Machine learning–driven advances in geophysical data assimilation

Marc Bocquet,[*]
Tobias Finn,[*] Sibo Cheng,[*] Chaopeng Ji,[*] Wenbo Wu,[*]
Charlotte Durand,[*] Yumeng Chen,[‖] Ivo Pasmans,[‖] Alberto Carrassi,[§]
Massimo Bonavita,[†] Alban Farchi,[†] Marcin Chrust,[†] Patrick Laloyaux,[†]
Julien Brajard,[‡] Laurent Bertino.[‡]

[*] CEREA, ENPC, EDF R&D, Institut Polytechnique de Paris, Île-de-France, France
[‖] Department of Meteorology and National Centre for Earth Observation, University of Reading, United-Kingdom
[§] Department of Physics and Astronomy, University of Bologna, Italy
[†] ECMWF, Reading, United Kingdom,
[‡] NERSC, Bergen, Norway,

Outline

1. Machine learning for data assimilation: literature overview through the lens of methods
   - Machine learning emulators within data assimilation
   - Machine learning-aided data assimilation
   - Machine learning-induced latent space data assimilation
   - Generative AI for data assimilation
   - Holistic machine learning for data assimilation
2. Learning data assimilation from artificial intelligence
   - Sequential data assimilation for chaotic dynamics
   - Learning data assimilation from AI
   - DAN learns to infer the error covariances
   - DAN learns non-Gaussian priors
3. Conclusions

Outline

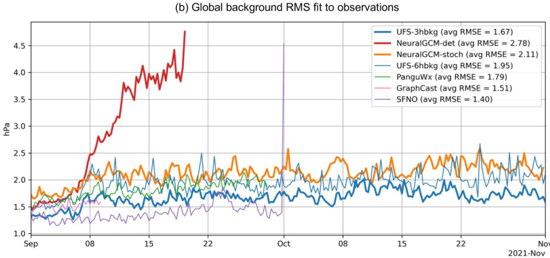
1. Machine learning for data assimilation: literature overview through the lens of methods
   - Machine learning emulators within data assimilation
   - Machine learning-aided data assimilation
   - Machine learning-induced latent space data assimilation
   - Generative AI for data assimilation
   - Holistic machine learning for data assimilation
2. Learning data assimilation from artificial intelligence
   - Sequential data assimilation for chaotic dynamics
   - Learning data assimilation from AI
   - DAN learns to infer the error covariances
   - DAN learns non-Gaussian priors
3. Conclusions

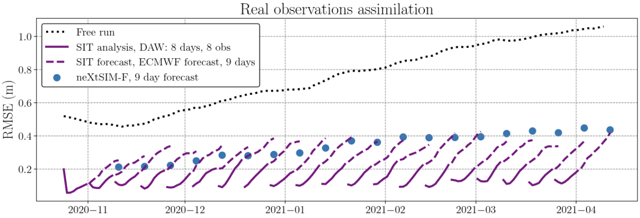
Outline

Machine learning emulators within data assimilation (1/2)

▶ Using machine learning emulators (dyn. and obs.) into classical data assimilation cycle

Hatfield et al. 2021; Chattopadhyay et al. 2022; Li et al. 2024; Slivinski et al. 2025; Durand et al. 2025; Li et al. 2025; Tian et al. 2026; Kotsuki et al. 2025

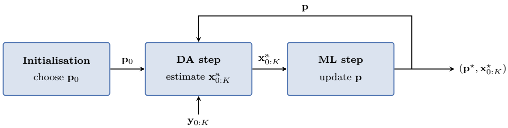




Machine learning emulators within data assimilation (2/2)

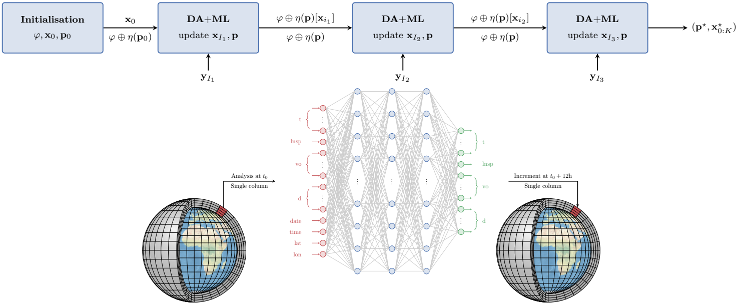
▶ Learning dynamics through Bayesian data assimilation

Hsieh et al. 1998; Abarbanel et al. 2018; Bocquet et al. 2019; Brajard et al. 2020; Bocquet et al. 2020; Brajard et al. 2021; Farchi et al. 2021b; G. Revach et al. 2022.



▶ Learning model error within hybrid models through data assimilation

Brajard et al. 2021; Farchi et al. 2021b; Farchi et al. 2021a; Legler et al. 2022; Farchi et al. 2023; Farchi et al. 2025.

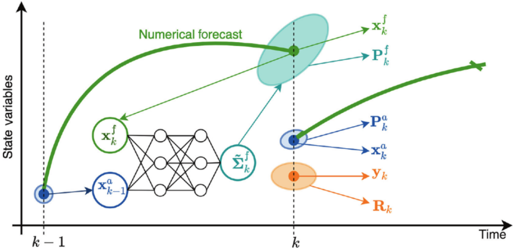


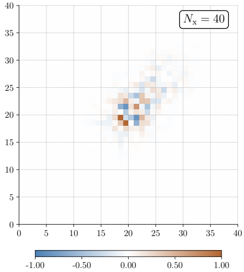
Outline

Machine learning-assisted data assimilation (1/2)

▶ For uncertainty quantification

Grooms 2021; L. M. Yang et al. 2021; Sacco et al. 2022; Sacco et al. 2024; Lu 2025; Bocquet et al. 2026.



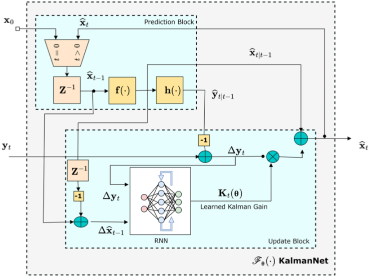


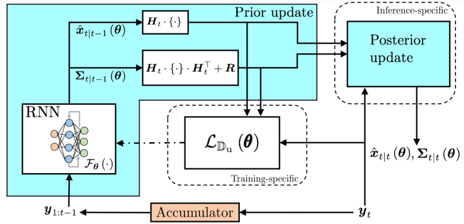
▶ For physical balances, tuning hybrid data assimilation systems

Ruckstuhl et al. 2021; Dong et al. 2023.

Machine learning-assisted data assimilation (2/2)

- ▶ Learning the (static) gain of the (ensemble) Kalman filter
  H. Hoang et al. 1994; S. Hoang et al. 1998; Luk et al. 2024.
- ▶ Tuning (hyper-)parameters through auto-differentiable (ensemble) Kalman filters
  Haarnoja et al. 2016; Y. Chen et al. 2022; Luk et al. 2024; Shlezinger et al. 2024.





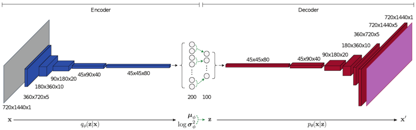
- ▶ Signal processing literature: Kalman filter/low dimensional oriented – still very impressive!
  Coskun et al. 2017; Shlezinger et al. 2023; Garcia Satorras et al. 2019; Klushyn et al. 2021; Pratik et al. 2021; Gedon et al. 2021; Shlezinger et al. 2022; Guy Revach et al. 2022; Cheng et al. 2023b; Choi et al. 2023; A. et al. 2024; Buchnik et al. 2024; Imbiriba et al. 2024; Ghosh et al. 2024

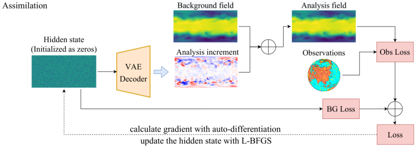
Outline

Machine learning-induced latent space data assimilation (1/2)

▶ Better representation of (possibly non-Gaussian) error background statistics

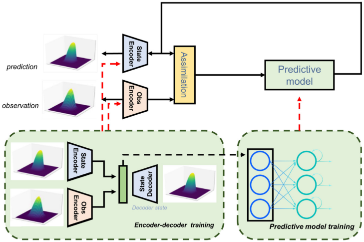
Mack et al. 2020; Melinc et al. 2024; Xiao et al. 2025

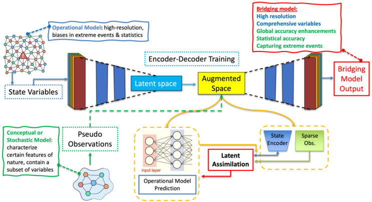




▶ Better handling heterogeneity of observations/states and multimodality

Cheng et al. 2022; Cheng et al. 2023a; Cheng et al. 2024; Behnoudfar et al. 2026

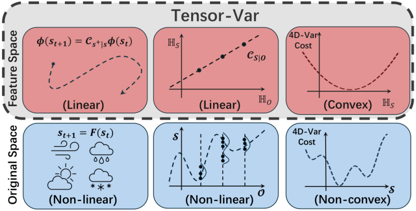


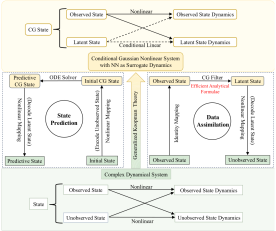


Machine learning-induced latent space data assimilation (2/2)

▶ Enforcing linearity/Gaussianity within the latent space

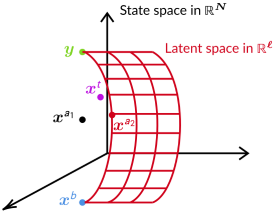
Tarumi et al. 2025; Y. Yang et al. 2025; C. Chen et al. 2025; Pasmans et al. 2026

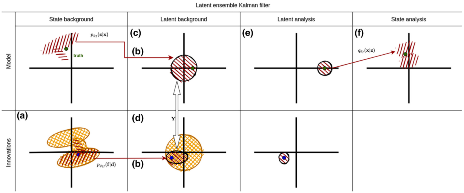




▶ Ensemble latent data assimilation

Haarnoja et al. 2016; Becker et al. 2019; Peyron et al. 2021; Penny et al. 2022; Tarumi et al. 2025; Pasmans et al. 2026



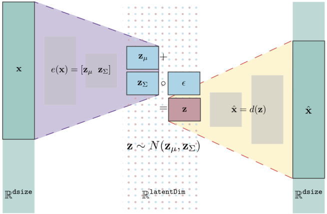


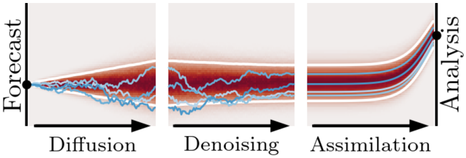
Outline

Generative AI for data assimilation (1/2)

▶ For building ensemble priors, from 3D-Var, EnOI, . . .

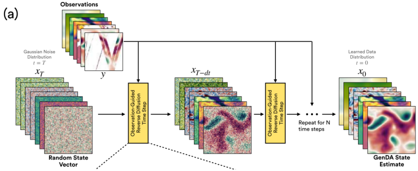
Grooms 2021; L. M. Yang et al. 2021; Finn et al. 2024.

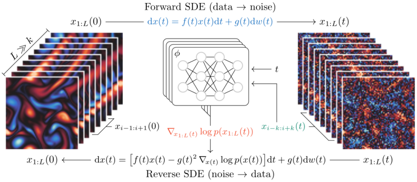




▶ . . . to 4DEnVar, . . .

Rozet et al. 2023; Rozet et al. 2024.

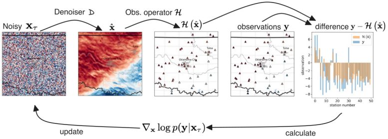


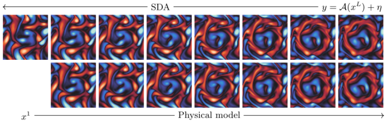


Generative AI for data assimilation (2/2)

▶ For an incremental score-based analysis (SDA)

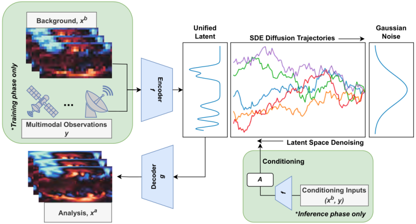
Chung et al. 2023; Rozet et al. 2023; Rozet et al. 2024; Manshausen et al. 2025; Martin et al. 2025.

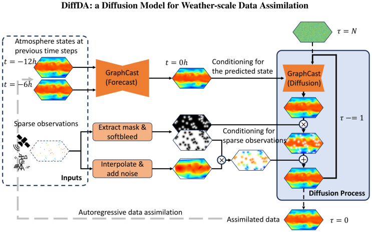




▶ As a replacement for the analysis

Chung et al. 2023; Huang et al. 2024; Qu et al. 2024; Hodyss et al. 2026.



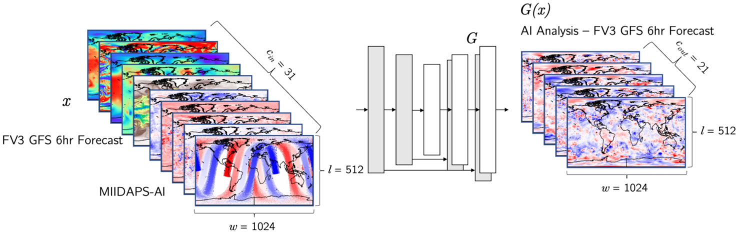


Outline

Holistic machine learning for data assimilation (1/4)

▶ Supervised learning of the analysis from data assimilation runs

R. S. Cintra et al. 2012; F. P. Härter et al. 2008; T. P. Härter et al. 2012; R. Cintra et al. 2016; R. S. Cintra et al. 2018; Maddy et al. 2024; Xu et al. 2025



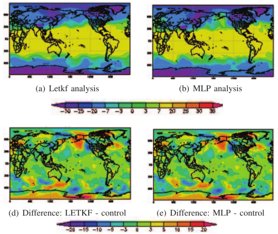
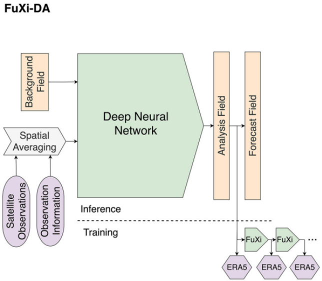


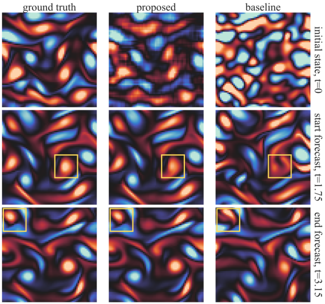
Fig. 6. Temperature (°C) Fields at layer 500 hPa to 08/01/2004 at 12 UTC. (a) LETKF analysis (b) MLP-DA analysis (d) diferrences between LETKF and MLP-DA analyses.

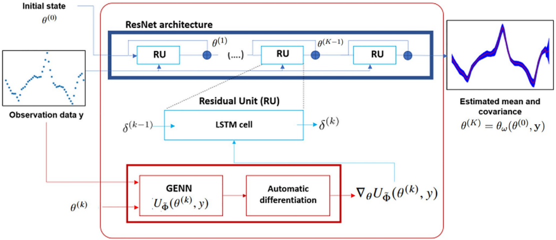


Holistic machine learning for data assimilation (2/4)

▶ Learning variational solver

Fablet et al. 2021; Frerix et al. 2021; Lafon et al. 2023; Filoche et al. 2023; Keller et al. 2024

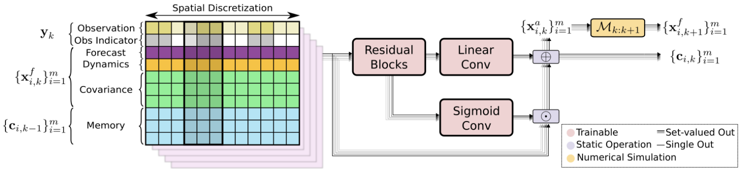


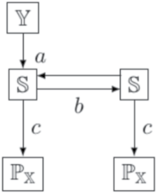


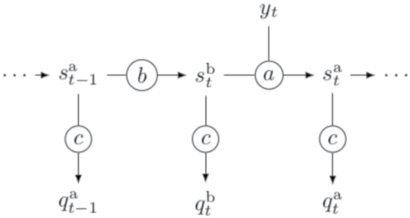
Holistic machine learning for data assimilation (3/4)

▶ Learning sequential data assimilation through the cycles

McCabe et al. 2021; Boudier et al. 2023; Bocquet et al. 2024; Ghosh et al. 2024; Bocquet et al. 2026



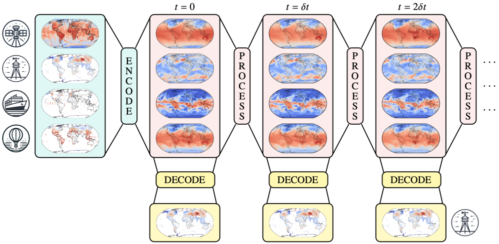


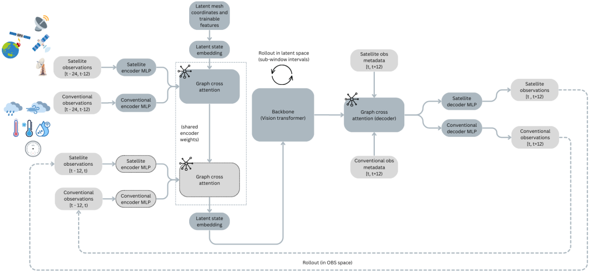


Holistic machine learning for data assimilation (4/4)

- ▶ End-to-end numerical weather prediction

McNally et al. 2024; Vaughan et al. 2024; Sun et al. 2025; Alexe et al. 2024; Allen et al. 2025; Lean et al. 2025





Outline

Sequential data assimilation for chaotic dynamics: primer

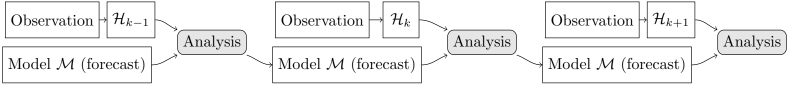
- Here, data assimilation (DA) methods are formulated from

$$\mathbf{x}_{k+1} = \mathcal{M}(\mathbf{x}_k), \tag{1a}$$

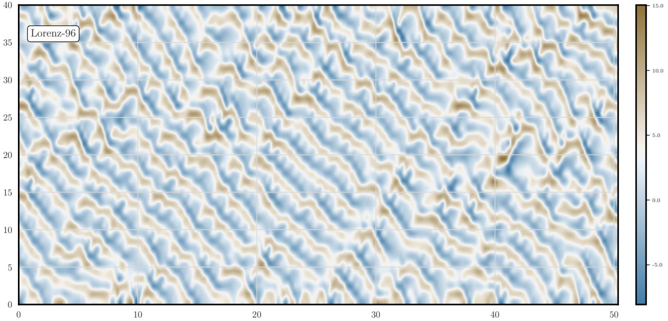
$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \boldsymbol{\varepsilon}_k, \qquad \boldsymbol{\varepsilon}_k \sim N(\mathbf{0}, \mathbf{R}_k), \tag{1b}$$

where $\mathcal{M}$ is the *autonomous* evolution model, $\mathbf{x}_k$ is the state vector at time τ_k, $\mathbf{y}_k$ is the observation vector, $\mathcal{H}_k$ is the observation operator, $\boldsymbol{\varepsilon}_k$ is the observation error, assumed to be additive, unbiased, white in time, and Gaussian of covariance matrix $\mathbf{R}_k$.

- DA for geofluids has to be *sequential* in time because (i) observations need to be assimilated *as they arrive* to update the state estimation, (ii) applied to *chaotic dynamics*, typical errors have an exponential growth.



Low-order chaotic dynamics – Lorenz 96 model



- It represents a mid-latitude zonal circle of the global atmosphere.
- Set of $N_\mathrm{x} = 40$ ordinary differential equations [Lorenz et al. 1998]:

$$\frac{\mathrm{d}x_n}{\mathrm{d}t} = (x_{n+1} - x_{n-2})x_{n-1} - x_n + F\,, \tag{2}$$

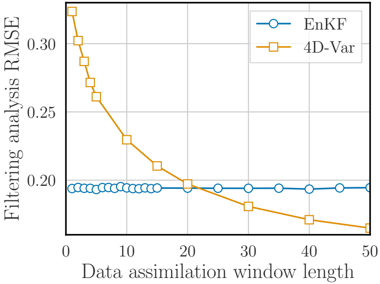
where $F = 8$, and the boundary is cyclic.

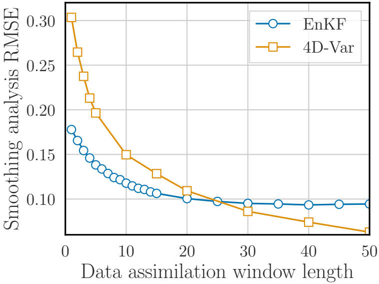
The edge of ensemble filtering methods

▶ The variational methods (3D–Var, 4D–Var): can handle nonlinearity of the operators, asynchronous observations, but *cannot handle the errors of the day.*

▶ The ensemble filtering methods (EnKFs): can only handle weak nonlinearity of the operators, cannot handle asynchronous observations, *can handle the errors of the day* through the ensemble.

▶ Testing the EnKF ($N_\mathrm{e} = 20$), 4D–Var, and IEnKS ($N_\mathrm{e} = 20$) variants with the chaotic 40–variable Lorenz 96 model [Bocquet et al. 2013; Asch et al. 2016]:





▶ In mild nonlinear regime, the EnKF significantly outperforms the (basic) 4D–Var with moderately large DA windows because it captures the *errors of the day*.

Learning data assimilation from AI

- Data Assimilation Network principle (DAN): replace the analysis with a neural network [McCabe et al. 2021; Boudier et al. 2023; Bocquet et al. 2024; Bocquet et al. 2026].

- [Implementing the idea] Let us assume that *$\mathcal{M}$ is known*, that the Jacobian of $\mathcal{H}_k$ is $\mathbf{H}_k$, and that we wish to learn an *incremental analysis operator $a_{\boldsymbol{\theta}}$*, typically a neural network parametrised by $\boldsymbol{\theta}$.

- If $\mathbf{E}_k^{\mathrm{a}}, \mathbf{E}_k^{\mathrm{f}} \in \mathbb{R}^{N_{\mathrm{x}} \times N_{\mathrm{e}}}$ are the analysis and forecast ensemble matrices at time τ_k, $a_{\boldsymbol{\theta}}$ is defined via the (ensemble) update:

$$\mathbf{E}_k^{\mathrm{a}} = \mathbf{E}_k^{\mathrm{f}} + a_{\boldsymbol{\theta}} \left(\mathbf{E}_k^{\mathrm{f}}, \mathbf{H}_k^{\intercal} \mathbf{R}_k^{-1} \boldsymbol{\delta}_k \right), \tag{3a}$$

where $\boldsymbol{\delta}_k$, the innovation at time τ_k, is defined by

$$\boldsymbol{\delta}_k \triangleq \mathbf{y}_k - \mathcal{H}_k \left(\bar{\mathbf{x}}_k^{\mathrm{f}} \right), \quad \bar{\mathbf{x}}_k^{\mathrm{f}} \triangleq \frac{1}{N_{\mathrm{e}}} \sum_{i=1}^{N_{\mathrm{e}}} \mathbf{x}_k^{\mathrm{f},i}. \tag{3b}$$

- The DA forecast step propagates the analysis ensemble, member-wise:

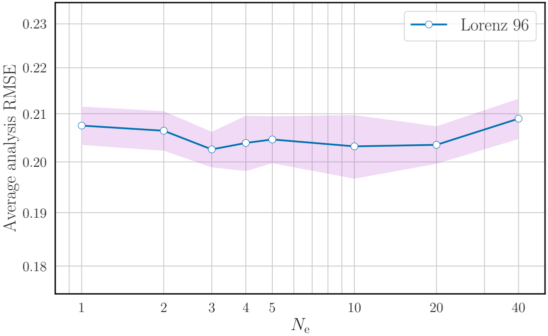
$$\mathbf{E}_{k+1}^{\mathrm{f}} = \mathcal{M} \left(\mathbf{E}_k^{\mathrm{a}} \right). \tag{4}$$

- The *loss function* is defined by

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{r=1}^{N_{\mathrm{r}}} \sum_{k=1}^{N_{\mathrm{c}}} \left\| \mathbf{x}_k^{\mathrm{t},r} - \bar{\mathbf{x}}_k^{\mathrm{a},r}(\boldsymbol{\theta}) \right\|^2, \quad \bar{\mathbf{x}}_k^{\mathrm{a},r} \triangleq \frac{1}{N_{\mathrm{e}}} \sum_{i=1}^{N_{\mathrm{e}}} \mathbf{x}_k^{\mathrm{a},i,r}. \tag{5}$$

DAN learns the uncertainty of its own estimate

▶ Key observation: The performance of $a_{\boldsymbol{\theta}}$ barely depends on the ensemble size N_e, and yet as good as that of the EnKF[1].

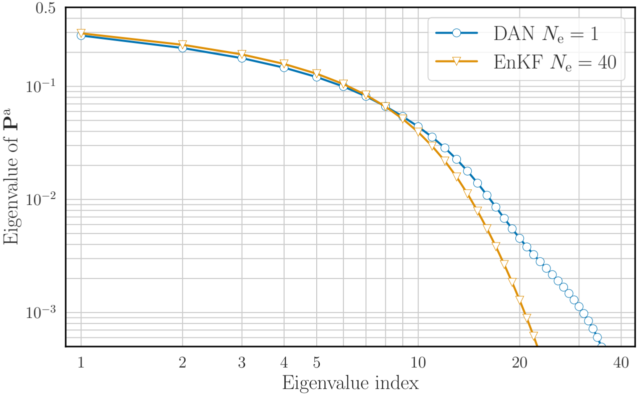


▶ Conclusion: $a_{\boldsymbol{\theta}}$ essentially implicitly learns the *error covariance matrix* associated to the analysis!

[1] [Bocquet et al. 2024]

Cautiously learning the error covariances

- The time-averaged eigenspectra of $\mathbf{P}^{\mathrm{a}}_{\mathrm{DAN}}$ and $\mathbf{P}^{\mathrm{a}}_{\mathrm{EnKF}}$:

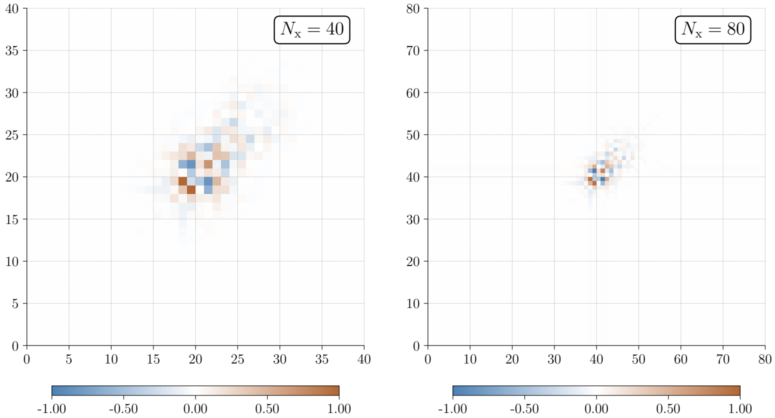


- They are remarkably close to each other for the first 10 modes. Beyond these modes the $a_{\boldsymbol{\theta}}$ operator is likely to selectively apply *some multiplicative inflation*, as one would expect from such stable DA runs.

Locality and scalability

▶ DAN scales perfectly when $N_\mathrm{x} \nearrow \infty$ and does not need to be retrained: it must learn a fixed number of *local patterns to diagnose the covariances*[2].

▶ These *local patterns* can be pictured via the diagnosed *marginal Kalman gain*[3].

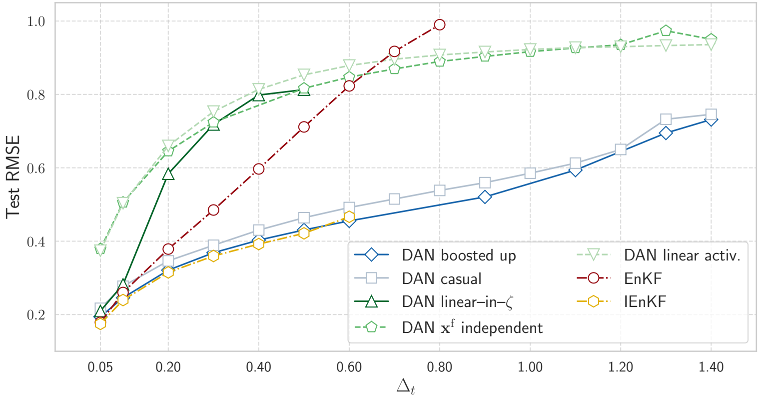


[2] [Bocquet et al. 2024]

[3] [Bocquet et al. 2026]

Data assimilation networks in stronger nonlinear conditions

- Testing DANs as *the update time-step Δ_t is increased*, with the L96 model[4].
- Comparison with well tuned *EnKF*, *IEnKF*, and well-tuned static background DA methods.



- Conclusion: $a_{\boldsymbol{\theta}}$ implicitly learns a *non-Gaussian prior* associated to the analysis!

[4] [Bocquet et al. 2026]

Conclusions

- A significantly growing literature on *machine learning and data assimilation*.

⟶ More mature (e.g., latent space), and still jumbled (e.g., generative AI).

- Many promising proofs of concepts over the past couple of years.

⟶ Now seems to be the time to thoroughly assess them.

- AI can also be leveraged as a *discovery tool* to learn new algorithms.

⟶ Can also shed light on our classical data assimilation schemes.

References I

[1] Gu. A. and Dao. T. "Mamba: Linear-Time Sequence Modeling with Selective State Spaces". In: *First Conference on Language Modeling*. 2024.

[2] H. D. I. Abarbanel, P. J. Rozdeba, and S. Shirman. "Machine Learning: Deepest Learning as Statistical Data Assimilation Problems". In: *Neural Computation* 30 (2018), pp. 2025–2055. eprint: https://doi.org/10.1162/neco_a_01094.

[3] M. Alexe et al. *GraphDOP: Towards skilful data-driven medium-range weather forecasts learnt and initialised directly from observations*. 2024. arXiv: 2412.15687 [physics.ao-ph].

[4] A. Allen et al. "End-to-end data-driven weather prediction". In: *Nature* (2025).

[5] M. Asch, M. Bocquet, and M. Nodet. *Data Assimilation: Methods, Algorithms, and Applications*. Fundamentals of Algorithms. SIAM, Philadelphia, 2016, p. 324.

[6] P. Becker et al. "Recurrent Kalman Networks: Factorized Inference in High-Dimensional Deep Feature Spaces". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, Sept. 2019, pp. 544–552.

[7] P. Behnoudfar et al. "Bridging idealized and operational models: an explainable AI framework for Earth system emulators". In: *npj Clim Atmos Sci* 9 (2026), p. 65.

[8] M. Bocquet and P. Sakov. "Joint state and parameter estimation with an iterative ensemble Kalman smoother". In: *Nonlin. Processes Geophys.* 20 (2013), pp. 803–818.

[9] M. Bocquet et al. "Accurate deep learning-based filtering for chaotic dynamics by identifying instabilities without an ensemble". In: *Chaos* 29 (2024), p. 091104.

[10] M. Bocquet et al. "Bayesian inference of chaotic dynamics by merging data assimilation, machine learning and expectation-maximization". In: *Foundations of Data Science* 2 (2020), pp. 55–80.

[11] M. Bocquet et al. "Data assimilation as a learning tool to infer ordinary differential equation representations of dynamical models". In: *Nonlin. Processes Geophys.* 26 (2019), pp. 143–162.

[12] M. Bocquet et al. "Elucidating the performance of data assimilation neural networks for chaotic dynamics". In: *EGUsphere* 2026 (2026), pp. 1–34.

[13] P. Boudier et al. "Data Assimilation Networks". In: *J. Adv. Model. Earth Syst.* 15 (2023), e2022MS003353.

[14] J. Brajard et al. "Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: a case study with the Lorenz 96 model". In: *J. Comput. Sci.* 44 (2020), p. 101171.

References II

[15] J. Brajard et al. "Combining data assimilation and machine learning to infer unresolved scale parametrisation". In: *Phil. Trans. R. Soc. A* 379 (2021), p. 20200086. eprint: arXiv:2009.04318.

[16] I. Buchnik et al. "Latent-KalmanNet: Learned Kalman Filtering for Tracking From High-Dimensional Signals". In: *IEEE Transactions on Signal Processing* 72 (2024), pp. 352–367.

[17] A. Chattopadhyay et al. "Towards physics-inspired data-driven weather forecasting: integrating data assimilation with a deep spatial-transformer-based U-NET in a case study with ERA5". In: *Geosci. Model Dev.* 15 (2022), pp. 2221–2237.

[18] C. Chen et al. "CGKN: A deep learning framework for modeling complex dynamical systems and efficient data assimilation". In: *J. Comp. Phys.* 532 (2025), p. 113950.

[19] Y. Chen, D. Sanz-Alonso, and R. Willett. "Autodifferentiable Ensemble Kalman Filters". In: *SIAM J. Math. Data Sci.* 4 (2022), pp. 801–833.

[20] S. Cheng et al. "Data-driven surrogate model with latent data assimilation: Application to wildfire forecasting". In: *J. Comp. Phys.* 464 (2022), p. 111302.

[21] S. Cheng et al. "Generalised Latent Assimilation in Heterogeneous Reduced Spaces with Machine Learning Surrogate Models". In: *J. Sci. Comput.* 94 (2023), pp. 1573–7691.

[22] S. Cheng et al. "Machine learning with data assimilation and uncertainty quantification for dynamical systems: a review". In: *IEEE/CAA J. Autom. Sin.* 10 (2023), pp. 1361–1387.

[23] S. Cheng et al. "Multi-domain encoder–decoder neural networks for latent data assimilation in dynamical systems". In: *Computer Methods in Applied Mechanics and Engineering* 430 (2024), p. 117201.

[24] G. Choi et al. "Split-KalmanNet: A Robust Model-Based Deep Learning Approach for State Estimation". In: *IEEE Transactions on Vehicular Technology* 72 (2023), 12326–=12331.

[25] H. Chung et al. "Diffusion Posterior Sampling for General Noisy Inverse Problems". In: *The Eleventh International Conference on Learning Representations*. 2023.

[26] R. Cintra, H. de Campos Velho, and S. Cocke. "Tracking the model: Data assimilation by artificial neural network". In: *2016 International Joint Conference on Neural Networks (IJCNN)*. 2016, pp. 403–410.

[27] R. S. Cintra and H. F. de Campos Velho. "Data assimilation by artificial neural networks for an atmospheric general circulation model". In: *Advanced applications for artificial neural networks*. Ed. by A. ElShahat. IntechOpen, 2018. Chap. 17, pp. 265–286.

[28] R. S. Cintra and H. F. de Campos Velho. "Global data assimilation using artificial neural networks in SPEEDY model". In: *1st International Symposium Uncertainty Quantification and Stochastic Modeling*. Maresias, são Sebastião (SP), Brazil, 2012, pp. 648–654.

References III

[29] H. Coskun et al. "Long Short-Term Memory Kalman Filters: Recurrent Neural Estimators for Pose Regularization". In: *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017, pp. 5525–5533.

[30] R. Dong et al. "A hybrid data assimilation system based on machine learning". In: *Frontiers in Earth Science* 10 (2023).

[31] C. Durand et al. "Four-dimensional variational data assimilation with a sea-ice thickness emulator". In: *EGUsphere* 2025 (2025), pp. 1–36.

[32] R. Fablet et al. "Learning Variational Data Assimilation Models and Solvers". In: *J. Adv. Model. Earth Syst.* 13 (2021), e2021MS002572.

[33] A. Farchi et al. "A comparison of combined data assimilation and machine learning methods for offline and online model error correction". In: *J. Comput. Sci.* 55 (2021), p. 101468.

[34] A. Farchi et al. "Development of an offline and online hybrid model for the Integrated Forecasting System". In: *Q. J. R. Meteorol. Soc.* (2025), e4934.

[35] A. Farchi et al. "Online Model Error Correction With Neural Networks in the Incremental 4D-Var Framework". In: *J. Adv. Model. Earth Syst.* 15 (2023), e2022MS003474.

[36] A. Farchi et al. "Using machine learning to correct model error in data assimilation and forecast applications". In: *Q. J. R. Meteorol. Soc.* 147 (2021), pp. 3067–3084.

[37] A. Filoche et al. "Learning 4DVAR Inversion Directly from Observations". In: *Computational Science – ICCS 2023*. Ed. by Jiří Mikyška et al. Cham: Springer Nature Switzerland, 2023, pp. 414–421.

[38] T. S. Finn et al. "Representation learning with unconditional denoising diffusion models for dynamical systems". In: *Nonlin. Processes Geophys.* 31 (2024), pp. 409–431.

[39] T. Frerix et al. "Variational Data Assimilation with a Learned Inverse Observation Operator". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by Marina Meila and Tong Zhang. Vol. 139. Proceedings of Machine Learning Research. PMLR, 18–24 Jul 2021, pp. 3449–3458.

[40] V. Garcia Satorras, Z. Akata, and M. Welling. "Combining Generative and Discriminative Models for Hybrid Inference". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019.

[41] D. Gedon et al. "Deep State Space Models for Nonlinear System Identification". In: *IFAC-PapersOnLine* 54 (2021). 19th IFAC Symposium on System Identification SYSID 2021, pp. 481–486.

[42] A. Ghosh, A. Honoré, and S. Chatterjee. "DANSE: Data-Driven Non-Linear State Estimation of Model-Free Process in Unsupervised Learning Setup". In: *IEEE Transactions on Signal Processing* 72 (2024), pp. 1824–1838.

References IV

[43] I. Grooms. "Analog ensemble data assimilation and a method for constructing analogs with variational autoencoders". In: *Q. J. R. Meteorol. Soc.* 147 (2021), pp. 139–149.

[44] T. Haarnoja et al. "Backprop KF: Learning Discriminative Deterministic State Estimators". In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee et al. Vol. 29. Curran Associates, Inc., 2016.

[45] F. P. Härter and H. F. de Campos Velho. "New approach to applying neural network in nonlinear dynamic model". In: *Appl. Math. Model.* 32 (2008), pp. 2621–2633.

[46] T. P. Härter and H. F. de Campos Velho. "Data Assimilation Procedure by Recurrent Neural Network". In: *Eng. Appl. Comput. Fluid Mech.* 6 (2012), pp. 224–233.

[47] S. Hatfield et al. "Building Tangent-Linear and Adjoint Models for Data Assimilation With Neural Networks". In: *J. Adv. Model. Earth Syst.* 13 (2021), e2021MS002521.

[48] H.S. Hoang, P. De Mey, and O. Talagrand. "A simple adaptive algorithm of stochastic approximation type for system parameter and state estimation". In: *Proceedings of 1994 33rd IEEE Conference on Decision and Control*. Vol. 1. 1994, 747–752 vol.1.

[49] S. Hoang et al. "Adaptive filtering: application to satellite data assimilation in oceanography". In: *Dynam. Atmos. Ocean* 27 (1998), pp. 257–281.

[50] D. Hodyss and M. Morzfeld. "Using Diffusion Models to do Data Assimilation". In: *Mon. Wea. Rev.* 154 (2026), pp. 165–182.

[51] W. W. Hsieh and B. Tang. "Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography". In: *Bull. Amer. Meteor. Soc.* 79 (1998), pp. 1855–1870.

[52] L. Huang et al. "DiffDA: a diffusion model for weather-scale data assimilation". In: *Proceedings of the 41st International Conference on Machine Learning*. ICML'24. Vienna, Austria: JMLR.org, 2024.

[53] T. Imbiriba et al. "Augmented Physics-Based Machine Learning for Navigation and Tracking". In: *IEEE Transactions on Aerospace and Electronic Systems* 60 (2024), pp. 2692–2704.

[54] J. D. Keller and R. Potthast. *AI-based data assimilation: Learning the functional of analysis estimation*. 2024. arXiv: `2406.00390 [physics.ao-ph]`.

[55] A. Klushyn et al. "Latent Matters: Learning Deep State-Space Models". In: *Advances in Neural Information Processing Systems*. Ed. by A. Beygelzimer et al. 2021.

[56] S. Kotsuki, K. Shiraishi, and A. Okazaki. "Ensemble data assimilation to diagnose AI-based weather prediction models: a case with ClimaX version 0.3.1". In: *Geosci. Model Dev.* 18 (2025), pp. 7215–7225.

References V

[57] N. Lafon, R. Fablet, and P. Naveau. "Uncertainty Quantification When Learning Dynamical Models and Solvers With Variational Methods". In: *J. Adv. Model. Earth Syst.* 15 (2023), e2022MS003446.

[58] P. Lean et al. *Learning from nature: insights into GraphDOP's representations of the Earth System*. 2025. arXiv: 2508.18018 [physics.ao-ph].

[59] S. Legler and T. Janjić. "Combining data assimilation and machine learning to estimate parameters of a convective-scale model". In: *Q. J. R. Meteorol. Soc.* 148 (2022), pp. 860–874.

[60] Y. Li et al. "A Machine Learning-Based Observation Operator for FY-4B GIIRS Brightness Temperatures Considering the Uncertainty of Label Data". In: *Journal of Geophysical Research: Machine Learning and Computation* 2 (2025), e2024JH000449.

[61] Y. Li et al. "FuXi-En4DVar: An Assimilation System Based on Machine Learning Weather Forecasting Model Ensuring Physical Constraints". In: *Geophys. Res. Lett.* 51 (2024), e2024GL111136.

[62] E. N. Lorenz and K. A. Emanuel. "Optimal sites for supplementary weather observations: simulation with a small model". In: *J. Atmos. Sci.* 55 (1998), pp. 399–414.

[63] F. Lu. "U-Net Kalman Filter (UNetKF): An Example of Machine Learning-Assisted Data Assimilation". In: *J. Adv. Model. Earth Syst.* 17 (2025), e2023MS003979.

[64] E. Luk et al. *Learning Optimal Filters Using Variational Inference*. 2024. arXiv: 2406.18066 [cs.LG].

[65] J. Mack et al. "Attention-based Convolutional Autoencoders for 3D-Variational Data Assimilation". In: *Computer Methods in Applied Mechanics and Engineering* 372 (2020), p. 113291.

[66] E. S. Maddy, S. A. Boukabara, and F. Iturbide-Sanchez. "Assessing the Feasibility of an NWP Satellite Data Assimilation System Entirely Based on AI Techniques". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 17 (2024), pp. 9828–9845.

[67] P. Manshausen et al. "Generative Data Assimilation of Sparse Weather Station Observations at Kilometer Scales". In: *J. Adv. Model. Earth Syst.* (2025), e2024MS004505.

[68] S. A. Martin, G. E. Manucharyan, and P. Klein. "Generative Data Assimilation for Surface Ocean State Estimation From Multi-Modal Satellite Observations". In: *J. Adv. Model. Earth Syst.* 17 (2025), e2025MS005063.

[69] M. McCabe and J. Brown. "Learning to Assimilate in Chaotic Dynamical Systems". In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato et al. Vol. 34. Curran Associates, Inc., 2021, pp. 12237–12250.

[70] T. McNally et al. *Red sky at night... producing weather forecasts directly from observations*. Jan. 2024.

[71] B. Melinc and Z. Zaplotnik. "3D-Var data assimilation using a variational autoencoder". In: *Q. J. R. Meteorol. Soc.* 150 (2024), pp. 2273–2295.

References VI

[72] I. Pasmans et al. “Ensemble Kalman filter in latent space using a variational autoencoder pair”. In: *Q. J. R. Meteorol. Soc.* 152 (2026), e70070.

[73] S. G. Penny et al. “Integrating recurrent neural networks with data assimilation for scalable data-driven state estimation”. In: *J. Adv. Model. Earth Syst.* 14 (2022), e2021MS002843.

[74] M. Peyron et al. “Latent space data assimilation by using deep learning”. In: *Q. J. R. Meteorol. Soc.* 147 (2021), pp. 3759–3777.

[75] K. Pratik et al. “Neural Augmentation of Kalman Filter with Hypernetwork for Channel Tracking”. In: *2021 IEEE Global Communications Conference (GLOBECOM)*. 2021, pp. 1–6.

[76] Y. Qu et al. “Deep Generative Data Assimilation in Multimodal Setting”. In: *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Los Alamitos, CA, USA: IEEE Computer Society, June 2024, pp. 449–459.

[77] G. Revach et al. “KalmanNet: Neural Network Aided Kalman Filtering for Partially Known Dynamics”. In: *IEEE Trans. Signal Process.* 70 (2022), pp. 1532–1547.

[78] Guy Revach et al. “Unsupervised Learned Kalman Filtering”. In: *2022 30th European Signal Processing Conference (EUSIPCO)*. 2022, pp. 1571–1575.

[79] F. Rozet and G. Louppe. “Score-based Data Assimilation”. In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh et al. Vol. 36. Curran Associates, Inc., 2023, pp. 40521–40541.

[80] F. Rozet et al. “Learning Diffusion Priors from Observations by Expectation Maximization”. In: *The Thirty-eighth Annual Conference on Neural Information Processing Systems*. 2024.

[81] Y. Ruckstuhl, T. Janjić, and S. Rasp. “Training a convolutional neural network to conserve mass in data assimilation”. In: *Nonlin. Processes Geophys.* 28 (2021), pp. 111–119.

[82] M. A. Sacco et al. “Evaluation of machine learning techniques for forecast uncertainty quantification”. In: *Q. J. R. Meteorol. Soc.* 148 (2022), pp. 3470–3490.

[83] M. A. Sacco et al. “On-line machine-learning forecast uncertainty estimation for sequential data assimilation”. In: *Q. J. R. Meteorol. Soc.* 150 (2024), pp. 2937–2954.

[84] N. Shlezinger, Y. C. Eldar, and S. P. Boyd. “Model-Based Deep Learning: On the Intersection of Deep Learning and Optimization”. In: *IEEE Access* 10 (2022), pp. 115384–115398.

[85] N. Shlezinger and T. Routtenberg. “Discriminative and Generative Learning for the Linear Estimation of Random Signals [Lecture Notes]”. In: *IEEE Signal Processing Magazine* 40 (2023), pp. 75–82.

References VII

[86] N. Shlezinger et al. 2024. eprint: 2410.12289 (cs.LG).

[87] L. C. Slivinski et al. "Assimilating Observed Surface Pressure into ML Weather Prediction Models". In: *Geophys. Res. Lett.* 52 (2025), e2024GL114396.

[88] X. Sun et al. "A data-to-forecast machine learning system for global weather". In: *Nature Communications* 16 (2025), p. 6658.

[89] y. Tarumi, K. Fukuda, and S.-I. Maeda. "Deep Bayesian Filter for Bayes-Faithful Data Assimilation". In: *Forty-second International Conference on Machine Learning*. 2025.

[90] X. Tian, D. Holdaway, and D. Kleist. "Evaluating Machine Learning Weather Models for Data Assimilation: Fundamental Limitations in Tangent Linear and Adjoint Properties". In: *Geophys. Res. Lett.* 53 (2026), e2025GL119402.

[91] A. Vaughan et al. *Aardvark weather: end-to-end data-driven weather forecasting*. 2024. arXiv: 2404.00411 [physics.ao-ph].

[92] Y. Xiao et al. "VAE-Var: Variational Autoencoder-Enhanced Variational Methods for Data Assimilation in Meteorology". In: *The Thirteenth International Conference on Learning Representations*. 2025.

[93] X. Xu et al. "FuXi-DA: a generalized deep learning data assimilation framework for assimilating satellite observations". In: *npj Clim Atmos Sci* (2025), p. 156.

[94] L. M. Yang and I. Grooms. "Machine learning techniques to construct patched analog ensembles for data assimilation". In: *J. Comput. Sci.* 443 (2021), p. 110532.

[95] Y. Yang et al. "Tensor-Var: Efficient Four-Dimensional Variational Data Assimilation". In: *Forty-second International Conference on Machine Learning*. 2025.