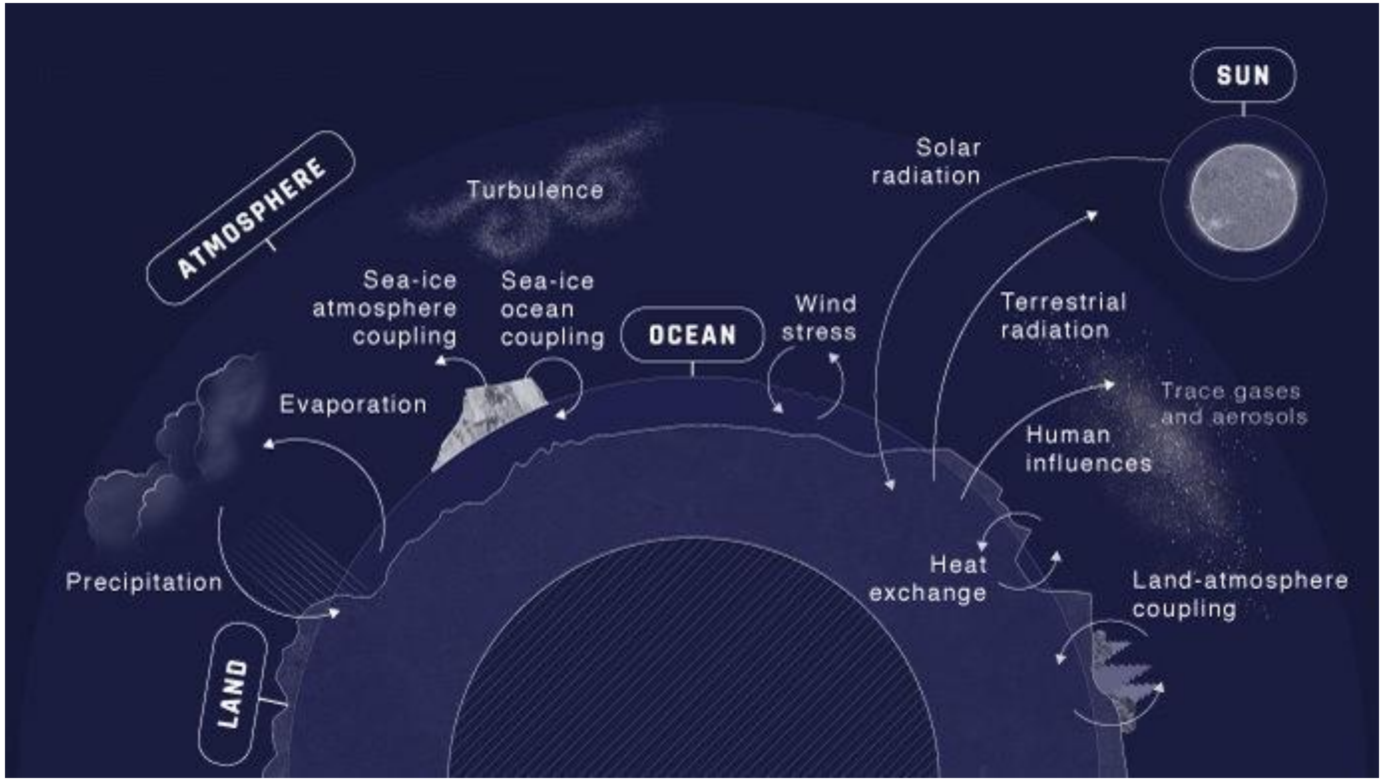
Bias aware data assimilation

Patrick Laloyaux (with the help of many colleagues)

ECMWF

Modelling the Earth system

ECMWF has been developing a comprehensive Earth system model which forms the basis for all our data assimilation and forecasting activities. In this talk we concentrate on the atmospheric component (IFS)



46r1 Improvements in the convection and radiation schemes

47r1 Quintic vertical interpolation in semi-Lagrangian advection

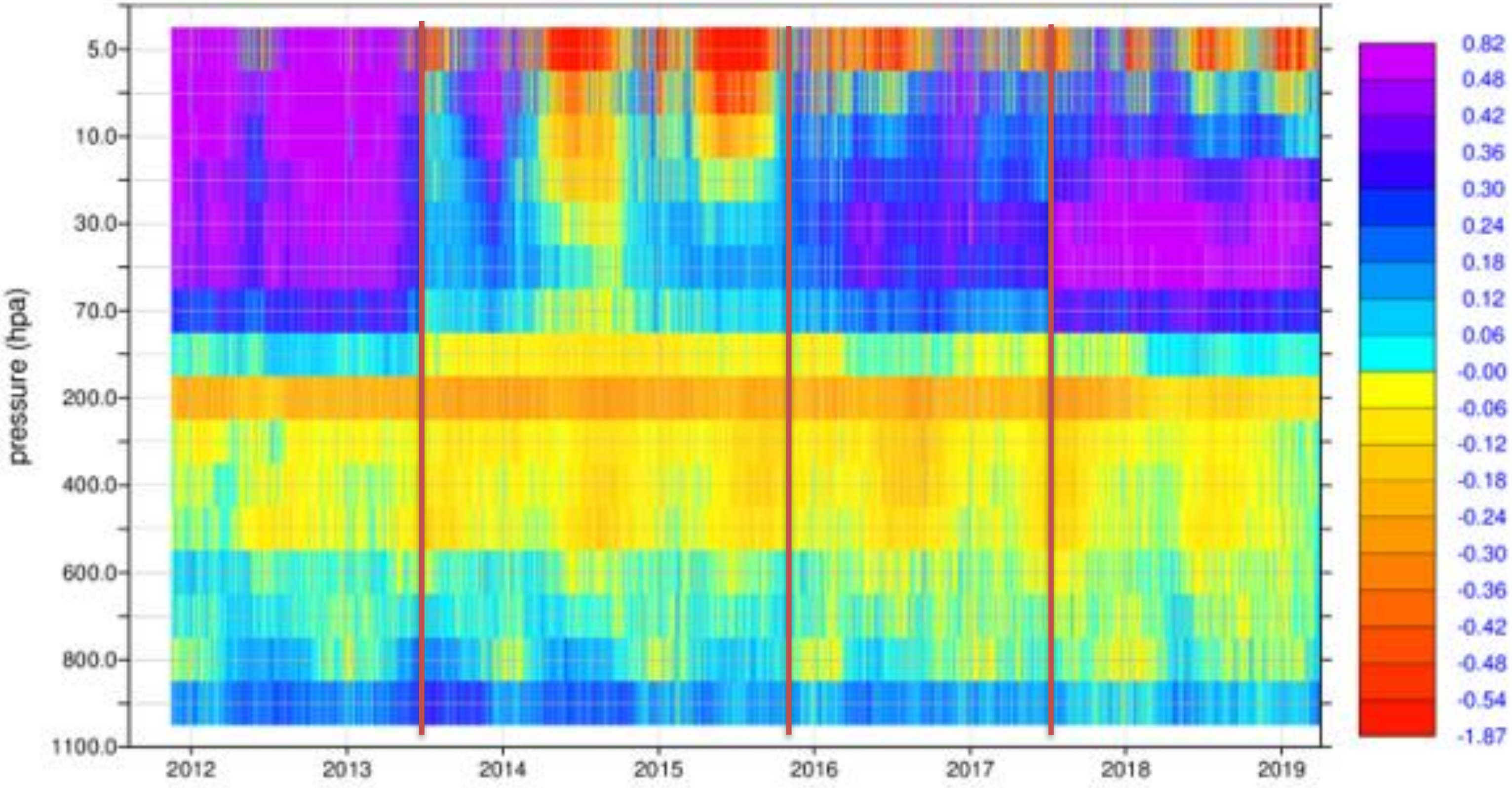
47r3 Moist physics upgrade

48r1 Hybrid linear ozone, semi-lagrangian vertical filter and new solar spectrum?

One of the best models in the world, but it still contains some residual biases that must be taken into account in DA

Monitoring the quality of the atmospheric model for DA

Difference between the 12-hour model trajectory with reference observations (radiosondes)



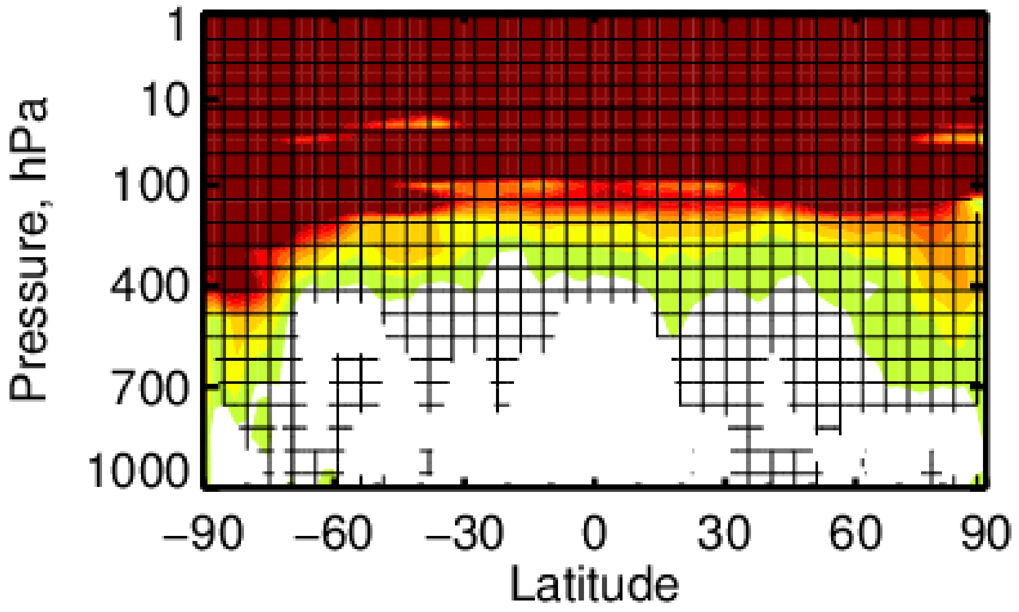
Systematic error when the atmospheric model is integrated over 12 hours

➔ Cold bias in the mid/lower stratosphere (>0.5C)

➔ Warm bias in the upper stratosphere (>0.5C)

Monitoring the quality of the atmospheric model for DA

Stratospheric biases can travel through the atmosphere and impact the troposphere

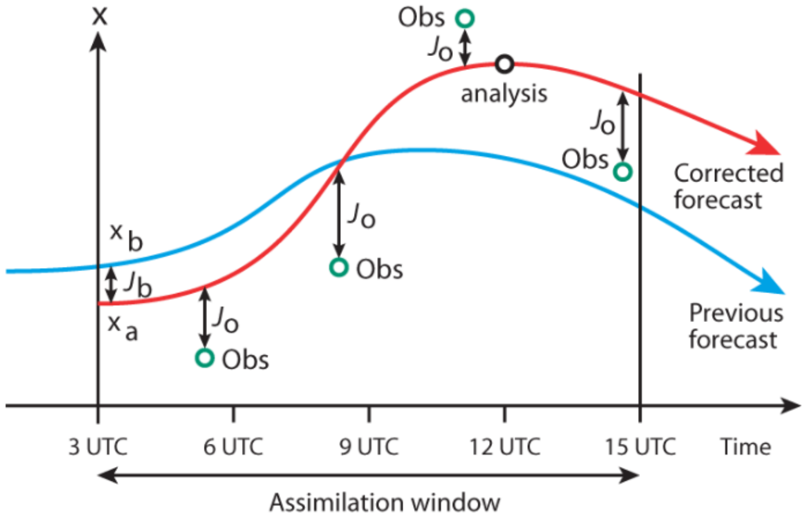


RMSE change when stratospheric observations (radiosondes, RO, infrared, microwave, …) are withheld.

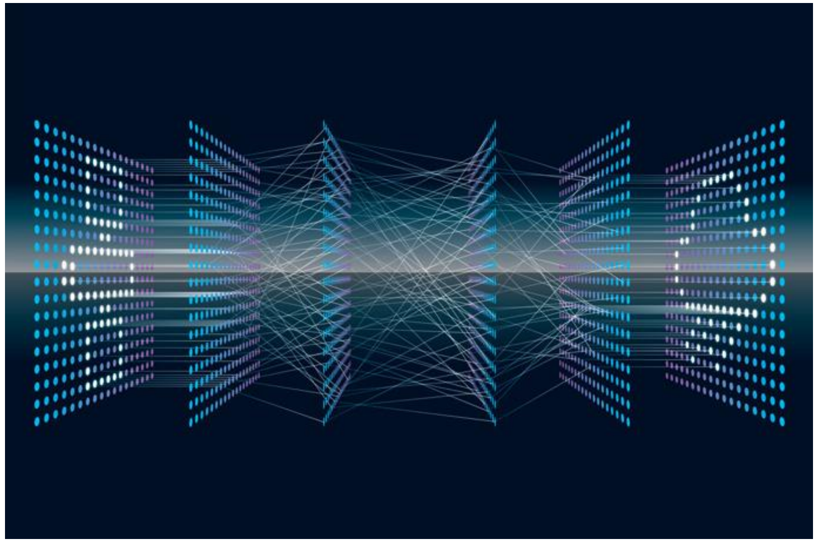
Verified against operations between 25/01/2020 and 25/03/2020

What is the best way to handle model biases? Data Assimilation approach? Machine Learning approach? What are the links between both of them?

- Data Assimilation approach

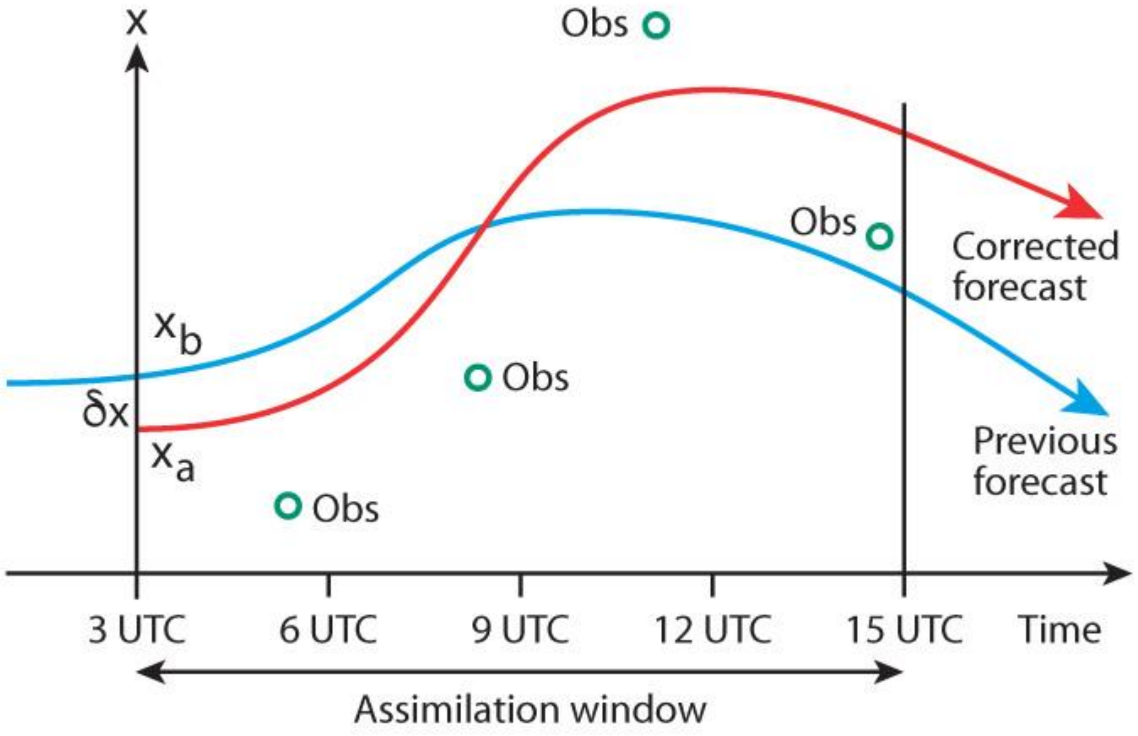


- Machine Learning approach



Standard 4D-Var formulation

4D-Var is one of the most popular algorithm to find the optimal initial state by minimising the discrepancies with the prior estimate and the observations



Model's equation

$$x_k = \mathcal{M}_k(x_{k-1})$$

4D-Var cost function

$$\begin{aligned} J(x_0) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\ &+ \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \end{aligned}$$

➔ Standard formulation assumes that the model is perfect

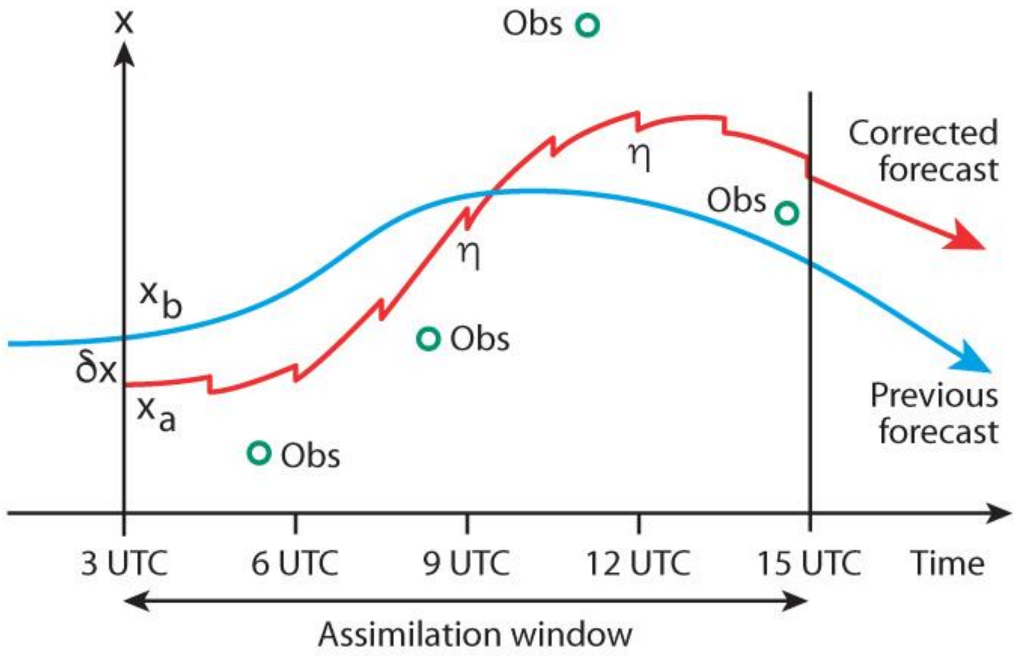
➔ A model trajectory is entirely determined by its initial condition

Weak-constraint 4D-Var formulation

We assume that the model is not perfect, adding an error term η in the model equation

$$x_k = \mathcal{M}_k(x_{k-1}) + \eta \quad \text{for } k = 1, 2, \cdots, K$$

The model error estimate η contains 3 physical fields (temperature, vorticity and divergence)



➔ Introduce additional degrees of freedom to fit background and observations

➔ Constant model error forcing over the assimilation window

➔ A model trajectory is entirely determined by its initial condition and the model error forcing

➔ Concept of scale separation introduced between background and model errors

Laloyaux et al., Exploring the potential and limitations of weak-constraint 4D-Var, 2020

Weak-constraint 4D-Var formulation

We assume that the model is not perfect, adding an error term η in the model equation

$$x_k = \mathcal{M}_k(x_{k-1}) + \eta \quad \text{for } k = 1, 2, \cdots, K$$

The model error estimate η contains 3 physical fields (temperature, vorticity and divergence)

$$\begin{aligned} J(x_0, \eta) &= \frac{1}{2}(x_0 - x_b)^T \mathbf{B}^{-1}(x_0 - x_b) \\ &+ \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k)]^T \mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\ &+ \frac{1}{2}(\eta - \eta_b)^T \mathbf{Q}^{-1}(\eta - \eta_b) \end{aligned}$$

Model initial condition

Model bias correction

➔ Introduce additional degrees of freedom to fit background and observations

➔ Constant model error forcing over the assimilation window

➔ A model trajectory is entirely determined by its initial condition and the model error forcing

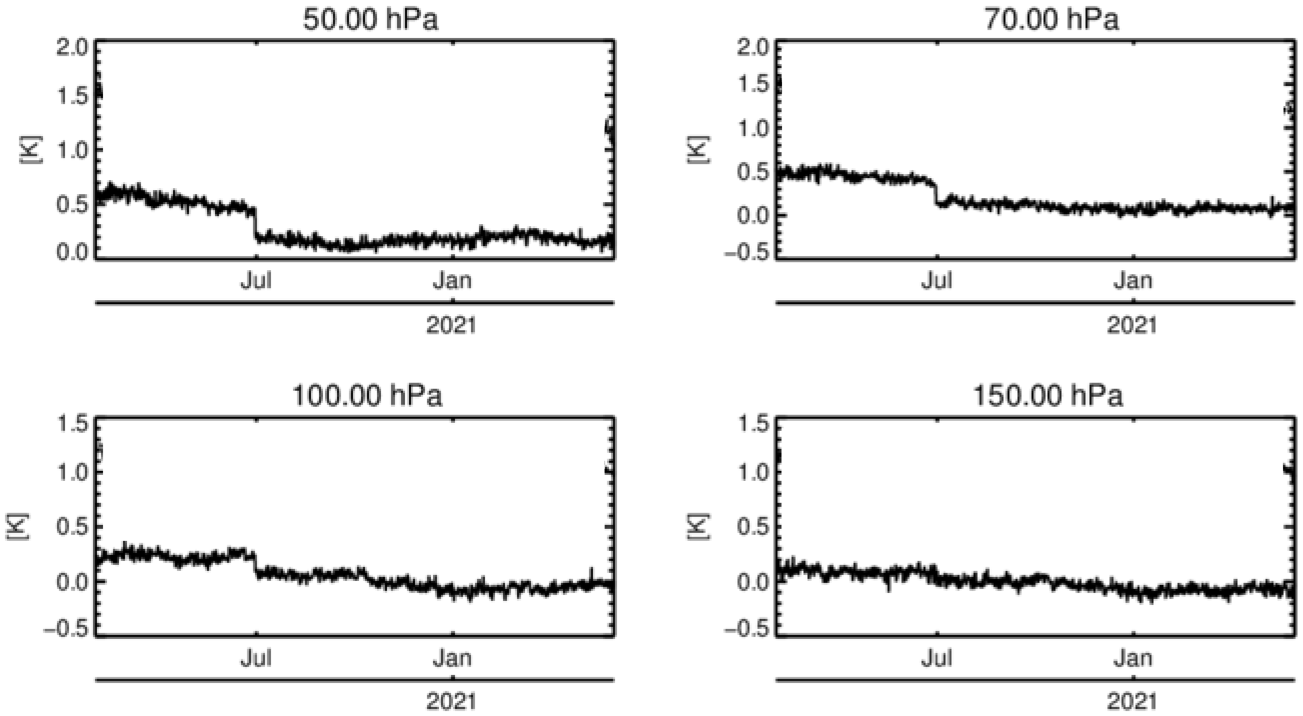
➔ Concept of scale separation introduced between background and model errors

Laloyaux et al., Exploring the potential and limitations of weak-constraint 4D-Var, 2020

Weak-constraint 4D-Var formulation in operations

This technique is used operationally since 30 June 2020 to correct the stratospheric biases in the HRES system

Mean first-guess departure with respect to temperature measurements from radiosondes



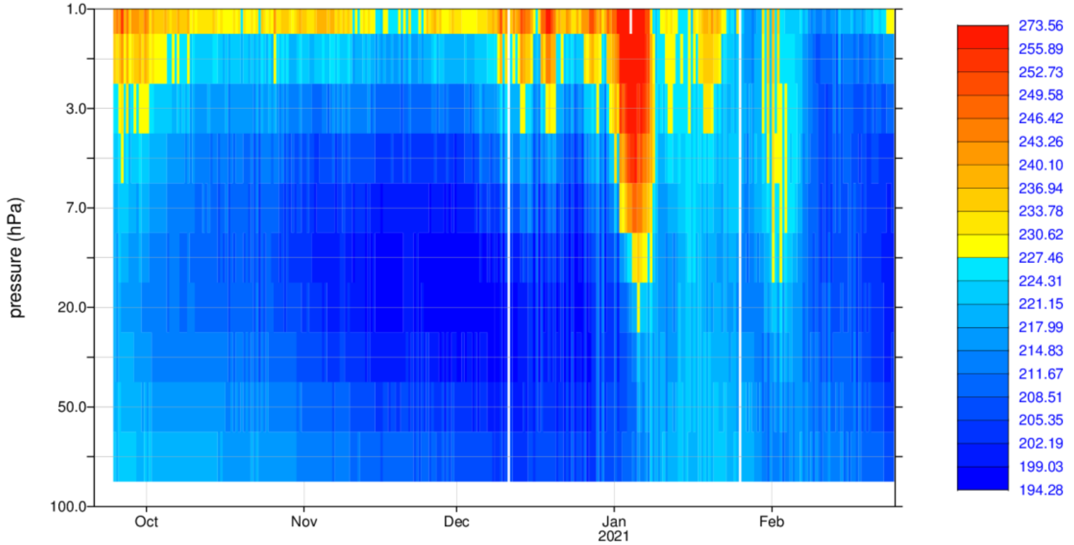
➔ bias reduced up to 50%

➔ weak-constraint 4D-Var in EDA will be implemented in 47r3 (similar reduction in biases)

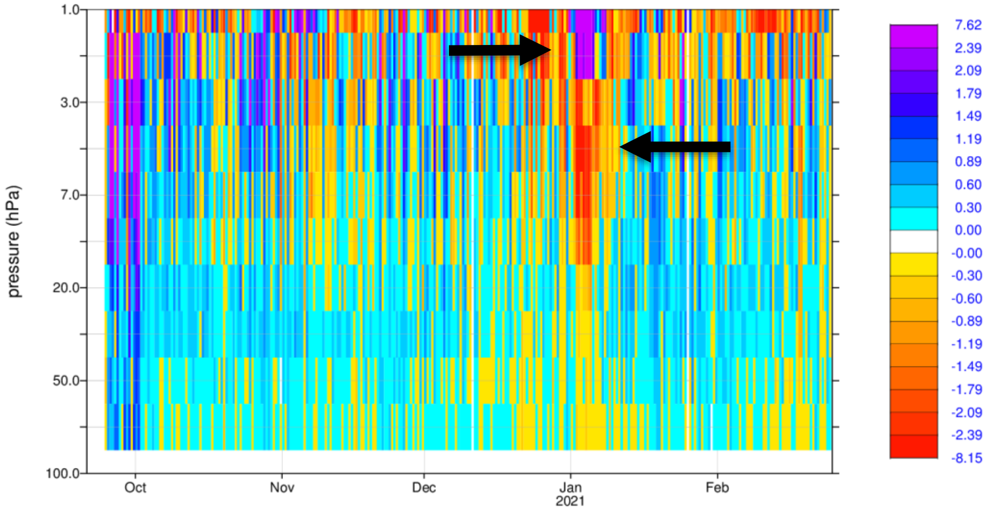
P. Laloyaux et al., Towards an unbiased stratospheric analysis, 2020

Extreme events in the stratosphere (SSW)

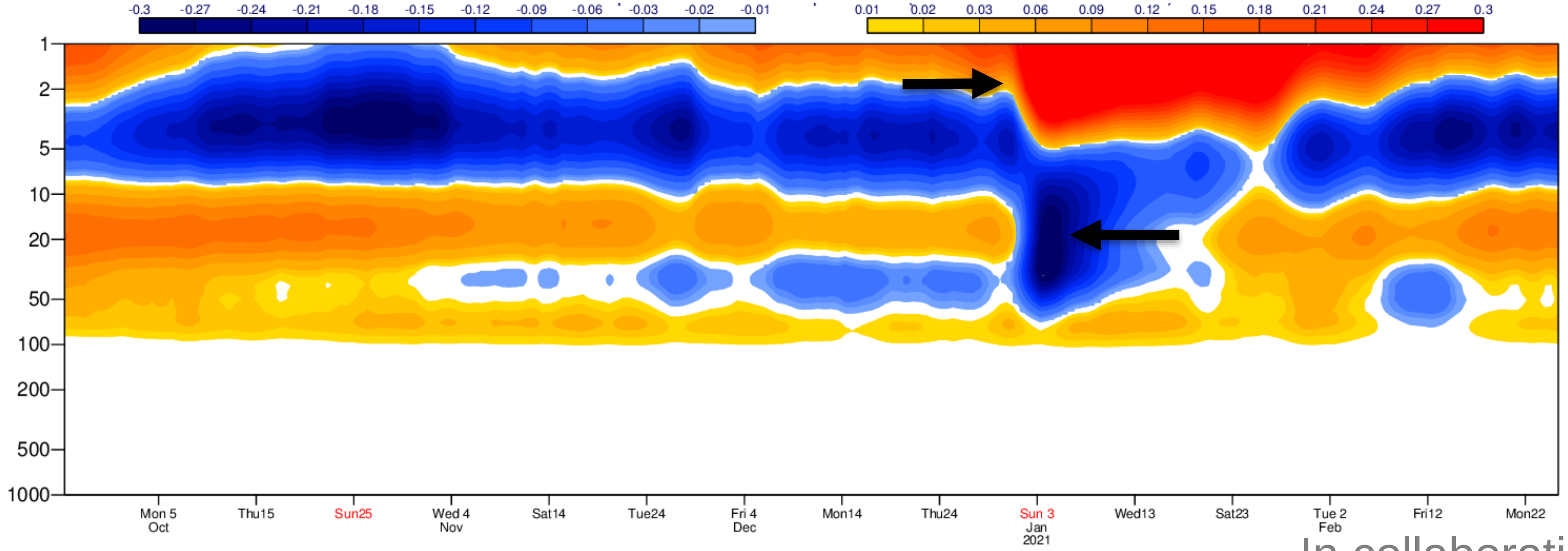
Rapid warming in the stratosphere (70N-90N)



Model bias dipole



Captured and partially corrected by weak-constraint 4D-Var



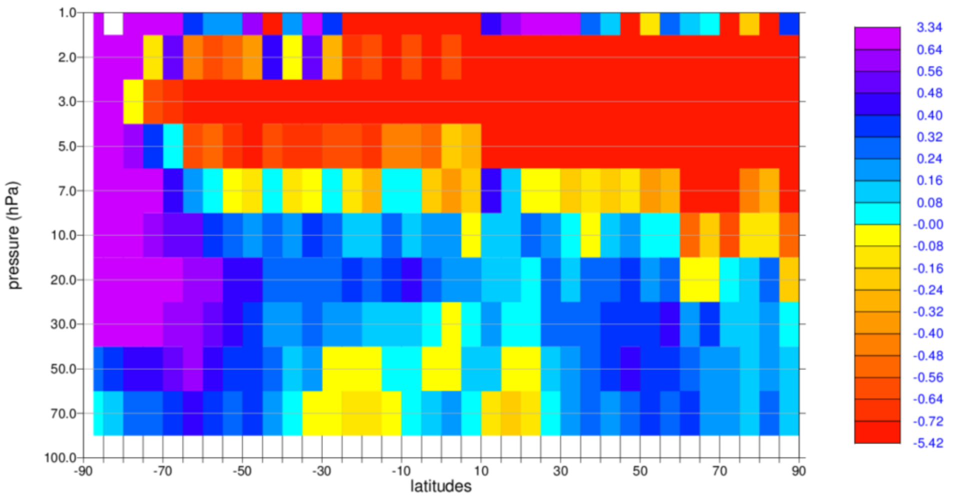
Quick and consistent response to an unexpected extreme situation

In collaboration with Inna Polichtchouk

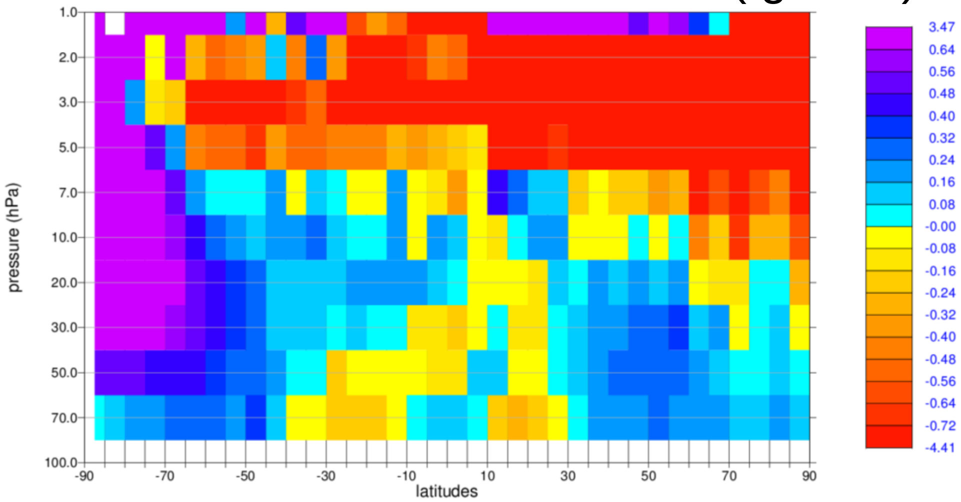
Model upgrade at ECMWF

The IFS model is upgraded on a regular basis. The bias of the new model version may be different and needs to be estimated

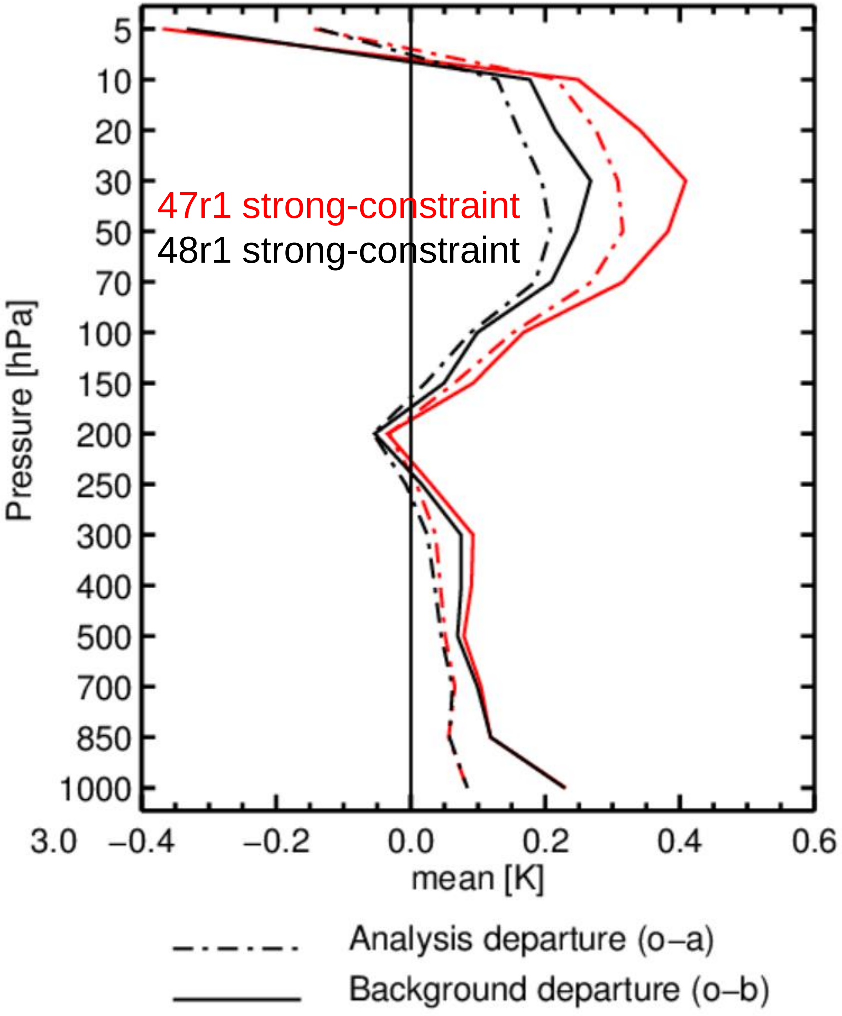
Model bias estimation in 47r1 (fg - RO)



Model bias estimation in 48r1 (fg - RO)



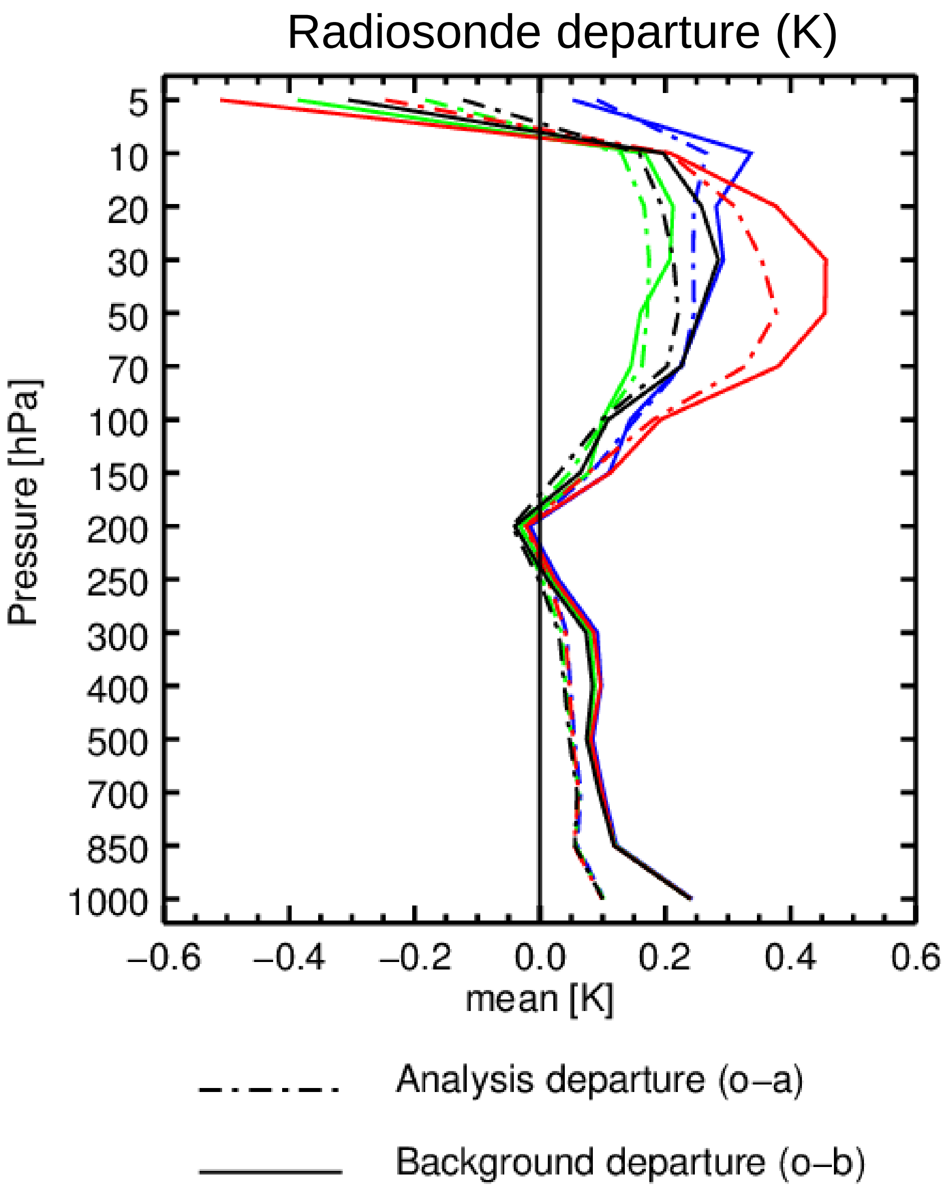
Radiosonde departure (K)

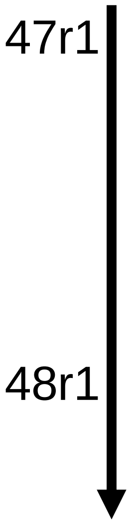


In collaboration with Robin Hogan

Model upgrade at ECMWF

The ECMWF model is upgraded on a regular basis. The bias of the new model is different and need to be estimated





Weak-constraint 4D-Var reduced the model bias

- strong-constraint
- weak-constraint

Weak-constraint 4D-Var is all set

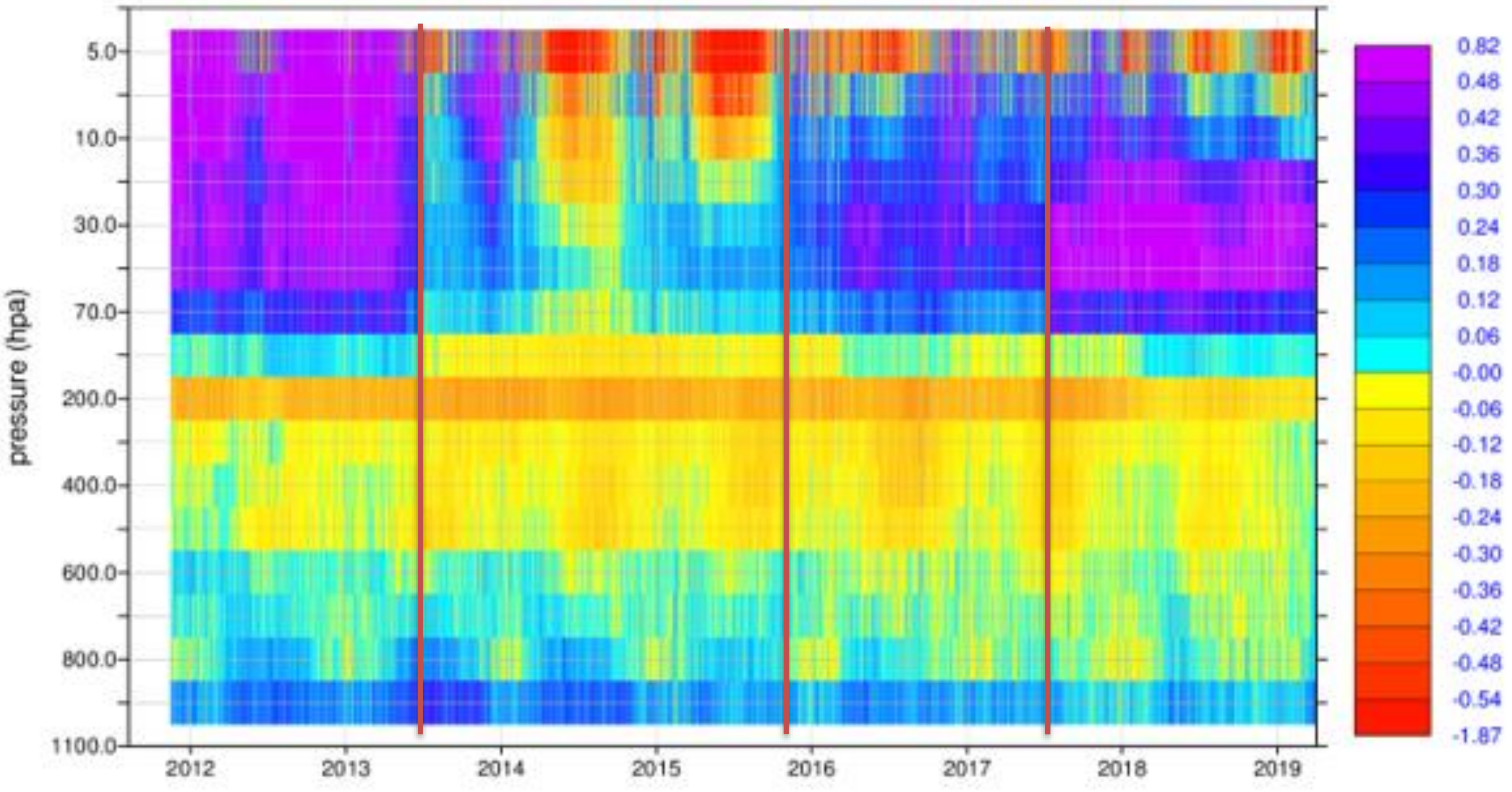
- strong-constraint (a possible candidate for 48r1)
- weak-constraint

No need to retune weak-constraint 4D-Var in 48r1 as it learns the new model bias on its own (only a careful monitoring).

Weak-constraint 4D-Var does not prevent model developments. They work hand in hand!

Monitoring the quality of the atmospheric model for DA

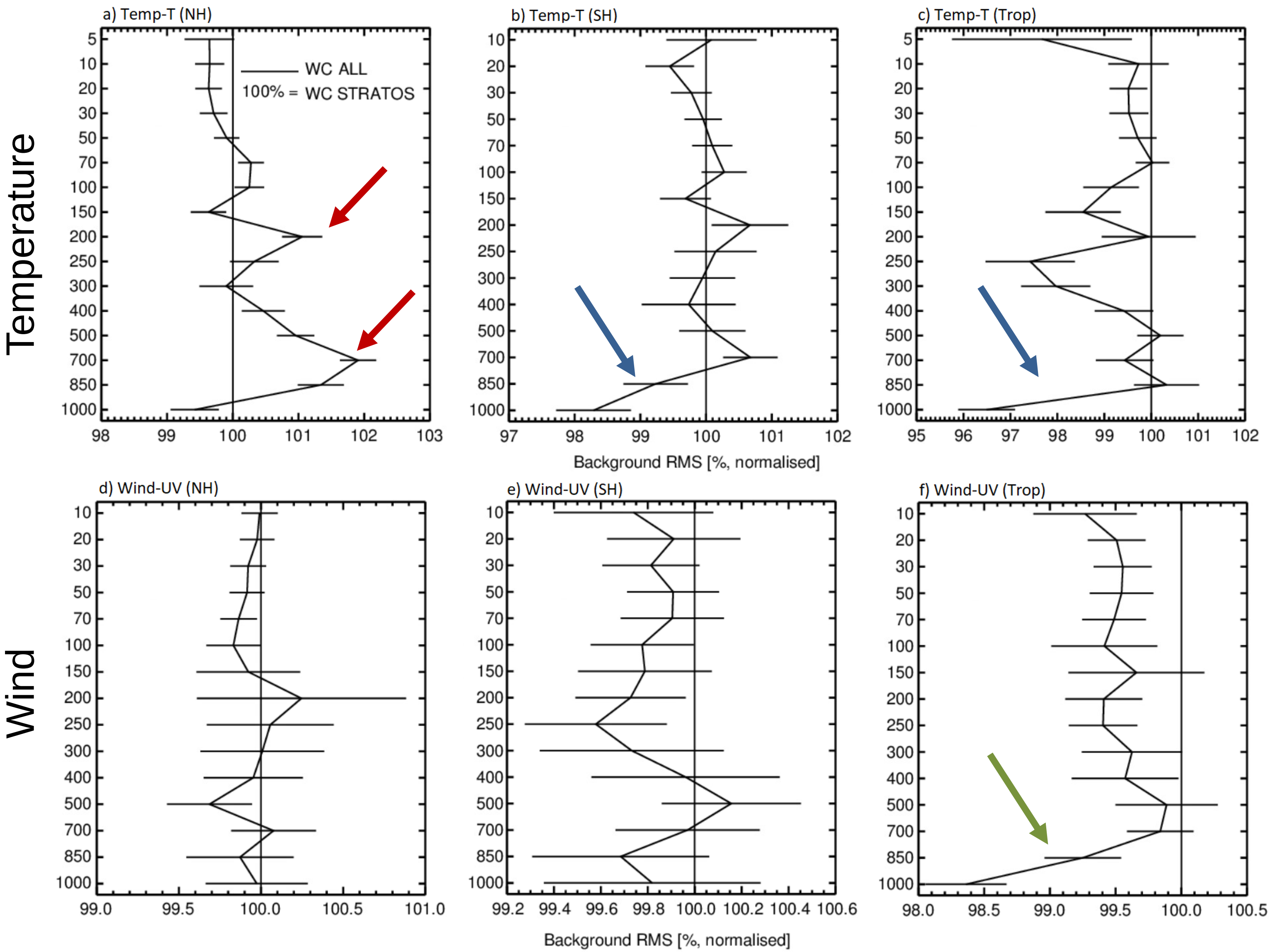
Difference between the 12-hour model trajectory with reference observations (radiosondes)



Systematic error when the atmospheric model is integrated over 12 hours

➔ The troposphere contains also some biases (temperature, wind, humidity, …)

Challenges to extend to the troposphere



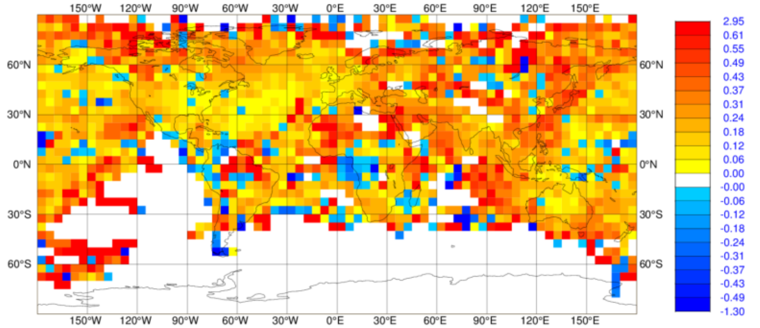
25/08/2019 to 25/10/2019

Extension of weak-constraint 4D-Var to the troposphere

Significant wind improvements in Tropics

Significant temperature improvements in Tropics and SH

Issue in the NH at 200hPa and 700hPa



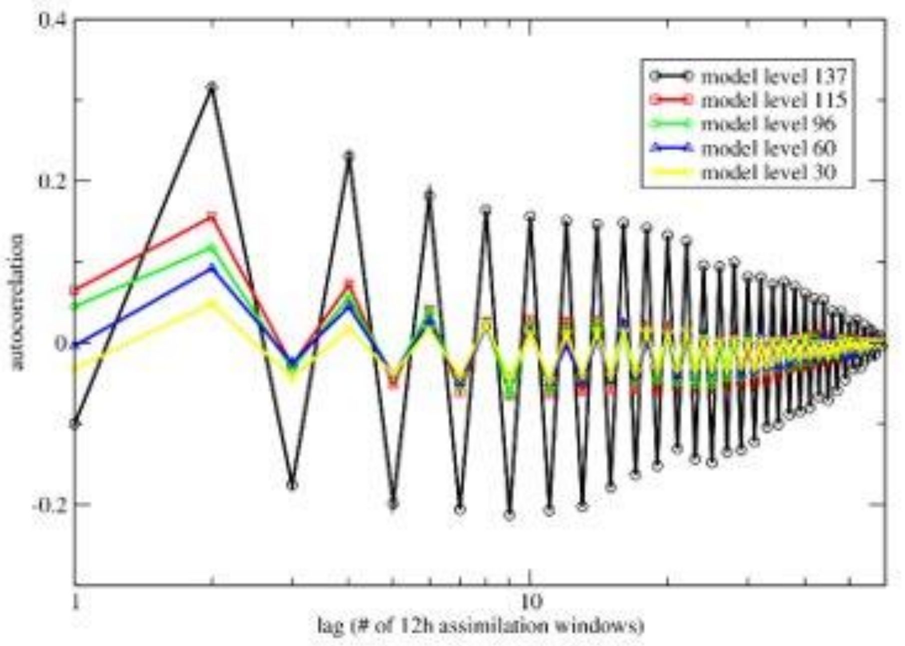
Aircraft background departures at 200hPa show a residual bias that is absorbed by WC-4DVar

Challenges to extend to the troposphere

The current version of WC-4D-Var has two crucial simplifications in the operational implementation:

1) constant error forcing during the assimilation window

2) persisting the model error estimate from one assimilation window to the following one.



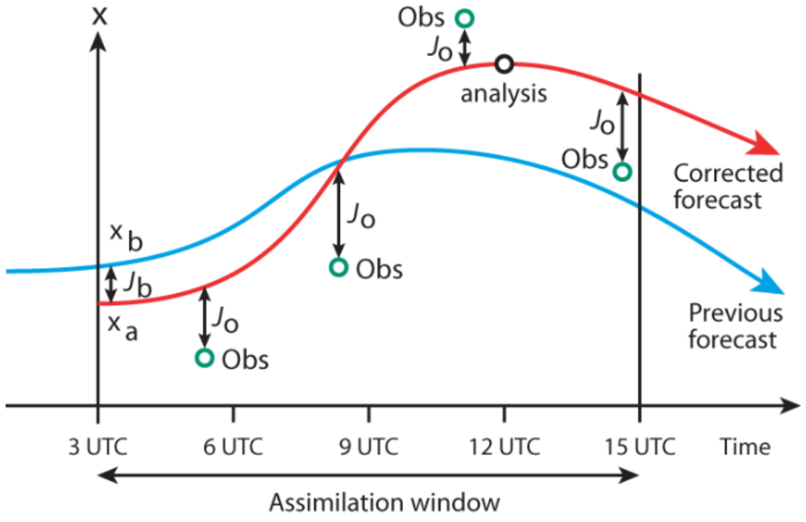
Diagnosed model error (analysis increment) time correlations reveal the presence of a significant diurnal cycle, most notably in the boundary layer

Humidity bias has to be taken into account

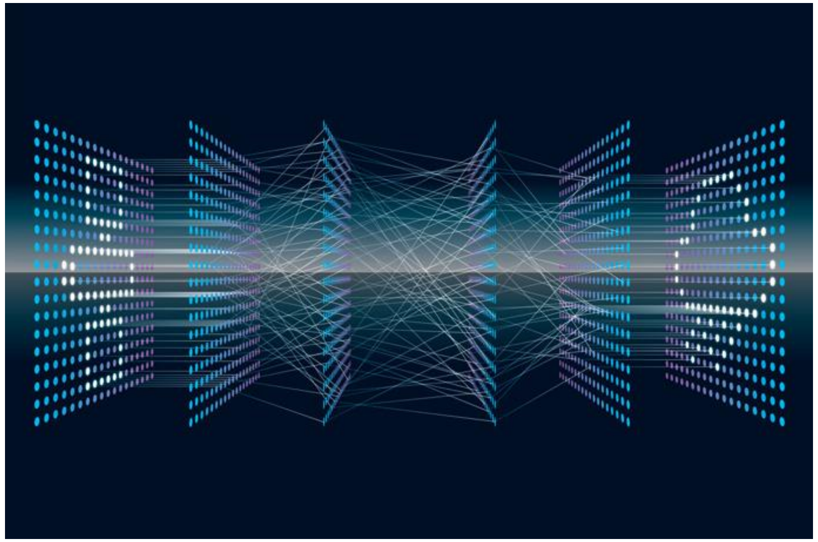
$$x_k = \mathcal{M}_k(x_{k-1}) + \eta \quad \text{for } k = 1, 2, \cdots, K$$

M. Bonavita, Exploring the structure of time-correlated model errors in the ECMWF data assimilation system, 2021

- Data Assimilation approach

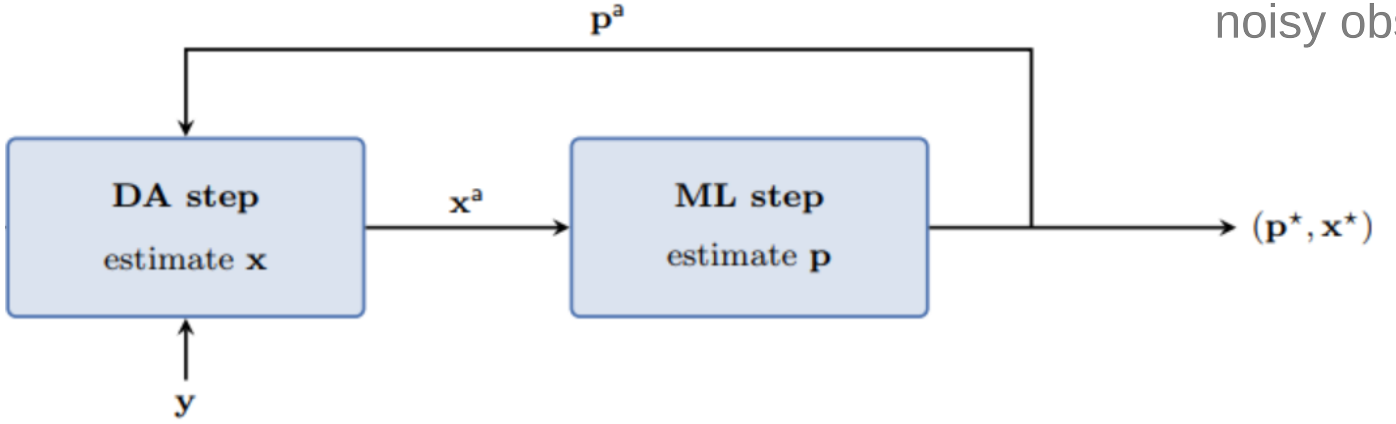


- Machine Learning approach

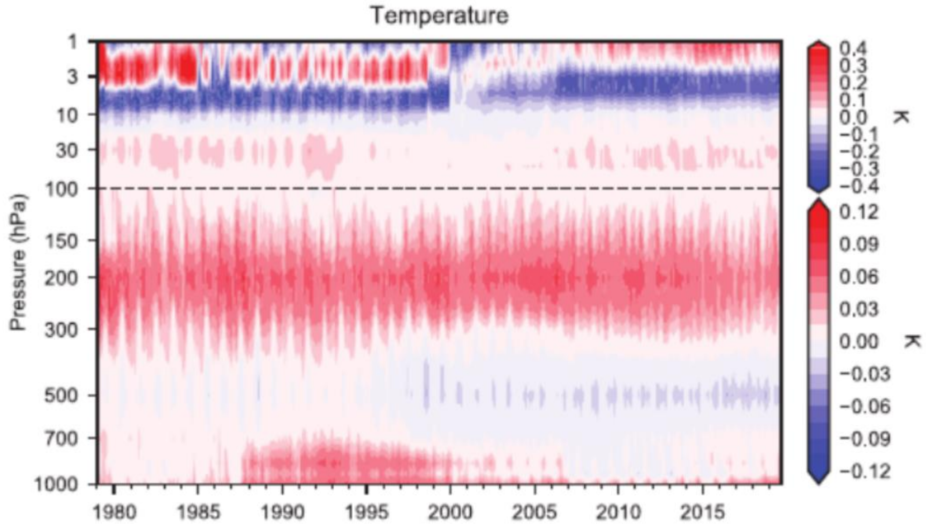


How to estimate model bias with a Neural Network

DA/NN framework



J. Brajar et al., Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations, 2020



- Construct a dataset in **model space** (e.g. increments)
- Train NN on the dataset
- Correct the model resolvent

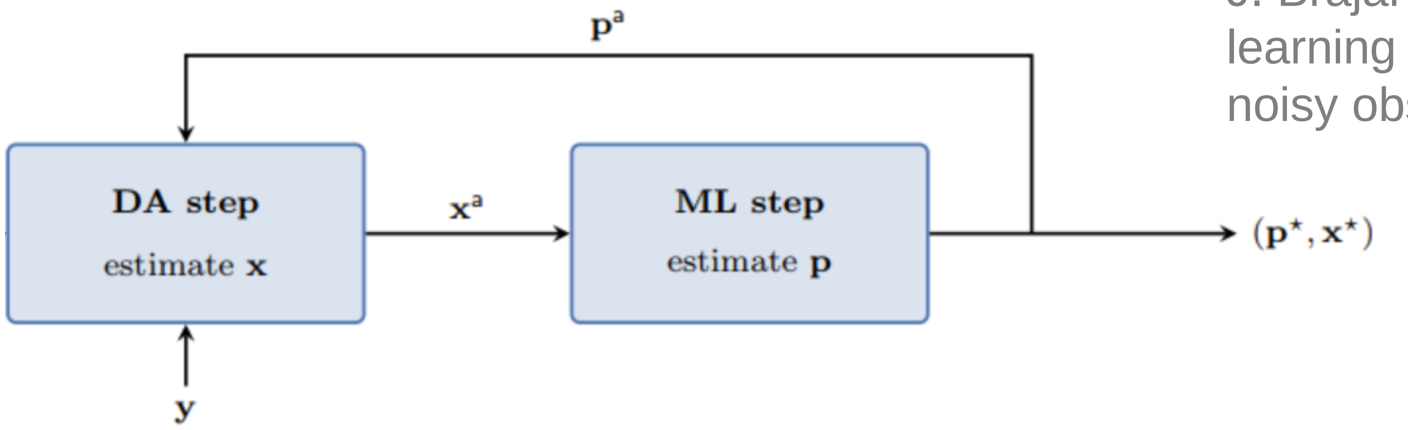
$$\mathcal{M}_k(\mathbf{p},\mathbf{x}) \triangleq \mathcal{M}_k^{\circ}(\mathbf{x}) + \mathcal{M}_k^{\mathrm{ml}}(\mathbf{p},\mathbf{x})$$

Timeseries of increments can highlight model biases but has some limitations as well

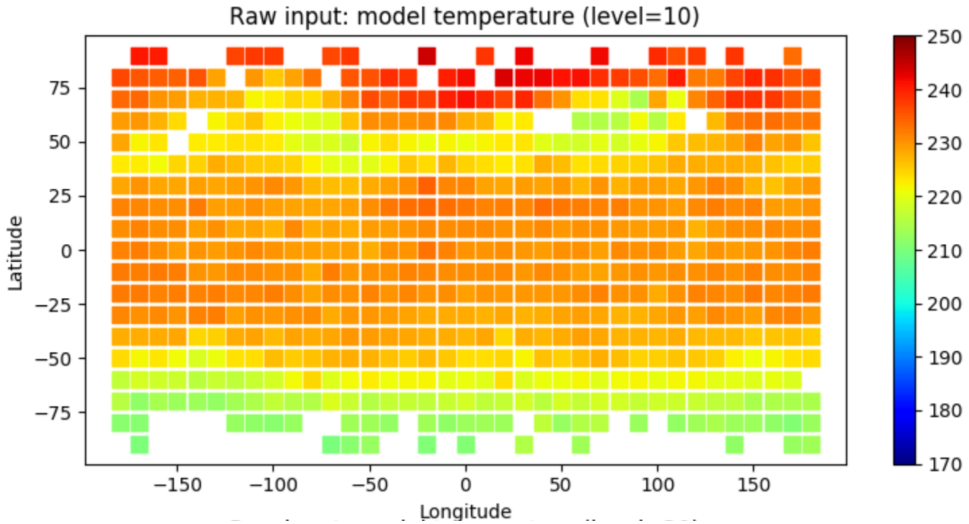
A. Farchi et al., Using machine learning to correct model error in data assimilation and forecast applications, 2020
M. Bonavita et al., Machine Learning for Model Error Inference and Correction, 2020

How to estimate model bias with a Neural Network

DA/NN framework



J. Brajar et al., Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations, 2020



- Construct a dataset in **observation space** (e.g. departure)
- Train NN on the dataset
- Correct the model resolvent

$$\mathcal{M}_k(\mathbf{p},\mathbf{x}) \triangleq \mathcal{M}_k^{\circ}(\mathbf{x}) + \mathcal{M}_k^{\mathrm{ml}}(\mathbf{p},\mathbf{x})$$

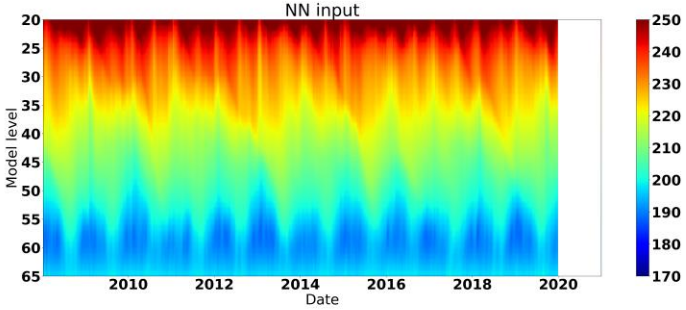
Temperature bias estimated from RO observations. The atmospheric state is never fully observed in NWP

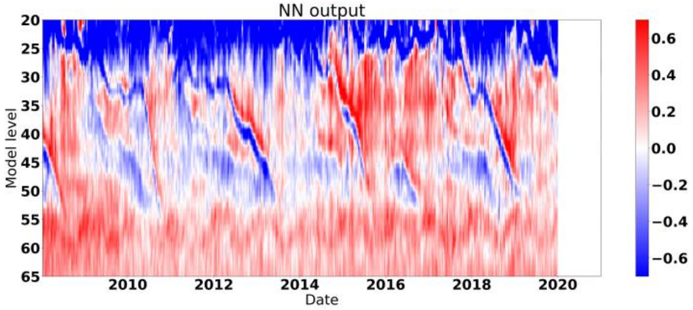
➔ average measurements on a 10-degree grid every 10 days

➔ interpolate to fill the gaps

How to estimate model bias with a Neural Network

12 years of ERA5: first-guess and departures with RO temperature retrievals (250 millions of RO observations)



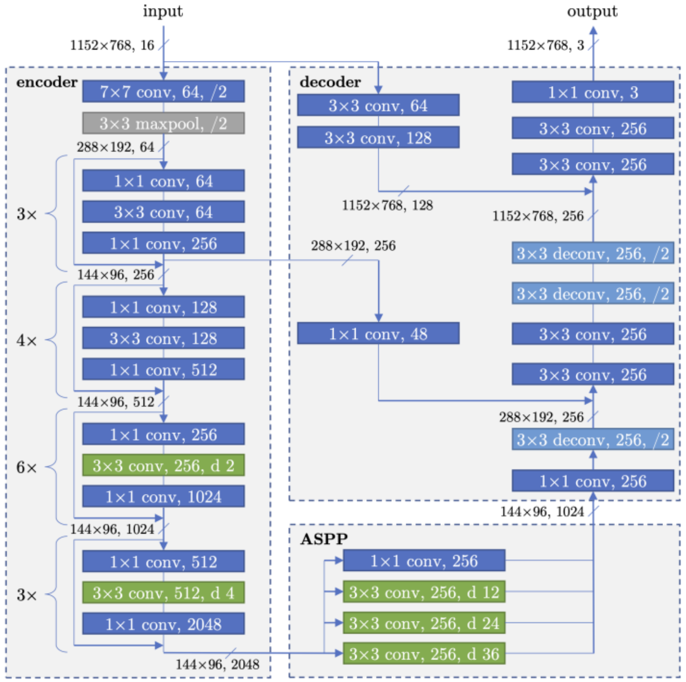


Dataset size

- input, output: 19x37x45 (31635)
- training set: 2008-2018 (2300 samples)

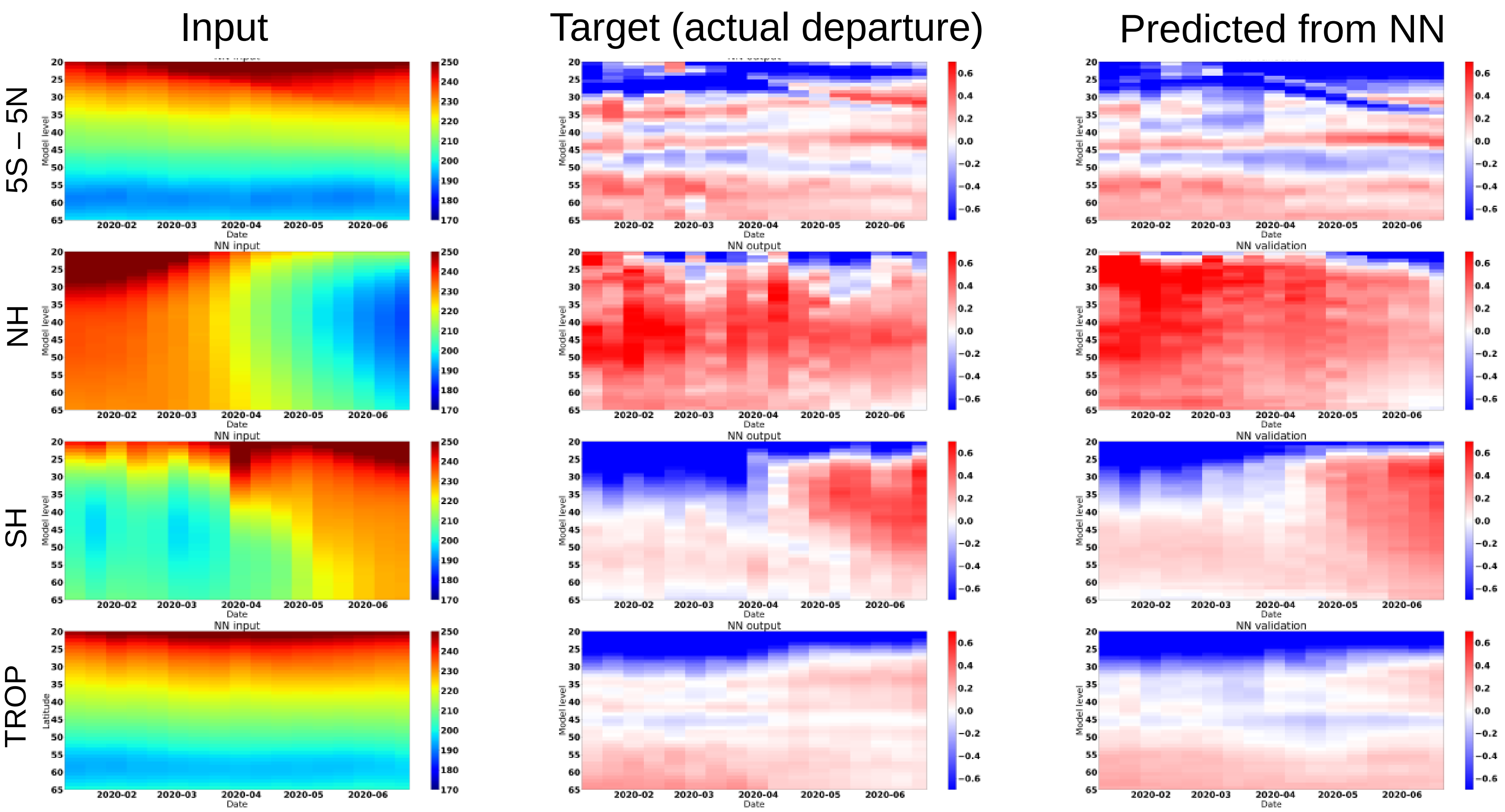
Convolutional neural network (CNN) are the best to learn computer-vision task.

The usual 3 channels (RGB) have been replaced by 45 channels (vertical levels in the stratosphere)



Kurth et al., doi:10.5555/3291656.3291724

Results from the NVIDIA CNN

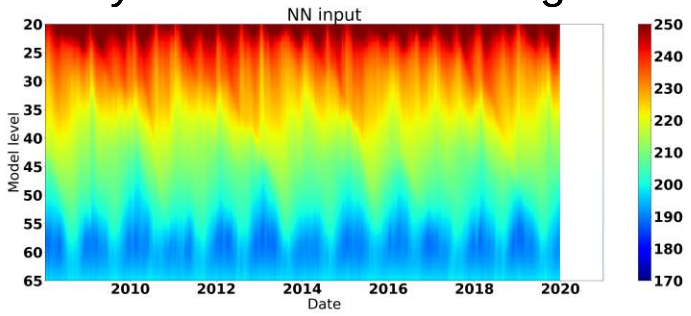


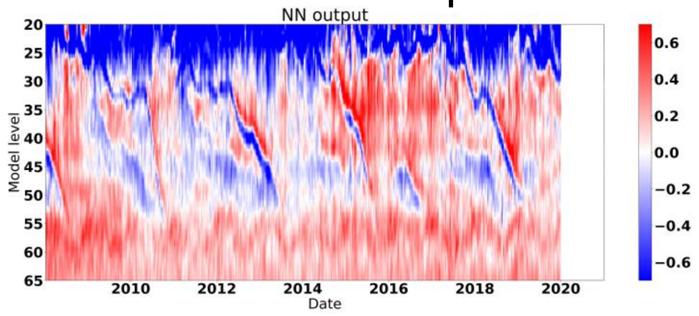
NN results are really good but it required a large dataset for training (2008-2018)

How to estimate model bias with a Neural Network

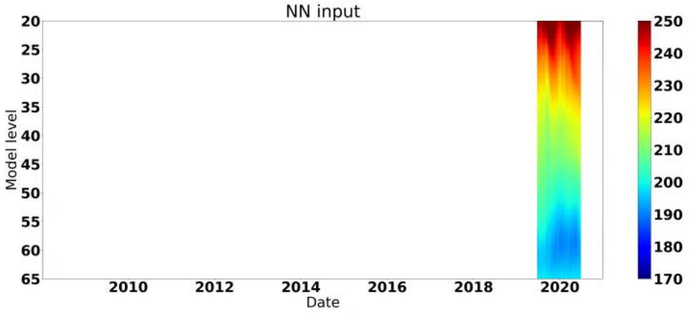
The ultimate goal is to learn the model bias from the latest IFS cycle: NN is retrained on a new smaller dataset (running 4D-Var is expensive)

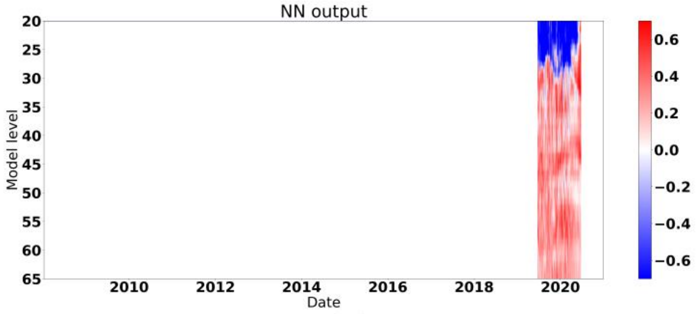
12 years of ERA5: first-guess and departures with RO temperature retrievals





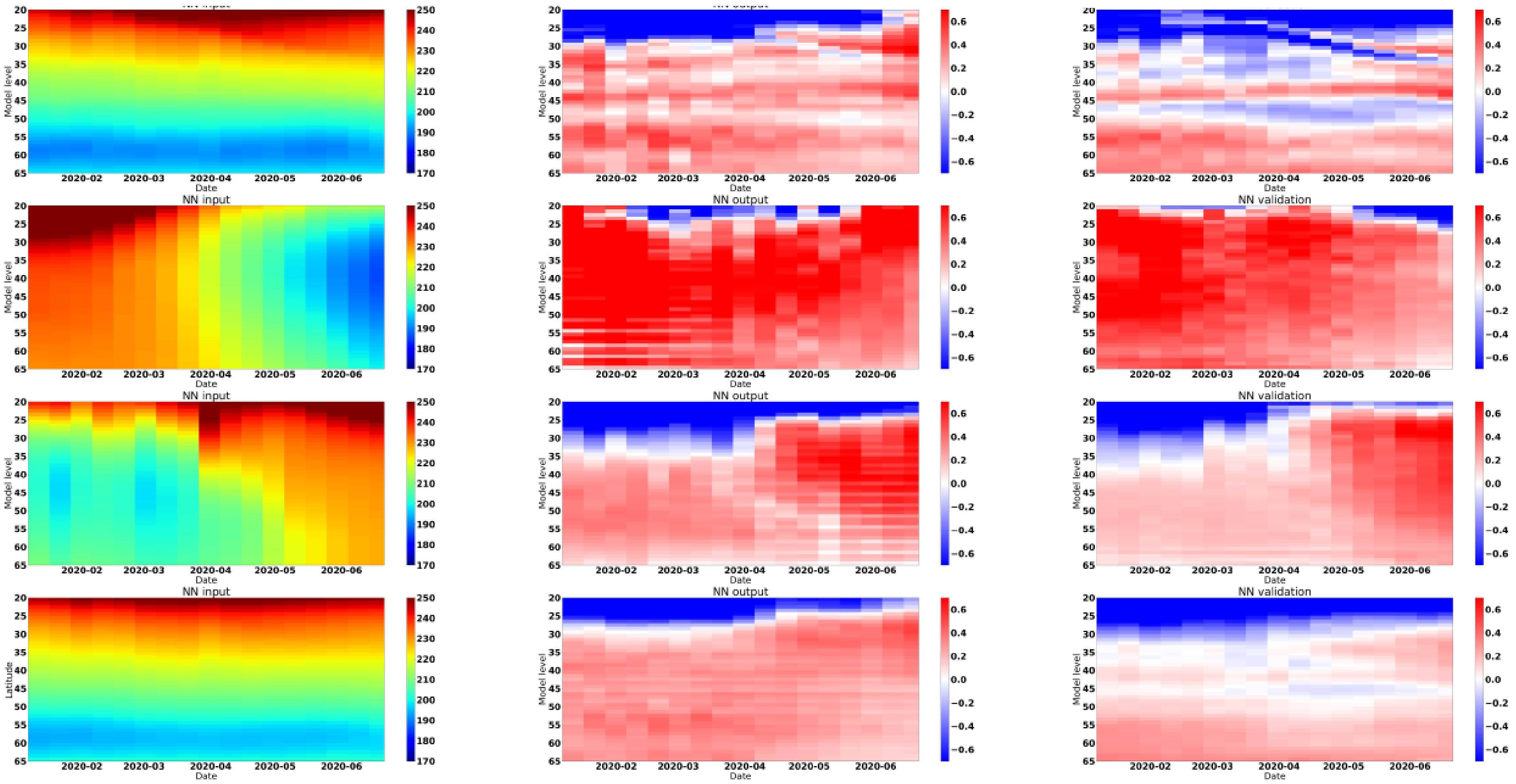
1 years of latest IFS cycle: first-guess and departures with RO temperature retrievals





Results from the NVIDIA CNN

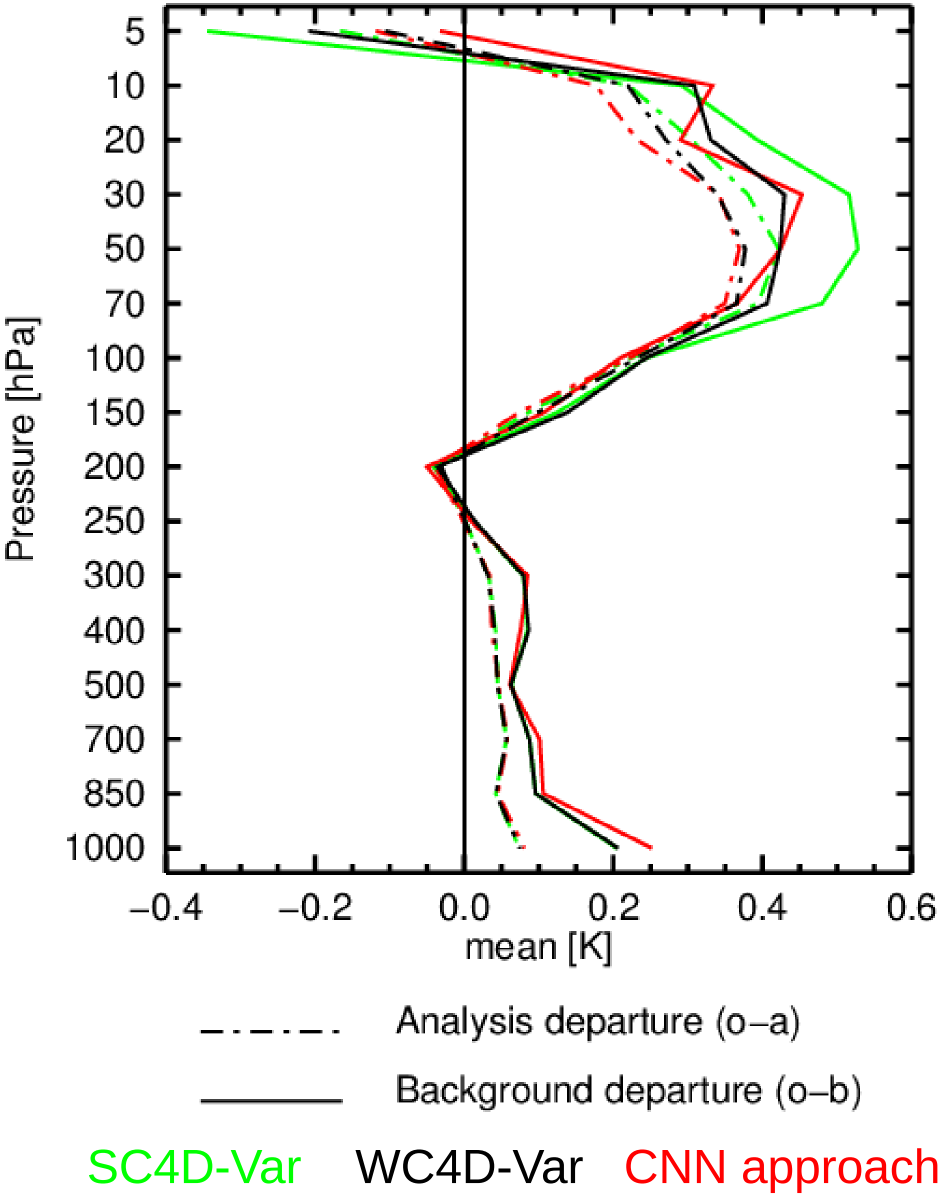
NEW Input | New Target (actual departure) | Predicted from the NN



NN results are not as good but satisfactory (mainly due to the small number of samples)

Results from the NVIDIA CNN

Radiosonde departure
(01/01/2020 – 20/01/2020)



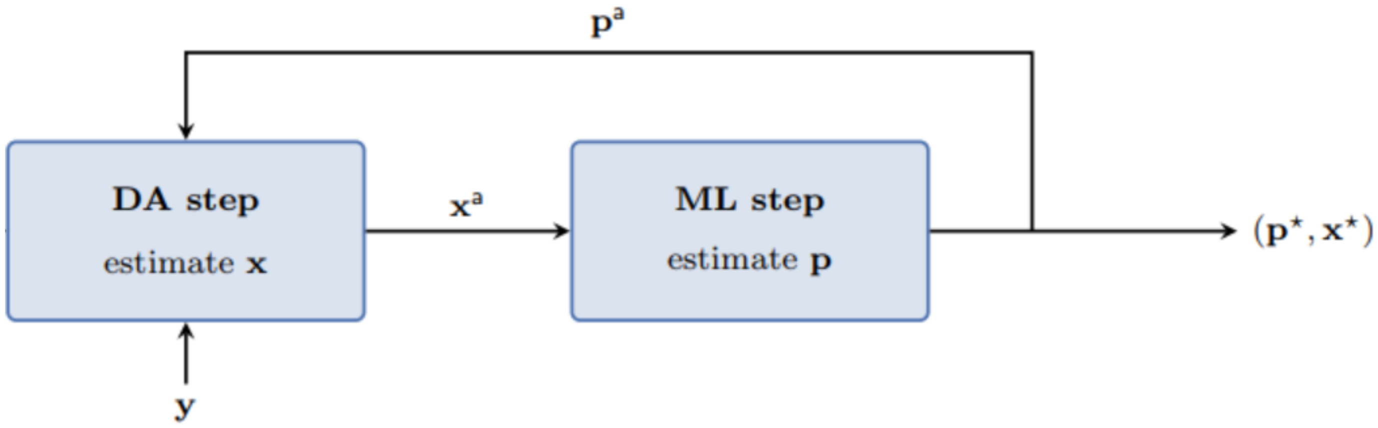
CNN performs similarly to WC4D-Var but WC4D-Var used only 20 days of data instead of 13 years for CNN!

Many ways to improve the NN approach:

- observation sparsity with GCN
- more observations
- better regularisation terms
- online learning

How to estimate model bias with a Neural Network

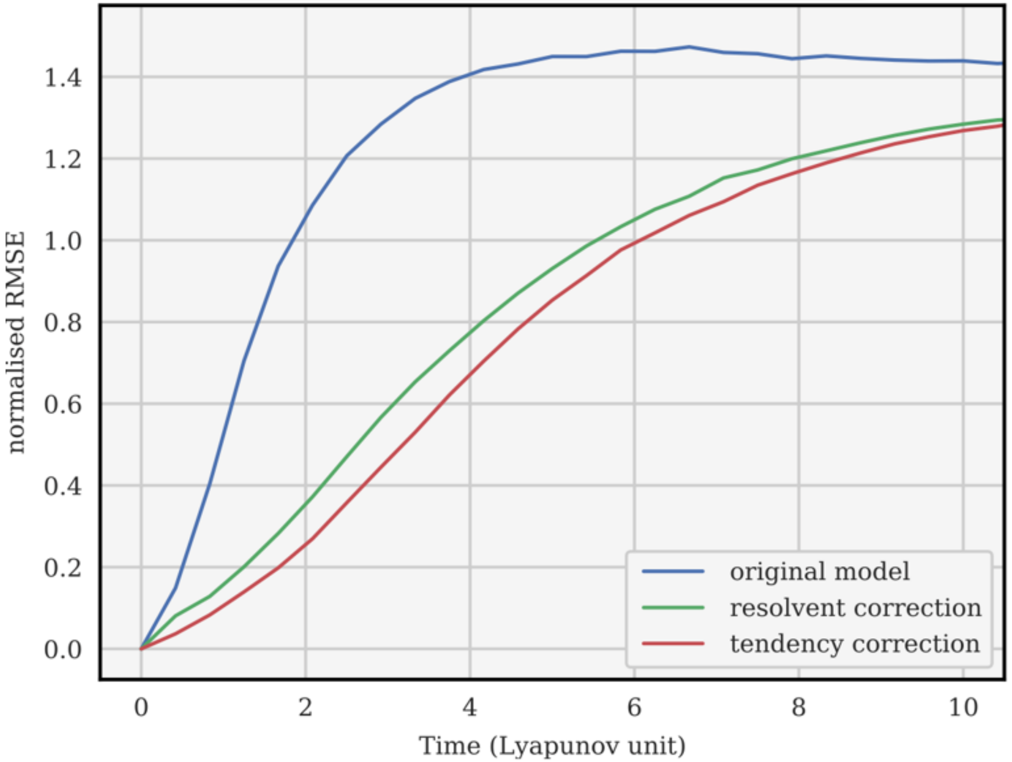
DA/NN framework



- Construct a dataset in **model space** (e.g. analysis)
- Train NN on the dataset
- Correct the model tendency

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathcal{F}^{\mathrm{o}}(\mathbf{x}) + \mathcal{F}^{\mathrm{ml}}(\mathbf{p}, \mathbf{x})$$

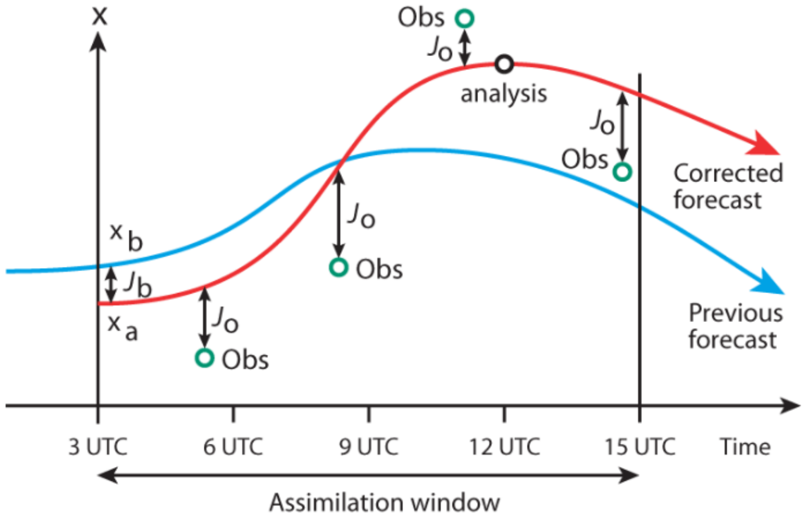
Forecast skill
(2-scale Lorenz system)



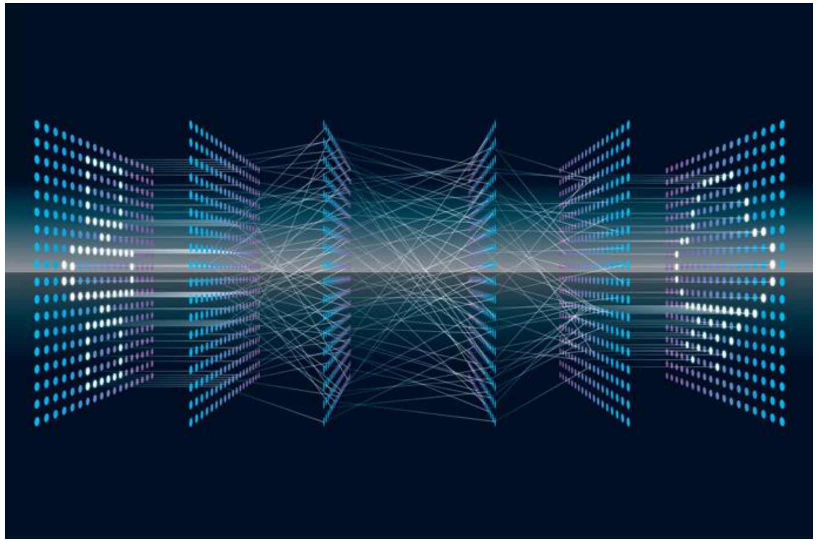
Need the adjoint of the total model (physical + statistical) for the training step.

A. Farchi et al., A comparison of combined data assimilation and machine learning methods for offline and online model error correction, 2020

- Data Assimilation approach



- Machine Learning approach



Cost / loss function equivalence of ML and variational DA

A. Geer (2021) Learning earth system models from observations: machine learning or data assimilation?

ML	Loss function	Basic loss function	Feature error?	Weights regularisation
DA	Cost function	Observation term	Prior knowledge of state	Prior knowledge of model

$$J(x,w) = \underbrace{\frac{(y - h(x,w))^2}{(\cancel{\sigma^y})^2}}_{J^y} + \underbrace{\cancel{\frac{(x^b - x)^2}{(\sigma^x)^2}}}_{J^x} + \underbrace{\frac{(\cancel{w^b} - w)^2}{(\cancel{\sigma^w})^2}}_{J^w}$$

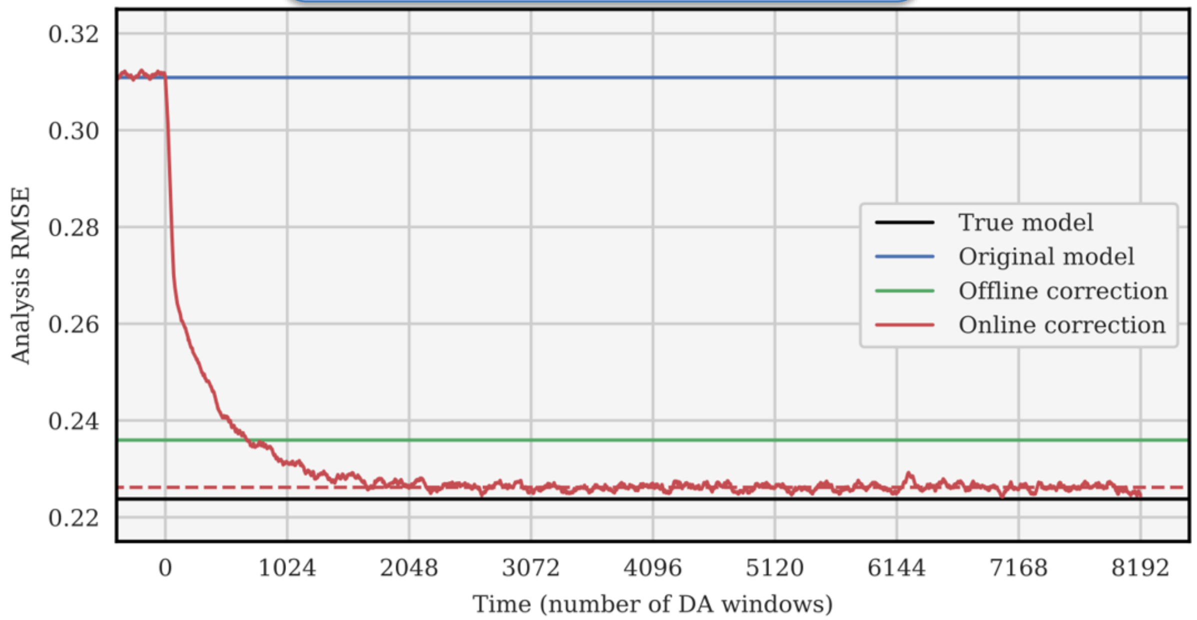
Online learning and links with weak-constraint 4D-Var

NN online loss function

$$\mathcal{J}\left(\mathbf{p}_k, \mathbf{x}_k\right) = \frac{1}{2}\left\|\mathbf{p}_k - \mathbf{p}_k^{\mathrm{b}}\right\|^2_{\mathbf{B}_{\mathrm{p}}^{-1}} + \frac{1}{2}\left\|\mathbf{x}_k - \mathbf{x}_k^{\mathrm{b}}\right\|^2_{\mathbf{B}_{\mathrm{x}}^{-1}} + \frac{1}{2}\sum_{l=0}^{L-1}\left\|\mathbf{y}_{k+l} - \mathcal{H}\circ\mathcal{M}_{l\Delta t}\left(\mathbf{p}_k, \mathbf{x}_k\right)\right\|^2_{\mathbf{R}^{-1}}$$

- learn both model state and NN parameters from observations.
- the online correction steadily improves the model, and eventually gets more accurate than the offline correction

Data assimilation score
(2-scale Lorenz system)



Weak-constraint 4D-Var cost function

$$\begin{aligned} J(x_0, \eta) &= \frac{1}{2}(x_0 - x_b)^T\mathbf{B}^{-1}(x_0 - x_b) \\ &+ \frac{1}{2}\sum_{k=0}^{K}[y_k - \mathcal{H}(x_k)]^T\mathbf{R}_k^{-1}[y_k - \mathcal{H}(x_k)] \\ &+ \frac{1}{2}(\eta - \eta_b)^T\mathbf{Q}^{-1}(\eta - \eta_b) \end{aligned}$$

A. Farchi et al., A comparison of combined data assimilation and machine learning methods for offline and online model error correction, 2020

ECMWF Strategy

‘Science and Technology’ strategic actions

The ‘Science and Technology’ strategic actions are linked to enhancements in the exploitation of observations, data assimilation, modelling and exploitation of new technologies, computational science and operational processes.

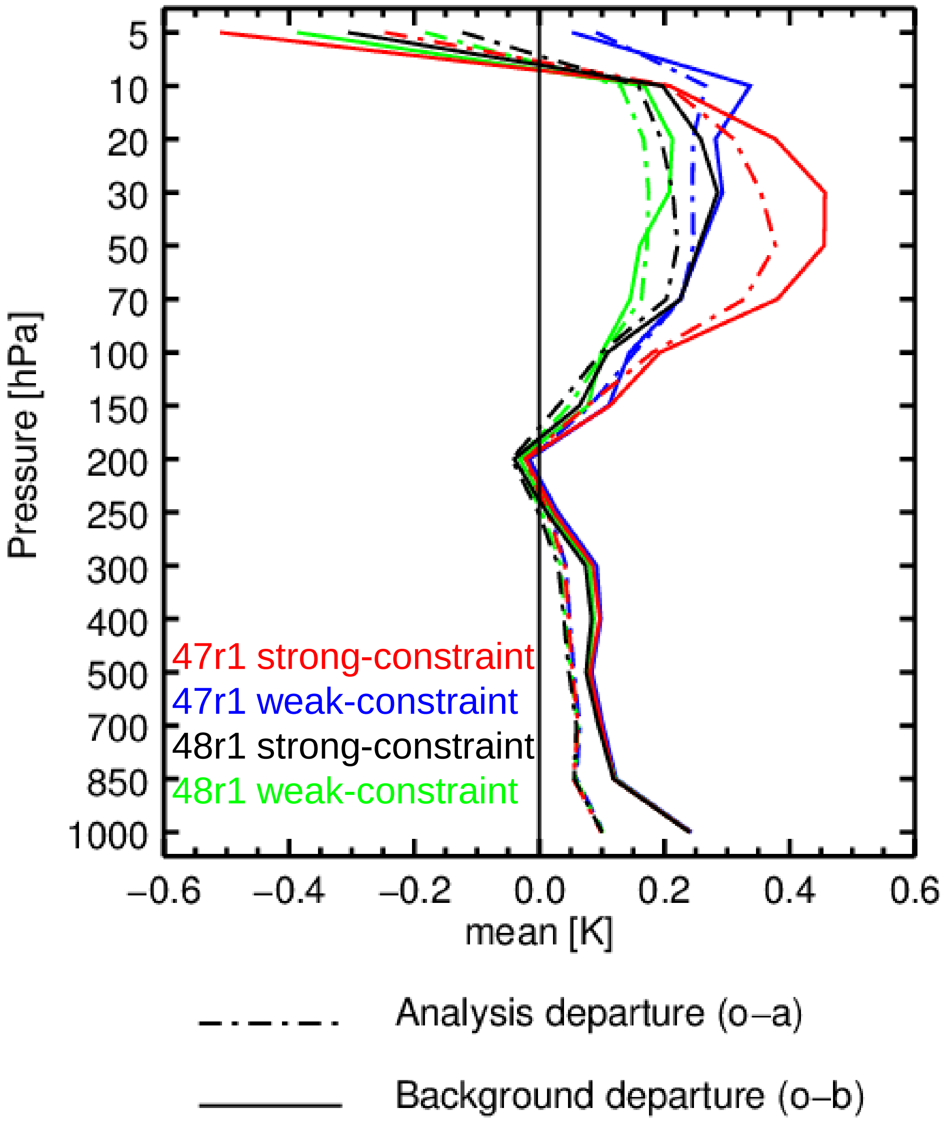
Strengthen leadership in Earth system data assimilation

ECMWF will strengthen its leadership position in data assimilation by progressing in coupled assimilation, algorithmic development and integration of approaches. This will include the incorporation of machine learning, with 4D-Var data assimilation being uniquely positioned to benefit from integrating machine learning technologies because the two fields share a common theoretical foundation and use similar computational tools.

Conclusion and future work

ECMWF has implemented a weak-constraint 4D-Var in operations that learns and correct model biases in the stratosphere

- online learning from all the observations
- dealing with extreme events (SSW)
- dealing with model upgrade
- untangle model and observation biases still challenging (eg troposphere)



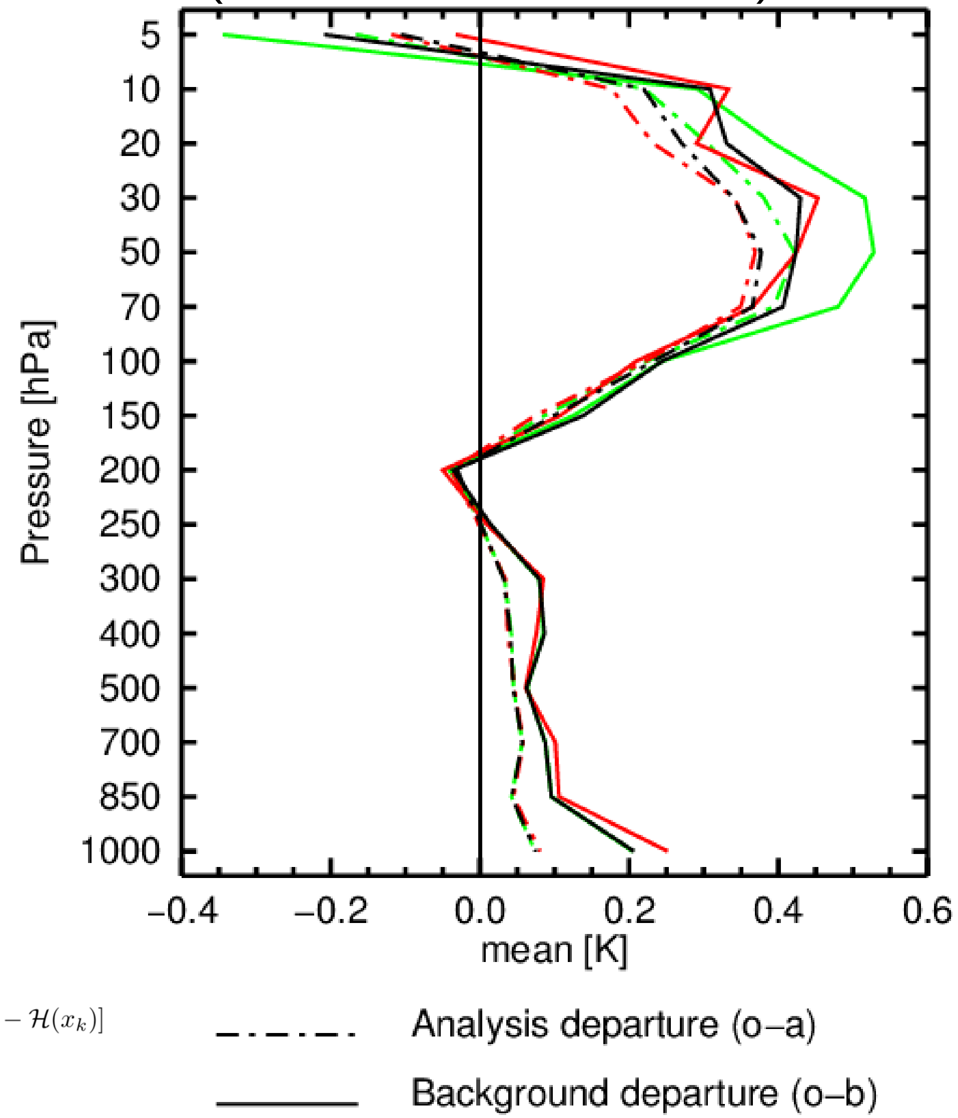
Conclusion and future work

ECMWF investigates the machine learning approach to correct for model biases (among other things). Preliminary results are encouraging

- Large datasets are required
- Retraining is challenging as limited availability of samples
- NN face the same issues to untangle model and observation biases (limited by the accuracy of the observations)
- Study on extreme events is required

As NN approach is getting more sophisticated, it gets closer to weak-constraint 4D-Var (two sides of the same coin)

Radiosonde departure
(01/01/2020 – 20/01/2020)



SC4D-Var WC4D-Var CNN approach

$$
\begin{aligned}
J(x_0,\eta) &= \frac{1}{2}(x_0-x_b)^T\mathbf{B}^{-1}(x_0-x_b)\\
&+ \frac{1}{2}\sum_{k=0}^{K}[y_k-\mathcal{H}(x_k)]^T\mathbf{R}_k^{-1}[y_k-\mathcal{H}(x_k)]\\
&+ \frac{1}{2}(\eta-\eta_b)^T\mathbf{Q}^{-1}(\eta-\eta_b)
\end{aligned}
$$

$$
\mathcal{J}(\mathbf{p}_k,\mathbf{x}_k)=\frac{1}{2}\left\|\mathbf{p}_k-\mathbf{p}_k^{\mathrm{b}}\right\|^2_{\mathbf{B}_{\mathrm{p}}^{-1}}+\frac{1}{2}\left\|\mathbf{x}_k-\mathbf{x}_k^{\mathrm{b}}\right\|^2_{\mathbf{B}_{\mathrm{x}}^{-1}}+\frac{1}{2}\sum_{l=0}^{L-1}\left\|\mathbf{y}_{k+l}-\mathcal{H}\circ\mathcal{M}_{l\Delta t}(\mathbf{p}_k,\mathbf{x}_k)\right\|^2_{\mathbf{R}^{-1}}
$$