Harmonizing Knowledge: Machine Learning Meets Data Assimilation

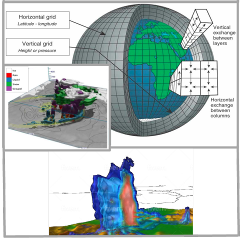
Tijana Janjić

Mathematical Institute for Machine Learning and Data Science,
KU Eichstätt-Ingolstadt

MIDS
Mathematical Institute for
Machine Learning and Data Science

Power of data assimilartion

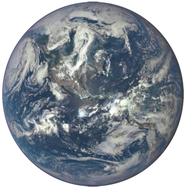
High-resolution simulation and noisy observations



- Numerical models of the atmosphere
- Use ensemble of simulations to learn uncertainty
- Use of noisy observations
- Detailed observation operator

→ Uncertainty-aware physics-informed estimate of the state and estimate of its uncertainty

Challenges of data assimilation

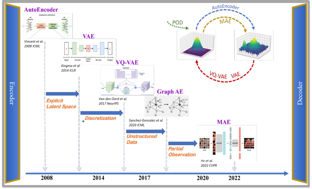


- Huge state vector
- Difficulties in estimating errors of the numerical model from data
- Difficulties in estimating observation error statistics
- Statistics-based DA algorithms that lose physical properties

→ Learn from ML

Power of ML

Reduced order models



- Fully data driven
- Able to find reduced order approximation
- Able to find hidden relations
- No prior assumptions
- VERY FAST

→

DA in latent space

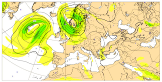
Model error estimation

Parameter estimation

Cheng et al. 2023: Machine Learning With Data Assimilation and Uncertainty Quantification for Dynamical Systems: A Review. IEEE/CAA Journal of Automatica Sinica, 10 (6), 1361-1387.

Challenges of ML

ML models



- Simpler ML models
- Physics not well represented (Bonavita 2024, Selz and Craig 2023)
- Ensemble under-dispersive
- Difficulties in use of noisy observations (Brajard et al. 2019)

→ Learn from DA

Learning from observations

DA + ML

- Learning the state and its uncertainty
- Learning parameters of physical models and their uncertainty

Learning from observations

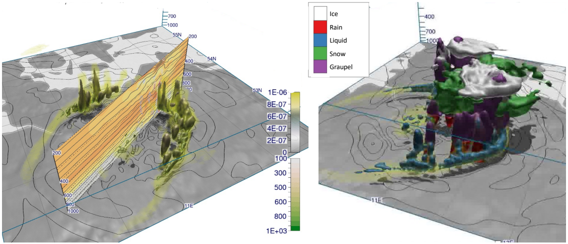
DA + ML

- Learning the state and its uncertainty
- Learning parameters of physical models and their uncertainty

Emphasis on:

- physical properties
- uncertainty quantification

Convective/Storm scale application

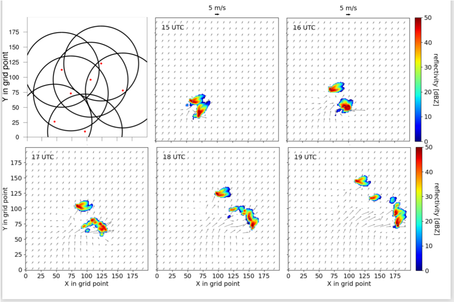


Met3D

- High resolution NWP models of atmosphere that incorporate our knowledge of the dynamics and physics.
- Accurate initial conditions are crucial for prediction, even more in the the future due to climate change and intensification of the water cycle.

Experimental setup

Idealized setup for radar DA



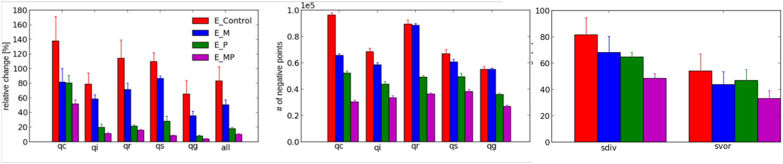
Zeng et al. 2021: Assimilating radar radial wind and reflectivity data in an idealized setup of the COSMO-KENDA system, Atmospheric Research, 249, 105282.

- COSMO model with a 2-km horizontal resolution
- Efficient Modular VOlume scanning RADar Operator (EMVORADO, Zeng et al., 2014, 2016)
- Both radial wind and reflectivity data are assimilated
- Ensemble size is 80
- Assimilated observations are perturbation of nature run with Gaussian noise with a standard deviation of 5.0 dBZ and 1.0 m/s

Results: Impact of the different constraints

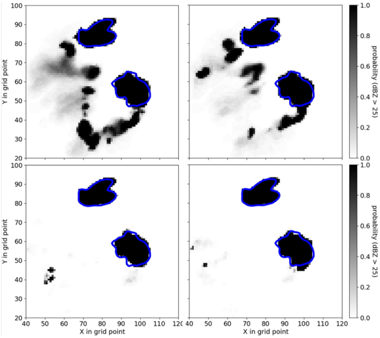
1. $E_{Control}$ Radar reflectivity and wind assimilated
2. E_M With mass constraint
3. E_P With positivity constraint (clear-air reflectivity data)
4. E_{MP} Both constraint

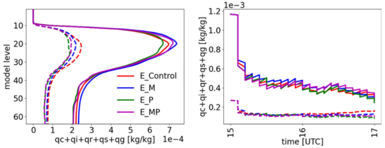
If clear-air reflectivity data are assimilated, a threshold value of 5 dBZ is set, that is, all reflectivity values smaller than 5 dBZ are set to 5 dBZ. If clear-air reflectivity data are not assimilated, all reflectivity values smaller than 5 dBZ are set to missing values.



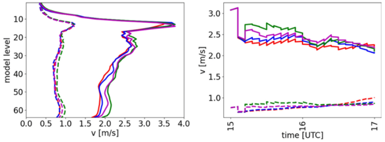
Janjic, T. and Y. Zeng, 2021, Weakly constrained LETKF for estimation of hydrometeor variables in convective-scale data assimilation, Geophysical Research Letters, 48, e2021GL094962.

Results: Impact of the different constraints



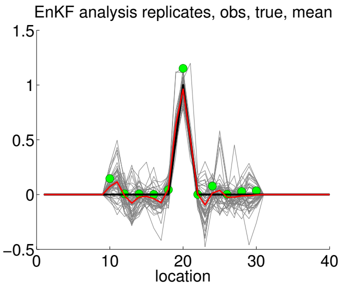


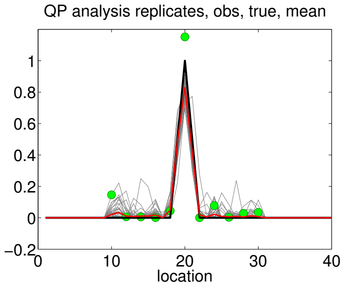
(a) (b)



(c) (d)

Physical properties and the analysis step





Analysis ensemble obtained with EnKF algorithm (left) and QPEns (right). Observations (green) are the true state plus log normal noise.

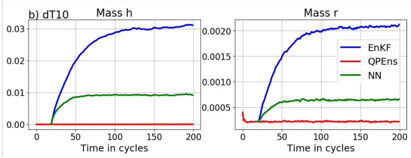
Janjic, T., D. McLaughlin, S. E. Cohn, M. Verlaan, 2014: Conservation of mass and preservation of positivity with ensemble-type Kalman filter algorithms, Mon. Wea. Rev., 142, No. 2, 755-773.

QPEns in high dimension

- EnKF —> QPEns after each cycle.

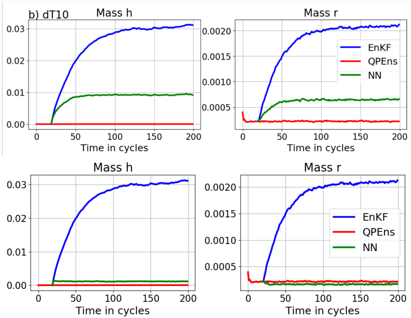
QPEns in high dimension

- EnKF —> QPEns after each cycle.
- With help of machine learning



QPEns in high dimension

- EnKF —> QPEns after each cycle.
- With help of machine learning



NN improves on QPEns result when physical constraint is introduced.

Ruckstuhl, Y., T. Janjić, S. Rasp, 2021: Training a convolutional neural network to conserve mass in data assimilation, Nonlin. Processes Geophys., 28, 111–119.

— Learning model parameters —

Augmented EnKF

Augment state vector x with parameters θ

$$\mathrm{w}_k^a = \left[\begin{array}{c} \mathrm{x}_k^a \\ \theta_k^a \end{array} \right]$$

Augmented EnKF

Augment state vector x with parameters θ

$$\mathrm{w}_k^a = \begin{bmatrix} \mathrm{x}_k^a \\ \theta_k^a \end{bmatrix}$$

Stochastic model for parameters

$$\theta_k^{b,i} = \theta_{k-1}^{a,i} + \mathrm{D}_{k-1}\mathrm{C}^{\frac{1}{2}}\eta^i$$

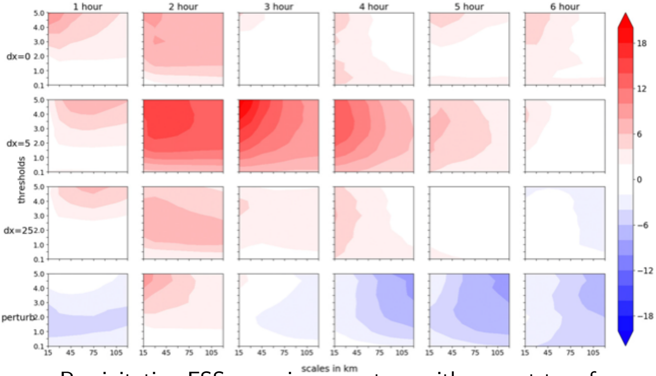
$\theta_{k-1}^{a,i}$ is the raw analysis value after applying the EnKF
$\theta_k^{b,i}$ the perturbed value that is passed to the model

D_{k-1} is a diagonal matrix that locally controls the ensemble spread

$\mathrm{C}^{\frac{1}{2}}$ is the error correlation matrix that specifies the correlations within parameter field

$\eta^i \sim \mathcal{N}(0, I)$ is the random realization of the stochastic model.

Example



Precipitation FSS score in percentage with respect to ref

Accounting for model error by allowing and estimating parameters (roughness length) can reduce state error.

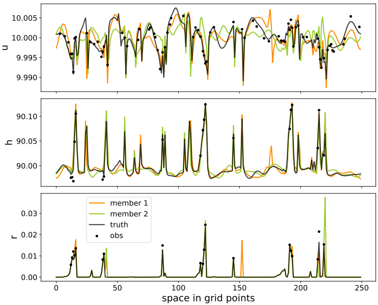
Ruckstuhl, Y. and T. Janjić, 2020, Combined State-Parameter Estimation with the LETKF for Convective-Scale Weather Forecasting, Mon. Wea. Rev., 148, 1607–1628.

Modified shallow water model

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{\partial(\phi + \gamma^2 r)}{\partial x} = \beta_u + D_u\frac{\partial^2 u}{\partial x^2}, \phi = \begin{cases} \phi_c & \text{if } h > h_c \\ gh & \text{otherwise}, \end{cases}$$

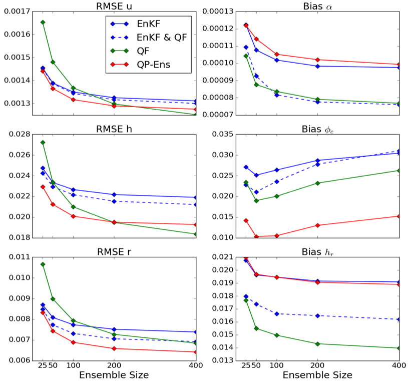
$$\frac{\partial r}{\partial t} + u\frac{\partial r}{\partial x} = D_r\frac{\partial^2 r}{\partial x^2} - \alpha r - \begin{cases} \delta\frac{\partial u}{\partial x}, & \text{if } h > h_r \text{ and } \frac{\partial u}{\partial x} < 0 \\ 0 & \text{otherwise}, \end{cases}$$

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = D_h\frac{\partial^2 h}{\partial x^2}.$$



Wuersch and Craig 2014: A simple dynamical model of cumulus convection for data assimilation research., Meteorol. Z., 23, 483-490.

Non-Gaussian aspects



QPEns (Janjic et al 2014), QF (Hodyss 2011,2012), EnKF (Evensen 2003).

Ruckstuhl Y. and T. Janjic, 2018, Parameter and state estimation with ensemble Kalman filter based approaches for convective scale data assimilation, Q. J. R. Meteorol. Soc., 144:712, 826–841.

NNs

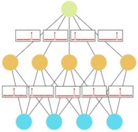
Input : state x_t
- wind u
- water level h
- rain r

$\longrightarrow$

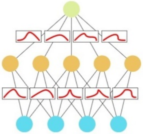
Neural Network
- fully connected
- feedforward

$\longrightarrow$

Output : model parameters θ_t
- rain removal rate α
- const. geopotential ϕ_c
- rain threshold h_r



Ensemble of Point Estimate Neural Networks
- train 15 neural networks
- 4 linear layers
- batch – norm + dropout
- activation : ReLU
- optimizer : Adam

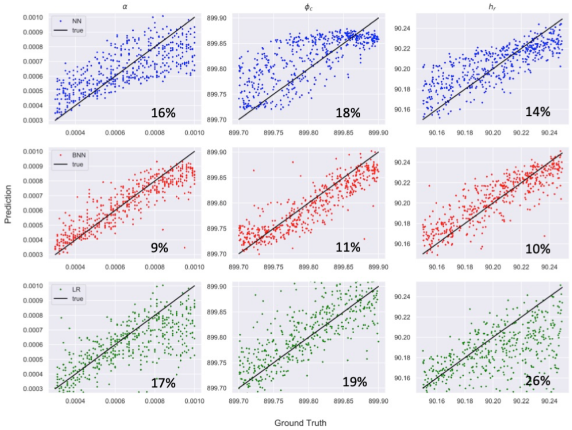


Bayesian Neural Network
- $N(\mu, \sigma)$ for each weight
- 4 linear layers
- batch – norm
- activation : LeakyReLU
- optimizer : SVI

(figure from Jospin et al. 2020)

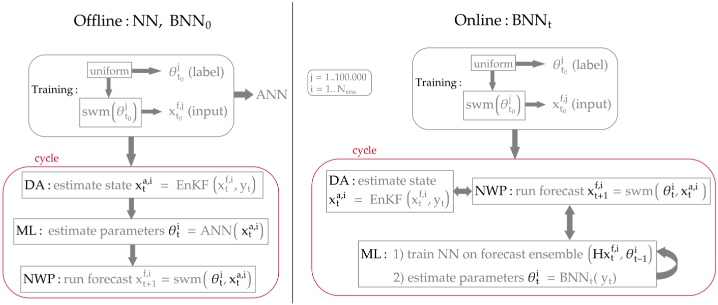
Trained on 100 000 model simulations using random parameter values from the uniform distributions.

NN accuracy



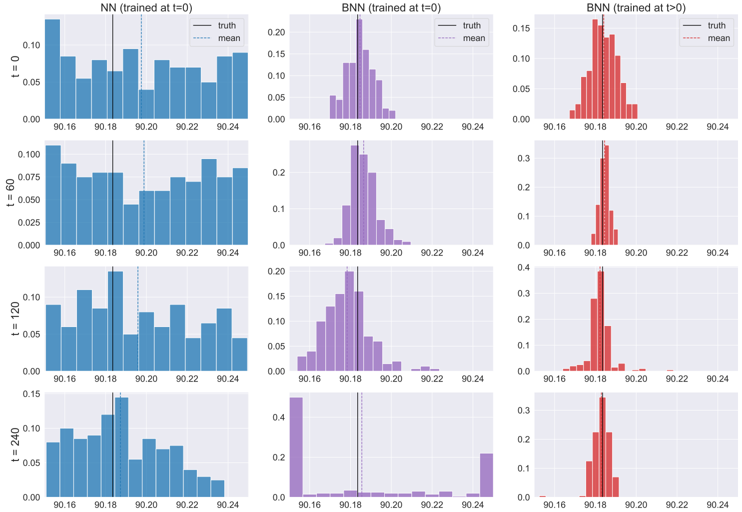
Output of NN (blue dots), BNN (red dots), and LR (green dots) against corresponding ground truths and ideal output (black lines) of 500 samples

DA+NNs for parameters



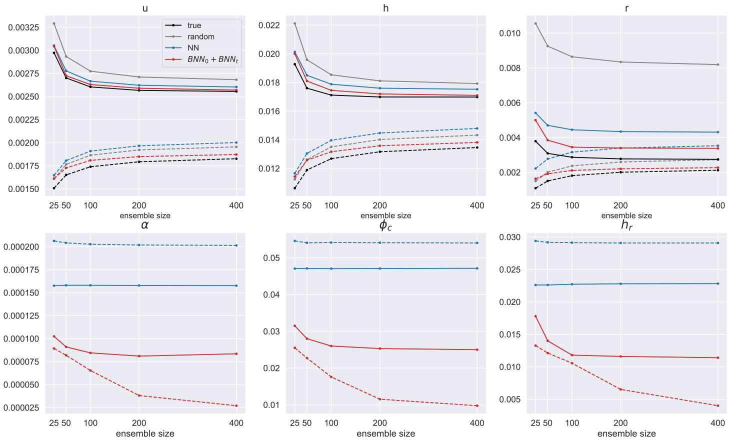
Legler and Janjic, 2022: Combining data assimilation and machine learning to estimate parameters of a convective-scale model. Q.J.R. Meteorol. Soc., 148, 860-874.

Combination with machine learning



Legler and Janjic, 2022: Combining data assimilation and machine learning to estimate parameters of a convective-scale model. Q.J.R. Meteorol. Soc., 148, 860-874.

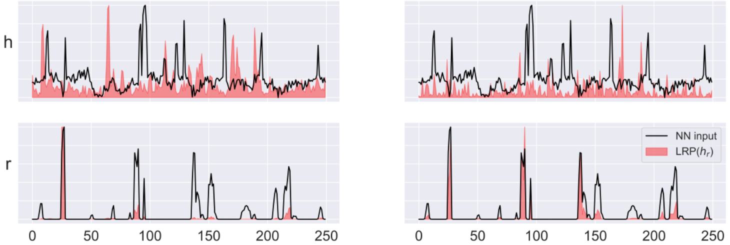
Sensitivity to Ensemble Size



Results are averaged over 100 individual experiments with different ground truth values.

True parameter is used in simulation (black), wrong (gray).

LRP

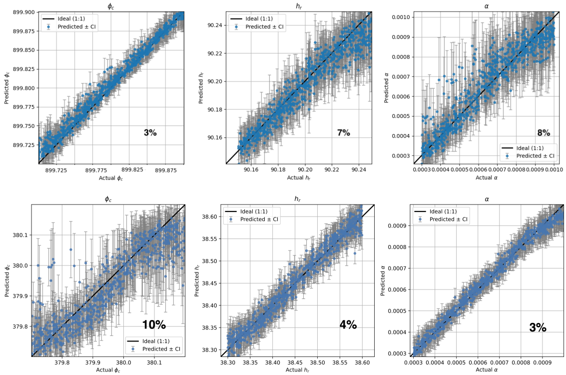


parameter	u	h	r
α	20%	57%	23%
ϕ_c	23%	61%	16%
h_r	21%	63%	16%

parameter	u	h	r
α	16%	59%	25%
ϕ_c	19%	44%	37%
h_r	21%	38%	41%

We follow Toms et al. 2020 to calculate Layer-wise relevance propagation (LRP) map for h_r in case Left: 3 parameters are estimated simultaneously. Right: LRP map when only h_r is estimated.

Uncertainty quantification



Offline prediction of parameters in different physical settings (rows). Conformal prediction for 90% confidence intervals for representation of uncertainty (figures from Maryam Ramezani Ziarani).

Conclusion

- We can learn from observations, both the state and parameters of geophysical model.
- Proper specification of initial conditions is crucial for prediction with any model.
- The state estimation can be improved by imposing conservation laws/physical constraints both for ML and DA.
- For parameters this is not the case.

Conclusion

- We can learn from observations, both the state and parameters of geophysical model.
- Proper specification of initial conditions is crucial for prediction with any model.
- The state estimation can be improved by imposing conservation laws/physical constraints both for ML and DA.
- For parameters this is not the case.

- Great potential in combination of ML with DA using powers of each method!