allowing complementary information from each modality to be integrated during training. Such multimodal fusion strategies have been widely explored in Earth observation applications, including the joint analysis of optical, SAR, and LiDAR observations[65].

Finally, the community anticipates that ML will help address long-standing limitations of traditional DA. These include assumptions of linearity in variational solvers, the perfect model hypothesis, challenges with high-dimensional error modeling, and the non-scalability of sequential processing. For example, ML may help operationalize fully non-Gaussian DA methods such as particle filters, which have historically been limited by computational cost in high-dimensional systems. While scalable variants such as local particle filters have alleviated some of these challenges, ML-based approaches, including latent-space data assimilation, may provide additional pathways to improve scalability[66]. Similarly, the integration of ML into background error modeling could lead to the development of more sophisticated, flow-dependent, and adaptive correlation structures. In systems that assimilate observations from rapidly evolving or irregular networks, such as citizen sensors and Internet of Things (e.g., crowdsourced measurements from smartphones or personal weather stations) or satellite constellations, ML may offer the flexibility needed to continuously update and recalibrate DA components in near real time.

In summary, while there is no consensus that end-to-end ML systems will replace traditional DA in the next five years, there is widespread agreement that ML will significantly enhance and reshape the DA landscape. Hybrid approaches are emerging as the most practical and promising path forward, enabling the Earth system science community to benefit from both physical insight and data-driven adaptability.

Feedback from data assimilation to machine learning: opportunities for integration and advancement

While DA has long drawn upon ML for tools such as pattern recognition, surrogate modeling, and data emulation, a reverse flow of influence is increasingly underway. In recent years, DA has begun to contribute meaningfully to the development of ML methodologies, particularly in areas such as uncertainty quantification, physics-constrained learning, and the treatment of observational data. As hybrid DA-ML systems mature and data-driven forecasting continues to expand, the knowledge and structure inherent in DA offer valuable insights for addressing foundational challenges in ML.

One of the most prominent contributions of DA to ML is in the domain of uncertainty quantification (UQ). Traditional DA frameworks, especially variational methods and Kalman filters, are inherently Bayesian and offer principled approaches to propagating and quantifying uncertainty in high-dimensional, dynamic systems. This rigor contrasts with many ML methods, where UQ has often been ad hoc or underdeveloped[44]. As ML begins to adopt more structured probabilistic frameworks, such as Bayesian neural networks and ensemble learning, DA principles are informing efforts to design loss functions that explicitly incorporate uncertainty in both model parameters and observations[63,67]. In particular, the DA community's experience in separating background and observational error covariances provides a well-tested model for avoiding overfitting, an enduring issue in ML.

Understanding the predictability of physical models enhances DA algorithms by clarifying the error growth and associated uncertainty quantification. Combining classical DA with predictability analysis could also yield valuable perspectives on the predictability of ML models, and guide future refinements of DA methods for ML systems. Studies have suggested that ML-based forecasting systems may exhibit weaker error amplification from small-amplitude perturbations in their initial conditions than physical models, raising questions about how predictability manifests in such systems[56]. However, recent operational AI-NWP systems also show loss of predictability over time in ways broadly consistent with physics-based NWP models. The apparent differences may partly reflect the tendency of some ML forecasts to produce smoother fields and underrepresent small-scale variability, which can damp the growth of dynamical instabilities that normally drive rapid error amplification in atmospheric models. The extent to which these characteristics arise from the learned dynamics themselves, rather than from aspects of current training strategies or model design, remains an active area of research. Predictability characteristics also vary substantially across Earth system components; for example, the ocean and cryosphere typically exhibit longer intrinsic predictability timescales than the atmosphere, implying that the implications of ML-based dynamics may differ across the coupled Earth system. Furthermore, ML models may exhibit a kind of extrapolative predictability, enabling them to project extreme weather events not present in their training data. Indeed, they can potentially learn the dynamics of strong tropical cyclones in one geographic region and successfully forecast similar phenomena in another[68]. Thus, to leverage the foundation of NWP DA to ML models and further construct the optimal DA for ML models, the predictability of ML models needs be explored.

DA also offers a conceptual and technical foundation for physics-informed machine learning (PIML), an area of growing interest across environmental sciences and engineering. The DA field has a long history of embedding physical constraints into assimilation processes, whether through background term regularization, adjoint-based gradient methods, or model-constrained optimization. These techniques are now influencing the design of ML architectures that integrate differential equations or enforce conservation laws directly during training. The use of physical constraints in learning, particularly in systems governed by known dynamics, enables improved generalizability and interpretability, key limitations in conventional black-box ML approaches[57,67].

Another mechanism through which DA contributes to ML is in the generation of high-quality training data. Reanalysis datasets, produced using state-of-the-art DA systems, serve as essential ground truth for supervised learning models in weather and climate forecasting. These reanalyses, such as ERA5 and its upcoming successors, synthesize sparse and noisy observations with numerical models to yield dynamically consistent estimates of the atmosphere, ocean, and land. In addition, it is found that "fine tuning" by training with recent operational NWP analyses has an impact on ML forecast performance. Consequently, ML systems trained on these products inherently benefit from the physical and statistical rigor of DA. In turn, the growing use of ML in forecasting is motivating the development of DA systems tailored for ML training, with increased resolution, reduced latency, and enhanced handling of uncertainty[17].

DA is also influencing ML by promoting online learning and continual adaptation frameworks. In DA, especially in ensemble and variational systems, the assimilation of new data is formulated as a sequential or recursive optimization problem. This concept aligns with efforts in ML to develop models that can be updated continually as new data arrives, without requiring complete retraining. Such approaches are critical for real-time applications and dynamic environments, including Earth observation, autonomous systems, and operational forecasting[63]. The idea of treating ML model updates as state estimation problems, akin to DA, is gaining traction as a means to improve stability and data efficiency.

Moreover, DA offers ML a valuable testbed for evaluating the robustness, scalability, and interpretability of algorithms. Modern DA systems must assimilate a wide range of observational data, often heterogeneous, sparse, and uncertain, into complex numerical models with millions to billions of degrees of freedom. The sheer scale and heterogeneity of this challenge are unmatched in most standard ML benchmarks. By applying ML methods to DA problems, researchers can assess the performance of new architectures or optimization strategies under realistic and demanding conditions, thereby driving innovation in the broader ML field[58]. In practice, assessing such architectures in operational contexts requires bridging heterogeneous software and hardware ecosystems. Many operational DA systems are implemented in Fortran or C and run on CPU-based high-performance computing infrastructures, whereas ML models are often developed in Python using GPU-accelerated frameworks. Recent efforts therefore rely on interoperable interfaces that allow ML components to be exported and embedded within operational workflows. For example, libraries such as FTorch enable PyTorch models to be coupled directly with Fortran codes, facilitating hybrid systems that combine traditional numerical models with ML modules running on CPUs or GPUs[69]. Such developments highlight the importance of flexible software architectures and heterogeneous computing environments for future DA-ML systems.

The interface between DA and ML is also producing novel algorithmic synergies. For instance, techniques from variational DA, such as weak-constraint formulations, adjoint-based optimization, and multiscale error modeling, are being repurposed to enhance the training of ML models, particularly in hybrid systems where neural components are coupled to physical models. In these systems, the adjoint of a hybrid model may be constructed by combining the adjoint of the physical core with the backpropagation of the neural network, enabling end-to-end differentiability[57]. Similarly, ensemble-based DA methods are informing ensemble training and model averaging strategies in ML, offering probabilistic estimates that are time-evolving and flow-dependent. The fusion of DA and ML has been incorporated into data learning models that integrate physical knowledge with data-driven methods and enhance predictive capabilities[12,70,71].

Despite these opportunities, in operational applications, the feedback from DA into ML remains under-leveraged, partly due to cultural and terminological differences between the communities. DA practitioners often work within physics-based numerical modeling frameworks, while ML research is more data-centric and exploratory. As a result, DA contributions to ML are sometimes rediscovered under different names or overlooked entirely. Nevertheless, as ML increasingly moves into domains with sparse data, strong prior knowledge, and real-time demands, hallmarks of DA, the need for principled methods from the DA community is likely to grow.

Notions of observability, observation impact, and observation influence have long been central to DA, but their role in ML remains less clear. While related ideas are emerging in the experimental design literature and in initiatives such as the BAS/Turing 'deep sensor' work, these concepts have not yet fully translated into weather prediction. It remains an open question whether they should be treated within the current ML-DA convergence or deferred to future developmental work.

In summary, DA is beginning to shape key areas of ML development, offering tested frameworks for uncertainty quantification, physics-informed learning, and online data integration. The continued blending of the two fields promises to yield more interpretable, trustworthy, and efficient ML systems, especially for applications in weather, climate, and Earth system science. As differentiable programming becomes more common and as ML systems increasingly interact with the physical world, the influence of DA methodologies is expected to deepen, fostering a new generation of scientifically grounded ML.

Additional reflections and emerging considerations

As the boundaries between DA and ML continue to dissolve, the scientific community is facing a set of both promising and cautionary developments. Several experts noted that the convergence of the two fields is accelerating, with implications for methodology, collaboration, reproducibility, and the role of public and private actors in shaping the future of Earth system modeling.

One area of concern is the divergence in research culture between DA and ML. While DA traditionally operates within tightly peer-reviewed, physically grounded frameworks, many ML contributions, particularly those driven by large-scale commercial entities or rapid publication cycles, lack the same level of scientific scrutiny. The sheer volume and variability in quality of ML publications have made it increasingly difficult to assess scientific validity, especially when reproducibility is not always prioritized. On the other hand, many ML researchers and open-source communities contribute complete and reproducible code bases, which sets a standard that DA researchers may also benefit from embracing[67].

The continued evolution of DA will likely benefit from incorporating unsupervised and supervised ML techniques, where supervised methods learn mappings from labeled datasets (e.g., reanalyses or analyses), while unsupervised approaches aim to identify structure or patterns directly from unlabeled data, particularly under non-stationary and data-scarce conditions. In contrast to supervised learning, unsupervised and self-supervised approaches may be better suited for identifying structure in high-dimensional environmental data, which is especially relevant in a changing climate where labeled data from past conditions may be less representative of future scenarios. This complements the emerging literature on physics-informed neural networks (PINNs) and PIML, which seek to embed governing equations directly into neural architectures for greater fidelity and interpretability[57].

In terms of operational readiness, ML methods also offer modularity and application-specific customization that can be advantageous in time-critical workflows. While traditional DA systems can also be tuned to particular observation types or applications, ML models can often be trained for specific tasks or data streams with comparatively lower computational cost and development overhead. This flexibility makes ML attractive for industry adoption, particularly where speed and cost efficiency are paramount. Experts emphasized that operational systems may soon lean more heavily on ML, while traditional DA frameworks may serve as scientific reference tools or data-generation systems for model training and validation[58]. However, in an operational context, where a single system must be capable of delivering thousands of products, as would typically be the case in a large weather service, it is not feasible to tune and train each product separately. Managing such a fragmented approach quickly becomes unmanageable. Addressing this challenge will likely require more modular system architectures that allow ML components to be integrated into existing DA production pipelines without disrupting operational workflows, for example through well-defined interfaces, interoperable software libraries, and hybrid CPU-GPU computing environments.

However, this shift also raises questions about the role of public institutions in maintaining foundational infrastructure for Earth system science. As commercial stakeholders increasingly capitalize on the predictive power of DA-ML integration, there is a growing need for governments to proactively engage in this transformation. Collaborative frameworks between public agencies and private sector actors could ensure that critical systems remain transparent, equitable, and scientifically robust.

Several contributors stressed that despite rapid advances in ML, the future of DA may remain structurally similar to the present, albeit with substantial augmentation via ML tools. In this view, hybrid systems, where ML accelerates or complements components of a DA pipeline, are expected to persist as the dominant architecture in operational and research contexts alike[63]. Continued focus on explainability, physical consistency, and uncertainty quantification will be essential for fostering trust and enabling rigorous scientific interpretation of results.

Other questions are around how ML may reshape the requirements for future global observing systems. Traditional OSSEs based on DA currently guide system design, but ML-DA integration could alter the relative impact and importance of different observations, with some becoming more critical. It is often claimed that only ~5–10% of the available observations are assimilated in global NWP systems, although this figure is a rough estimate of the data volume reduction and it should be interpreted with caution. In practice, observation usage varies by data type, and data reduction (e.g., thinning or superobbing) is primarily designed to manage representativeness errors and spatial correlations rather than implying a proportional loss of information. Key challenges remain in representing correlated observation errors and handling nonlinear observation operators. ML methods may offer new opportunities to exploit a broader range of observations by learning nonlinear relationships between measurements and model states, potentially enabling more information to be extracted from complex observational datasets.

Finally, the convergence of DA and ML is increasingly seen not as a competition but as a synergistic paradigm shift. DA contributes its strengths in integrating sparse, noisy observations with physically consistent models, while ML brings speed, adaptability, and the ability to discover patterns in massive datasets. Together, they offer a powerful framework for building the next generation of predictive models and Earth system monitoring capabilities. As such, fostering interdisciplinary collaboration and sustained community engagement will be vital to navigating this rapidly evolving landscape.

Conclusions

The intersection of ML and DA is rapidly transforming how Earth system models are developed, trained, and operated. While ML is enhancing the efficiency, flexibility, and scalability of DA systems, DA continues to provide the physical grounding, uncertainty quantification, and methodological rigor needed to ensure reliable, interpretable results. Rather than replacing one another, ML and DA are proving to be complementary tools, each enriching the capabilities of the other.

This synthesis of expert perspectives underscores several key trends. Hybrid DA-ML systems are likely to remain the dominant architecture over the next five years, with ML contributing to observation operator emulation, bias correction[72,73], model error estimation, and real-time data handling. Fully data-driven forecasting systems, though still in early stages, are advancing quickly and may achieve competitive performance in certain applications. Meanwhile, DA continues to influence ML by offering strategies for physics-informed learning, continual adaptation, and principled handling of uncertainty.

As these fields continue to converge, it will be critical to maintain scientific transparency, foster reproducibility, and ensure that operational and research systems remain interpretable and physically consistent. Equally important is the role of the broader scientific community, including academia, public agencies, and the private sector, in shaping ethical, open, and collaborative frameworks for DA-ML integration.

Ultimately, the combined strengths of DA and ML offer a path toward more agile, accurate, and insightful Earth system prediction. This convergence marks not just a technological evolution, but a scientific shift, one that invites new questions, redefines best practices, and opens the door to more holistic and data-rich understanding of our planet.

Data availability

No datasets were generated or analysed during the current study.

Received: 18 December 2025; Accepted: 12 April 2026;
Published online: 23 April 2026

References

1. ECMWF Director-General. *Machine Learning to Play Growing Role in Weather Forecasting*. https://www.ecmwf.int/en/about/media-centre/news/2024/machine-learning-play-growing-role-weather-forecasting-says-dg (2024).
2. UK Met Office. *Artificial Intelligence for Numerical Weather Prediction*. https://www.metoffice.gov.uk/research/approach/collaboration/artificial-intelligence-for-numerical-weather-prediction (2025).
3. Sun, X. et al. A data-to-forecast machine learning system for global weather. *Nat. Commun.* **16**, 6658 (2025).
4. Allen, A. et al. End-to-end data-driven weather prediction. *Nature* **641**, 1172–1179 (2025).
5. Price, I. et al. Probabilistic weather forecasting with machine learning. *Nature* **637**, 84–90 (2025).
6. Ben Bouallègue, Z. 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).
7. de Burgh-Day, C. O. & Leeuwenburg, T. Machine learning for numerical weather and climate modelling: a review. *Geosci. Model Dev.* **16**, 6433–6466 (2023).
8. Lam, R. et al. Learning skillful medium-range global weather forecasting. *Science* **382**, 1416–1421 (2023).
9. Lang, S. et al. AIFS–ECMWF's data-driven forecasting system. Preprint at https://arxiv.org/abs/2406.01465 (2024).
10. Asch, M., Bocquet, M. & Nodet, M. *Data Assimilation: Methods, Algorithms, and Application*s (SIAM, 2016).
11. Bauer, P., Thorpe, A. & Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55 (2015).
12. Arcucci, R., Zhu, J., Hu, S. & Guo, Y.-K. Deep data assimilation: integrating deep learning with data assimilation. *Appl. Sci.* **11**, 1114 (2021).
13. Eyre, J., English, S. & Forsythe, M. Assimilation of satellite data in numerical weather prediction. Part I: The early years. *Q J. R. Meteorol. Soc.* **146**, 49–68 (2020).
14. Eyre, J. et al. Assimilation of satellite data in numerical weather prediction. Part II: Recent years. *Q. J. R. Meteorol. Soc.* **148**, 521–556 (2022).
15. Geer, A. Simultaneous inference of sea ice state and surface emissivity model using machine learning and data assimilation. *J. Adv. Model. Earth Syst.* **16**, e2023MS004080 (2024).
16. Geer, A. Joint estimation of sea ice and atmospheric state from microwave imagers in operational weather forecasting. *Q. J. R. Meteorol. Soc.* **150**, 3796–3826 (2024).
17. Dahoui, M. *Use of Machine Learning for the Detection and Classification of Observation Anomalies*. https://www.ecmwf.int/en/newsletter/174/earth-system-science/use-machine-learning-detection-and-classification (2023).
18. Bonavita, M. & Laloyaux, P. Machine learning for model error inference and correction. *J. Adv. Model. Earth Syst.* **12**, e2020MS002232 (2020).
19. Raynaud, L., Brochet, C. & Moldovan, G. ML for weather prediction at Météo-France : current status and future plans, EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-1675. https://doi.org/10.5194/egusphere-egu24-1675 (2024).
20. Hubbard, K. G. et al. *An AI-Driven Approach to Environmental Data Fusion and Assimilation*. Tech. Rep., NOAA (2021).
21. Lang, S. et al. AIFS-CRPS: ensemble forecasting using a model trained with a loss function based on the continuous ranked probability score. *NPJ Artif. Intell.* **2**, 18 (2026).
22. ECMWF & partners. *Anemoi: A New Framework for Forecasting Based on Machine Learning*. https://www.ecmwf.int/en/about/media-centre/news/2024/anemoi-new-framework-weather-forecasting-based-machine-learning (2024).
23. Nipen, T. N. et al. Regional data-driven weather modeling with a global stretched-grid. *Artif. Intell. Earth Syst.* https://doi.org/10.1175/AIES-D-25-0001.1 (2025).
24. Evensen, G. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367 (2003).
25. Liang, J., Terasaki, K. & Miyoshi, T. A machine learning approach to the observation operator for satellite radiance data assimilation. *J. Meteorol. Soc. Jpn.* **101**, 79–95 (2023).
26. Chattopadhyay, A., Mustafa, M., Hassanzadeh, P., Bach, E. & Kashinath, K. Towards physics-inspired data-driven weather forecasting: integrating data assimilation with a deep spatial-transformer-based U-NET in a case study with ERA5. *Geosci. Model Dev.* **15**, 2221–2237 (2022).
27. Chattopadhyay, A., Nabizadeh, E., Bach, E. & Hassanzadeh, P. Deep learning-enhanced ensemble-based data assimilation for high-dimensional nonlinear dynamical systems. *J. Comput. Phys.* **477**, 111918 (2023).
28. Bocquet, M. Surrogate modelling for the climate sciences dynamics with machine learning and data assimilation. *Front. Appl. Math. Stat.* **9**, 1133226 (2023).
29. Hatfield, S. et al. Building tangent-linear and adjoint models for data assimilation with neural networks. *J. Adv. Model. Earth Syst.* **13**, e2021MS002521 (2021).
30. Tian, X., Holdaway, D. & Kleist, D. Evaluating machine learning weather models for data assimilation: fundamental limitations in tangent linear and adjoint properties. *Geophys. Res. Lett.* **53**, e2025GL119402 (2026).
31. Slivinski, L., Whitaker, J. S., Frolov, S., Smith, T. A. & Agarwal, N. Assimilating observed surface pressure into ML weather prediction models. *Geophys. Res. Lett.* **52**, e2024GL114396 (2025).
32. Bonavita, M. On some limitations of current machine learning weather prediction models. *Geophys. Res. Lett.* **51**, e2023GL107377 (2024).
33. Wang, Z., Lei, L., Anderson, J. L., Tan, Z.-M. & Zhang, Y. Convolutional neural network-based adaptive localization for an ensemble Kalman filter. *J. Adv. Model. Earth Syst.* **15**, e2023MS003642 (2023).
34. McCabe, M. & Brown, J. Learning to assimilate in chaotic dynamical systems. *Adv. Neural Inf. Process. Syst.* **34**, 12237–12250 (2021).
35. Boudier, P., Fillion, A., Gratton, S., Gürol, S. & Zhang, S. Data assimilation networks. *J. Adv. Model. Earth Syst.* **15**, e2022MS003353 (2023).
36. Bach, E., Baptista, R., Luk, E. & Stuart, A. Learning optimal filters using variational inference. Preprint at https://arxiv.org/abs/2406.18066 (2025).
37. Bocquet, M. et al. Accurate deep learning-based filtering for chaotic dynamics by identifying instabilities without an ensemble. *Chaos: Interdiscip. J. Nonlinear Sci.* **34**, 091104 (2024).

38. Bach, E., Baptista, R., Calvello, E., Chen, B. & Stuart, A. Learning enhanced ensemble filters. Preprint at https://arxiv.org/abs/2504.17836 (2025).
39. Keller, J. D. & Potthast, R. AI-based data assimilation: learning the functional of analysis estimation. Preprint at https://arxiv.org/abs/2406.00390 (2024).
40. Alexe, M. et al. GraphDOP: Towards skilful data-driven medium-range weather forecasts learnt and initialised directly from observations. Preprint at https://arxiv.org/abs/2412.15687 (2024).
41. Huang, L., Gianinazzi, L., Yu, Y., Dueben, P. D. & Hoefler, T. Diffda: a diffusion model for weather-scale data assimilation. Preprint at https://arxiv.org/abs/2401.05932 (2024).
42. Casas, C. Q., Arcucci, R., Wu, P., Pain, C. & Guo, Y.-K. A reduced order deep data assimilation model. *Phys. D: Nonlinear Phenom.* **412**, 132615 (2020).
43. Arcucci, R., Mottet, L., Pain, C. & Guo, Y.-K. Optimal reduced space for variational data assimilation. *J. Comput. Phys.* **379**, 51–69 (2019).
44. van Leeuwen, P. J., Chiu, J.-Y. C. & Yang, C.-K. K. Uncertainty quantification for deep learning. *Environ. Data Sci.* **4**, e53 (2025).
45. 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).
46. Brajard, J., Carrassi, A., Bocquet, M. & Bertino, L. Combining data assimilation and machine learning to infer unresolved scale parametrisation. *Philos. Trans. R. Soc. A***379** (2021).
47. Bocquet, M., Brajard, J., Carrassi, A. & Bertino, L. Bayesian inference of chaotic dynamics by merging data assimilation, machine learning and expectation-maximization. *Found. Data Sci.* **2**, 55–80 (2020).
48. Bocquet, M., Farchi, A. & Malartic, Q. Online learning of both state and dynamics using ensemble Kalman filters. *Found. Data Sci.* **3**, 305–330 (2021).
49. Farchi, A., Bocquet, M., Laloyaux, P. & Bonavita, M. A comparison of combined data assimilation and machine learning methods for offline and online model error correction. *J. Comput. Sci.* **55**, 101468 (2021).
50. Farchi, A., Chrust, M., Bocquet, M., Laloyaux, P. & Bonavita, M. Online model error correction with neural networks in the incremental 4D-Var framework. *J. Adv. Model. Earth Syst.* **15**, e2022MS003474 (2023).
51. Boukabara, S.-A. et al. Leveraging modern artificial intelligence for remote sensing and NWP: Benefits and challenges. *Bull. Am. Meteorol. Soc.* **100**, ES473–ES491 (2019).
52. Boukabara, S.-A. et al. Outlook for exploiting artificial intelligence in the earth and environmental sciences. *Bull. Am. Meteorol. Soc.* **102**, E1016–E1032 (2021).
53. Chen, N. *Stochastic Methods for Modeling and Predicting Complex Dynamical Systems: Uncertainty Quantification, State Estimation, and Reduced-Order Models* (Springer International Publishing, 2023).
54. Bach, E., Baptista, R., Sanz-Alonso, D. & Stuart, A. Inverse problems and data assimilation: a machine learning approach. Preprint at https://arxiv.org/abs/2410.10523 (2024).
55. Xiao, Y. et al. Fengwu-4DVar: coupling the data-driven weather forecasting model with 4D variational assimilation. Preprint at https://arxiv.org/abs/2312.12455 (2023).
56. Selz, T. & Craig, G. C. Can artificial intelligence-based weather prediction models simulate the butterfly effect? *Geophys. Res. Lett.* **50**, e2023GL105747 (2023).
57. Fablet, R. et al. Learning variational data assimilation models and solvers. *J. Adv. Model. Earth Syst.* **13**, e2021MS002572 (2021).
58. Farchi, A., Chrust, M., Bocquet, M. & Bonavita, M. Development of an offline and online hybrid model for the integrated forecasting system. *Q. J. R. Meteorol. Soc.* **151**, e4934 (2025).
59. Skrunes, F. K. et al. Application of deep learning to the estimation of normalization coefficients in diffusion-based covariance models. Preprint at https://arxiv.org/abs/2312.05068 (2023).
60. Scheck, L. A neural network based forward operator for visible satellite images and its adjoint. *J. Quant. Spectrosc. Radiat. Transf.* **276**, 107841 (2021).
61. Abramovicz, D. et al. Bias correction for radiances using machine learning. *ITSC-25 Abstracts* (2025).
62. Cutraro, A. et al. Machine learning-based lightning forecast for convective resolving numerical weather prediction. In *13th Climate Informatics Conference, Zenodo* (2024).
63. Malartic, Q., Farchi, A. & Bocquet, M. State, global, and local parameter estimation using local ensemble Kalman filters: Applications to online machine learning of chaotic dynamics. *Q. J. R. Meteorol. Soc.* **148**, 2167–2193 (2022).
64. Goyal, A. & Bengio, Y. Inductive biases for deep learning of higher-level cognition. *Proc. R. Soc. A* **478**, 20210068 (2022).
65. Li, J. et al. Deep learning in multimodal remote sensing data fusion: a comprehensive review. *Int. J. Appl. Earth Obs. Geoinf.* **112**, 102926 (2022).
66. Poterjoy, J. A localized particle filter for high-dimensional nonlinear systems. *Monthly Weather Rev.* **144**, 59–76 (2016).
67. Geer, A. Learning Earth system models from observations: machine learning or data assimilation? *Philos. Trans. R. Soc. A* **379** (2021).
68. 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).
69. Atkinson, J. et al. FTorch: a library for coupling PyTorch models to Fortran. *J. Open Source Softw.* **10**, 107 (2025).
70. Buizza, C. et al. Data learning: Integrating data assimilation and machine learning. *J. Comput. Sci.* **58**, 101525 (2022).
71. Lever, J. et al. Facing & mitigating common challenges when working with real-world data: the data learning paradigm. *J. Comput. Sci.* **85**, 102523 (2025).
72. Sacco, M. A., Ruiz, J. J., Pulido, M. & Tandeo, P. Evaluation of machine learning techniques for forecast uncertainty quantification. *Q. J. R. Meteorol. Soc.* **148**, 3470–3490 (2022).
73. Sacco, M. A., Pulido, M., Ruiz, J. J. & Tandeo, P. On-line machine-learning forecast uncertainty estimation for sequential data assimilation. *Q. J. R. Meteorol. Soc.* **150**, 2937–2954 (2024).

Acknowledgements

We acknowledge the discussions and support of the Working Group on Data Assimilation and Observing Systems (DAOS) of the World Weather Research Programme (WWRP) of the World Meteorological Organization (WMO) that led to this article. Their contributions were essential in advancing scientific understanding and fostering international cooperation for weather research.

Author contributions

All authors contributed to the scientific discussions that informed this work. R.A. drafted the initial version of the manuscript summarizing these discussions. S.H. and R.A. prepared the first complete manuscript. R.A., S.H., S.D., L.L., N.C., E.B., A.W., T.M., M.E.D., C.D., R.S., S.L., P.D., N.B., P.L., A.G., M.B., P.J.v.L., S.C., M.Bo., N.Z., H.F.d.C.V., J.J.R., P.B., S.A.B., A.C., R.T., A.Co., D.K., A.Gh., X.W., N.S., G.R., A.M.M., K.L. and N.Ca. contributed to writing, editing, and revising the manuscript. G.R. prepared Fig. 1. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Rossella Arcucci.

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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/.

© The Author(s) 2026