

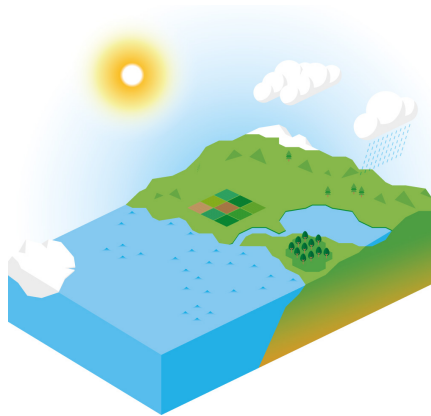
# ECMWF Data Assimilation Training Course 2023

## Coupled Data Assimilation

Phil Browne

2023-05-25

# Components of ECMWF forecasts



Components of ECMWF's Earth system.  
Along with the atmosphere, there are the  
ocean, wave, sea ice, land surface, snow, and lake models.

Introduction and basis of coupled assimilation

Mathematical and computational challenges in coupled data assimilation

Coupled Assimilation terminology

Physical and practical challenges in coupled data assimilation

Coupled assimilation at ECMWF

Coupled assimilation configurations in reanalysis and future NWP

Overview and summary of coupled assimilation at ECMWF

# Why do coupled assimilation?

## Observation impact

Observations of one component of the system can update another, e.g.

- ▶ Observing surface winds could directly impact on the land surface model, e.g. through updating both surface temperatures and evaporation rates
- ▶ Observations of passive (chemical) tracers can be used to update atmospheric winds

## More consistent analyses

By treating components together they may become more *balanced*.

- ▶ This can lead to a reduction in fast adjustments at the start of forecasts, known as *initialisation shock*

## Allows the use of a dynamical model for different components

## What are the alternatives to coupled assimilation?

- ▶ Completely ignore all processes in other components
- ▶ Use fixed climatology for other components
- ▶ Use someone else's analysis of the other components, e.g. Level 4 products as a boundary condition

One can write the 4D-Var cost function as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}_b - \mathbf{x})^T P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + \frac{1}{2} \sum_k (\mathbf{y}_k - \mathcal{G}_k(\mathbf{x}))^T R_k^{-1}(\mathbf{y}_k - \mathcal{G}_k(\mathbf{x}))$$

and its gradient

$$-\nabla J(\mathbf{x}) = P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + \sum_k M_k^T H_k^T R_k^{-1}(\mathbf{y}_k - \mathcal{G}_k(\mathbf{x})).$$

Cycling of 4D-Var can be written as:

$$\mathbf{x}_b = \mathcal{M}(\mathbf{x}_a)$$

Question: which terms in these equations can transfer information from one component of  $\mathbf{x}$  to another?

## A simple example 1

Suppose we model the temperature in the room, but we split the room in half and have one temperature for each half;  $x_1$  and  $x_2$ .

Now let us measure the temperature in each half of the room;  $y_1$  and  $y_2$ .

## A simple example 1

Suppose we model the temperature in the room, but we split the room in half and have one temperature for each half;  $x_1$  and  $x_2$ .

Now let us measure the temperature in each half of the room;  $y_1$  and  $y_2$ .

### Uncoupled assimilation (3D-Var)



## A simple example 1

Suppose we model the temperature in the room, but we split the room in half and have one temperature for each half;  $x_1$  and  $x_2$ .

Now let us measure the temperature in each half of the room;  $y_1$  and  $y_2$ .

### Uncoupled assimilation (3D-Var)

$$\mathbf{x} = x_1, \mathbf{y} = y_1, H = 1, R = \sigma_{y_1}^2, P_b = \sigma_{x_1}^2$$

$$J_1(\mathbf{x}) = (x_{b_1} - x_1)\sigma_{x_1}^{-2}(x_{b_1} - x_1) + (y_1 - x_1)\sigma_{y_1}^{-2}(y_1 - x_1)$$

## A simple example 1

Suppose we model the temperature in the room, but we split the room in half and have one temperature for each half;  $x_1$  and  $x_2$ .

Now let us measure the temperature in each half of the room;  $y_1$  and  $y_2$ .

### Uncoupled assimilation (3D-Var)

$$\mathbf{x} = x_1, \mathbf{y} = y_1, H = 1, R = \sigma_{y_1}^2, P_b = \sigma_{x_1}^2$$

$$J_1(\mathbf{x}) = (x_{b_1} - x_1)\sigma_{x_1}^{-2}(x_{b_1} - x_1) + (y_1 - x_1)\sigma_{y_1}^{-2}(y_1 - x_1)$$

$$\mathbf{x} = x_2, \mathbf{y} = y_2, H = 1, R = \sigma_{y_2}^2, P_b = \sigma_{x_2}^2$$

$$J_2(\mathbf{x}) = (x_{b_2} - x_2)\sigma_{x_2}^{-2}(x_{b_2} - x_2) + (y_2 - x_2)\sigma_{y_2}^{-2}(y_2 - x_2)$$

### Coupled assimilation (3D-Var)

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, H = I, R = \begin{bmatrix} \sigma_{y_1}^2 & 0 \\ 0 & \sigma_{y_2}^2 \end{bmatrix}, P_b = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_{12}}^2 \\ \sigma_{x_{12}}^2 & \sigma_{x_2}^2 \end{bmatrix}$$

$$J(\mathbf{x}) = \begin{bmatrix} x_{b_1} - x_1 \\ x_{b_2} - x_2 \end{bmatrix}^T \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_{12}}^2 \\ \sigma_{x_{12}}^2 & \sigma_{x_2}^2 \end{bmatrix}^{-1} \begin{bmatrix} x_{b_1} - x_1 \\ x_{b_2} - x_2 \end{bmatrix} \\ + (y_1 - x_1)\sigma_{y_1}^{-2}(y_1 - x_1) + (y_2 - x_2)\sigma_{y_2}^{-2}(y_2 - x_2)$$

## A simple example 3

Suppose we stop observing  $y_2$ .

- ▶ In uncoupled 3D-Var,  $x_2 = x_{b_2}$ , i.e. nothing happens for this variable
- ▶ In coupled 3D-Var  $x_2$  is still updated if  $\sigma_{x_{12}}^2 \neq 0$

## A simple example 3

Suppose we stop observing  $y_2$ .

- ▶ In uncoupled 3D-Var,  $x_2 = x_{b_2}$ , i.e. nothing happens for this variable
- ▶ In coupled 3D-Var  $x_2$  is still updated if  $\sigma_{x_{12}}^2 \neq 0$

If we never observe  $y_2$ , the *cross-covariance*  $\sigma_{x_{12}}^2$  allows us to constrain  $x_2$ .

Recall the 3D-Var cost function:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}_b - \mathbf{x})^T P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathbf{x}))^T R^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x}))$$

and its gradient

$$-\nabla J(\mathbf{x}) = P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + H^T R^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x}))$$

There are two ways  $x_2$  can influence  $x_1$ :

- ▶  $H$  is a function of both  $x_1$  and  $x_2$
- ▶  $P_{b12} \neq 0$

Introduction and basis of coupled assimilation

**Mathematical and computational challenges in coupled data assimilation**

Coupled Assimilation terminology

Physical and practical challenges in coupled data assimilation

Coupled assimilation at ECMWF

Coupled assimilation configurations in reanalysis and future NWP

Overview and summary of coupled assimilation at ECMWF

Received: 21 December 2021 | Revised: 11 May 2022 | Accepted: 12 May 2022 | Published on: 10 August 2022

DOI: 10.1002/qj.4330

Quarterly Journal of the  
Royal Meteorological Society 

## RESEARCH ARTICLE

# Coupled data assimilation at ECMWF: current status, challenges and future developments

Patricia de Rosnay<sup>1</sup>  | Philip Browne<sup>1</sup>  | Eric de Boissésou<sup>1</sup>  | David Fairbairn<sup>1</sup>  |  
Yoichi Hirahara<sup>1,2</sup>  | Kenta Ochi<sup>1,2</sup> | Dinand Schepers<sup>1</sup>  | Peter Weston<sup>1</sup>  |  
Hao Zuo<sup>1</sup>  | Magdalena Alonso-Balmaseda<sup>1</sup>  | Gianpaolo Balsamo<sup>1</sup>  |  
Massimo Bonavita<sup>1</sup>  | Niels Borman<sup>1</sup>  | Andy Brown<sup>1</sup> | Marcin Chrust<sup>1</sup> |  
Mohamed Dahoui<sup>1</sup>  | Giovanna Chiara<sup>1</sup>  | Stephen English<sup>1</sup>  | Alan Geer<sup>1</sup>  |  
Sean Healy<sup>1</sup>  | Hans Hersbach<sup>1</sup>  | Patrick Laloyaux<sup>1</sup>  |  
Linus Magnusson<sup>1</sup>  | Sébastien Massart<sup>1</sup>  | Anthony McNally<sup>1</sup> |  
Florian Pappenberger<sup>1</sup>  | Florence Rabier<sup>1</sup> 

Patricia de Rosnay et al. (2022). "Coupled data assimilation at ECMWF: current status, challenges and future developments". In: *Quarterly Journal of the Royal Meteorological Society* 148.747, pp. 2672–2702



## Reminder: 4D-Var equations

One can write the 4D-Var cost function as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}_b - \mathbf{x})^T P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + \frac{1}{2} \sum_k (\mathbf{y}_k - \mathcal{G}_k(\mathbf{x}))^T R_k^{-1}(\mathbf{y}_k - \mathcal{G}_k(\mathbf{x}))$$

and its gradient

$$-\nabla J(\mathbf{x}) = P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + \sum_k M_k^T H_k^T R_k^{-1}(\mathbf{y}_k - \mathcal{G}_k(\mathbf{x})).$$

Cycling of 4D-Var can be written as:

$$\mathbf{x}_b = \mathcal{M}(\mathbf{x}_a)$$

# Why coupled assimilation and not just multivariate assimilation?

- ▶  $P_b$  — background error covariance coupling
- ▶  $\mathcal{G} = \mathcal{H}(\mathcal{M})$ 
  - ▶  $\mathcal{H}$  — observation operator coupling
  - ▶  $\mathcal{M}$  — coupled nonlinear trajectories
- ▶  $M^T$  — coupled TL/AD
- ▶  $H^T$  — coupled Jacobian (of observation operator)
- ▶  $R$  — coupled observation error covariance matrix
- ▶  $\mathcal{M}$  — coupled model for cycling the analyses

# Why coupled assimilation and not just multivariate assimilation?

- ▶  $P_b$  — background error covariance coupling
- ▶  $\mathcal{G} = \mathcal{H}(\mathcal{M})$ 
  - ▶  $\mathcal{H}$  — observation operator coupling
  - ▶  $\mathcal{M}$  — coupled nonlinear trajectories
- ▶  $M^T$  — coupled TL/AD
- ▶  $H^T$  — coupled Jacobian (of observation operator)
- ▶  $R$  — ~~coupled observation error covariance matrix~~
- ▶  $\mathcal{M}$  — coupled model for cycling the analyses

# Coupled assimilation challenges 1: The state vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$$

The components  $\mathbf{x}_1$  and  $\mathbf{x}_2$  can

- ▶ Be on different grids
- ▶ Have different representations (eg. spectral vs grid point)
- ▶ Stored on different processors
- ▶ Stored in very different places in the code (i.e. different modules)
- ▶ Not even represented in the same code!

So there can be huge technical challenges to overcome before one can even write down a coupled state vector.

$$P_b = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix}$$

Assuming we can specify  $P_{11}$  and  $P_{22}$ , then we have to specify the cross covariance terms  $P_{12}$ .

We have all the same problems as for the state vector: grids, storage, etc.

$P_b$  also needs to be a covariance matrix, i.e. be positive definite.

These problems go away if  $P_{12} = \mathbf{0}$ .

Ensemble methods for representing  $P_{12}$  are promising.

$$\mathcal{H} = \mathcal{H} \left( \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \right)$$

If the problems of the state vector can be solved, then a coupled observation operator should be possible.

This then requires  $H^T$ , a coupled Jacobian, which is a technical exercise to ensure its availability.

$$\mathcal{M} = \mathcal{M} \left( \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \right)$$

This is typically the *easiest* form of coupling to use in coupled data assimilation. This is because the model is generally developed before the assimilation technique is put in place.

We will return to this later on.

## Coupled assimilation challenges 5: TL/AD coupling

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \quad \text{i.e.} \quad M\mathbf{x} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} M_{11}\mathbf{x}_1 + M_{12}\mathbf{x}_2 \\ M_{21}\mathbf{x}_1 + M_{22}\mathbf{x}_2 \end{bmatrix}$$

$$M^T = \begin{bmatrix} M_{11}^T & M_{21}^T \\ M_{12}^T & M_{22}^T \end{bmatrix} \quad \text{i.e.} \quad M^T\mathbf{x} = \begin{bmatrix} M_{11}^T & M_{21}^T \\ M_{12}^T & M_{22}^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} M_{11}^T\mathbf{x}_1 + M_{21}^T\mathbf{x}_2 \\ M_{12}^T\mathbf{x}_1 + M_{22}^T\mathbf{x}_2 \end{bmatrix}$$

The coupled TL/AD terms  $M_{12}$  and  $M_{21}$  may not be implemented.



Introduction and basis of coupled assimilation

Mathematical and computational challenges in coupled data assimilation

**Coupled Assimilation terminology**

Physical and practical challenges in coupled data assimilation

Coupled assimilation at ECMWF

Coupled assimilation configurations in reanalysis and future NWP

Overview and summary of coupled assimilation at ECMWF

Recall the gradient of the 4D-Var cost function

$$-\nabla J(\mathbf{x}) = P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + \sum_k M_k^T H_k^T R_k^{-1}(\mathbf{y}_k - \mathcal{G}_k(\mathbf{x})).$$

Then if any of  $P_b$ ,  $M$ ,  $H$ , or  $\mathcal{G} = \mathcal{H}(\mathcal{M})$  have coupling terms then an observation in one component will modify the state of the other *within the assimilation cycle*.

Recall the gradient of the 4D-Var cost function

$$-\nabla J(\mathbf{x}) = P_b^{-1}(\mathbf{x}_b - \mathbf{x}) + \sum_k M_k^T H_k^T R_k^{-1}(\mathbf{y}_k - \mathcal{G}_k(\mathbf{x})).$$

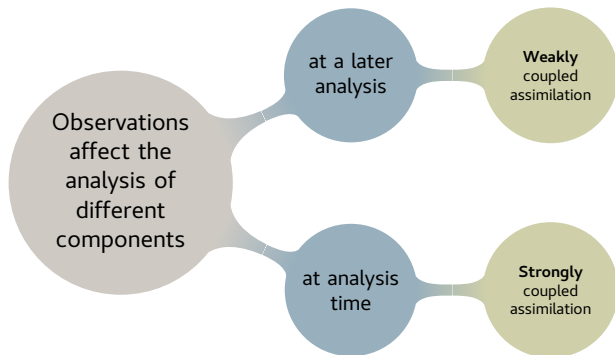
Then if any of  $P_b$ ,  $M$ ,  $H$ , or  $\mathcal{G} = \mathcal{H}(\mathcal{M})$  have coupling terms then an observation in one component will modify the state of the other *within the assimilation cycle*.

Suppose none of these terms have coupling terms, but the model which cycles the analysis,

$$\mathbf{x}_b = \mathcal{M}(\mathbf{x}_a),$$

is coupled. Then an observation in one component will modify the state of the other *in the next the assimilation cycle*.

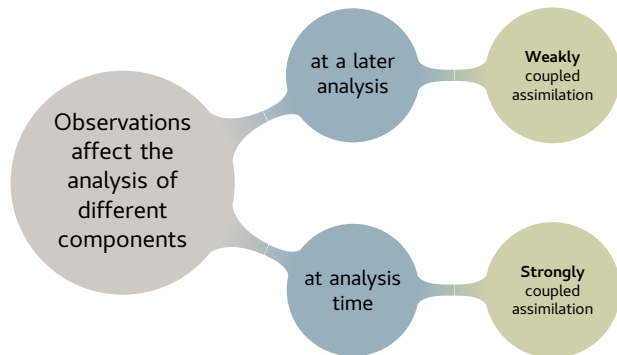
# Coupled DA categories



## Formal definitions are available in

Stephen G Penny et al. (2017). *Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations*. World Meteorological Organisation. URL: <https://community.wmo.int/wwrp-publications>

# Coupled DA categories



Formal definitions are available in

Stephen G Penny et al. (2017). *Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations*. World Meteorological Organisation. URL: <https://community.wmo.int/wwrp-publications>

Note that these are only broad categorisations, they tell you nothing about which terms are coupled.

# Coupled assimilation in the ensemble Kalman filter

One can write the EnKF update as

$$\begin{aligned}\mathbf{x}_t^a &= \mathbf{x}_t^b + P^b H^T (H P^b H^T + R)^{-1} (y - \mathcal{H}(\mathbf{x}_t^b)) \\ &= \mathcal{M}(\mathbf{x}_{t-1}^a) + P^b H^T (H P^b H^T + R)^{-1} (y - \mathcal{H}(\mathcal{M}(\mathbf{x}_{t-1}^a)))\end{aligned}$$

1. If the observation operator is coupled then we have immediate cross component impact.
2. If the observation operator is not coupled then for an observation to have immediate cross component impact  $P^b$  must have nonzero diagonal terms.
3. If neither of these conditions hold then the nonlinear model  $\mathcal{M}$  will mix the information as the ensemble members are propagated from one analysis time to the next.

Hence (1) and (2) are a form of strongly coupled assimilation.

The case (3) would fall under weakly coupled assimilation.

Note that localisation methods across the different components need to be specified!

Introduction and basis of coupled assimilation

Mathematical and computational challenges in coupled data assimilation

Coupled Assimilation terminology

Physical and practical challenges in coupled data assimilation

Coupled assimilation at ECMWF

Coupled assimilation configurations in reanalysis and future NWP

Overview and summary of coupled assimilation at ECMWF

## Long assimilation windows

- 😊 The longer the assimilation window, the more observations we get to put into our systems
- 😊 The longer the assimilation window, the more flow dependence we obtain in our solution - i.e. we become less reliant on the background error covariance that we specify at  $t = 0$ .
- 😞 The longer the assimilation window, the more nonlinearities make the tangent linear approximation worse.
- 😞 The longer the assimilation window, the more model error needs to be accounted for.



# Timescales in the Earth System

- ▶ Microscale turbulence minutes
- ▶ Mesoscale storms (tornadoes/thunderstorms) hours
- ▶ Synoptic scale cyclones days
- ▶ Planetary waves/blocking structures weeks
- ▶ Intraseasonal features months
- ▶ Seasonal cycles/ENSO years

# Timescales in the Earth System

▶ Microscale turbulence	minutes
▶ Mesoscale storms (tornadoes/thunderstorms)	hours
▶ Synoptic scale cyclones	days
▶ Planetary waves/blocking structures	weeks
▶ Intraseasonal features	months
▶ Seasonal cycles/ENSO	years
▶ Internal waves	hours
▶ Tides	days
▶ Mesoscale eddies	weeks/months
▶ ENSO	years
▶ Thermohaline circulation	centuries

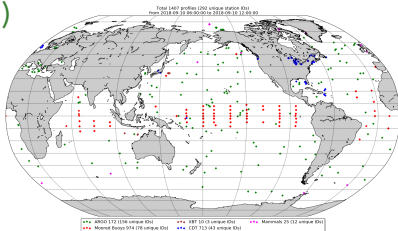
# Tangent linear model and approximation

- ▶ Observations of one system may not be available within the assimilation window for the faster components.
- ▶ You won't get impact from observations that have not arrived.

# Ocean observation latency

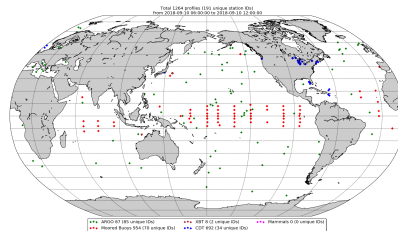
## 49 hour cut-off

172 (156)  
 974 (78)  
 10 (3)  
 713 (43)  
 25 (12)

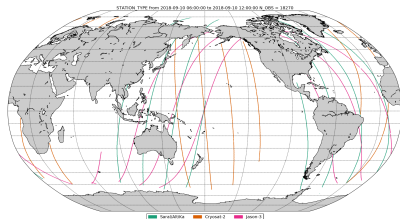


## 1 hour cut-off

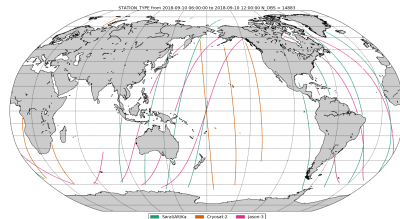
87 (85)  
 954 (70)  
 8 (2)  
 692 (34)  
 0 (0)



18270

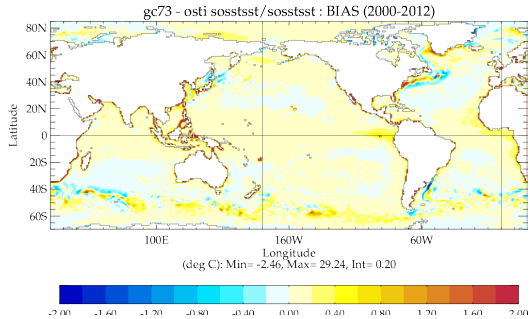


14883

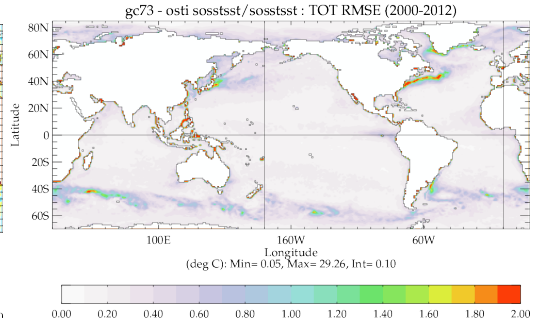


- ▶ Can you afford to devote the same number of resources to every component of your system?  
e.g. At ECMWF we have 50 perturbed members in our atmospheric EDA.  
We have 4 perturbed members in the ocean EDA!
- ▶ Do you have similar observation networks in all the components?  
These might make your assimilation more or less nonlinear, and more or less Gaussian in the different components.
- ▶ Can you use the same strategies for modelling your assimilation operators?  
e.g. spectral transforms that work in the atmosphere might not work in the ocean or on land
- ▶ Would you introduce a **bias** in the system going to coupled assimilation that uncoupled assimilation does not have?

## BIAS



## RMSE



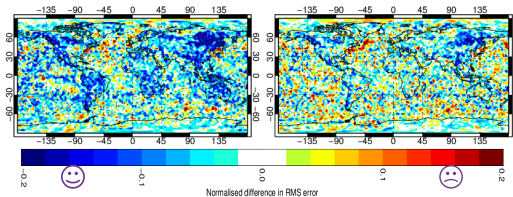
# Coupled DA done badly - SST biases in the gulf stream

Change in error in Z10U (Winter analysis coupled assimilation – control)

6–Dec–2016 to 28–Feb–2017 from 150 to 169 samples. Verified against own-analysis.

T+12

T+24

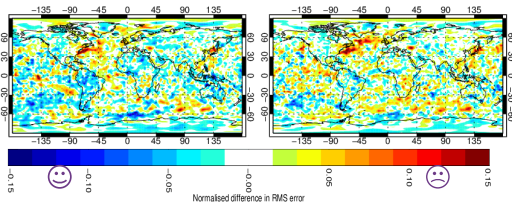


Change in error in VW (Winter analysis coupled assimilation – control)

6–Dec–2016 to 28–Feb–2017 from 150 to 169 samples. Verified against own-analysis.

T+12; 1000hPa

T+24; 1000hPa

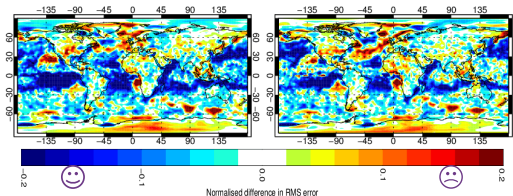


Change in error in T (Summer analysis coupled assimilation – control)

6–Jun–2017 to 31–Aug–2017 from 154 to 173 samples. Verified against own-analysis.

T+12; 1000hPa

T+24; 1000hPa

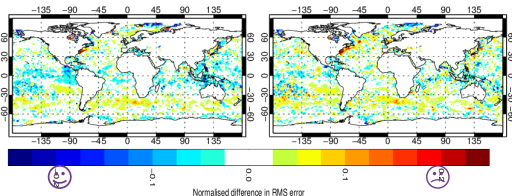


Change in error in SWH (Winter analysis coupled assimilation – control)

6–Dec–2016 to 28–Feb–2017 from 150 to 169 samples. Verified against own-analysis.

T+12

T+24





Introduction and basis of coupled assimilation

Mathematical and computational challenges in coupled data assimilation

Coupled Assimilation terminology

Physical and practical challenges in coupled data assimilation

**Coupled assimilation at ECMWF**

Coupled assimilation configurations in reanalysis and future NWP

Overview and summary of coupled assimilation at ECMWF

## Weakly coupled assimilation

- ▶ The concept of weakly coupled assimilation is very flexible.
- ▶ It allows different components to be assimilated separately and information mixed by the model.
- ▶ It also allows for different assimilation techniques to be used for each component.
- ▶ It doesn't actually require the assimilation windows of each component to be aligned.

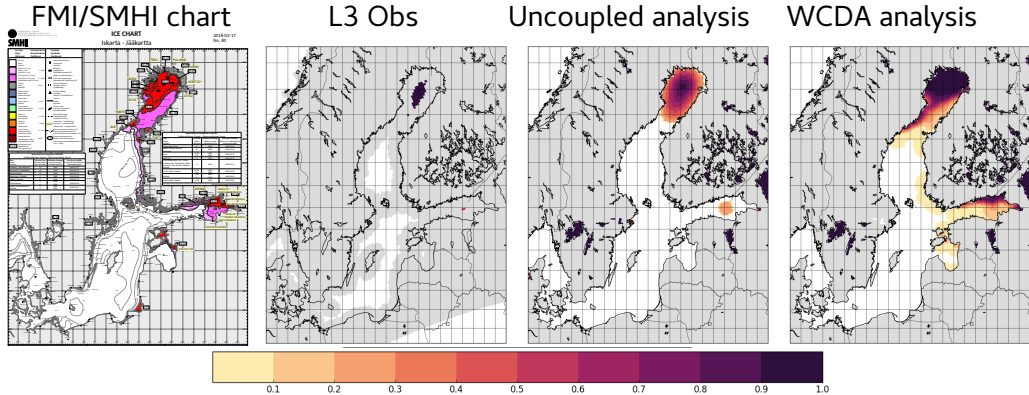
We use 2D-OI for the wave analysis and snow depth analysis.

We use 1D-OI for the soil temperature analysis and the SEKF for soil moisture analysis.

We use 3D-Var FGAT for the ocean and sea ice analysis.

See next lecture by Patricia and tomorrow's lectures by Hao and Richard

# WCDA example - sea ice in the Baltic Sea



Sea ice concentrations in the Baltic sea on 20180217. [FMI 2018]

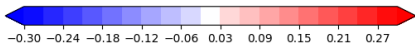
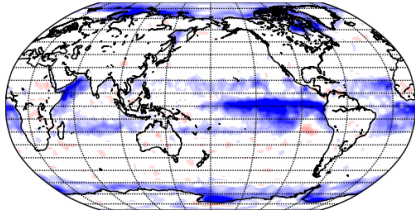
## New possible biases with coupled assimilation

- ▶ If a coupled  $\mathcal{M}$  has different biases compared to the uncoupled version, then these can be passed between the components of the analysis.
- ▶ In our ocean model we have more *effective* resolution in the tropics (due to latitudinal variability of the Rossby radius of deformation) than the extratropics.
- ▶ As such the model performs less well in the extratropics, and so to not degrade the atmospheric analysis we do not couple the ocean and atmosphere here.

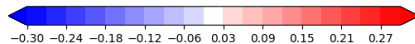
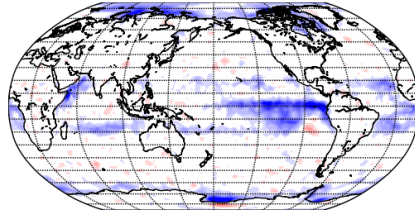
Philip A Browne et al. (2019). "Weakly Coupled Ocean-Atmosphere Data Assimilation in the ECMWF NWP System". In: *Remote Sensing* 11 (234), pp. 1-24. DOI: [10.3390/rs11030234](https://doi.org/10.3390/rs11030234)

# WCDA maps of surface temperature forecast errors

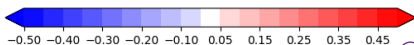
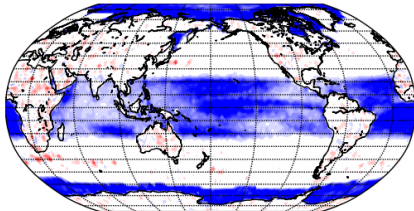
Normalised difference in rms error of T at 1000hPa T+12hrs



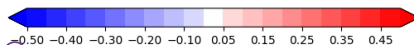
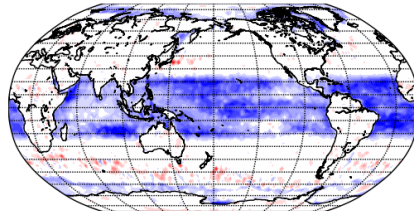
Normalised difference in rms error of T at 1000hPa T+48hrs



Normalised difference in rms error of SKT T+12hrs



Normalised difference in rms error of SKT T+120hrs



Introduction and basis of coupled assimilation

Mathematical and computational challenges in coupled data assimilation

Coupled Assimilation terminology

Physical and practical challenges in coupled data assimilation

Coupled assimilation at ECMWF

**Coupled assimilation configurations in reanalysis and future NWP**

Overview and summary of coupled assimilation at ECMWF

## Outer loop coupling (1)

One approach is known as **outer loop coupling** 4D-Var in which  $\mathcal{M}$  is coupled, i.e. we use coupled trajectories and the coupled model to cycle the analysis.

We neglect all coupling terms in the TL and AD:

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} := \begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix} \quad \text{i.e.} \quad M\mathbf{x} = \begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} M_{11}x_1 \\ M_{22}x_2 \end{bmatrix}$$

$$M^T = \begin{bmatrix} M_{11}^T & M_{21}^T \\ M_{12}^T & M_{22}^T \end{bmatrix} := \begin{bmatrix} M_{11}^T & 0 \\ 0 & M_{22}^T \end{bmatrix} \quad \text{i.e.} \quad M^T\mathbf{x} = \begin{bmatrix} M_{11}^T & 0 \\ 0 & M_{22}^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} M_{11}^T x_1 \\ M_{22}^T x_2 \end{bmatrix}$$

## Outer loop coupling (2)

Recall the **linearisation** state  $\mathbf{x}^{(m)}$  such that

$$\mathbf{x} = \mathbf{x}^{(m)} + \delta\mathbf{x}^{(m)}$$

Then the cost function becomes

$$\begin{aligned} J(\delta\mathbf{x}^{(m)}) &= \frac{1}{2}(\mathbf{x}_b - \mathbf{x}^{(m)} - \delta\mathbf{x}^{(m)})^T P_b^{-1}(\mathbf{x}_b - \mathbf{x}^{(m)} - \delta\mathbf{x}^{(m)}) \\ &\quad + \frac{1}{2}(\mathbf{d}^{(m)} - G\delta\mathbf{x}^{(m)})^T R^{-1}(\mathbf{d}^{(m)} - G\delta\mathbf{x}^{(m)}) \end{aligned}$$

where  $\mathbf{d}^{(m)} = \mathbf{y} - \mathcal{G}(\mathbf{x}^{(m)})$



## Outer loop coupling (3)

If we do not specify background error coupling terms (i.e.  $P_b$  block diagonal), the TL/AD is not coupled, and we have no coupled observation operators, then the cost function is separable:

$$J(\delta \mathbf{x}^{(m)}) = J \left( \begin{bmatrix} \delta \mathbf{x}_1^{(m)} \\ \delta \mathbf{x}_2^{(m)} \end{bmatrix} \right) = J_1(\delta \mathbf{x}_1^{(m)}) + J_2(\delta \mathbf{x}_2^{(m)})$$

Exercise: show this!

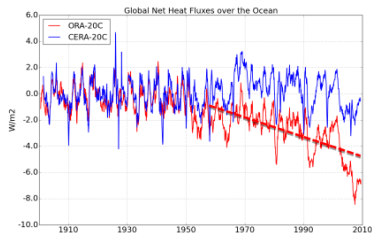
Thus the interaction between components happens through  $\mathcal{G}$  each **outer loop** of the minimisation.

This method formed the basis of the CERA system which produced the CERA-20C and CERA-SAT reanalyses.

Patrick Laloyaux et al. (2018). "CERA-20C: A Coupled Reanalysis of the Twentieth Century". In: *Journal of Advances in Modeling Earth Systems* 10 (5), pp. 1172–1195. DOI: [10.1029/2018MS001273](https://doi.org/10.1029/2018MS001273)

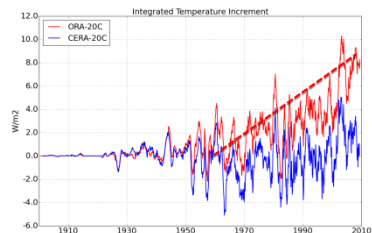
Dinand Schepers et al. (2018). CERA-SAT: A coupled satellite-era reanalysis. DOI: [10.21957/sp619ds74g](https://doi.org/10.21957/sp619ds74g)

# Balanced ocean-atmosphere analysis



Global net air-sea fluxes toward the ocean in CERA-20C and ORA-20C.

- ▶ Spurious trend in ORA-20C probably due to shift in wind forcing in ERA-20C (heat lost)



Ocean temperature increment in CERA-20C and ORA-20C.

- ▶ Increment in ORA-20C is trying to compensate for the heat lost
- ▶ CERA-20C fluctuates around zero suggesting a more balanced air-sea interface

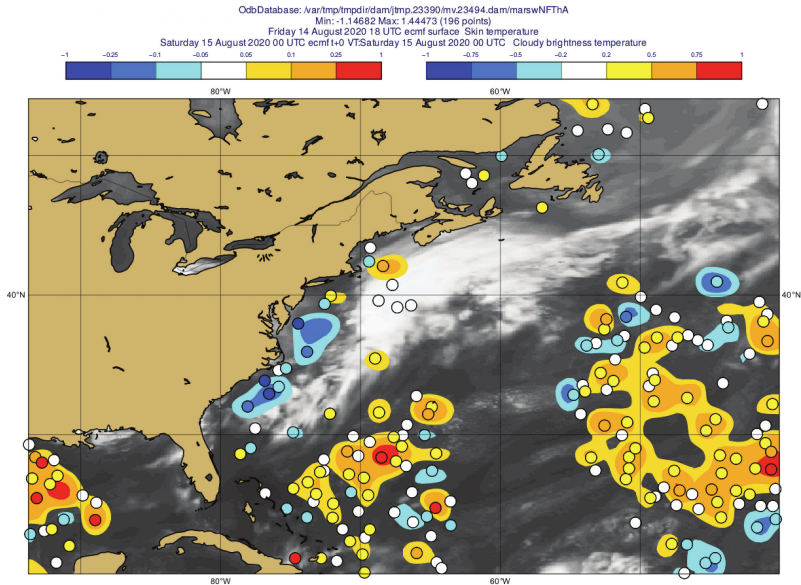
Courtesy of E. de Boisséson

# Outer loop coupling for Hurricanes Irma and Jose

Coupled assimilation

Uncoupled analysis (OSTIA)

# SKT assimilation with coupled observation operators



Introduction and basis of coupled assimilation

Mathematical and computational challenges in coupled data assimilation

Coupled Assimilation terminology

Physical and practical challenges in coupled data assimilation

Coupled assimilation at ECMWF

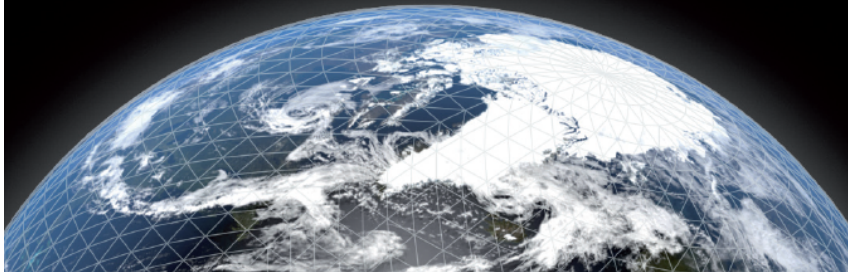
Coupled assimilation configurations in reanalysis and future NWP

Overview and summary of coupled assimilation at ECMWF



THE STRENGTH OF A COMMON GOAL

A ROADMAP TO 2025



“ECMWF has started to explore a new coupled assimilation system to initialise the numerical weather forecast in a more comprehensive and balanced manner. Such an approach has the potential to better use satellite measurements and to improve the quality of our forecasts. It will generate a reduction of initialisation shocks in coupled forecasts by fully accounting for interactions between the components. It will also lead to the generation of a consistent Earth-system state for the initialisation of forecasts across all timescales”,  
ECMWF Roadmap to 2025

ECMWF (2016). *The Strength of a Common Goal: A Roadmap To 2025*. ECMWF. URL: [https://www.ecmwf.int/sites/default/files/ECMWF\\_Roadmap\\_to\\_2025.pdf](https://www.ecmwf.int/sites/default/files/ECMWF_Roadmap_to_2025.pdf)

# ECMWF ROADMAP TO 2030



The strength of a common goal



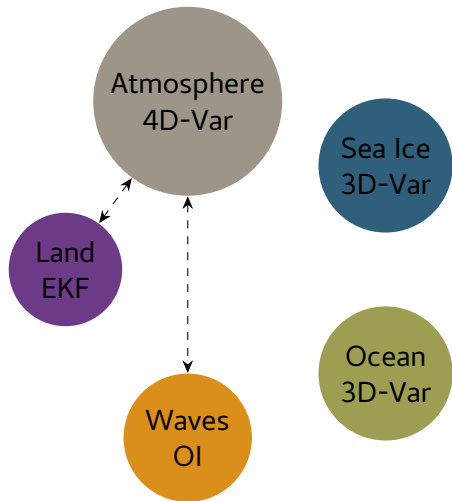


“The Centre will work towards overcoming challenges in the initialisation of ever more sophisticated Earth system components in its Integrated Forecasting System (IFS). It will do so through enhancing the consistency of the assimilation approach and optimal coupling between the various components of the Earth system and through advanced use of observations.”,  
ECMWF Roadmap to 2030

ECMWF (2021). *ECMWF: A Roadmap To 2030*. ECMWF. URL:  
<https://www.ecmwf.int/sites/default/files/medialibrary/2021-01/ecmwf-roadmap-to-2030.pdf>

# Components of the Earth System Analyses

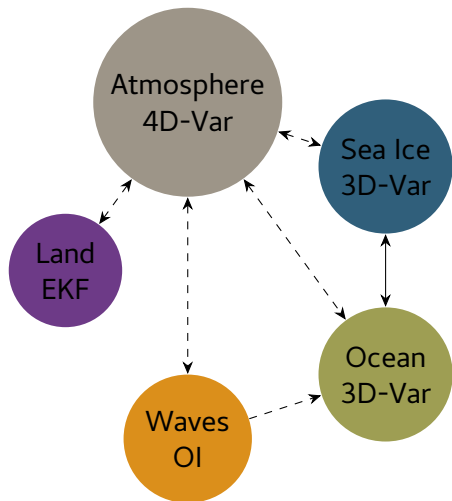
## Coupled Assimilation in ERA5



- ▶ ERA5: land and waves weakly coupled

# Components of the Earth System Analyses

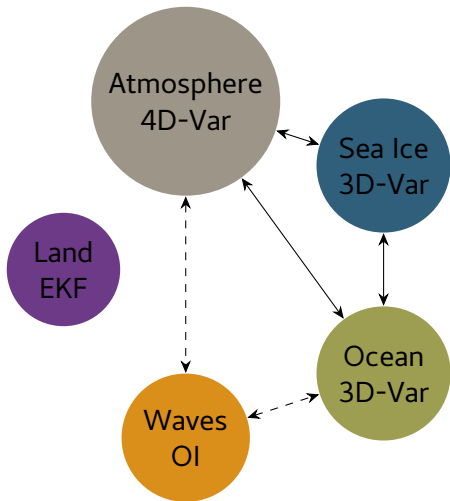
## Coupled Assimilation in NWP



► HRES NWP, EDA

# Components of the Earth System Analyses

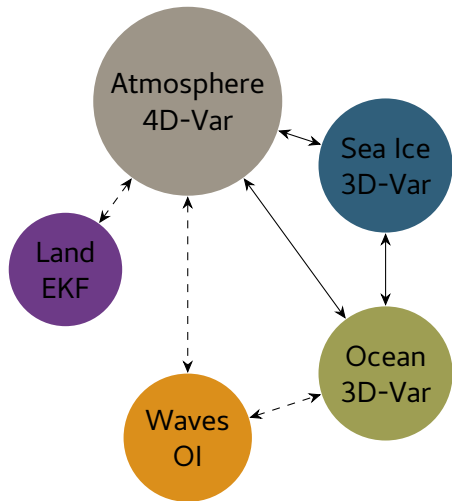
Coupled Assimilation in reanalyses and NWP research



- ▶ CERA-20C: outer loop coupling for atm-ocean, sea ice

# Components of the Earth System Analyses

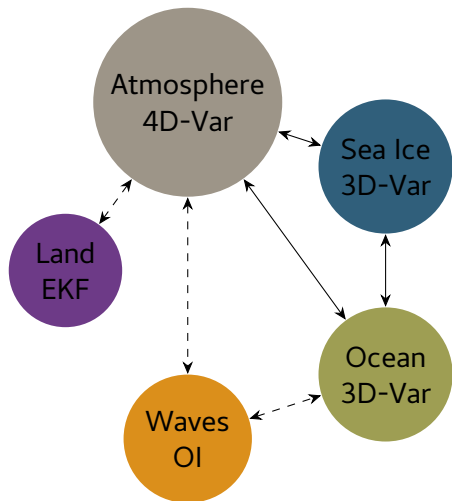
Coupled Assimilation in reanalyses and NWP research



- ▶ CERA-20C: outer loop coupling for atm-ocean, sea ice
- ▶ CERA-SAT with also land-atm weak coupling and full observing system

# Components of the Earth System Analyses

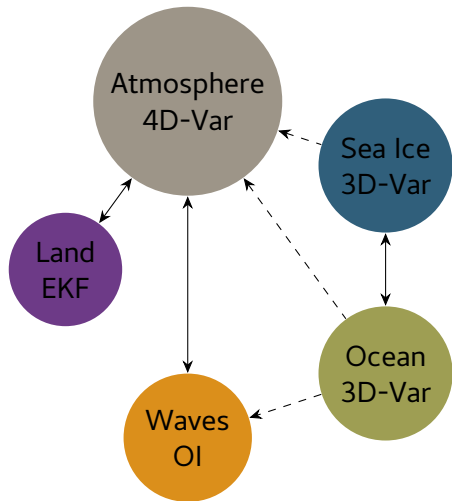
Coupled Assimilation in reanalyses and NWP research



- ▶ CERA-20C: outer loop coupling for atm-ocean, sea ice
- ▶ CERA-SAT with also land-atm weak coupling and full observing system
- ▶ Hence different coupling strategies are used for the different configurations

# Towards an Earth System Approach

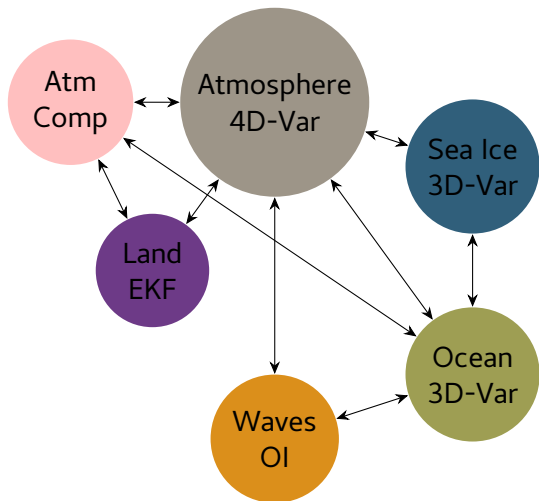
Coupled Assimilation - ERA6?



- ▶ Consistency of the coupling approaches across the different components of the Earth system
- ▶ Comprehensive Earth system approach; atmosphere, land, ocean, sea ice, waves

# Towards an Earth System Approach

Coupled Assimilation - ERA7?



- ▶ Consistency of the coupling approaches across the different components of the Earth system
- ▶ Comprehensive Earth system approach; atmosphere, land, ocean, sea ice, waves, atmospheric composition



- ▶ Coupled data assimilation is, **in theory**, the same as multivariate DA
- ▶ Coupled data assimilation can improve balance in analyses and can increase the use of, and information gained from, observations
- ▶ Issues arise from:
  - ▶ Varying timescales in the different components - leads to poor TL approximation for “long” windows
  - ▶ Various components of the Earth system running separate models/executables - the full adjoint is not always available
  - ▶ Observation availability for all the components
- ▶ Strongly coupled assimilation: observations can impact multiple components at **that** analysis
- ▶ Weakly coupled assimilation: observations can impact multiple components at **a later** analysis
- ▶ ECMWF is regularly doing coupled assimilation:
  - ▶ Which components are coupled depends on the system
  - ▶ We continue to increase the amount and strength of coupling between the Earth system components
- ▶ Specifying cross-covariances is a big future challenge

- Browne, Philip A et al. (2019). "Weakly Coupled Ocean-Atmosphere Data Assimilation in the ECMWF NWP System". In: *Remote Sensing* 11 (234), pp. 1-24. DOI: 10.3390/rs11030234.
- de Rosnay, Patricia et al. (2022). "Coupled data assimilation at ECMWF: current status, challenges and future developments". In: *Quarterly Journal of the Royal Meteorological Society* 148.747, pp. 2672-2702.
- ECMWF (2016). *The Strength of a Common Goal: A Roadmap To 2025*. ECMWF. URL: [https://www.ecmwf.int/sites/default/files/ECMWF\\_Roadmap\\_to\\_2025.pdf](https://www.ecmwf.int/sites/default/files/ECMWF_Roadmap_to_2025.pdf).
- (2021). *ECMWF: A Roadmap To 2030*. ECMWF. URL: <https://www.ecmwf.int/sites/default/files/medialibrary/2021-01/ecmwf-roadmap-to-2030.pdf>.
- FMI et al. (2018). *Baltic sea ice chart*. URL: <http://cdn.fmi.fi/marine-observations/products/ice-charts/20180217-full-color-ice-chart.pdf>.
- Laloyaux, Patrick et al. (2018). "CERA-20C: A Coupled Reanalysis of the Twentieth Century". In: *Journal of Advances in Modeling Earth Systems* 10 (5), pp. 1172-1195. DOI: 10.1029/2018MS001273.
- Penny, Stephen G et al. (2017). *Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations*. World Meteorological Organisation. URL: <https://community.wmo.int/wwrp-publications>.
- Schepers, Dinand et al. (2018). *CERA-SAT: A coupled satellite-era reanalysis*. DOI: 10.21957/sp619ds74g.