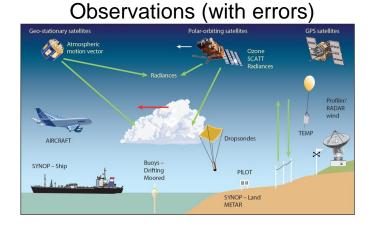
# Model error in data assimilation

**Patrick Laloyaux** 

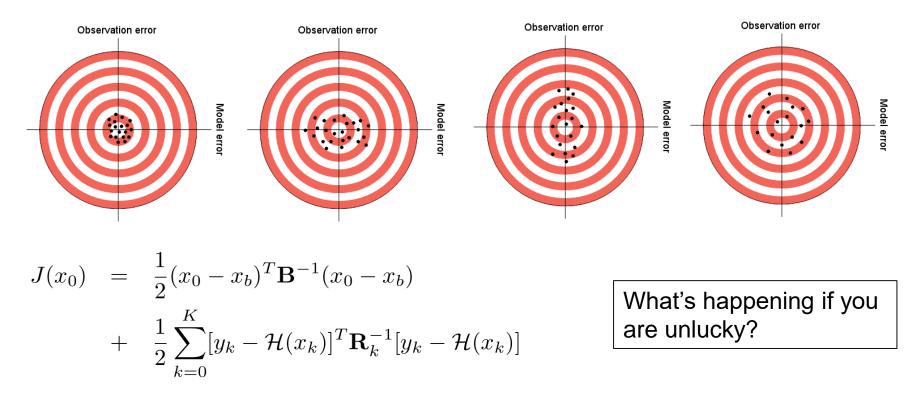
Learn about the typical issues that occur in 4D-Var Identify systematic errors in observations and model (biases) Develop bias correction methods

#### What you have seen so far on data assimilation

#### Model (with errors) Non-orographic wave drag O<sub>3</sub> Chemistry CH<sub>4</sub> Oxidation Long-wave Short-way radiation radiation Subgrid-scale orographic drag Turbulent diffusion Shallow convection Latent Sensible Long-wave Short-wave heat heat flux flux Wind waves Ocean mode

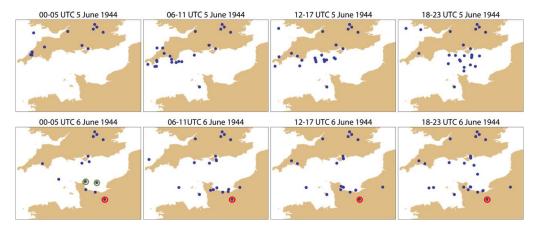


#### If you are lucky, model and observations are not biased



## Data assimilation and gross (obvious) errors

Observation error Model error



			From			GREENOCK					,	To 0	ES	TR	EHA	M	AN OUISTREHAM	
			×	Log (Stating type)	R	ance un sugh Nater	True Course	Mean volutions minute	Win Direc-	a becade	and Swell	Corrected Darometric Pressure	Tee	(Inter	-		E GRANTED TO SHIPS COMPANY	
			, nT	 	Miles	Vater Hiteo	Coure	21	Direc- tion (true)	104 Le	Sea	Millibars	Day	Wet	1		REMARKS	
essure	1010.5hF	Pa (mb)	0100	963.46	n	3	ver	157.9			-			_	0/00	( tail	glad a sof for antering several chand	
mperature	285.95K	• •	)200	975:83	11	6	115	161.5			-			-				
w point (wet bulb)		(54°F)	0300	9.88-79	13	0	170	167.5			-		-	-1				
nd direction	225 <sup>0</sup>	(SW)	ж	1001-49	ā	7	var	162.5	SW	40	1 33	1010-5	55	54 3	7 040	actim	stim	
nd speed	6.7ms <sup>-1</sup>	(Force 4)		1013-31										-	0547	gened	d in Conbardment out.	
eather/) Visibility	Code 97	(c/7)				1.1		1			1.		-	+	tal	1 Ent	195 gend from Eisel. 07062 store	
a temperature	285.35K	(54 <sup>0</sup> F)		19 20.84		t		1	1	1-1	1-		-	+	way	Filero	the farties.	
			0500	1026.43	1	2	var	97.5	W	40	33	10.11_	51	34	4 0924	S	wood direct-liton he in popular	
essure mperature	1014.8hF 285.35K			1031-47	1	1_1		1 -					-	-	Pert	tow ,	of and speeding regfor avoiding action of the Grany what a ince dia your	
w point (wet bulb)		(52 <sup>o</sup> F)		1036-0							-		-	-	- 1010 1	il du	ounded france LCICS ) 506. 1040 14.70. working party.	
nd direction	270°	(W)	1200	1637.7	24	0	ver	28.7	w	3 6	33	1511-6	55.5	4 5	6 120	7 aint	laid uprimy ked	
nd speed	6.7ms <sup>-1</sup>	(Force 4)	Dista	the Water		00	Lati	•	•	shede	1	pending on	-	co	PA & B & B / PA	ienced	ANCHOR DEARINGS	
eather/) Visibility	Code 96	(c/6)		252.7		200	49		0	เริ่ม	-	a #4		7	int.		2242. Shames the 1571' Chamber the 167' Commiss the 1854	
a temperature	284.25K	(52 <sup>0</sup> F)		at nots		_	49		• •	314	-		-				- "	
			$\vdash$	mat		1	49	1	0 1	6H 	1 1.	-x fin	-	Numb Sick	Lin	n		
	<b>`</b>		120		-	-							-	-				
olacklist,	)			10377		1.1		1	-				-	-	1=30	LLI(S)	516 ab-pidearth six asually the	

H.M.S. " FROBISHER

HMS Frobisher	Pressure	1010.5hP	a (mb)
at .	Temperature	285.95K	(55 <sup>o</sup> F)
+ 14	Dew point (wet bulb)	284.95K	(54 <sup>0</sup> F)
AI	Wind direction	225°	(SW)
	Wind speed	6.7ms <sup>-1</sup>	(Force 4)
	(Weather/) Visibility	Code 97	(c/7)
	Sea temperature	285.35K	(54 <sup>0</sup> F)
HMS Hawkins	Pressure	1014.8hP	'a (mb)
	Temperature	285.35K	(54°F)
	Dew point (wet bulb)	283.35K	(52 <sup>0</sup> F)
	Wind direction	270°	(W)
	Wind speed	6.7ms <sup>-1</sup>	(Force 4)
	(Weather/) Visibility	Code 96	(c/6)
	Sea temperature	284.25K	(52 <sup>0</sup> F)

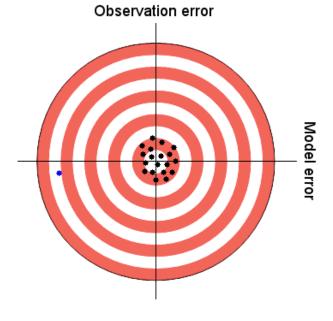
#### $\rightarrow$ Preliminary analysis (bl → Online Quality Control

	1031	_	24	0		29-1	w	3	ć	33	511-6	55,	4 5	-	164 Erhoffud	rking party. Thing party.
Donnes en Orașt treWiser 2527 Zote Time Leșt Atom Fi 57		~~	Boliton 0500		Latitude 49 20 rt		D ISW		1	l Depending on		Current			in experienced	ANCHOR BEARINGS 2242. { lagrans (L. 15]] 2242. { lagrans (L. 15]] } Berning (L. 1852.)
		1200			49		0 131			land for			7	ż	4	
			5000		49 124		0 15H			l	k fiz	Number Sck Li		Li	r 1	e de la composición d
1300	103	7.77	-	0	vas	51.2		-	-	-		-	-	-		· · · · · · · · · · · · · · · · · · ·
	1042			10		642		-	-	-		-	-		anboard.	16 ab-fidearth six awall
	1042			5	ver	229		-	-	-		- 1500 1 taking C 1504 Willing l		1500 1 taking	casualty died & was burie littlet festles in these	
1600	104	2.32	7	-	var	57.1	w.	3	26	-	1012-	57	5	-		

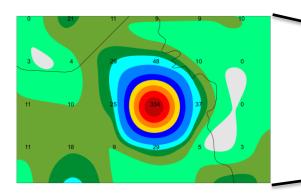
" JUESDAY THE 6 day of JUNE

. 19 44.

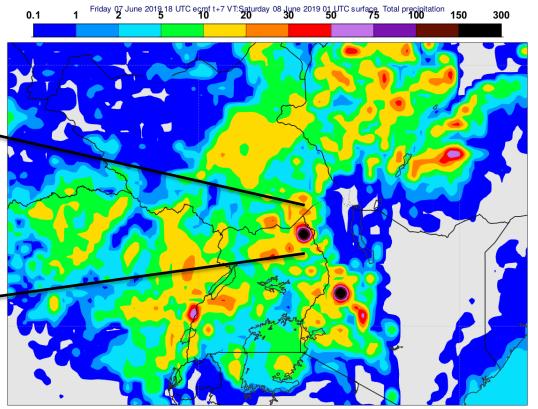
# Data assimilation and gross (obvious) errors



# Total precipitation on 07 June 2019 (accumulated over 6 hours)

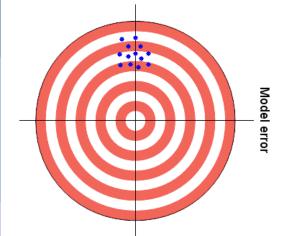


→ Continuous monitoring
→ Keep improving the model



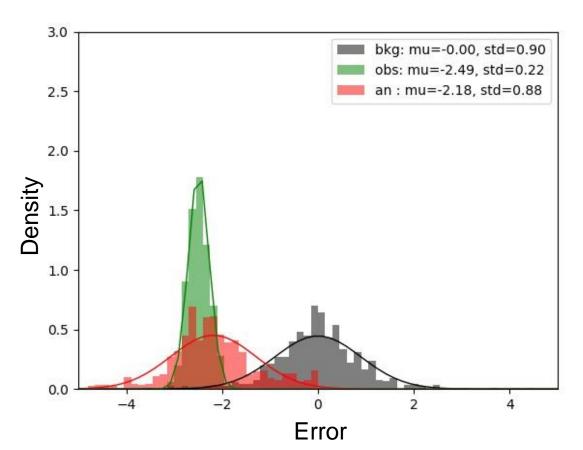
#### Data assimilation and biased observations

Observation error



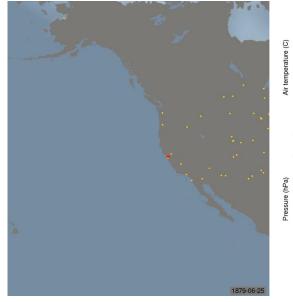
If standard 4D-Var is used to assimilate biased observations (systematic errors), the resulting analysis will be biased.

In this case the background is more accurate than the analysis!



### Data assimilation and biased observations

#### The USS Jeannette (1879, Artic, 33 crew members)



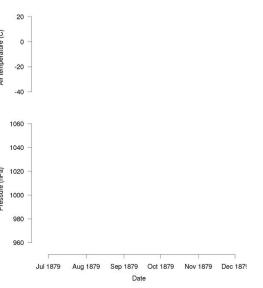


Photo # NH 52000 Steamer Jeannette sinking after being crushed by Arctic ice, June 1881



THE SINKING OF THE JEANNETTE.

hoto # NH 52002 Jeannette's crewmen drag their boats over the Arctic ice, June-August 1881



DRAGGING THE BOATS OVER THE ICE

Photo # NH 92142 LCdr. DeLong and his party wading ashore on the Lena Delta, Siberia, 17 Sept. 1881

WADING ASHORE.

SST measurements from standard buckets have a cold bias (~0.4C)



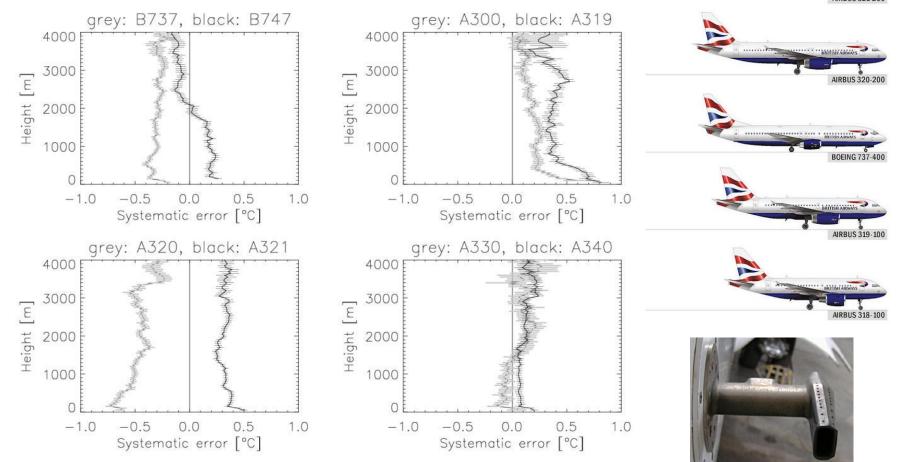
and a state and the the the the first and find the but one	that the shart the state of the far and the
The state of the s	The paper to good to day to the by the King to
1.5	The Share bear hade a mare for diday
. cost he marget up and -	have a of fator may lite - a dye dif to see by what's fit
· · · · start · · ·	their data as glow, whe constall for hald git have still have to
· · · · · ·	In that we the figs after the store particular interfaces to the static and cover at to theme the that is to be a the star in the factor and the starter and lighter approximate and a static of all of the starter to the starter and
An and a part of the second se	Mar Mar Anna Mar Mar Land A and Pho Real Anna Mar Mar Mar Mar Mar Mar Mar Real Anna Mar
100 000 og sodarft og til	I Part to an ever a list to the same of an hard and to the first the set of t
. as a seas	
of sellest and	2
. are the world and a	Cedelle 48
Inspire V. Summer Ro. (Sense Roman V ) 1 1 1 Departs V. Roman Row Romann V 1 1 1 Summer V. Roman Row Romann V 1 1 Summer V. Roman V 1 1 Summer V 1 Summer V 2	etay and



Data assimilation and biased observations

One year of measurements from aircrafts landing at Frankfurt

No external reference, observation bias is estimated using the hourly mean of all measured profiles



BRITISH AIRWAYS

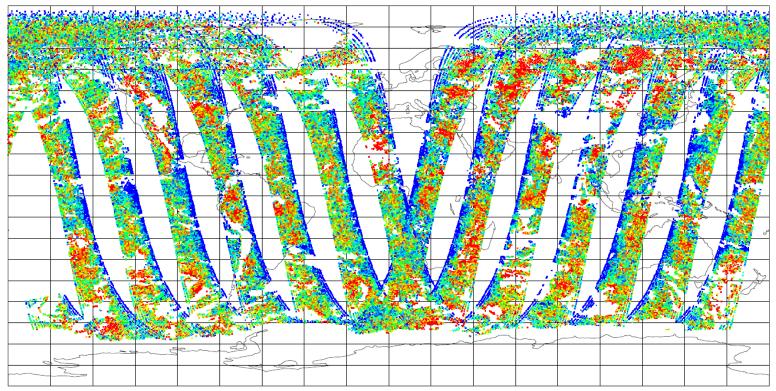
#### Data assimilation and biased observations



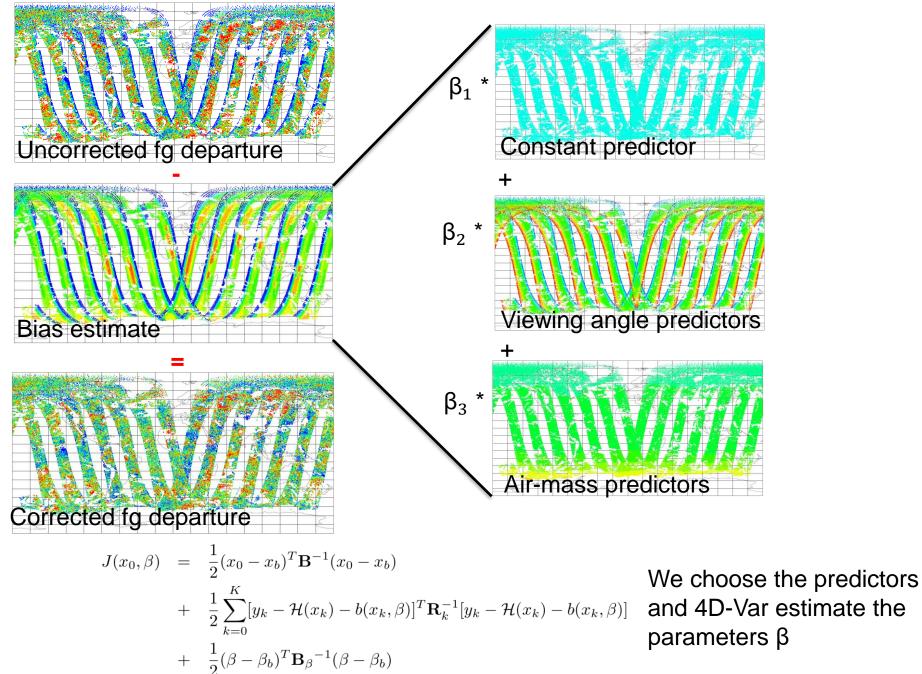
ECMWF is assimilating polar-orbiting Metop-C satellite (launched on 7 November 2018)

Observation bias is estimated inside 4D-Var
→ comparing measurements with model
→ specifying the structure of the model bias

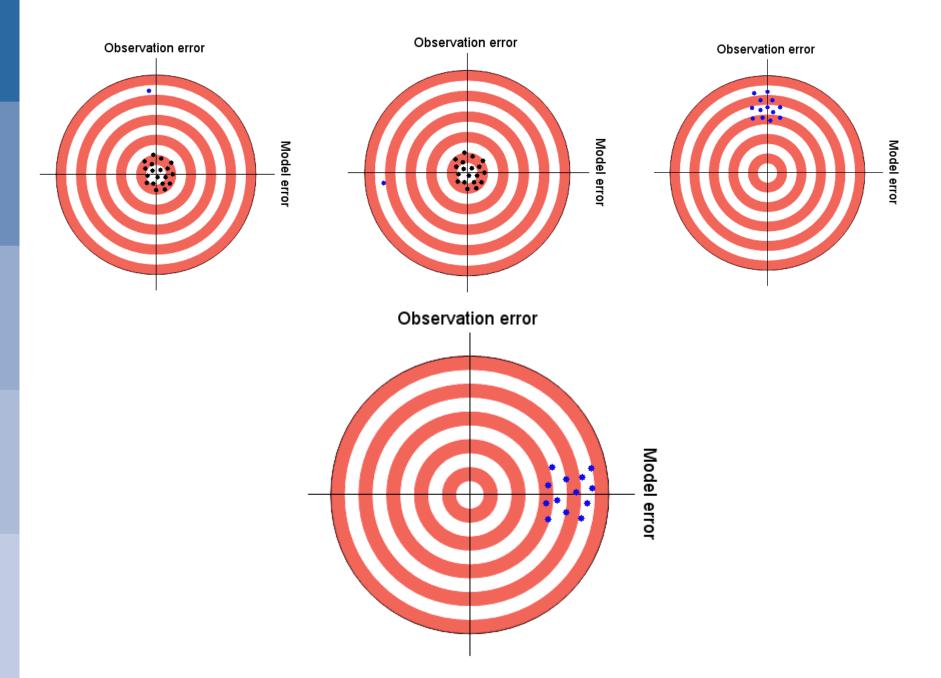
Metop-C AMSUA-A Channel 5 (obs-model)



#### Data assimilation and biased observations

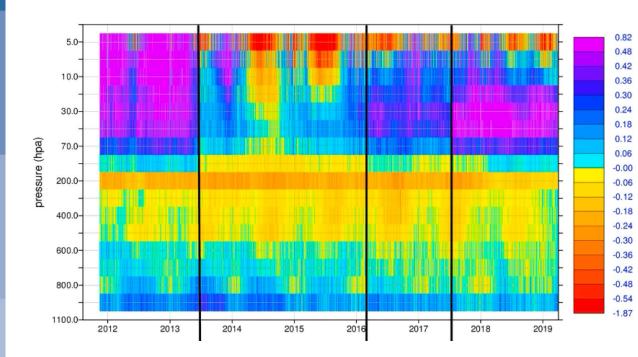


#### Data assimilation and biased model



### How to estimate model biases

The first-guess trajectory of the model can be compared to accurate observations



Difference between radiosonde temperature observations and the IFS first-guess trajectory (O-B)

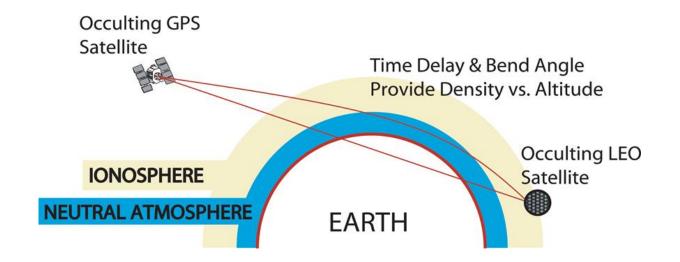


Errors in models are often systematic rather than random, zero-mean

- $\rightarrow$  Largest bias in the stratosphere
- $\rightarrow$  Model has a temperature cold bias in the lower/mid stratosphere
- $\rightarrow$  Model has a warm bias in the upper stratosphere

#### How to estimate model biases

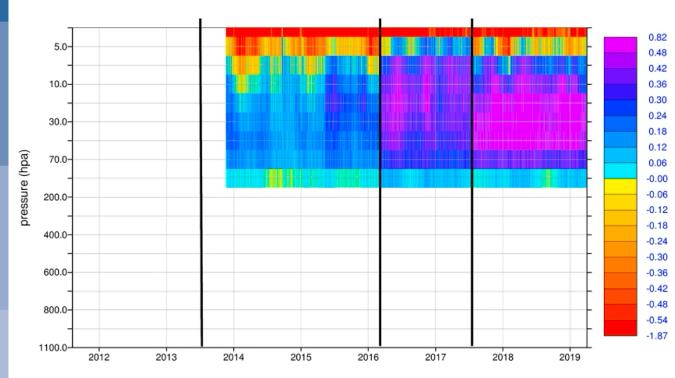
The GPS satellites are used for positioning and navigation. GPS-RO (Radio Occultation) is based on analysing the bending caused by the atmosphere along paths between a GPS satellite and a receiver placed on a low-earth-orbiting satellite.



- $\rightarrow$  As the LEO moves behind the earth, we obtain a profile of bending angles
- $\rightarrow$  Temperature profiles can then be derived
- → GPS-RO can be assimilated without bias correction. They are good for highlighting errors/biases

### How to estimate model biases

The first-guess trajectory of the model can be compared to accurate observations



Difference between GPS-RO temperature retrievals and the IFS first-guess trajectory

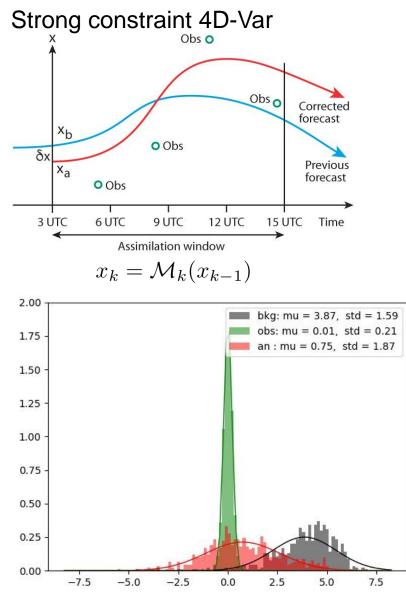
(O-B)



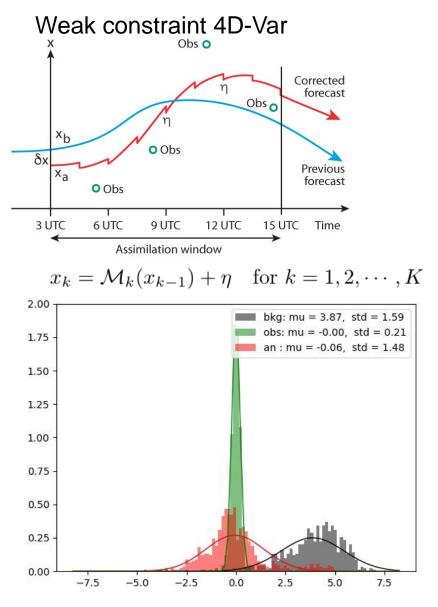
Errors in models are often systematic rather than random, zero-mean

- $\rightarrow$  Model has a temperature cold bias in the lower/mid stratosphere
- $\rightarrow$  Model has a warm bias in the upper stratosphere

### How to deal with model biases in data assimilation



→ Large bias and standard deviation in the analysis



→ Bias in the analysis has been reduced, standard deviation as well

#### Weak constraint 4D-Var

We assume that the model is not perfect, adding an error term  $\eta$  in the model equation

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

The model error estimate  $\eta$  contains 3 physical fields

- temperature
- vorticity
- divergence

Constant model error forcing over the assimilation window to correct the model bias

- $\rightarrow$  Introduce additional controls to target an unbiased analysis
- $\rightarrow$  The model error covariance matrix Q constrains the model error field

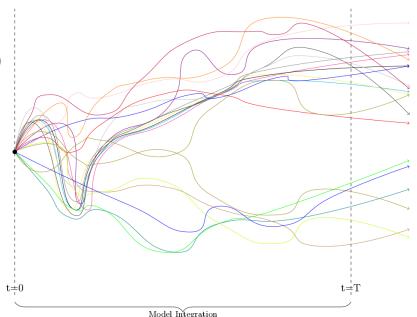
### How to estimate the model error covariance matrix (Q)

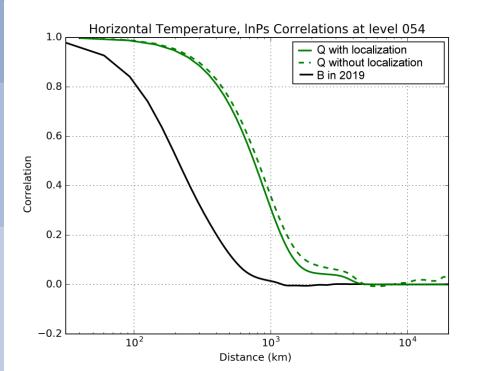
#### Estimate the model error covariance matrix

➔ run the ensemble forecasting system (ENS) with perturbed physics (51 members with the same initial condition for different days)

→ differences after 12 hours are used to compute Q

$$Q_{\rm f} = \frac{1}{N-1} \sum_{i=1}^{N} \left( f_i^{12} - f_{i+1}^{12} \right) \left( f_i^{12} - f_{i+1}^{12} \right)^{\rm T}$$

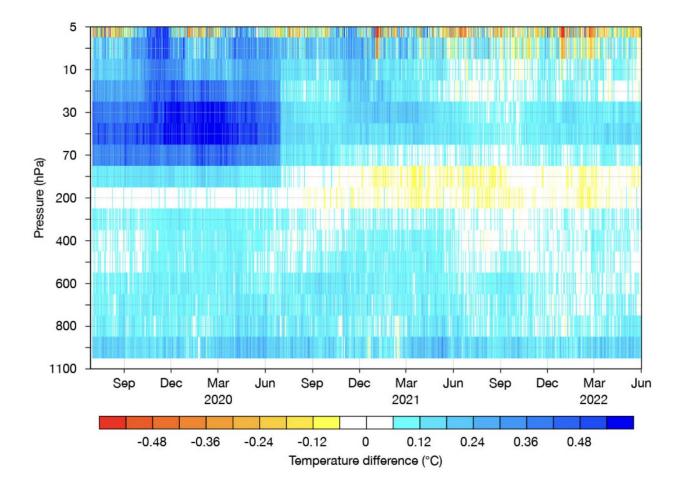




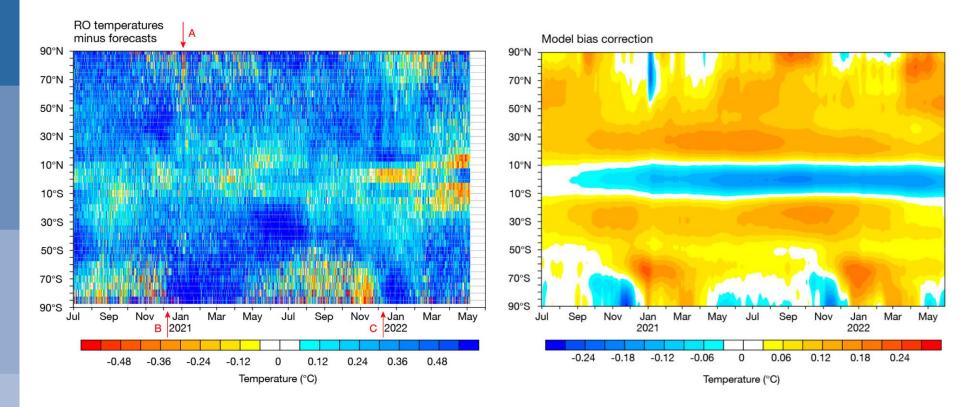
4D-Var corrects small scale errors (background errors) by changing the initial condition and large scale errors (model errors) by changing the model forcing

#### Weak-constraint 4D-Var in operations for the stratosphere

Time series of the difference between radiosonde temperature observations and model first-guess (47r1 implemented on 30 June 2020)



#### Weak-constraint 4D-Var in operations for the stratosphere

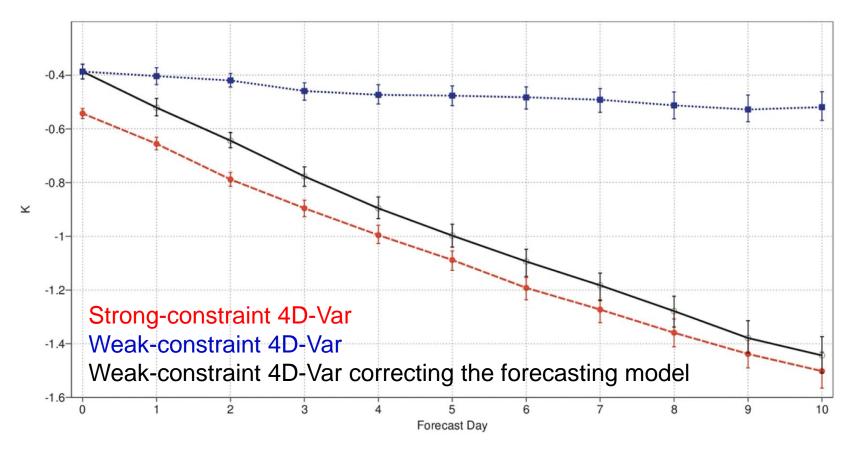


A) On 31 December 2020, a Sudden Stratospheric Warming (SSW) event started over the northern hemisphere

B&C) Clear seasonal cycle in the model bias over the southern hemisphere with a sharp transition in early December 2020 and 2021

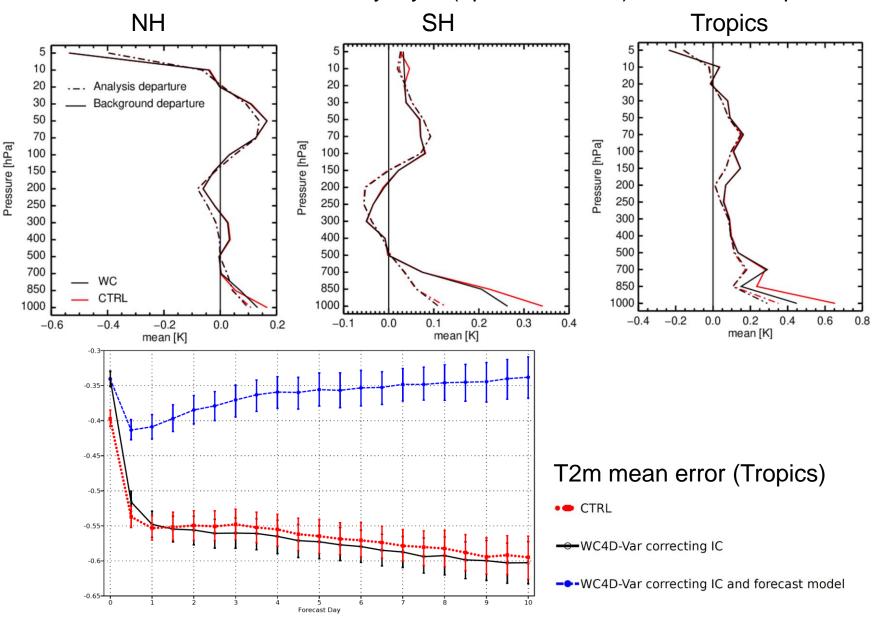
#### Weak-constraint 4D-Var and medium-range forecast

Mean error of the 10-day forecast at 50hPa with respect to the radiosonde observations

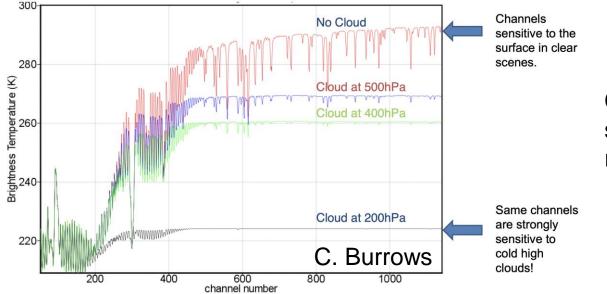


#### Weak-constraint 4D-Var in the boundary layer

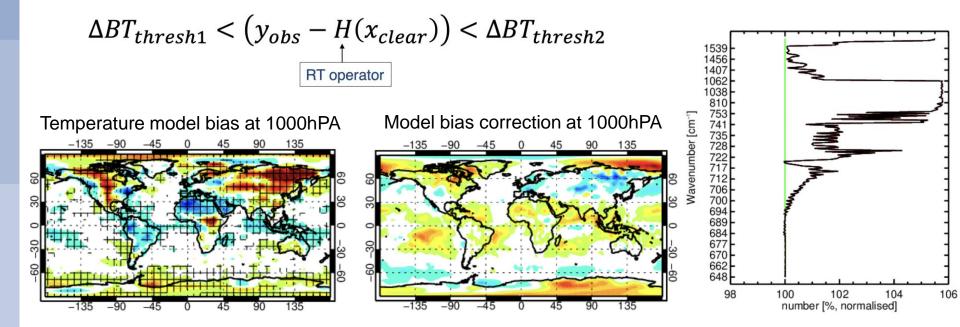
Activate WC4D-Var in the boundary layer (up to ~ 900hPa) to correct temperature



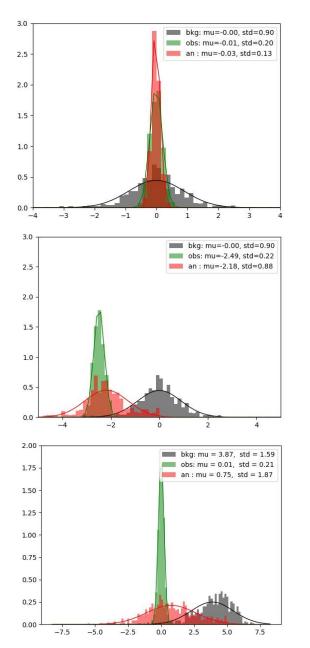
#### Positive feedback for satellite infrared assimilation



Clouds have a very strong impact on infrared radiance measurements



#### Summary 1/3



Background: unbiased (only random errors) Observation: unbiased (only random errors) Standard 4D-Var

Background: unbiased (only random errors) Observation: biased Standard 4D-Var & Variational Bias Control (VarBC)

Background: biased Observation: unbiased (only random errors) Weak constraint 4D-Var



How do I know if my observations are biased? How do I know if my model is biased? You don't know the truth, but you have to trust something

Reference observations are used



Radiosondes



**GPS-RO** 

#### Summary 3/3

From bias-blind to bias-aware data assimilation

$$J(x_{0},\beta,\eta) = \frac{1}{2}(x_{0} - x_{b})^{T}\mathbf{B}^{-1}(x_{0} - x_{b}) \\ + \frac{1}{2}\sum_{k=0}^{\mathrm{Radiosonde}} [y_{k} - \mathcal{H}(x_{k})]^{T}\mathbf{R}_{k}^{-1}[y_{k} - \mathcal{H}(x_{k})] \\ + \frac{1}{2}\sum_{k=0}^{\mathrm{GPSRO}} [y_{k} - \mathcal{H}(x_{k})]^{T}\mathbf{R}_{k}^{-1}[y_{k} - \mathcal{H}(x_{k})] \\ + \frac{1}{2}\sum_{k=0}^{\mathrm{Others}} [y_{k} - \mathcal{H}(x_{k}) - b(x_{k},\beta)]^{T}\mathbf{R}_{k}^{-1}[y_{k} - \mathcal{H}(x_{k}) - b(x_{k},\beta)] \\ + \frac{1}{2}(\beta - \beta_{b})^{T}\mathbf{B}_{\beta}^{-1}(\beta - \beta_{b}) \\ + \frac{1}{2}(\eta - \eta_{b})^{T}\mathbf{Q}^{-1}(\eta - \eta_{b})$$

Any questions? Feel free to contact me patrick.laloyaux@ecmwf.int