Data Assimilation: A fusion of knowledge and the rise of the machines

Tony McNally ECMWF

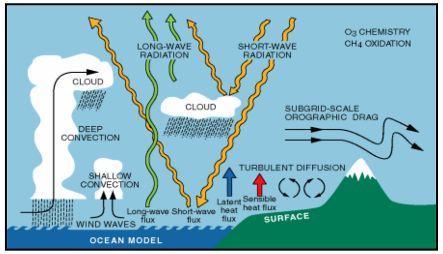
QJRMS 150^{th} Anniversary

Overview:

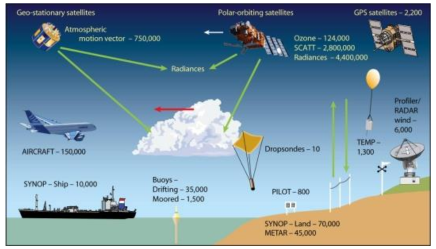
- A brief introduction to Data Assimilation
- Where has Data Assimilation been used ?
- Where is Data Assimilation heading next ?
- Summary
- A challenge for the established literature

A brief introduction to Data Assimilation

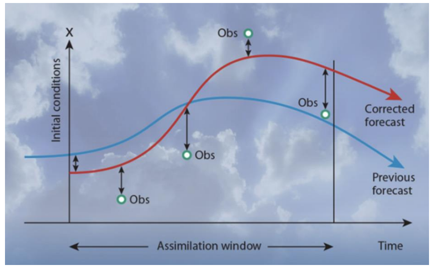
- **Models** give a **complete** description of the atmospheric state, but errors can grow rapidly in time



- **Observations** provide an incomplete description of the atmospheric state, but do bring **accurate** up to date information

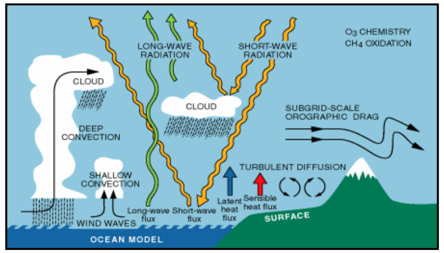


- The **Data Assimilation algorithm** combines these two sources of information to produce an **optimal** (best) estimate of the atmospheric state

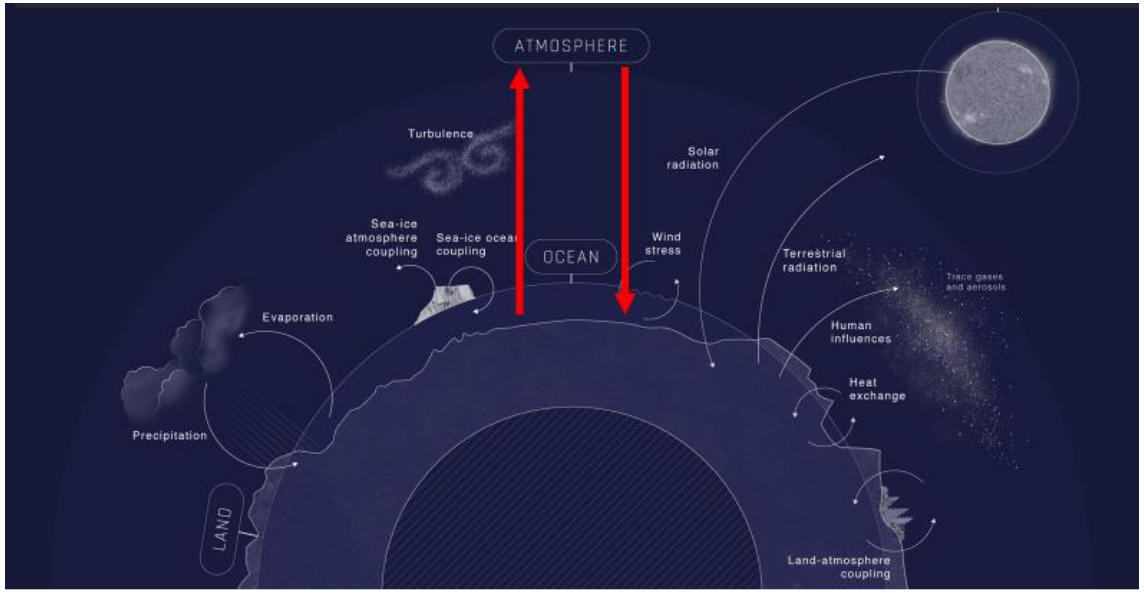


Data Assimilation: The model

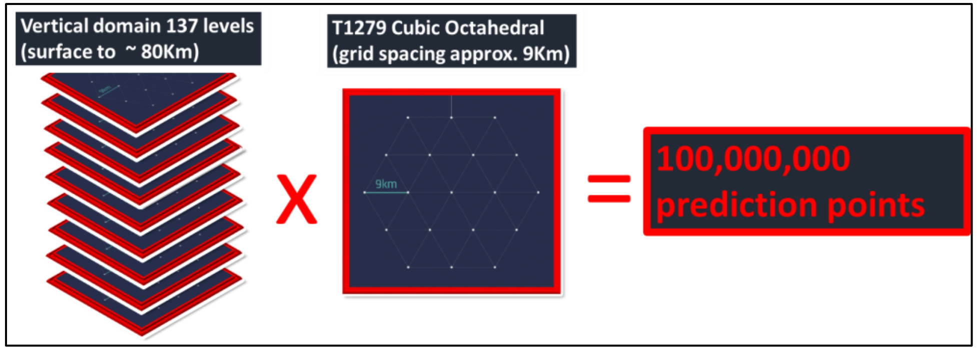
- **Models** give a **<u>complete</u>** description of the atmospheric state, but errors can grow rapidly in time



<u>Coupled</u> Earth System Simulators / Digital Twins

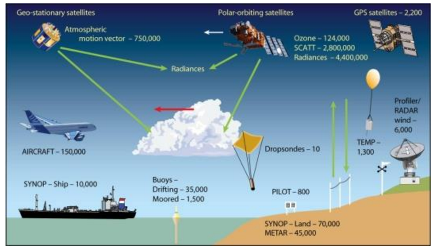


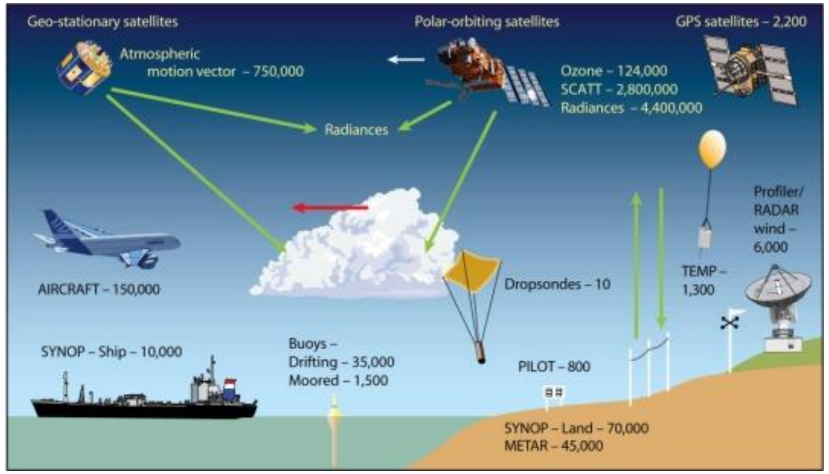
Atmosphere Component

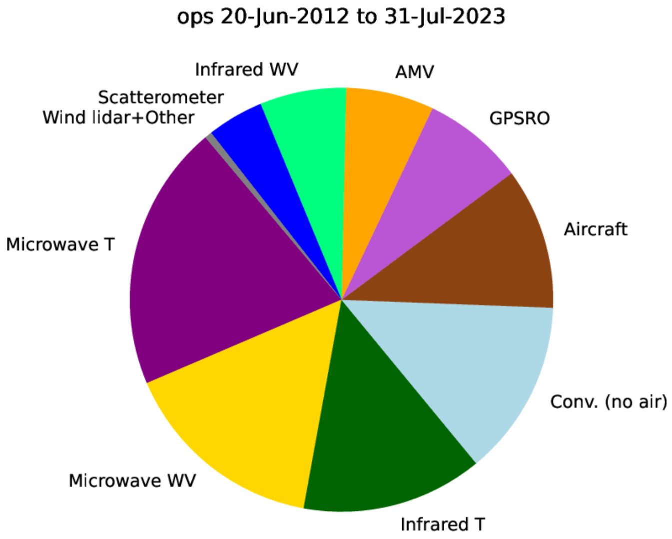


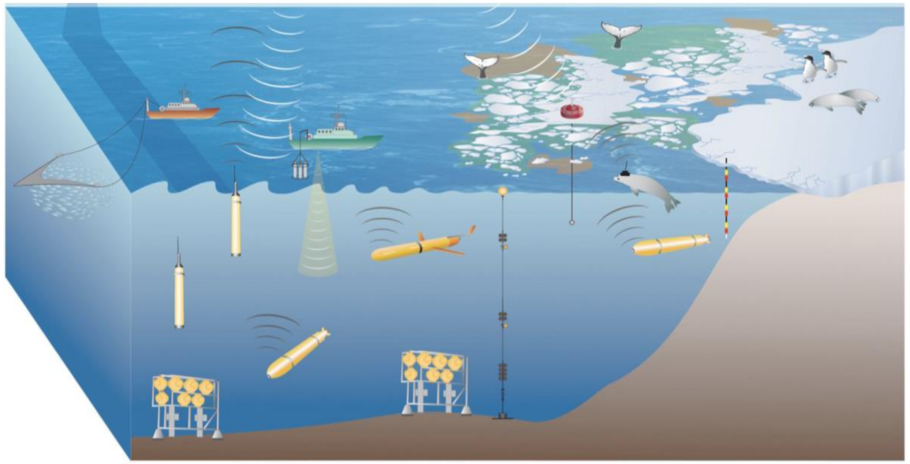
Data Assimilation: The observations

- **Observations** provide an incomplete description of the atmospheric state, but do bring **<u>accurate</u>** up to date information



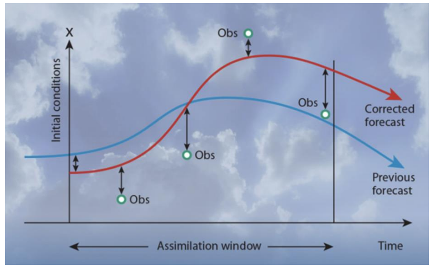






Data Assimilation: The algorithm

- The **Data Assimilation algorithm** combines these two sources of information to produce an <u>**optimal**</u> (best) estimate of the atmospheric state



(1) There are lots of approaches:

- Optimal Interpolation (OI)
- Variational (Var)
- Ensemble Kalman Filter (ENKF)
- Local Ensemble Transform KF (LETKF)
- Simplified Extended Kalman Filter (SEKF)
- Any combination of the above…

(2) The are theoretically equivalent

A method of deriving the OI formula originates from the cost function. Even though h in (2) can be non-linear, here we will first approximate it by linearization about the $\vec{x}_B$.

$$\text{Let} \quad \vec{x} = \vec{x}_B + \delta\vec{x}, \tag{4}$$

$$\text{then} \quad \vec{h}[\vec{x}_B + \delta\vec{x}] \approx \vec{h}[\vec{x}_B] + \mathbf{H}\delta\vec{x}. \tag{5}$$

$\mathbf{H}$ is a matrix which represents the linearization of $\vec{h}$ about $\vec{x}_B$. (5) is a Taylor expansion of $\vec{h}$ about $\vec{x}_B$ to first order where $\mathbf{H}$ is the first derivative (called the 'Jacobian'),

$$\mathbf{H} = \left.\frac{\partial\vec{h}}{\partial\vec{x}}\right|_{\vec{x}_B}, \tag{6}$$

$$\text{which is a matrix notation for the elements} \quad \mathbf{H}_{ij} = \frac{\partial h_i}{\partial x_j} \quad (1 \leqslant i \leqslant p, \quad 1 \leqslant j \leqslant n). \tag{7}$$

Substitute (4)-(5) into (2), and rearrange

$$\begin{aligned} J &= \frac{1}{2}\delta\vec{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\vec{x} + \frac{1}{2}(\vec{y} - \vec{h}[\vec{x}_B] - \mathbf{H}\delta\vec{x})^{\mathrm{T}}\mathbf{R}^{-1}(\vec{y} - \vec{h}[\vec{x}_B] - \mathbf{H}\delta\vec{x}), \\ &= \frac{1}{2}\delta\vec{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\vec{x} + \frac{1}{2}(\mathbf{H}\delta\vec{x} - \{\vec{y} - \vec{h}[\vec{x}_B]\})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\vec{x} - \{\vec{y} - \vec{h}[\vec{x}_B]\}). \end{aligned}$$

J is minimized at the analysis, $\vec{x}_A$, where $\nabla_x J = 0$

$$\nabla_x J[\delta\vec{x} = \delta\vec{x}_A] = \mathbf{B}^{-1}\delta\vec{x}_A + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\vec{x}_A - \{\vec{y} - \vec{h}[\vec{x}_B]\}) = 0,$$

(see §D.1 and §D.2 to derive this gradient expression), where $\vec{x}_A = \vec{x}_B + \delta\vec{x}_A$. This expression can be rearranged for $\delta\vec{x}_A$

$$(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})\delta\vec{x}_A = \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\vec{y} - \vec{h}[\vec{x}_B]),$$

$$\delta\vec{x}_A = \vec{x}_A - \vec{x}_B = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\vec{y} - \vec{h}[\vec{x}_B]). \tag{8}$$

This equation can be written in a different way by using the following Sherman-Morrison-Woodbury formula (see the problem sheet, Q5)

$$(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})\mathbf{B}\mathbf{H}^{\mathrm{T}} = \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}), \tag{9}$$

which can be proven easily. It is straightforward to rearrange (9) to resemble the string of matrix operators that are present in (8)

$$(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}})^{-1}, \tag{10}$$

making (8) into an equivalent form

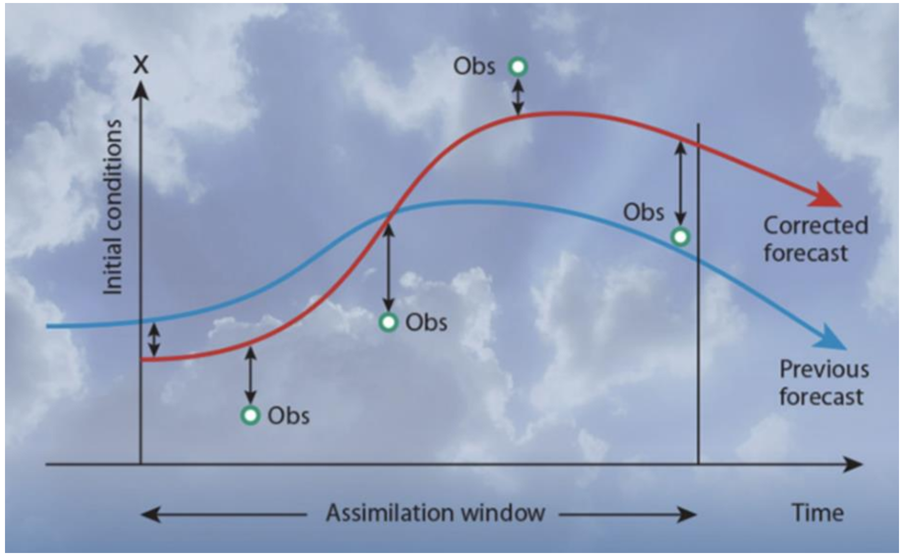
$$\vec{x}_A - \vec{x}_B = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}})^{-1}(\vec{y} - \vec{h}[\vec{x}_B]). \tag{11}$$

(11) is the Optimal Interpolation (OI) or Best Linear Unbiased Estimator (BLUE) formula, derived using the 'max. likelihood' (or 'min. cost') method. Since OI and Var. are equivalent when the forward model is linear (ie when (5) holds exactly), (11) can be used to understand how Var. works.

(3) Implementation is critical

- Application appropriate (global / regional)
- How accurate is your prior knowledge (model)
- How well constrained by observations ?
- Computer resources
- Human resources / sectorial skill

Data Assimilation: The 4D-Var algorithm



State vector

background error covariance

$$J(x) = (x - x_b)^T \mathbf{B}^{-1}(x - x_b) + (y - \mathrm{H}[x])^T \mathbf{R}^{-1}(y - \mathrm{H}[x])$$

observations

observation* error covariance

observation operator (maps the model state to the observation space)


Quarterly Journal of the Royal Meteorological Society
RMetS

Article | Free to Read

The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics

F. Rabier, H. Järvinen, E. Klinker, J.-F. Mahfouf, A. Simmons

First published: April 2000 Part A | https://doi.org/10.1002/qj.49712656415 | Citations: 534

- It has a physical model at its core – the analysis is **physical**
- Global DA is solved **globally** (no spatial boxing or localization)
- Accelerates the use of satellite **radiance** observations

Data Assimilation: Where has it been used ?

Data Assimilation: Where has it been used ?

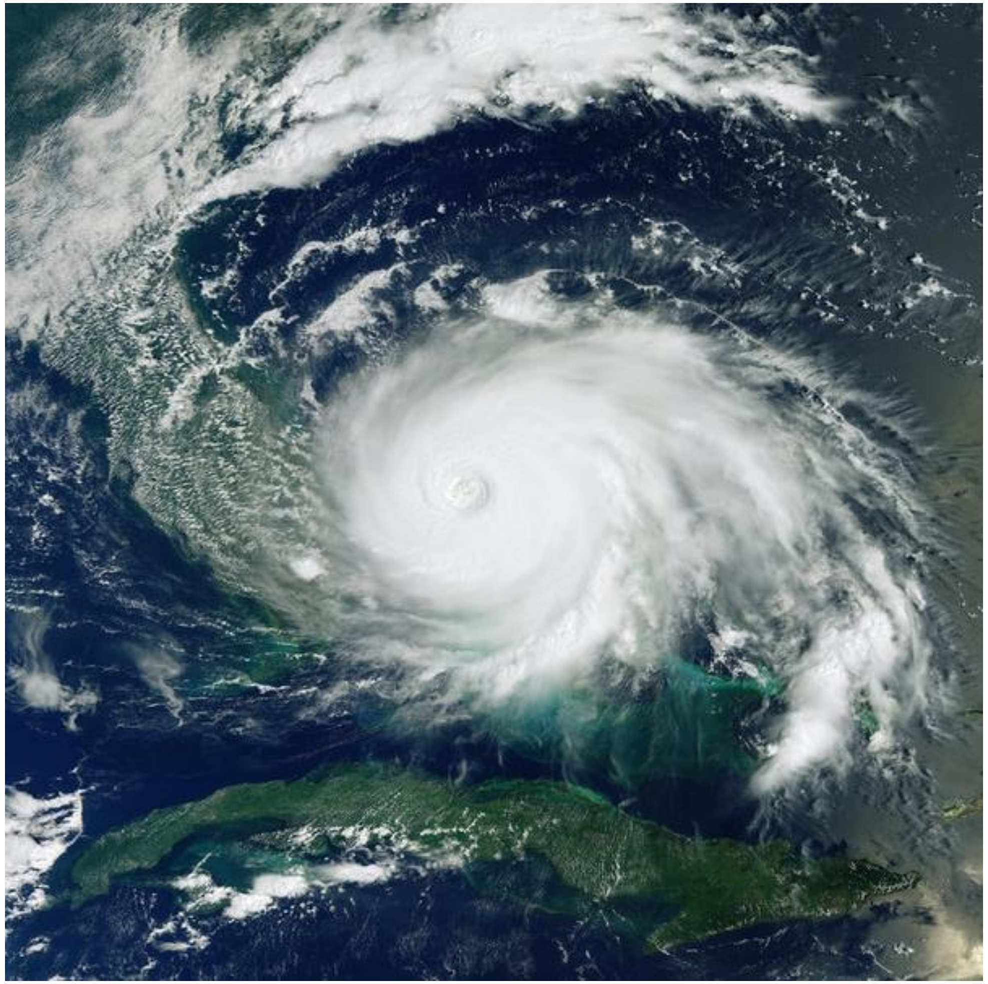
- Numerical weather prediction (weather forecasting)
- Climate reanalysis

Data Assimilation: Where has it been used ?

- **Numerical weather prediction (weather forecasting)**
- Climate reanalysis

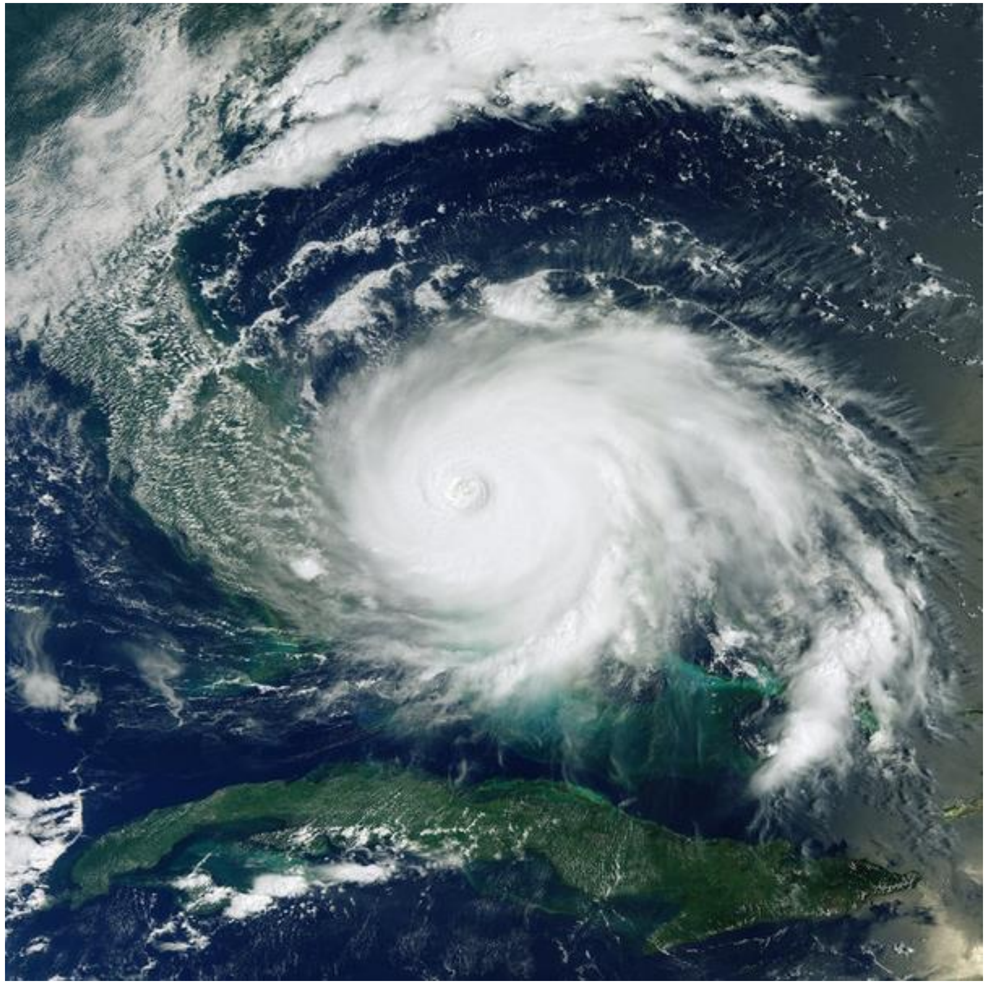
Data Assimilation: Numerical Weather Prediction

Dorian viewed from the Sentinel-3 satellite



Data Assimilation: Numerical Weather Prediction

Dorian viewed from the Sentinel-3 satellite



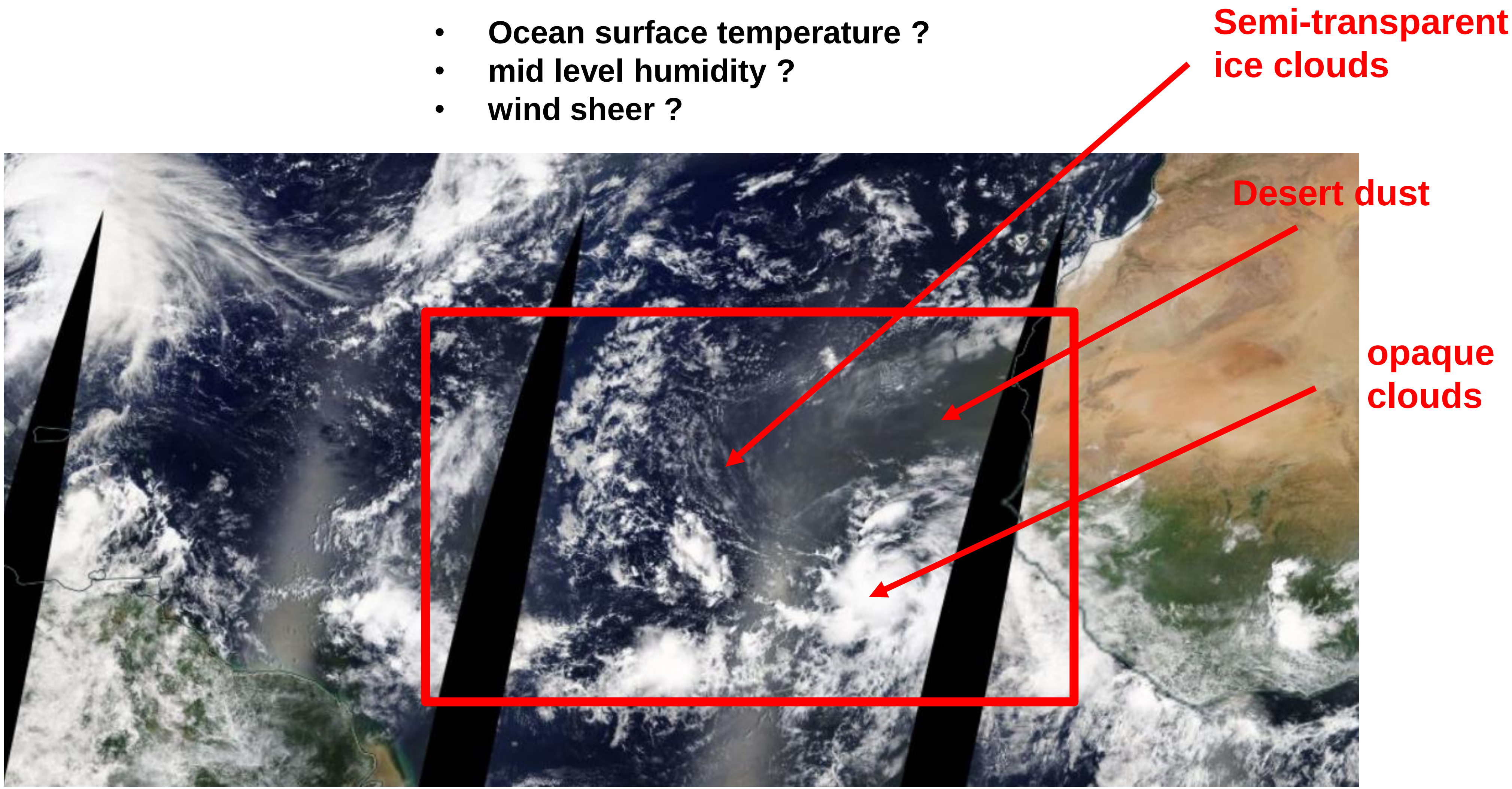
Dorian viewed from the Bahamas

Good forecasts and excellent evacuation plans <u>significantly mitigated storm human impact</u>

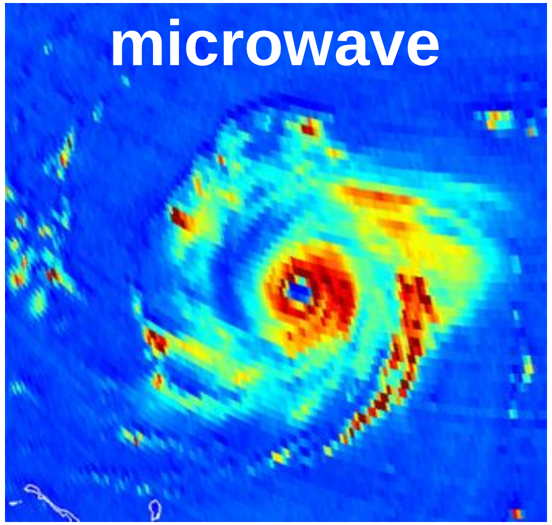
Data Assimilation: The power of data fusion

Early identification of storm genesis…in a challenging environment

- **Ocean surface temperature ?**
- **mid level humidity ?**
- **wind sheer ?**

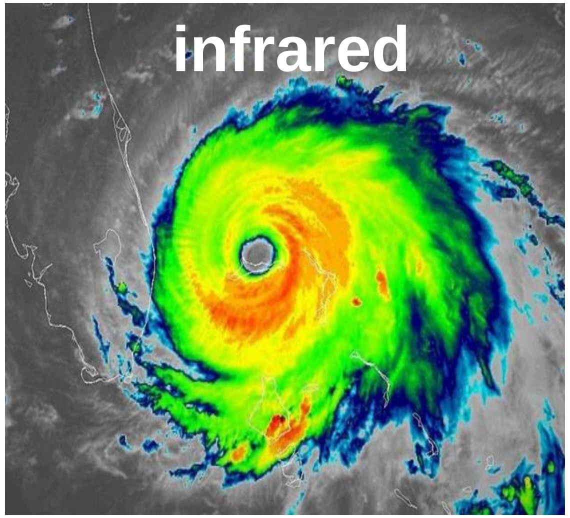


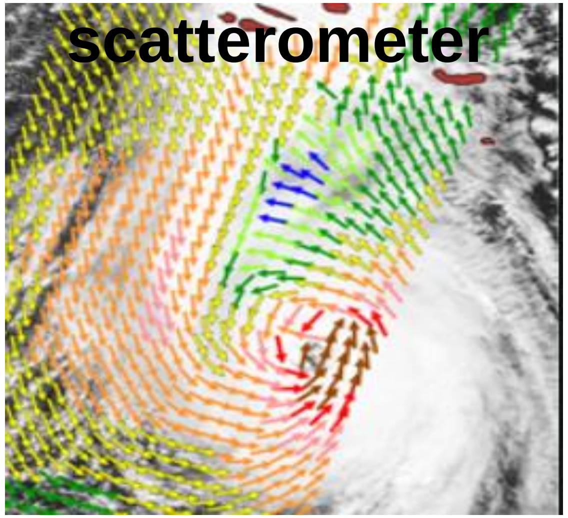
Data Assimilation: The power of data fusion

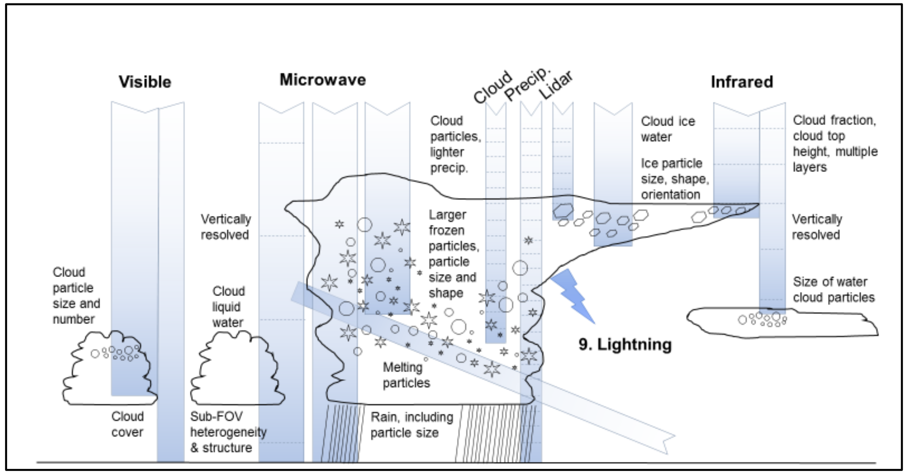


DA systems are able to effectively combine information from many highly heterogenous sources…

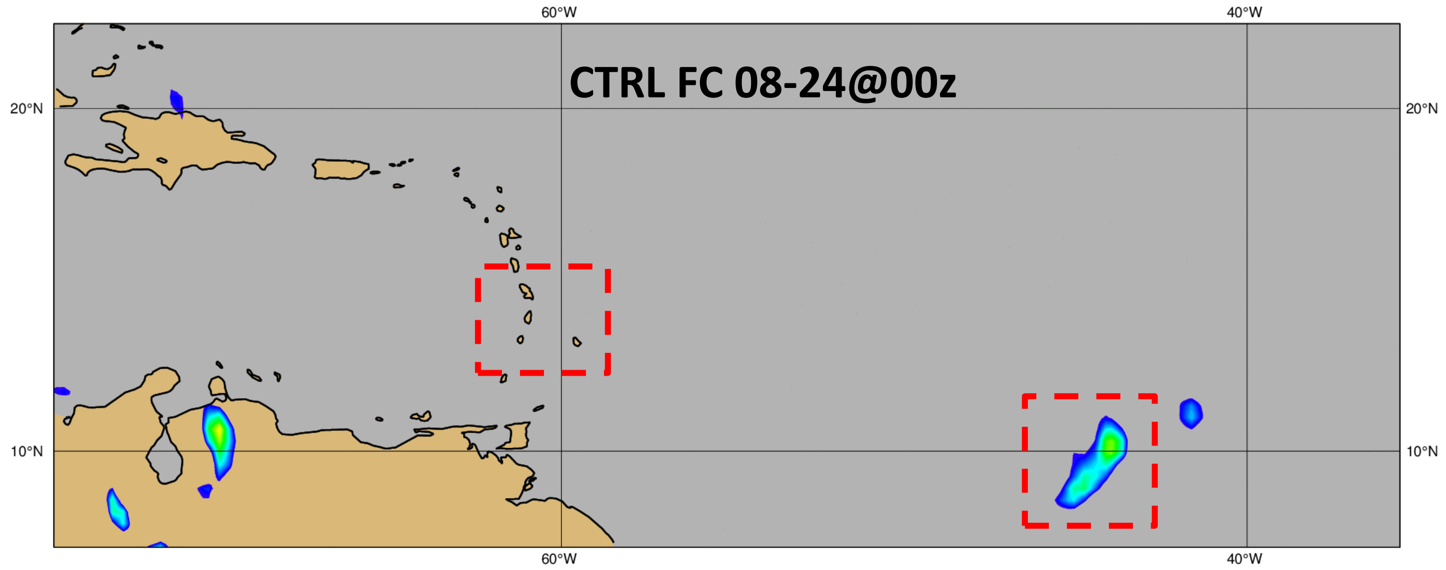
… to build up a multi-dimensional and multi-parameter view of the atmosphere



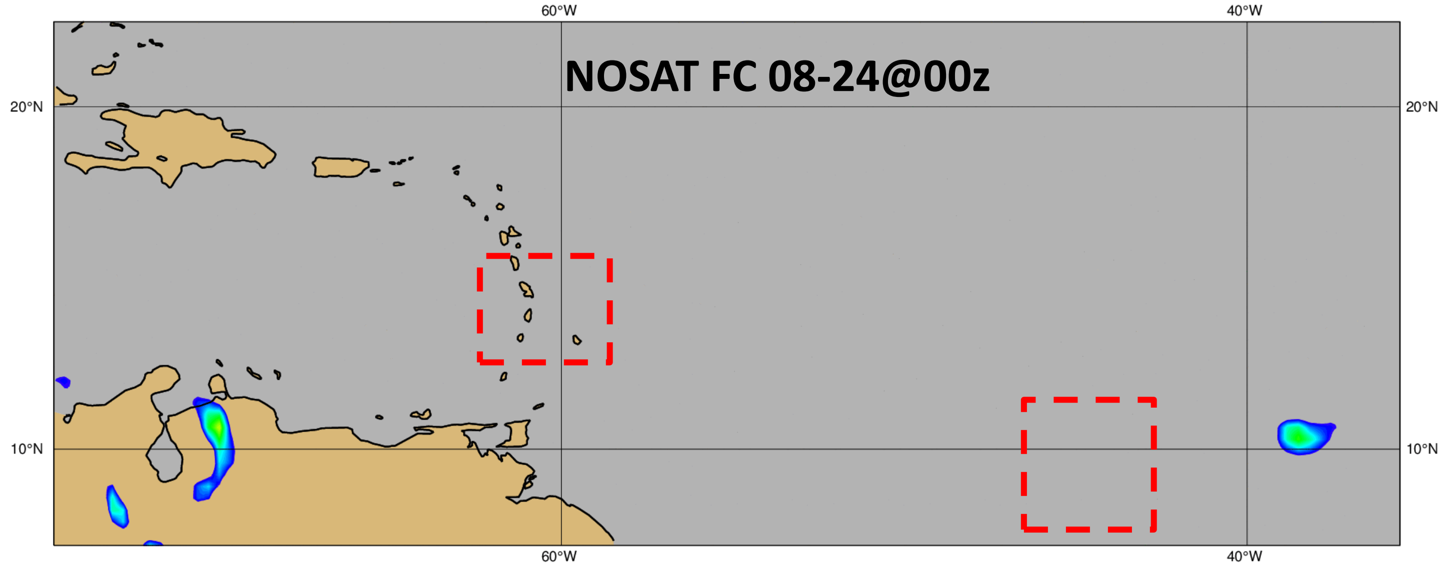




Data Assimilation: The power of data fusion



Control system with satellites identifies storm genesis on 24th August and provides **4 days warning** of direct strike on Windward Islands



System with **satellites denied** (for 36hrs prior to forecast) misses the storm genesis and provides **no warning of strike** on Windward Islands

Data Assimilation: Where has it been used ?

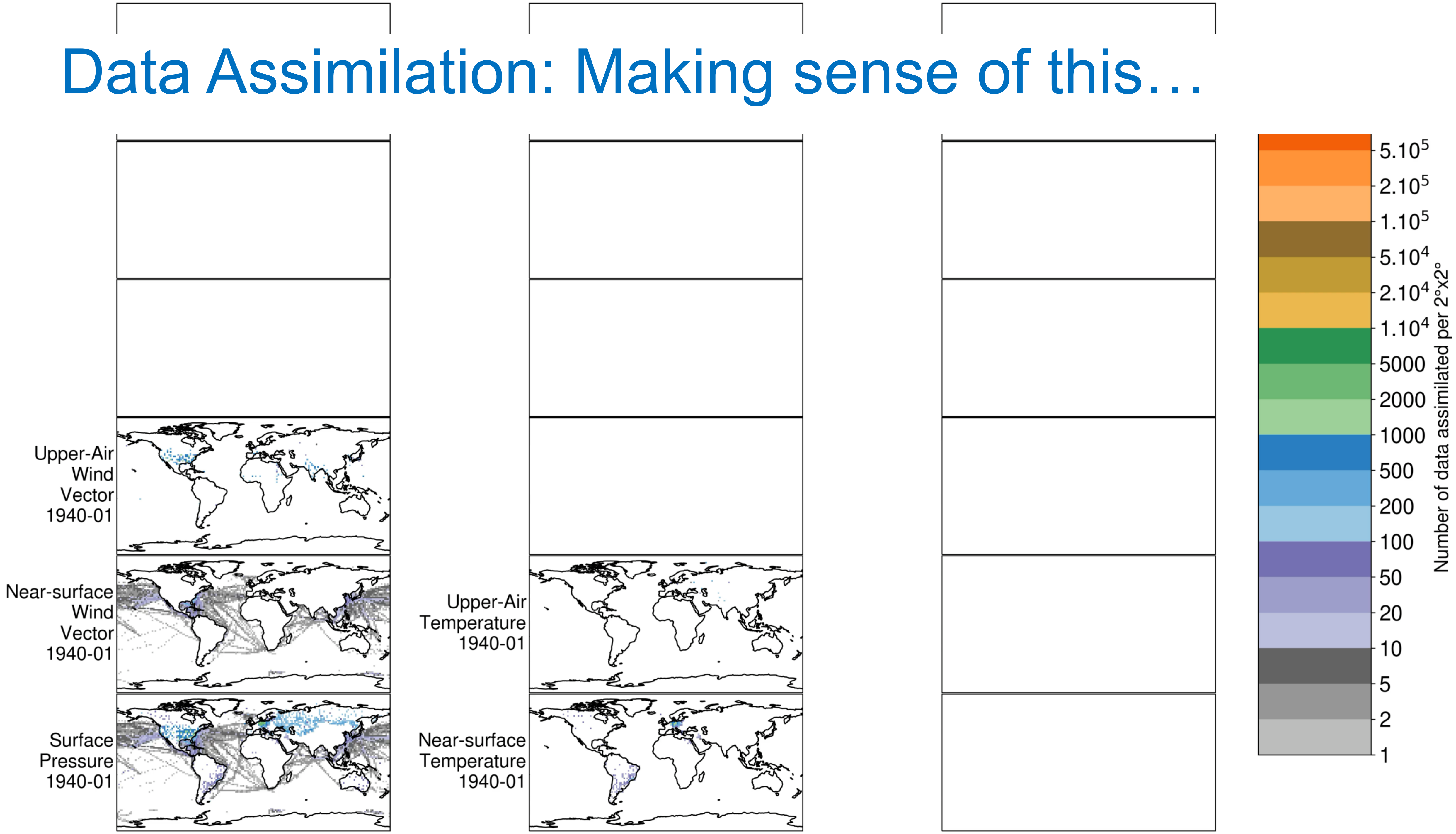
- Numerical weather prediction (weather forecasting)

- **Climate reanalysis**

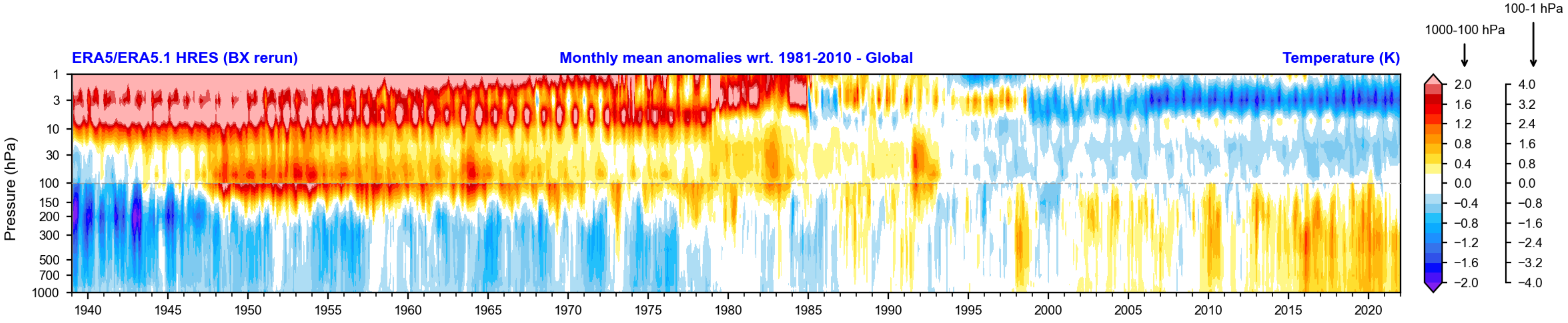
Data Assimilation: Climate reanalysis

- Reanalysis is indispensable for research, climate science and climate services.
- Most cited datasets in the scientific literature.
- ERA5 has 240 citations in the IPCC AR6 WGI report.
- Reanalysis is the backbone for Copernicus services.
- Reanalysis provides fundamental training data for machine learning applications (e.g. weather forecasting).

Data Assimilation: Making sense of this…



…to produce this




Quarterly Journal of the Royal Meteorological Society

RMetS

RESEARCH ARTICLE | Open Access

The ERA5 global reanalysis

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons ... See all authors

First published: 17 May 2020 | https://doi.org/10.1002/qj.3803 | Citations: 7,139

Data Assimilation: Where has it been used ?

- Numerical weather prediction (weather forecasting)
- Climate reanalysis
- **Atmospheric composition**

Data Assimilation: Where next ?

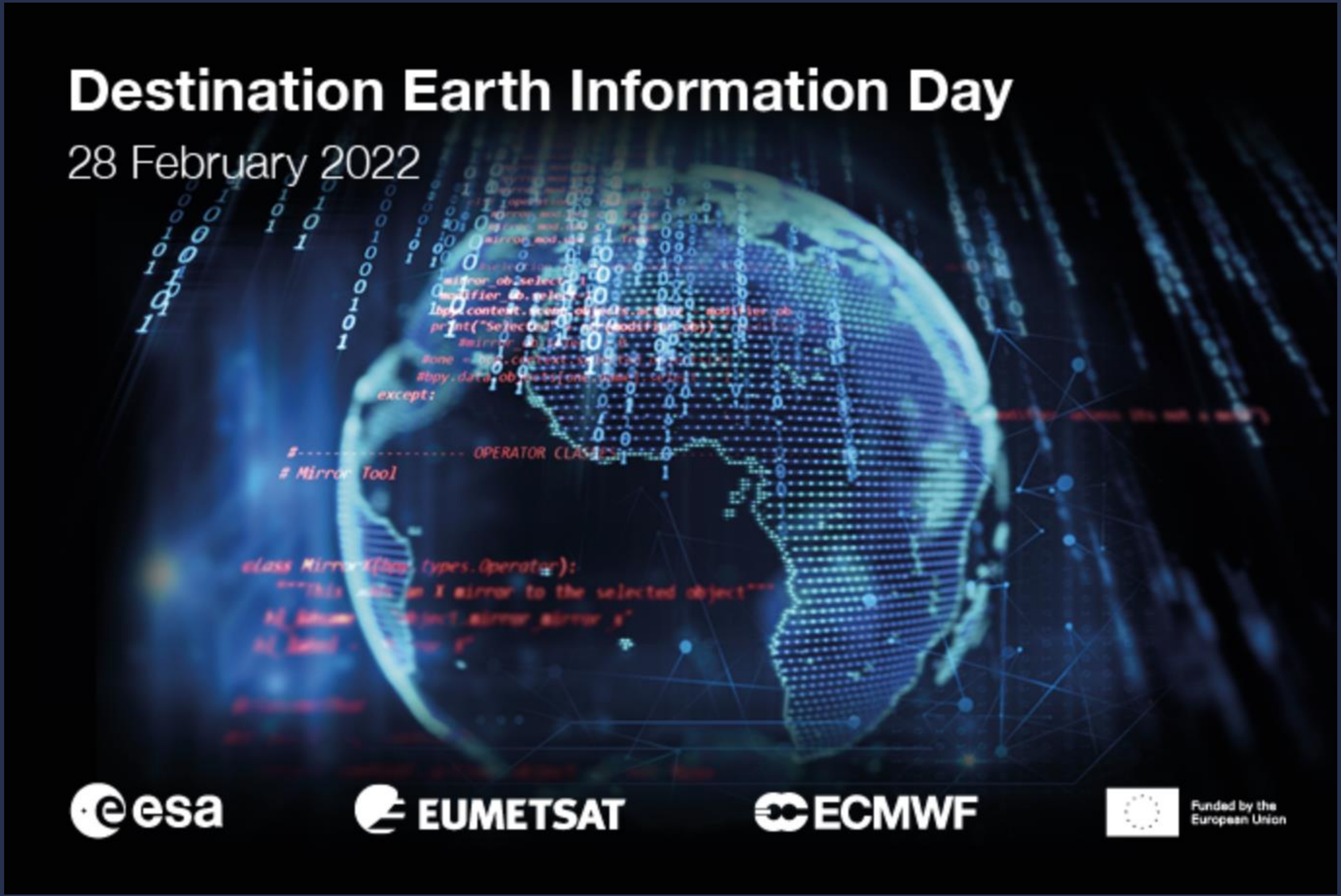
Data Assimilation: Where next ?

- Km Scale DA systems
- Using DA to improve models (weather and climate)
- The rise of the machines

Data Assimilation: Where next ?

- **Km Scale DA systems**
- Using DA to improve models (weather and climate)
- The rise of the machines

Exploring the limits of high-resolution…DestinE

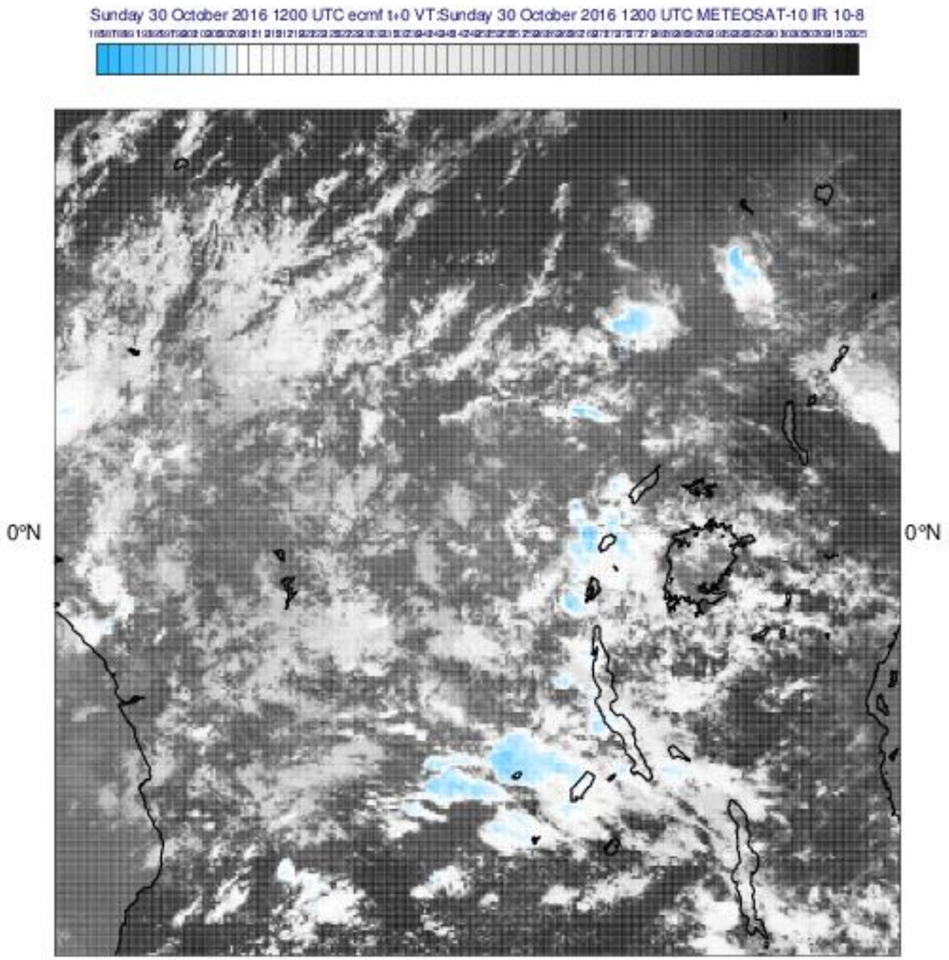


High-resolution DA…but how high ?

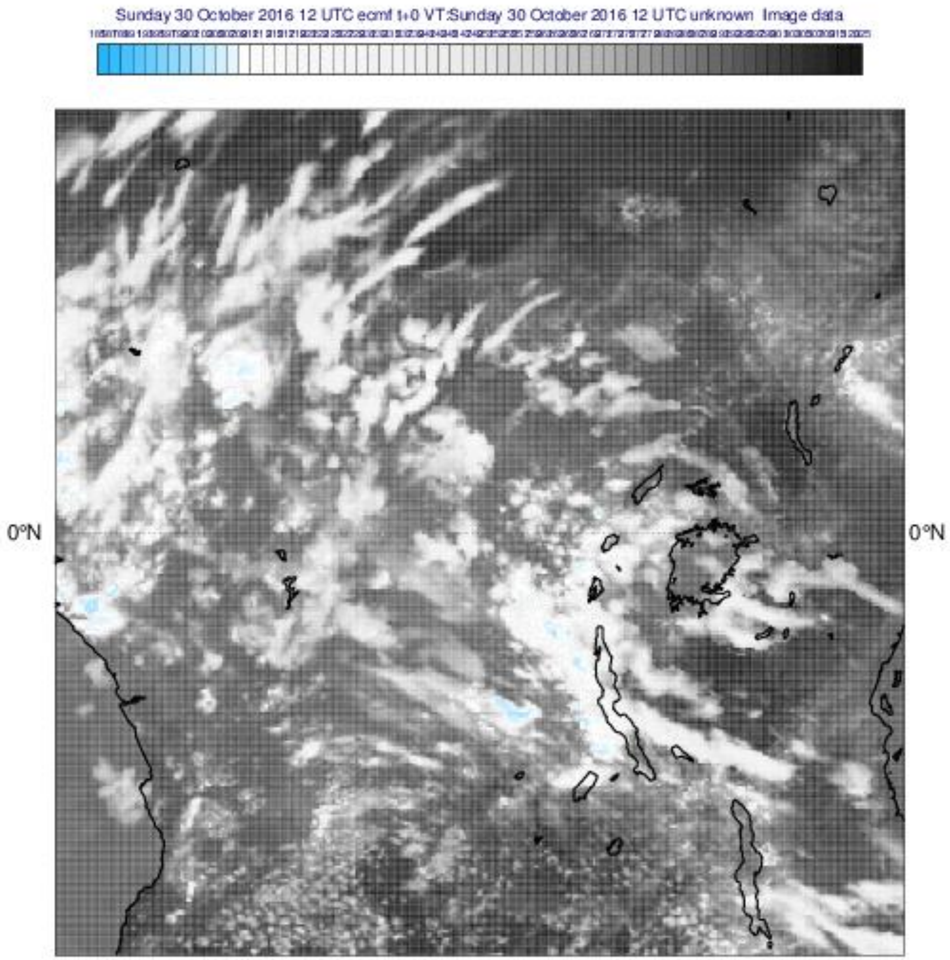
Scales required to exploit observation to full potential

Scales required to initialize our forecast models

MET-11 SEVIRI real Observations



Simulated from TCO7999 model (~1.25Km)



Data Assimilation: Where next ?

- Km Scale DA systems
- **Using DA to improve models (weather and climate)**
- The rise of the machines

Data Assimilation: Improving models

$$
\begin{aligned}
J(x) = {} & (x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \\
& (y - \underline{\mathrm{H}[x]})^T \mathbf{R}^{-1}(y - \underline{\mathrm{H}[x]}) + \underline{(\beta - \beta_b)^T \mathbf{B}_\beta{}^{-1}(\beta - \beta_b)} \\
& + \underline{(\eta - \eta_b)^T \mathbf{Q}^{-1}(\eta - \eta_b)}
\end{aligned}
$$

Quarterly Journal of the Royal Meteorological Society
RMetS

Research Article

Model-error estimation in 4D-Var

Yannick Trémolet

First published: 13 July 2007 | https://doi.org/10.1002/qj.94 | Citations: 97

Quarterly Journal of the Royal Meteorological Society
RMetS

RESEARCH ARTICLE

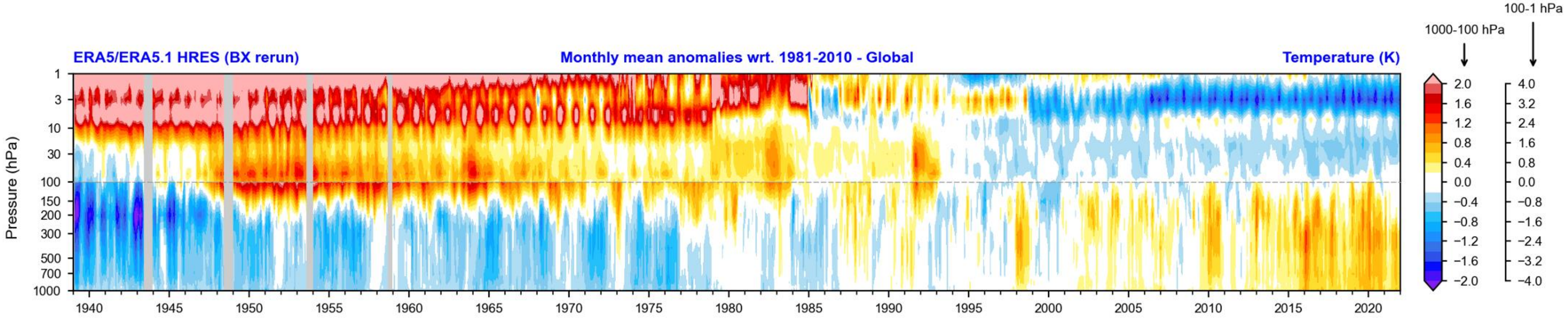
Exploring the potential and limitations of weak-constraint 4D-Var

P. Laloyaux, M. Bonavita, M. Chrust, S. Gürol

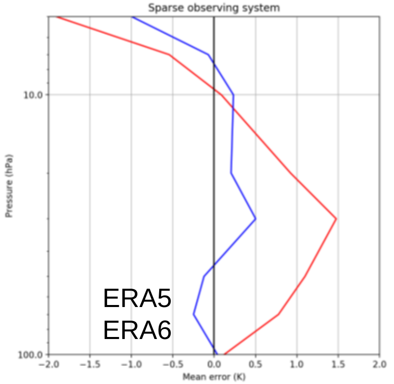
First published: 15 August 2020 | https://doi.org/10.1002/qj.3891 | Citations: 9

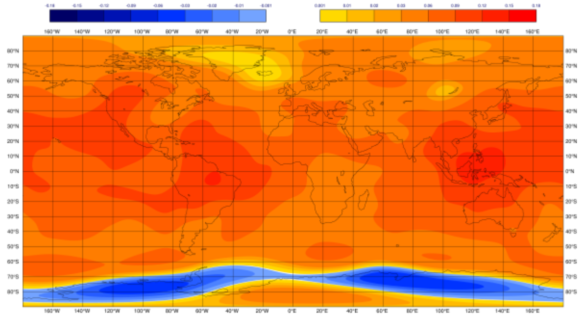
…and can be used to constrain historical periods

Reanalysis during periods poorly constrained by observations (e.g. pre-satellite) *inherit* systematic model error, causing shocks when major observing systems come and go which can compromise climate trends



Mean fit to radiosonde data in pre-satellite era



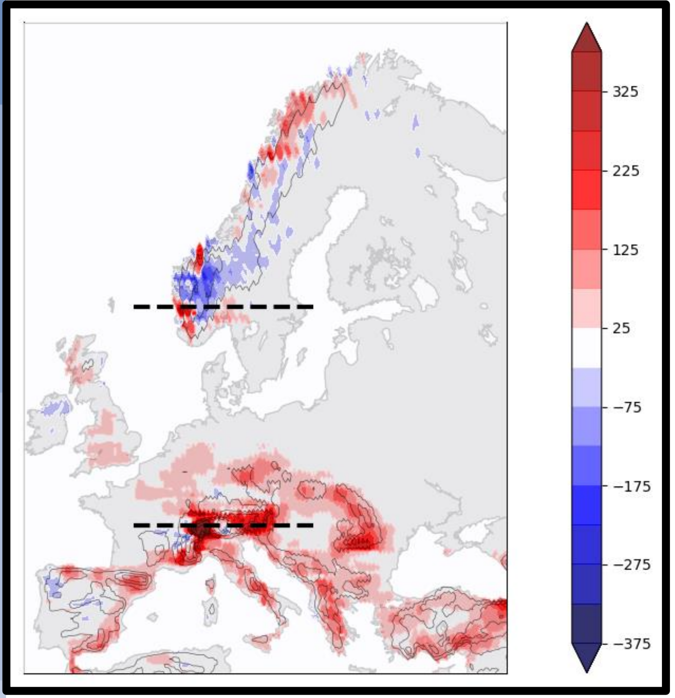


Running weak constraint 4D-Var during current <u>well observed</u> periods provides an accurate estimate of systematic model error

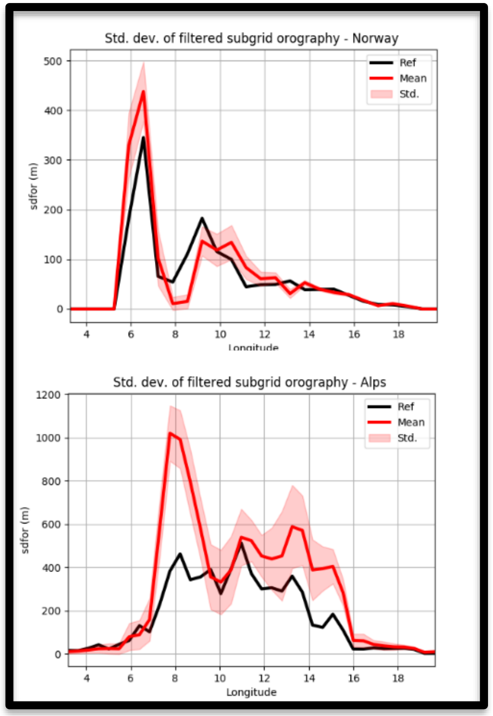
Which can be applied back during <u>poorly</u> observed periods to improve the reanalysis

Data Assimilation: Improving models

Application of parameter estimation to improving the standard deviation of model sub-grid orography



Mean analysis increments applied to SGO



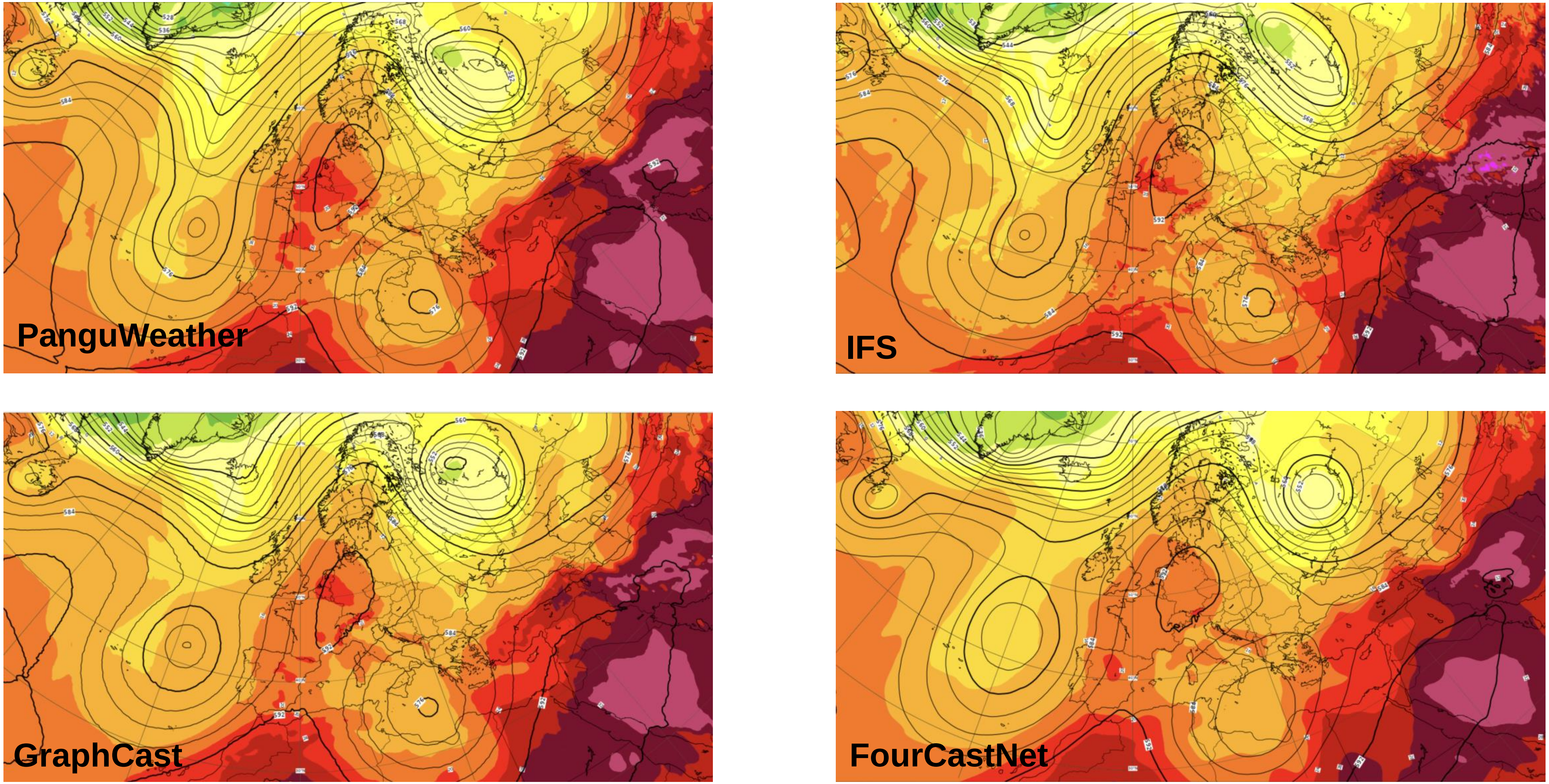
Original (black) and optimized (red) model parameter values for Norway (top) and Alps (lower)

- Surface pressure observations are very sensitive to changes in the assumed model sub-grid orography (SGO) which is <u>part of the observation operator</u>
- Adding SGO as an <u>augmentation</u> of the 4D-Var control vector (XCV formulation) allows the observations to improve this parameter of the model
- Use of the 4D-Var <u>optimized</u> SGO parameter in medium-range forecasts improves skill

Data Assimilation: Where next ?

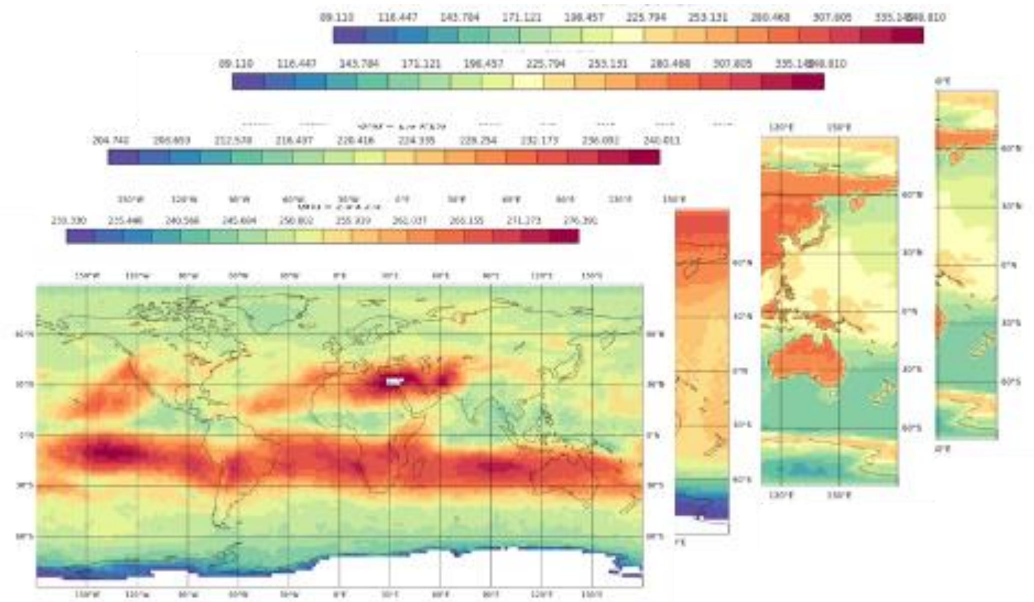
- Km Scale DA systems
- Using DA to improve models (weather and climate)
- **The rise of the machines**

Data Assimilation: The backbone of ML forecasts

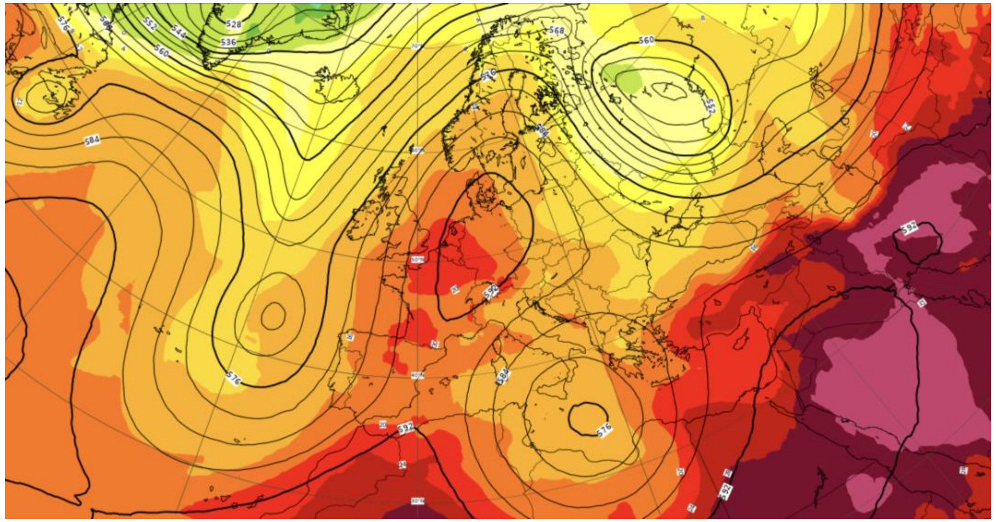


Data Assimilation…will machines take over ?

Observations



Forecasts



Summary

- Data Assimilation is an incredibly powerful technique to bring together different sources of knowledge

- It has played a key role in advancing **weather** forecast accuracy, understanding a changing **climate** (but also many other applications I did not mention)

- Data Assimilation heading to ever **higher resolution** to support models and extract the most from satellite observing systems

- Machine learning is accelerating but also **challenging** our approach to NWP and Data Assimilation

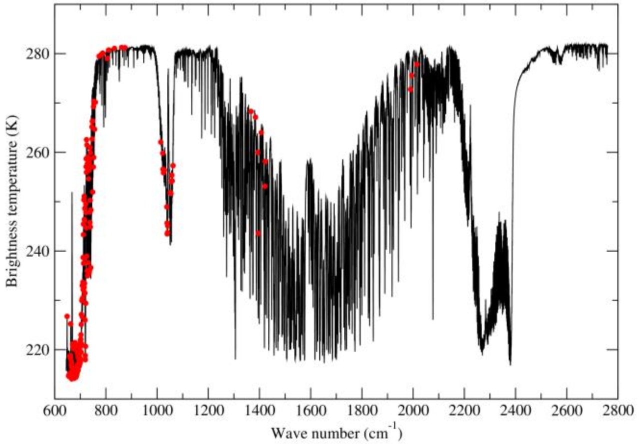
But…How does the QJRMS adapt ?

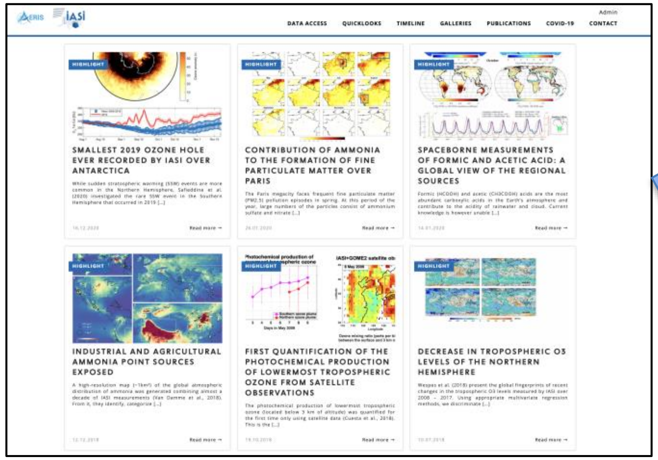
- ML is advancing at an astonishing rate
- It is disruptive technology in the very best sense of the word
- It has its own literature base (at least for the time being) but this literature looks very different to a science publication
- ML does science …without understanding the science!!
- Many commercial technology players are deeply involved in this area
- **How does the established scientific literature adapt, and does it have a role to play in the future ?**

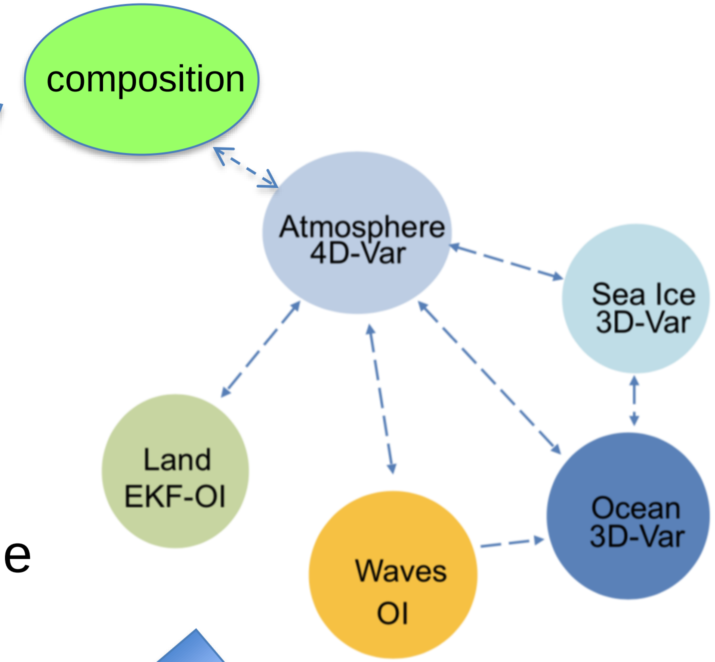
Quarterly Journal of the
Royal Meteorological Society

RMetS

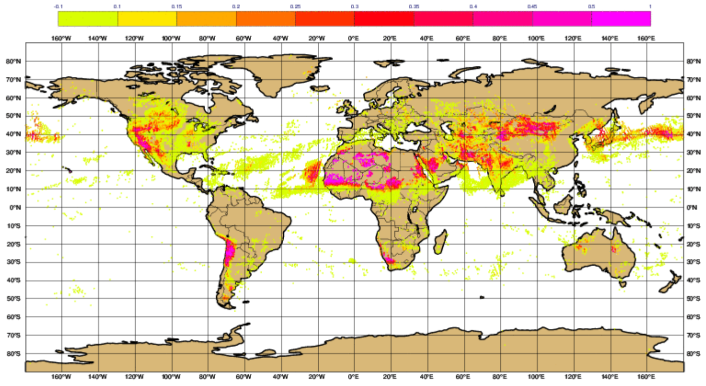
Chemical species providing wind information



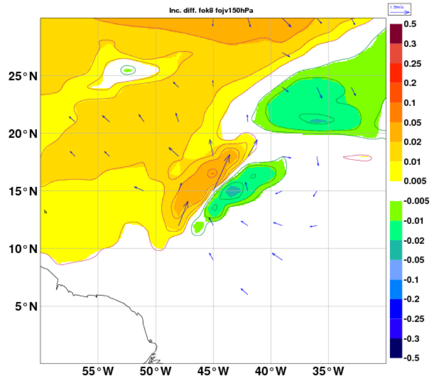




Advection wind tracing from dust aerosol ?

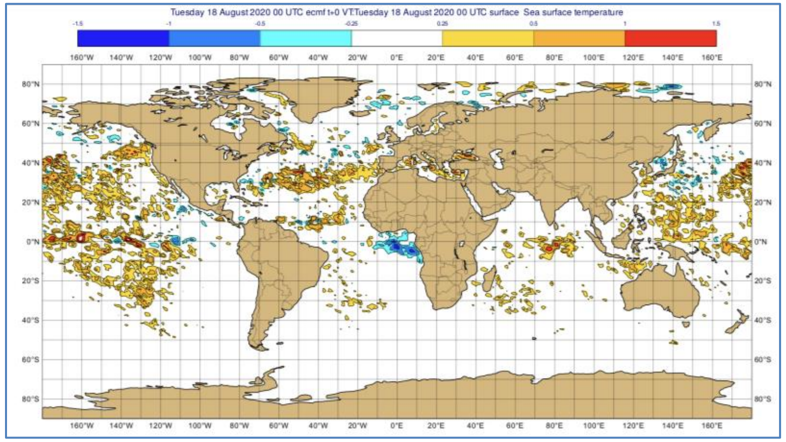


Advection wind tracing from ozone in the stratosphere. *(C Lupu)*

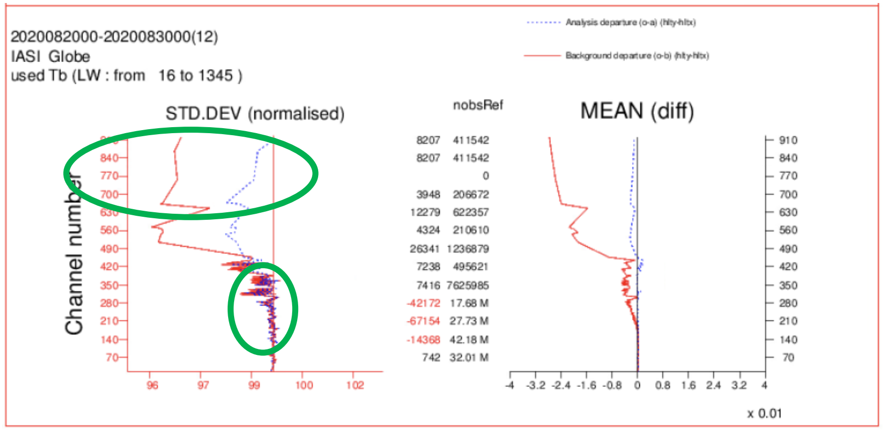


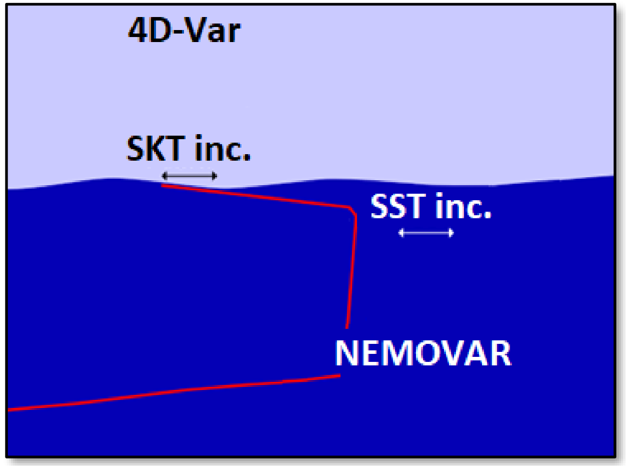
Coupled radiance based SST analysis (RADSST)

NEMOVAR SST changes forced by IASI

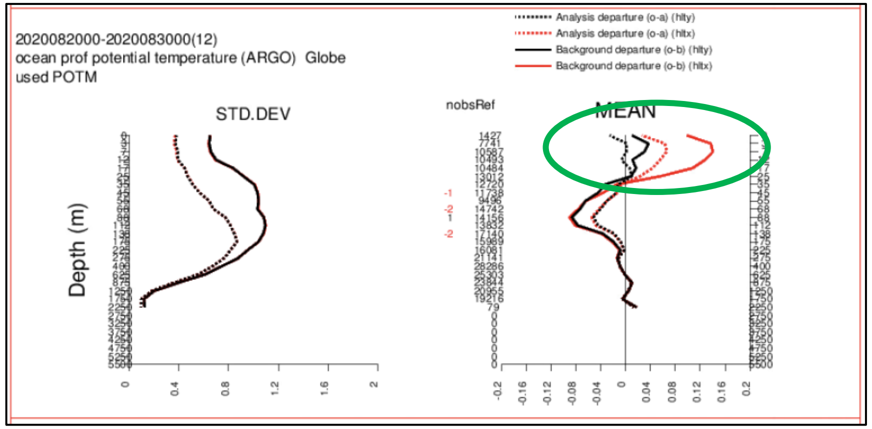


Changes have ***memory*** in the ocean and feed back to improve IASI use in the atmosphere





ARGO floats



Assimilating IASI in RADSST produces a better fit to surface **and** sub-surface **in situ ocean observations** which simultaneously **anchor the IASI assimilation**