



**Barcelona  
Supercomputing  
Center**

*Centro Nacional de Supercomputación*



EXCELENCIA  
SEVERO  
OCHOA

# Data assimilation in dust models

## Part 1: Introduction and overview

Enza Di Tomaso  
Earth Sciences Department





# Introduction to data assimilation (DA)

**Data assimilation provides a framework for combining model simulations and observations in an optimal way**

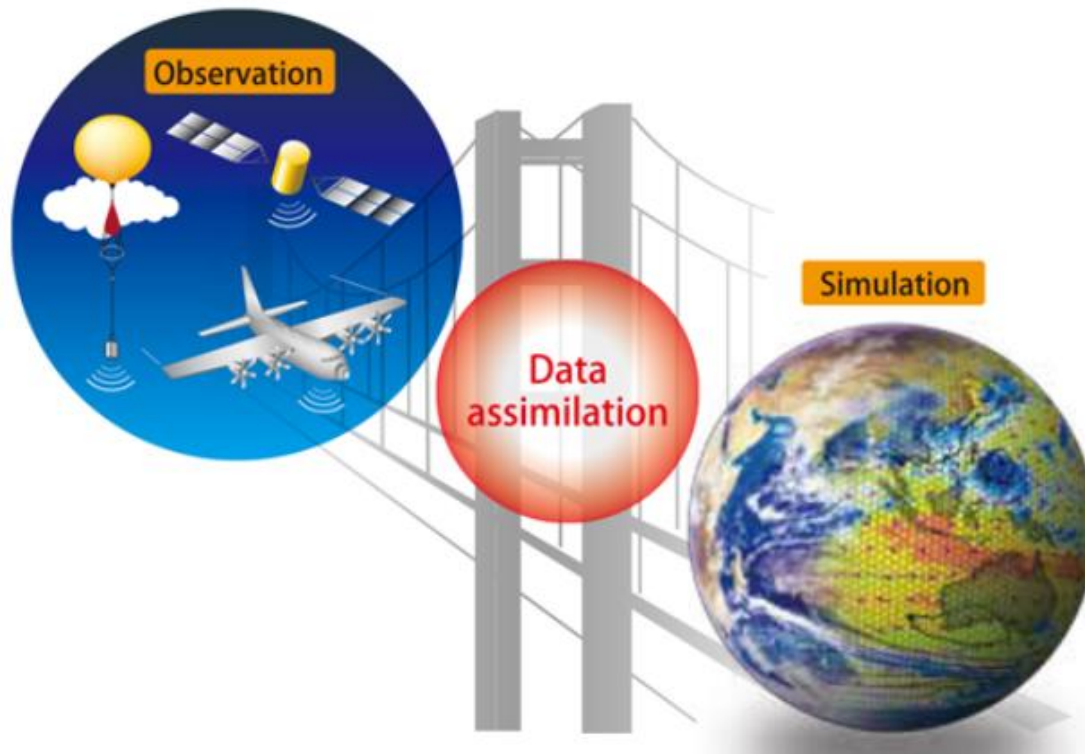


Figure credit: <http://daconf15.umd.edu/>

Data assimilation provides a framework for combining model simulations and observations in an optimal way

**Why would we want to do that?**

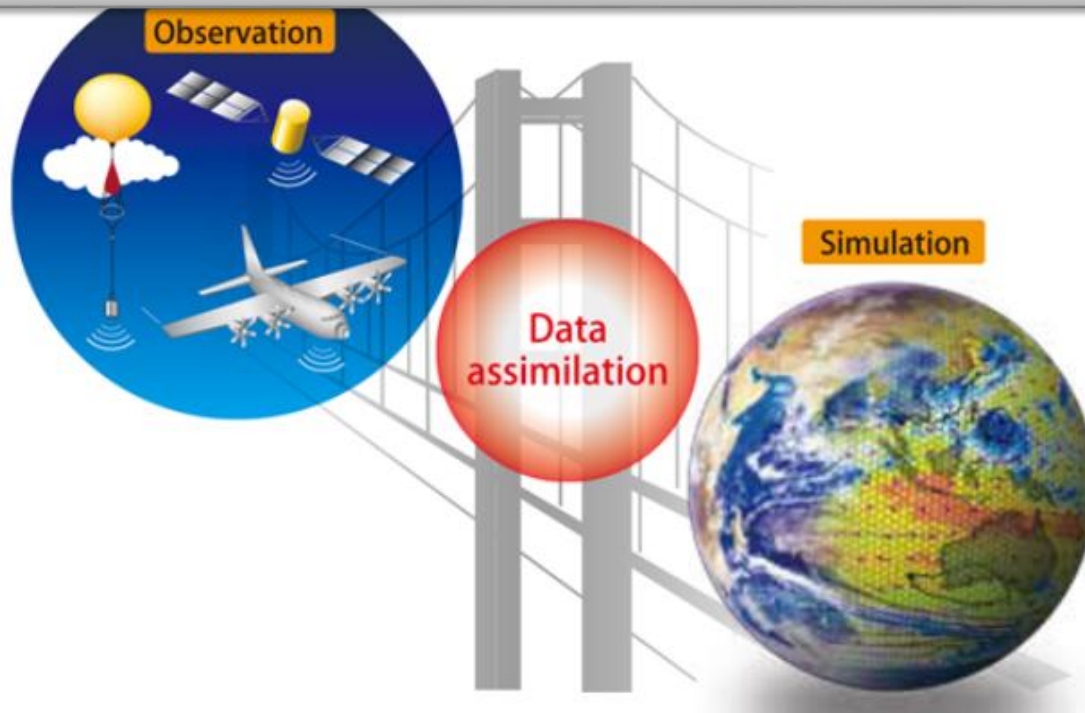
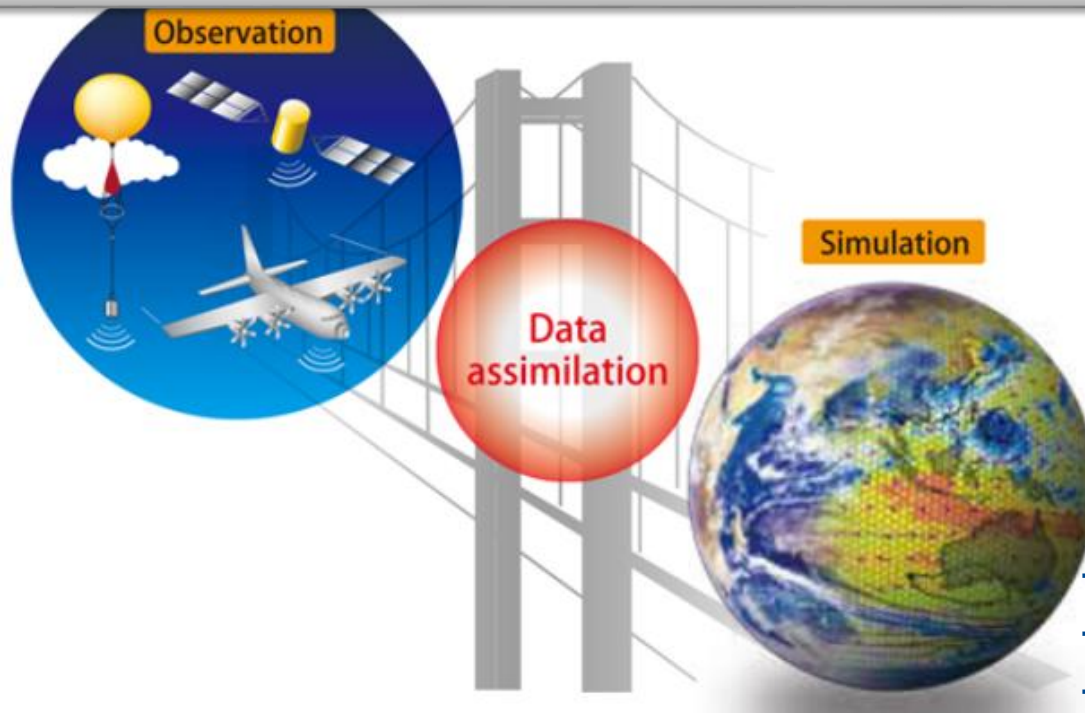


Figure credit: <http://daconf15.umd.edu/>

Data assimilation provides a framework for combining model simulations and observations in an optimal way

**Why would we want to do that?**



- predictive skills
- speciation
- spatial and temporal resolution

Figure credit: <http://daconf15.umd.edu/>

# DA in forecast mode



**Barcelona  
Supercomputing  
Center**  
Centro Nacional de Supercomputación



time

now



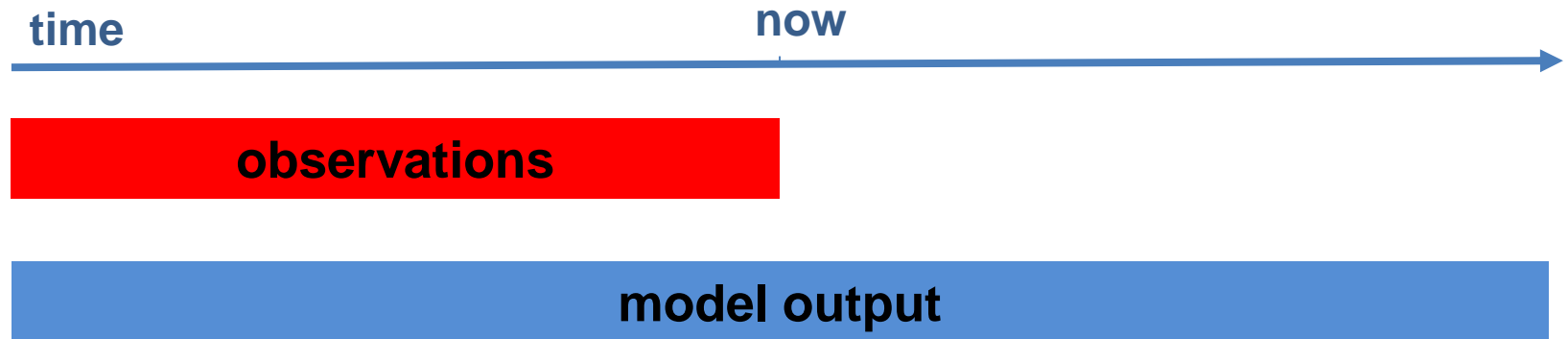
The diagram consists of a horizontal blue arrow pointing to the right, representing the flow of time. Above the left end of the arrow is the word 'time', and above a point further along is the word 'now'. Below the arrow, starting from the left and extending to the 'now' point, is a solid red rectangular block. Inside this red block, the word 'observations' is written in black, bold, lowercase letters.

**observations**

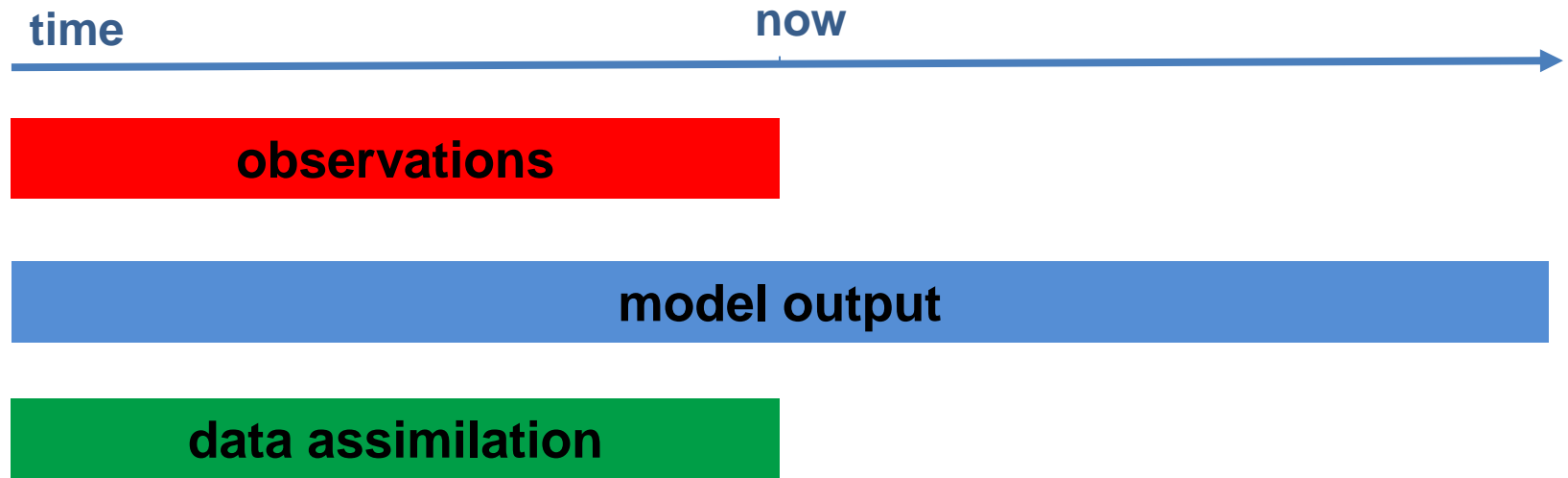
# DA in forecast mode



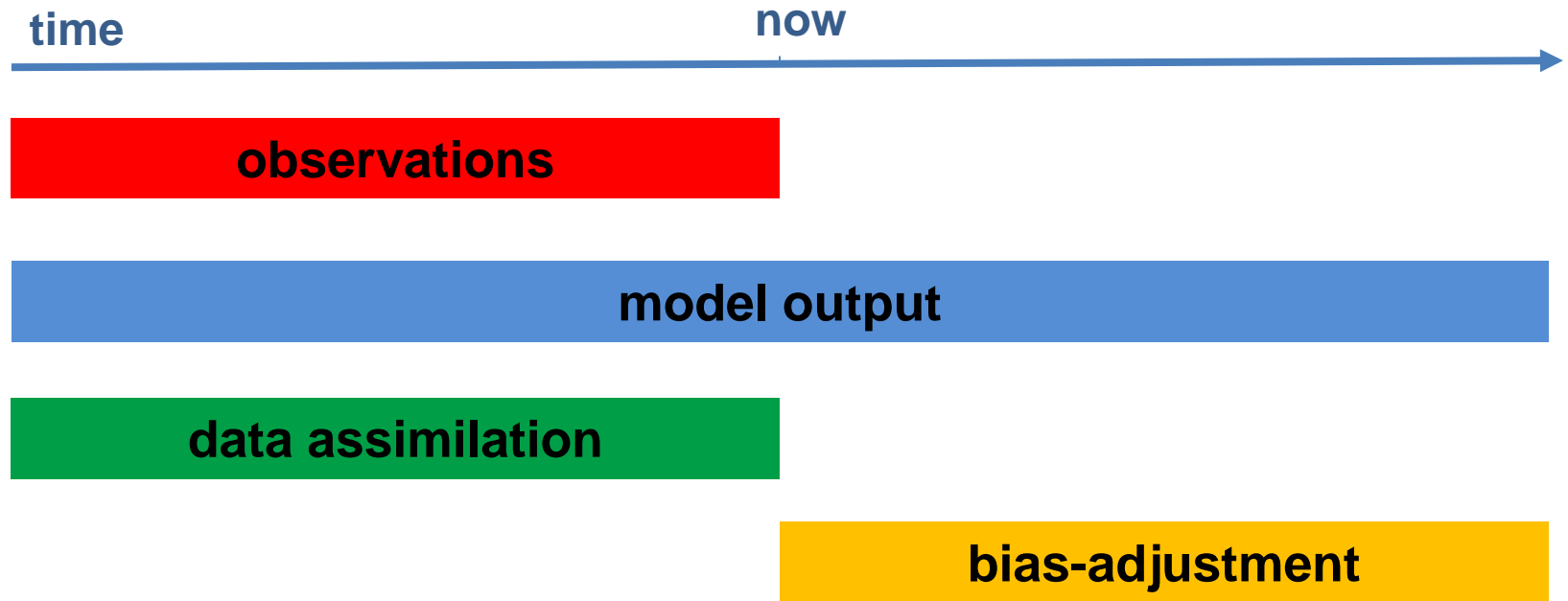
**Barcelona  
Supercomputing  
Center**  
Centro Nacional de Supercomputación



# DA in forecast mode





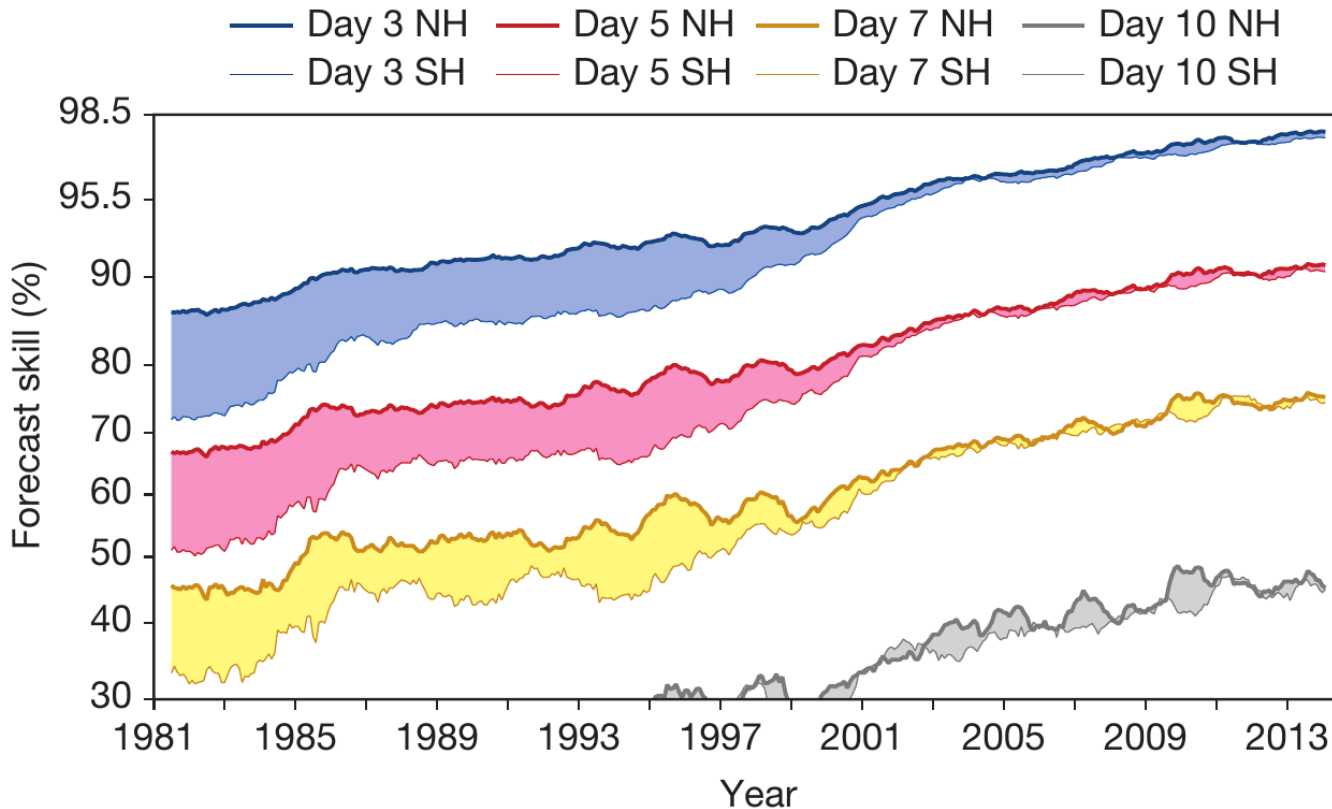


Obtain the ‘best’ estimate of current atmospheric conditions **(analysis)**

Obtain the ‘best’ estimate of current atmospheric conditions **(analysis)**

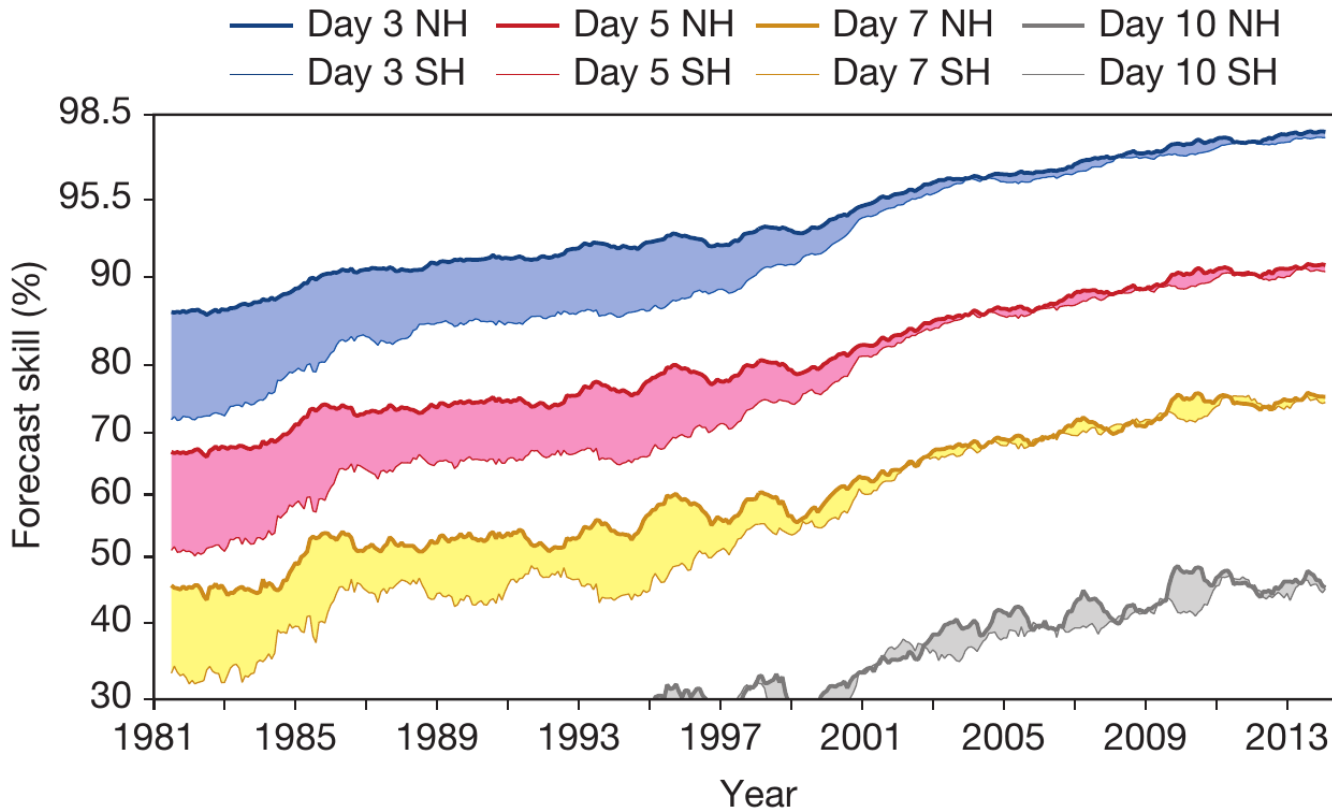
*-> useful to initialise models and improve predictions*

**A measure of forecast skill at three-, five-, seven- and ten-day ranges, computed over the extra-tropical northern and southern hemispheres.**



(Bauer et al. 2015)

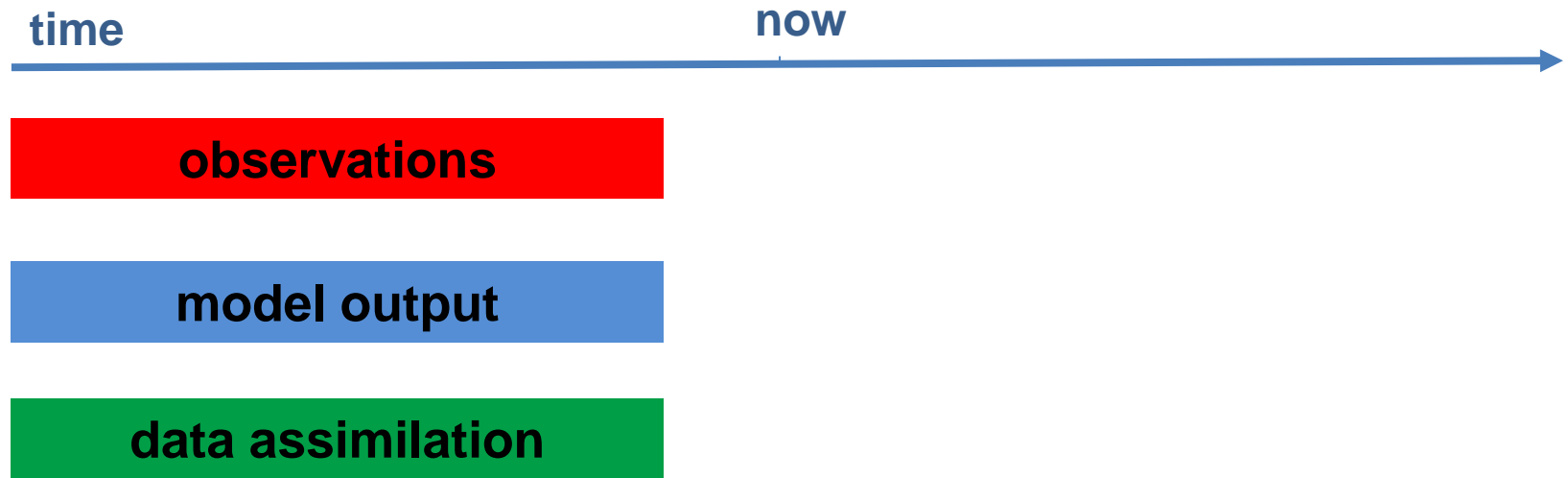
**A measure of forecast skill at three-, five-, seven- and ten-day ranges, computed over the extra-tropical northern and southern hemispheres.**



(Bauer et al. 2015)

**The convergence of the curves for NH and SH indicates the breakthrough in exploiting satellite data**

# DA in retrospective mode



**Create datasets describing the recent history of the  
atmosphere (reanalysis)**

Create datasets describing the recent history of the atmosphere **(reanalysis)**

*-> useful to monitor climate change and for research*



- NAAPS reanalysis at 1 degree resolution  
(Lynch et al., 2016)
- MACC-II reanalysis at 78 km resolution  
(Cuevas et al., 2015; Inness et al., 2013)
- NASA MERRAaero reanalysis at 0.5 degree resolution  
(Buchard et al., 2015; Buchard et al., 2016).

**Optimise uncertain model parameters  
(parameter estimation)**

*-> useful to calibrate models*

**Examine observation impacts (OSEs and OSSE)**

*-> useful to design observing systems*

# Why might you be interested in it?



**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*



# Why might you be interested in it?



- **Improve forecasting skill**

# Why might you be interested in it?



- **Improve forecasting skill**
- **Understanding analyses and reanalysis**

# Why might you be interested in it?



- **Improve forecasting skill**
- **Understanding analyses and reanalysis**
- **Gain insight into observations**

# Why might you be interested in it?



- **Improve forecasting skill**
- **Understanding analyses and reanalysis**
- **Gain insight into observations**
- **Tailor your observation dataset for DA and provide feedback on their use**



# Why might you be interested in it?



- **Improve forecasting skill**
- **Understanding analyses and reanalysis**
- **Gain insight into observations**
- **Tailor your observation dataset for DA and provide feedback on their use**
- **Application of how we reason**
- **...**

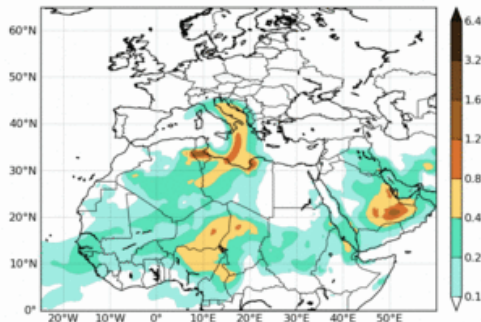


# DA in aerosol forecast models

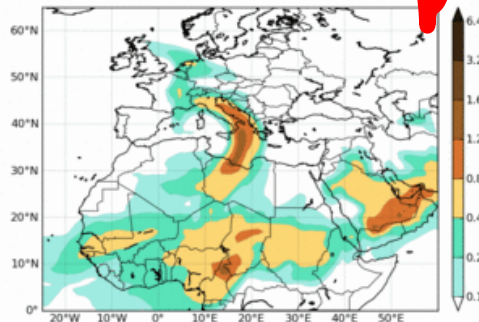
# SDS-WAS intercomparison

Dust optical depth: 2014 3 Apr FC+24

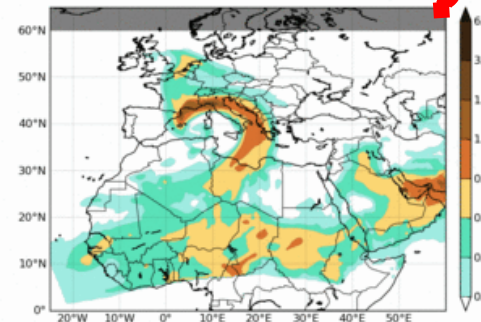
WMO SDS-WAS N.Africa-Middle East-Europe RC  
BSC-DREAM8b Dust AOD  
Run: 12h 03 APR 2014 Valid: 12h 04 APR 2014 (H+24)



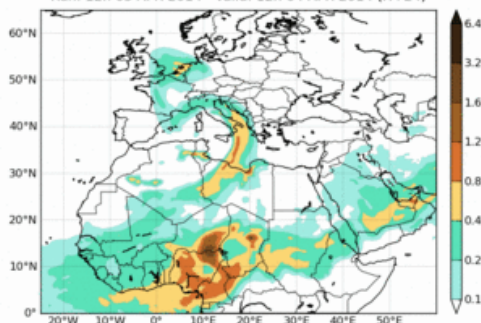
WMO SDS-WAS N.Africa-Middle East-Europe RC  
MACC-ECMWF Dust AOD  
Run: 00h 03 APR 2014 Valid: 12h 04 APR 2014 (H+36)



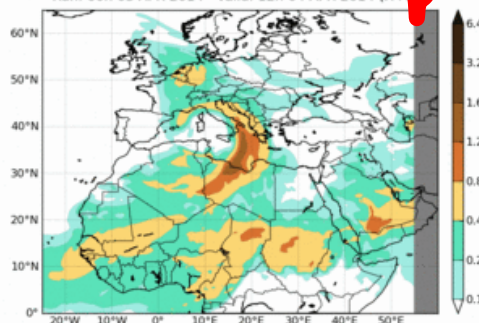
WMO SDS-WAS N.Africa-Middle East-Europe RC  
DREAM8-NMME-MACC Dust AOD  
Run: 12h 03 APR 2014 Valid: 12h 04 APR 2014 (H+24)



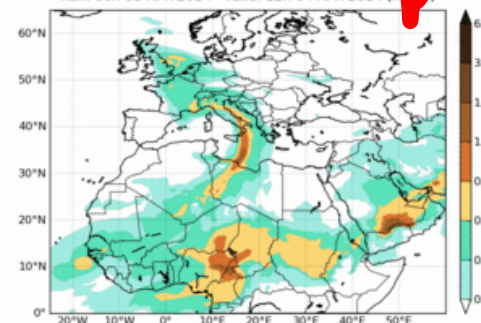
WMO SDS-WAS N.Africa-Middle East-Europe RC  
NMMB/BSC-Dust Dust AOD  
Run: 12h 03 APR 2014 Valid: 12h 04 APR 2014 (H+24)



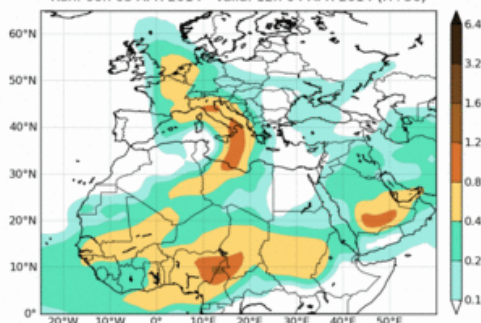
WMO SDS-WAS N.Africa-Middle East-Europe RC  
U.K. MetOffice Dust AOD  
Run: 00h 03 APR 2014 Valid: 12h 04 APR 2014 (H+36)



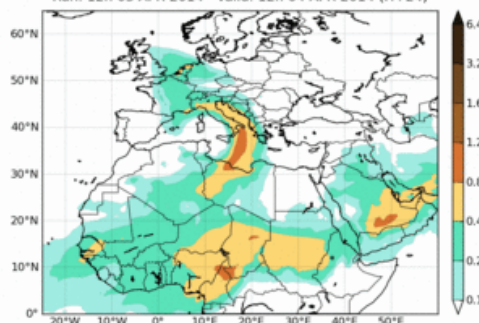
WMO SDS-WAS N.Africa-Middle East-Europe RC  
NASA GEOS-5 Dust AOD  
Run: 00h 03 APR 2014 Valid: 12h 04 APR 2014 (H+36)



WMO SDS-WAS N.Africa-Middle East-Europe RC  
NCEP NGAC Dust AOD  
Run: 00h 03 APR 2014 Valid: 12h 04 APR 2014 (H+36)



WMO SDS-WAS N.Africa-Middle East-Europe RC  
MEDIAN Dust AOD  
Run: 12h 03 APR 2014 Valid: 12h 04 APR 2014 (H+24)



V = data assimilation

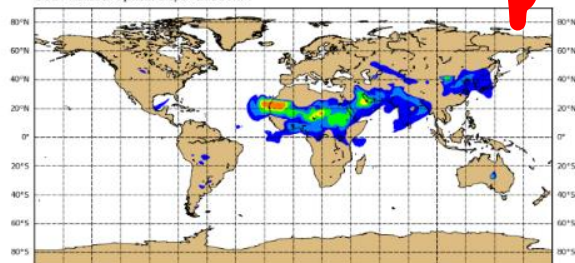
# ICAP intercomparison



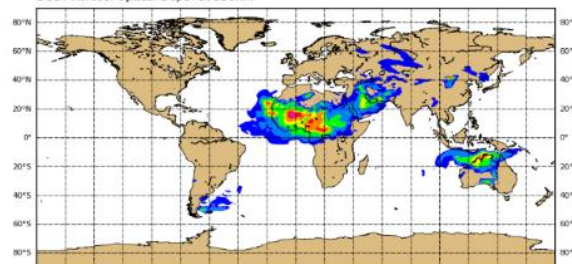
**Barcelona  
Supercomputing  
Center**  
Centro Nacional de Supercomputación

EXCELENCIA  
SEVERO  
OCHOA

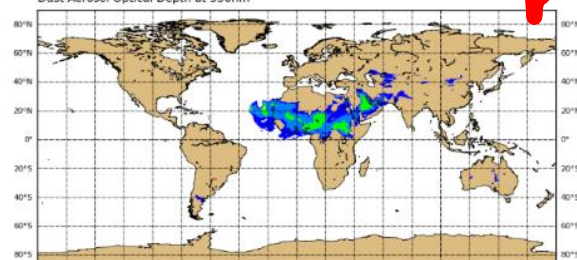
Thursday 31 January 2013 00UTC NAAPS Forecast t+006  
Thursday 31 January 2013 06UTC Valid Time  
DUST Aerosol Optical Depth at 550nm



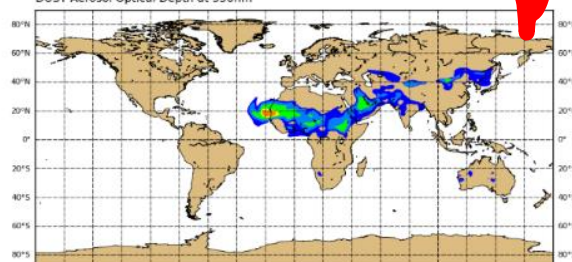
Thursday 31 January 2013 00UTC NMMB Forecast t+006  
Thursday 31 January 2013 06UTC Valid Time  
DUST Aerosol Optical Depth at 550nm



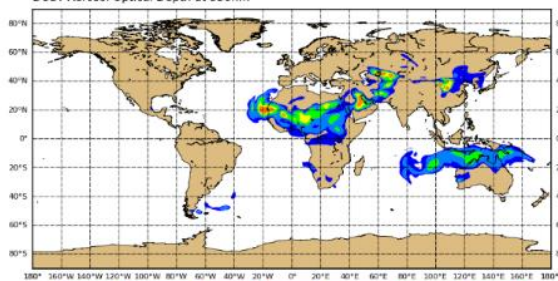
Thursday 31 January 2013 00UTC GEOS-5 Forecast t+006  
Thursday 31 January 2013 06UTC Valid Time  
DUST Aerosol Optical Depth at 550nm



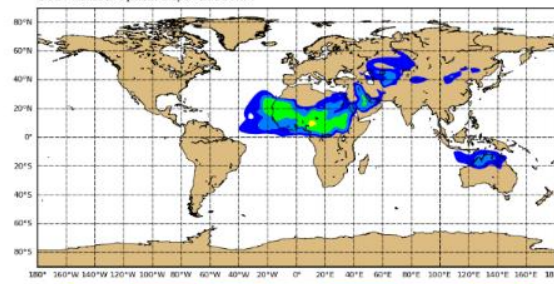
Thursday 31 January 2013 00UTC MACC Forecast t+006  
Thursday 31 January 2013 06UTC Valid Time  
DUST Aerosol Optical Depth at 550nm



Thursday 31 January 2013 00UTC MASINGAR Forecast t+006  
Thursday 31 January 2013 06UTC Valid Time  
DUST Aerosol Optical Depth at 550nm



Thursday 31 January 2013 00UTC NGAC Forecast t+006  
Thursday 31 January 2013 06UTC Valid Time  
DUST Aerosol Optical Depth at 550nm



0.1 0.2 0.4 0.8 1.2 2.5 5.0 9.0

Plots Generated Sunday 24 February 2013 05UTC NRL/Monterey Aerosol Modeling

0.1 0.2 0.4 0.8 1.2 2.5 5.0 9.0

Plots Generated Sunday 24 February 2013 05UTC NRL/Monterey Aerosol Modeling

[http://www.nrlmry.navy.mil/aerosol/icap\\_date.1087.php?date=2014121200&field=aod&spec=dust&regc=global&icap=2&modir=icap\\_01&quad=0](http://www.nrlmry.navy.mil/aerosol/icap_date.1087.php?date=2014121200&field=aod&spec=dust&regc=global&icap=2&modir=icap_01&quad=0)

**✓ = data assimilation**



## CURRENT ICAP MODELS

Organization	BSC	Copernicus / ECMWF	JMA	Meteo France	NASA	US Navy	NOAA	UK Met Office
Model	NMMB/BSC-CTM	CAMS	MASINGAR	MOCAGE	GEOS-5	NAAPS	NGAC	UKMO
Status	QO	O-24 hrs	QO	O	QO	O	O	O
Meteorology	Offline NMMB	Inline IFS	inline AGCM	Offline ARPEGE	Inline GEOS-5	Offline NAVGEM	Inline GFS	Inline UM
Resolution	1.4x1	0.4x0.4	0.56x0.56	2x2	0.25x0.31	0.33x0.33	1x1	0.35x0.23
levels	24	60	40	47	72	60	64	70
DA	LETKF <sup>P</sup>	4DVar	EnKF <sup>P</sup>	2018	2DVar +LDE	2DVar 3DVar, EnKF <sup>P</sup>	NA	4DVar
Assimilated Obs	DAQ MODIS+DB	DAQ MODIS+DB	CALIOP, MODIS, Himawari-8	NA	Neural Net MODIS	DAQ MODIS, CALIOP	NA	MODIS Dust AOT
Species	Dust Sea Salt BC, OC (POA,SOA) Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	Anthro+bio B. Burn Dust Sea Salt	Dust BC OC Sea Salt Sulfate	Dust
Size Bins	8 (dust, salt) Bulk (BC, OC, Su)	3	10	6	5	1	5	2
Bio. Burn. Emissions	NA	GFAS	GFAS	GFAS	QFED	FLAMBE	GBBEPx	NA

- The ICAP-MME is run daily w/ 1x1 deg res at 00Z for 6 hrly fcasts out to 120 hrs w/ a 1-day latency.
- Modal AOT (550nm) and dust AOT (550nm) data in NetCDF is available at [http://usgodae.org/cgi-bin/datalist.pl?dset=nrl\\_icap\\_mme&summary=Go](http://usgodae.org/cgi-bin/datalist.pl?dset=nrl_icap_mme&summary=Go)

## CURRENT ICAP MODELS

Organization	BSC	Copernicus / ECMWF	JMA	Meteo France	NASA	US Navy	NOAA	UK Met Office
Model	NMMB/BSC-CTM	CAMS	MASINGAR	MOCAGE	GEOS-5	NAAPS	NGAC	UKMO
Status	QO	O-24 hrs	QO	O	QO	O	O	O
Meteorology	Offline NMMB	Inline IFS	inline AGCM	Offline ARPEGE	Inline GEOS-5	Offline NAVGEM	Inline GFS	Inline UM
Resolution	1.4x1	0.4x0.4	0.56x0.56	2x2	0.25x0.31	0.33x0.33	1x1	0.35x0.23
levels	24	60	40	47	72	60	64	70
DA	LETKF <sup>P</sup>	4DVar	EnKF <sup>P</sup>	2018	2DVar +LDE	2DVar 3DVar, EnKF <sup>P</sup>	NA	4DVar
Assimilated Obs	DAQ MODIS+DB	DAQ MODIS+DB	CALIOP, MODIS, Himawari-8	NA	NeuralNet MODIS	DAQ MODIS, CALIOP	NA	MODIS Dust AOT
Species	Dust Sea Salt BC, OC (POA,SOA) Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	Anthro+bio B. Burn Dust Sea Salt	Dust BC OC Sea Salt Sulfate	Dust
Size Bins	8 (dust, salt) Bulk (BC, OC, Su)	3	10	6	5	1	5	2
Bio. Burn. Emissions	NA	GFAS	GFAS	GFAS	QFED	FLAMBE	GBBEPx	NA

- The ICAP-MME is run daily w/ 1x1 deg res at 00Z for 6 hrly fcasts out to 120 hrs w/ a 1-day latency.
- Modal AOT (550nm) and dust AOT (550nm) data in NetCDF is available at [http://usgodae.org/cgi-bin/datalist.pl?dset=nrl\\_icap\\_mme&summary=Go](http://usgodae.org/cgi-bin/datalist.pl?dset=nrl_icap_mme&summary=Go)

## CURRENT ICAP MODELS

Organization	BSC	Copernicus / ECMWF	JMA	Meteo France	NASA	US Navy	NOAA	UK Met Office
Model	NMMB/BSC-CTM	CAMS	MASINGAR	MOCAGE	GEOS-5	NAAPS	NGAC	UKMO
Status	QO	O-24 hrs	QO	O	QO	O	O	O
Meteorology	Offline NMMB	Inline IFS	inline AGCM	Offline ARPEGE	Inline GEOS-5	Offline NAVGEM	Inline GFS	Inline UM
Resolution	1.4x1	0.4x0.4	0.56x0.56	2x2	0.25x0.31	0.33x0.33	1x1	0.35x0.23
levels	24	60	40	47	72	60	64	70
DA	LETKF <sup>P</sup>	4DVar	EnKF <sup>P</sup>	2018	2DVar +LDE	2DVar 3DVar, EnKF <sup>P</sup>	NA	4DVar
Assimilated Obs	DAQ MODIS+DB	DAQ MODIS+DB	CALIOP, MODIS, Himawari-8	NA	Neural Net MODIS	DAQ MODIS, CALIOP	NA	MODIS Dust AOT
Species	Dust Sea Salt BC, OC (POA,SOA) Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	BC Dust OC Sea Salt Sulfate	Anthro+biogenic B. Burn Dust Sea Salt	Dust BC OC Sea Salt Sulfate	Dust
Size Bins	8 (dust, salt) Bulk (BC, OC, Su)	3	10	6	5	1	5	2
Bio. Burn. Emissions	NA	GFAS	GFAS	GFAS	QFED	FLAMBE	GBBEPx	NA

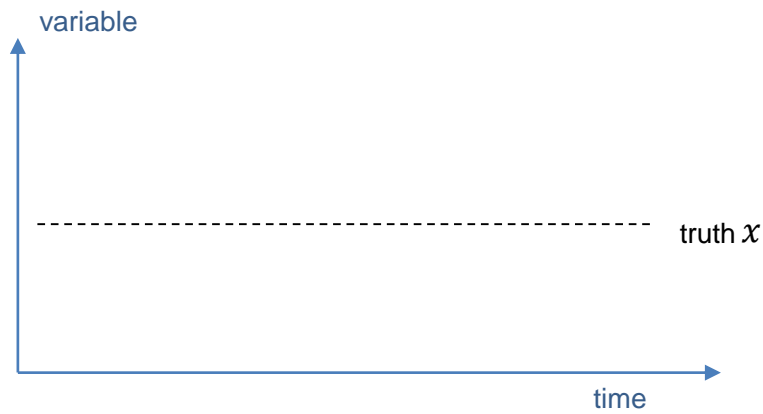
- The ICAP-MME is run daily w/ 1x1 deg res at 00Z for 6 hrly fcasts out to 120 hrs w/ a 1-day latency.
- Modal AOT (550nm) and dust AOT (550nm) data in NetCDF is available at [http://usgodae.org/cgi-bin/datalist.pl?dset=nrl\\_icap\\_mme&summary=Go](http://usgodae.org/cgi-bin/datalist.pl?dset=nrl_icap_mme&summary=Go)



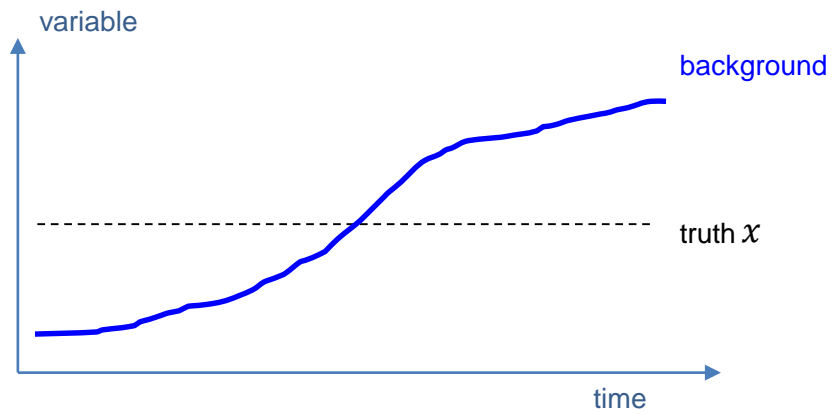
# Basic principles



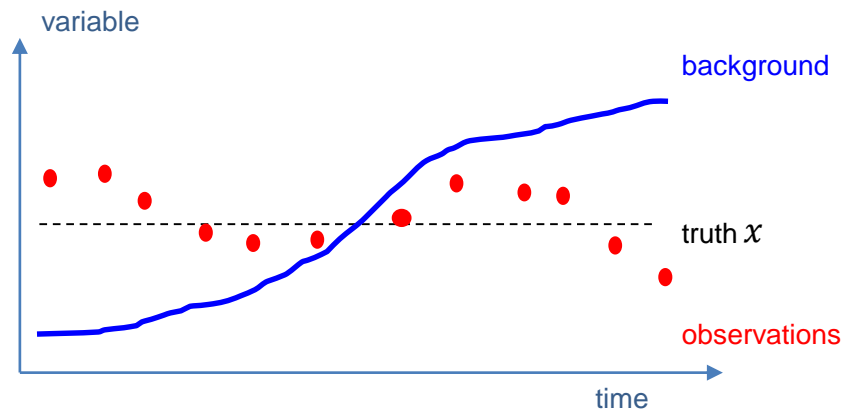
# Optimal Estimation Approach



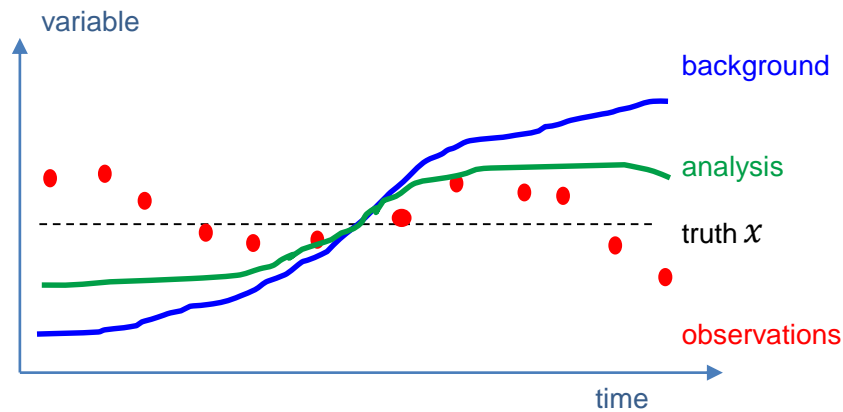
# Optimal Estimation Approach



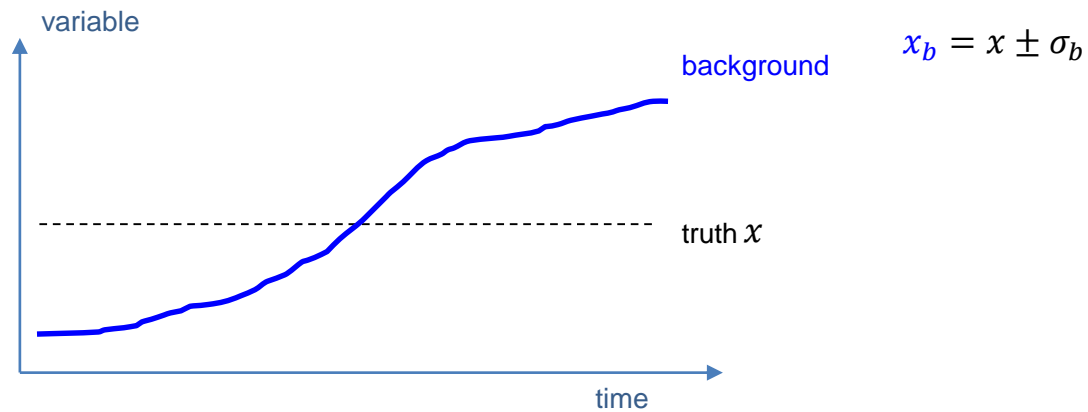
# Optimal Estimation Approach



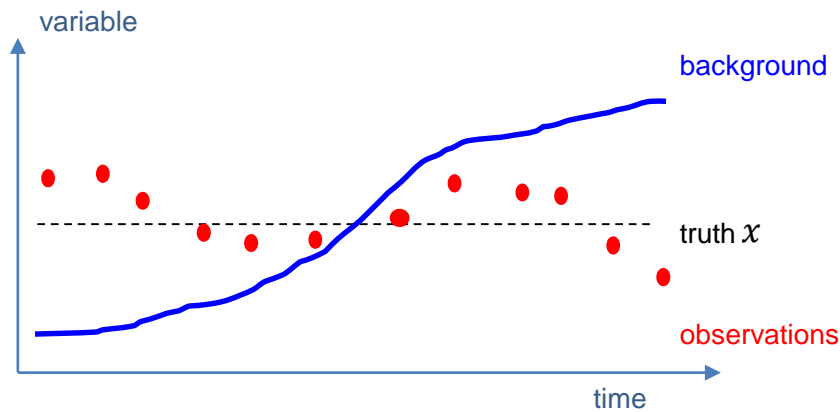
# Optimal Estimation Approach



# Optimal Estimation Approach



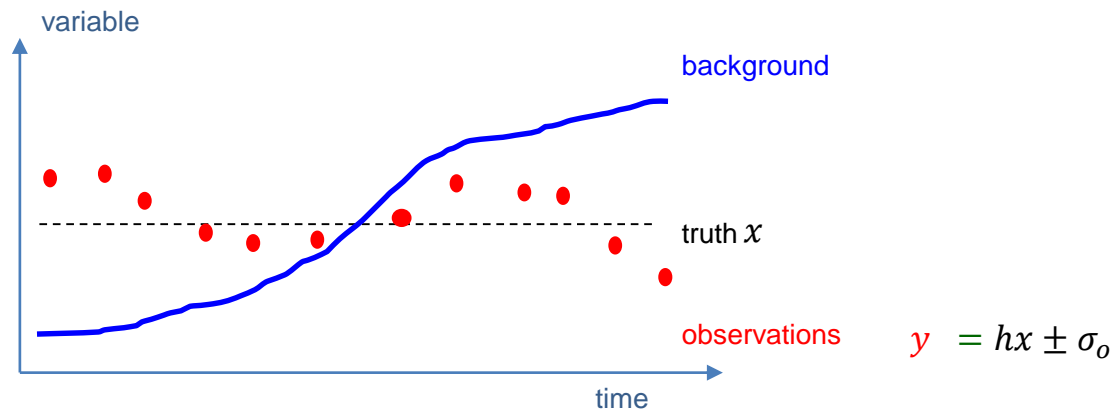
# Optimal Estimation Approach



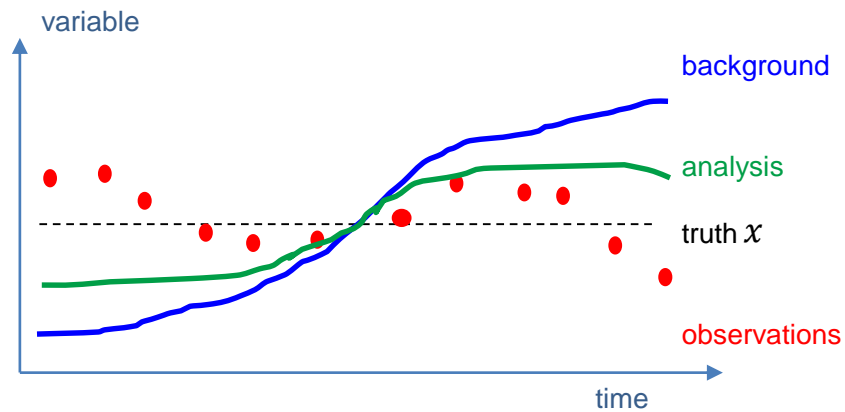
$$y = H(x) \pm \sigma_o$$

$$\sigma_o^2 = \sigma_i^2 + \sigma_m^2$$

# Optimal Estimation Approach



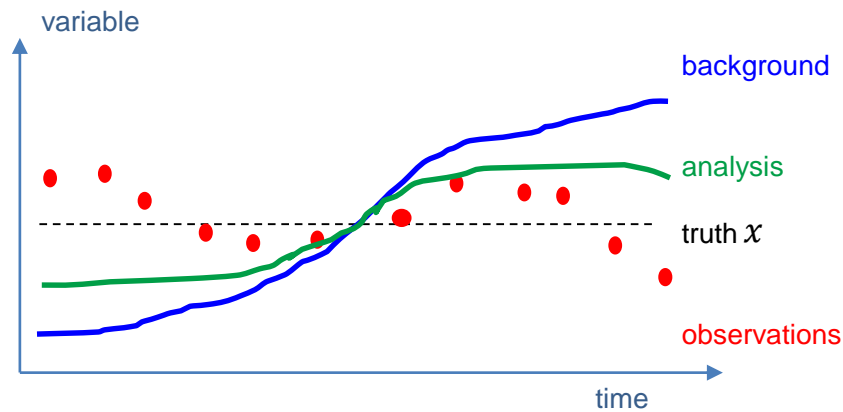
# Optimal Estimation Approach



$$x_a = x_b + g(y - hx_b)$$



# Optimal Estimation Approach

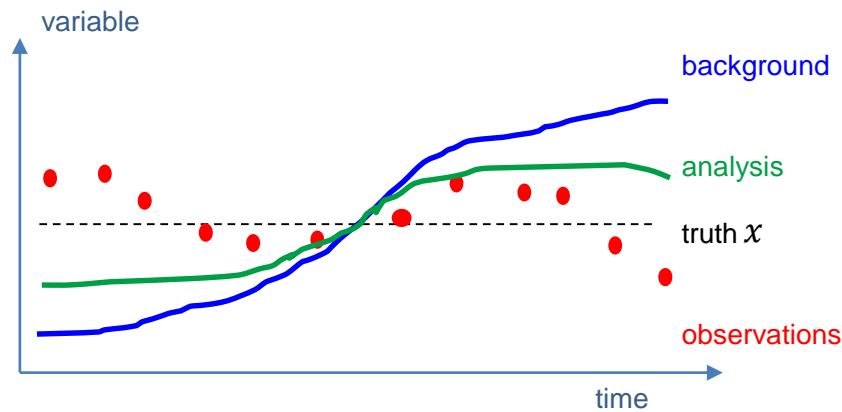


$$x_a = x_b + g(y - hx_b)$$

where

$$g = \frac{h\sigma_b^2}{h^2\sigma_b^2 + \sigma_o^2}$$

# Optimal Estimation Approach

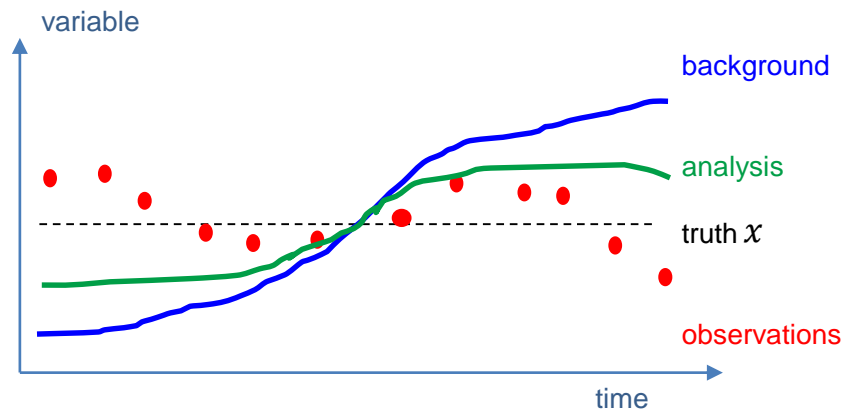


$$x_a = x_b + g(y - hx_b)$$

where  $g = \frac{h\sigma_b^2}{h^2\sigma_b^2 + \sigma_o^2}$

$$\sigma_b \ll \sigma_o/h \quad g \rightarrow 0 \quad x_a \rightarrow x_b$$

# Optimal Estimation Approach



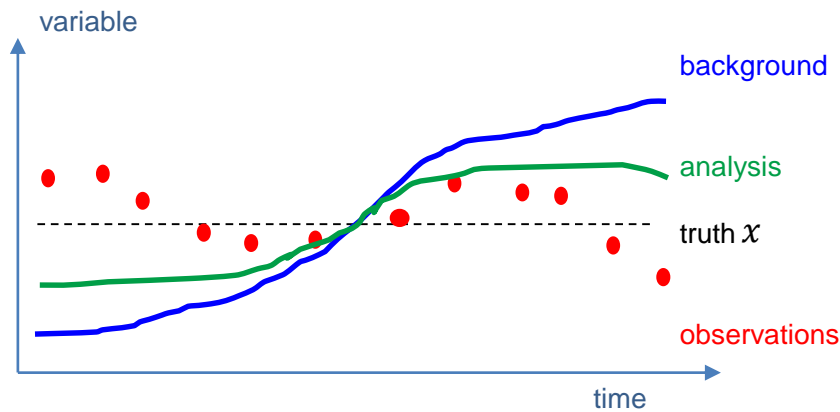
$$x_a = x_b + g(y - hx_b)$$

where  $g = \frac{h\sigma_b^2}{h^2\sigma_b^2 + \sigma_o^2}$

$$\sigma_b \ll \sigma_o/h \quad g \rightarrow 0 \quad x_a \rightarrow x_b$$

$$\sigma_b \gg \sigma_o/h \quad g \rightarrow 1/h \quad x_a \rightarrow y/h$$

# Optimal Estimation Approach

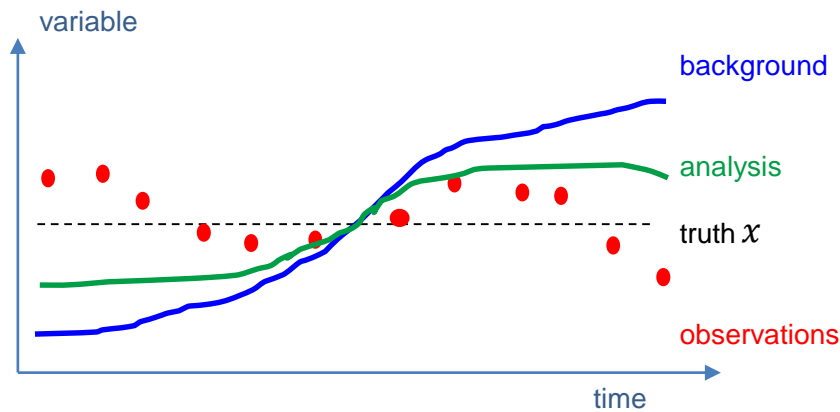


$$x_a = x_b + g(y - hx_b)$$
$$\sigma_a^2 = \left( \frac{1}{\sigma_b^2} + \frac{1}{(\sigma_o/h)^2} \right)^{-1}$$

where

$$g = \frac{h\sigma_b^2}{h^2\sigma_b^2 + \sigma_o^2}$$

# Optimal Estimation Approach

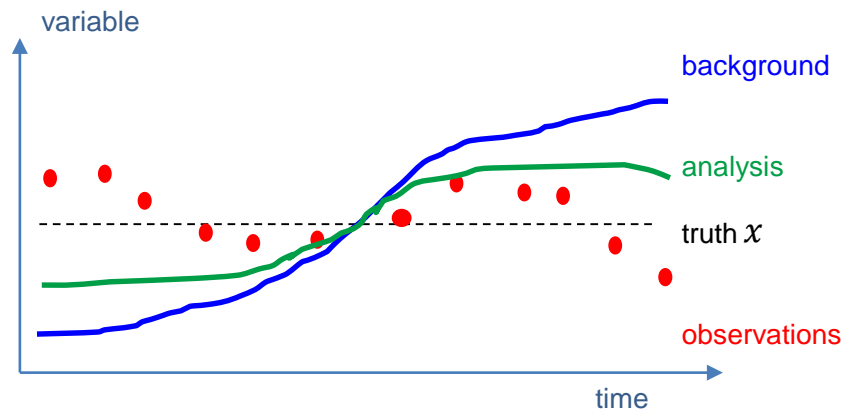


$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{G}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

$$\mathbf{S}_a = (\mathbf{H}^T \mathbf{S}_o^{-1} \mathbf{H} + \mathbf{S}_b^{-1})^{-1}$$

where  $\mathbf{G} = \mathbf{S}_b \mathbf{H}^T (\mathbf{H} \mathbf{S}_b \mathbf{H}^T + \mathbf{S}_o)^{-1}$

# Optimal Estimation Approach



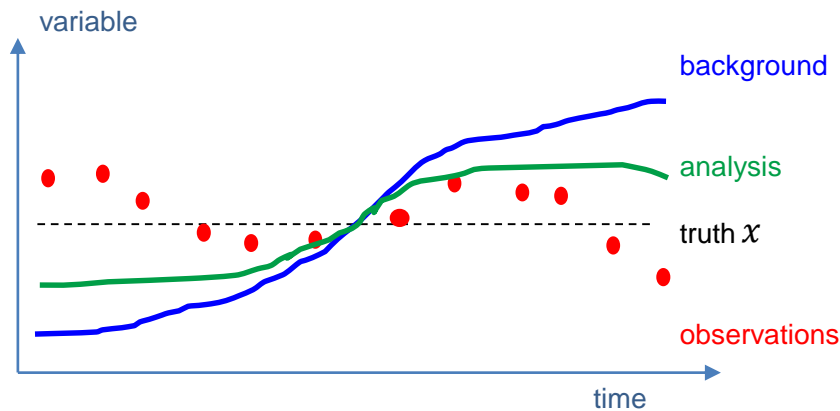
$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{G}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

$$\mathbf{S}_a = (\mathbf{H}^T \mathbf{S}_o^{-1} \mathbf{H} + \mathbf{S}_b^{-1})^{-1}$$

where  $\mathbf{G} = \mathbf{S}_b \mathbf{H}^T (\mathbf{H} \mathbf{S}_b \mathbf{H}^T + \mathbf{S}_o)^{-1}$

**observation error  
covariance matrix**

# Optimal Estimation Approach



$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{G}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

$$\mathbf{S}_a = (\mathbf{H}^T \mathbf{S}_o^{-1} \mathbf{H} + \mathbf{S}_b^{-1})^{-1}$$

where  $\mathbf{G} = \mathbf{S}_b \mathbf{H}^T (\mathbf{H} \mathbf{S}_b \mathbf{H}^T + \mathbf{S}_o)^{-1}$

**Background error  
covariance matrix**

- **spatial spreading of information from observations**

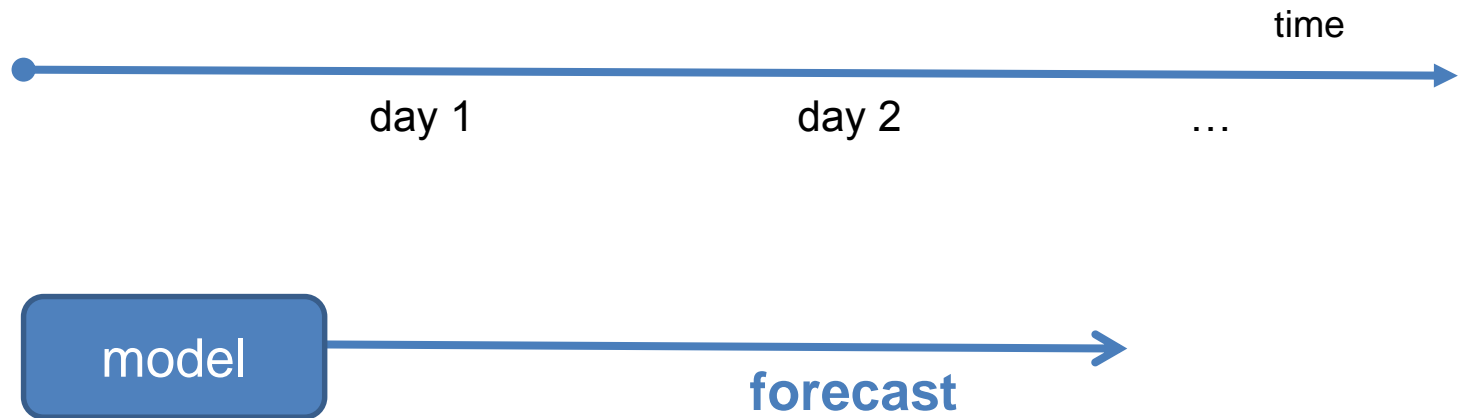


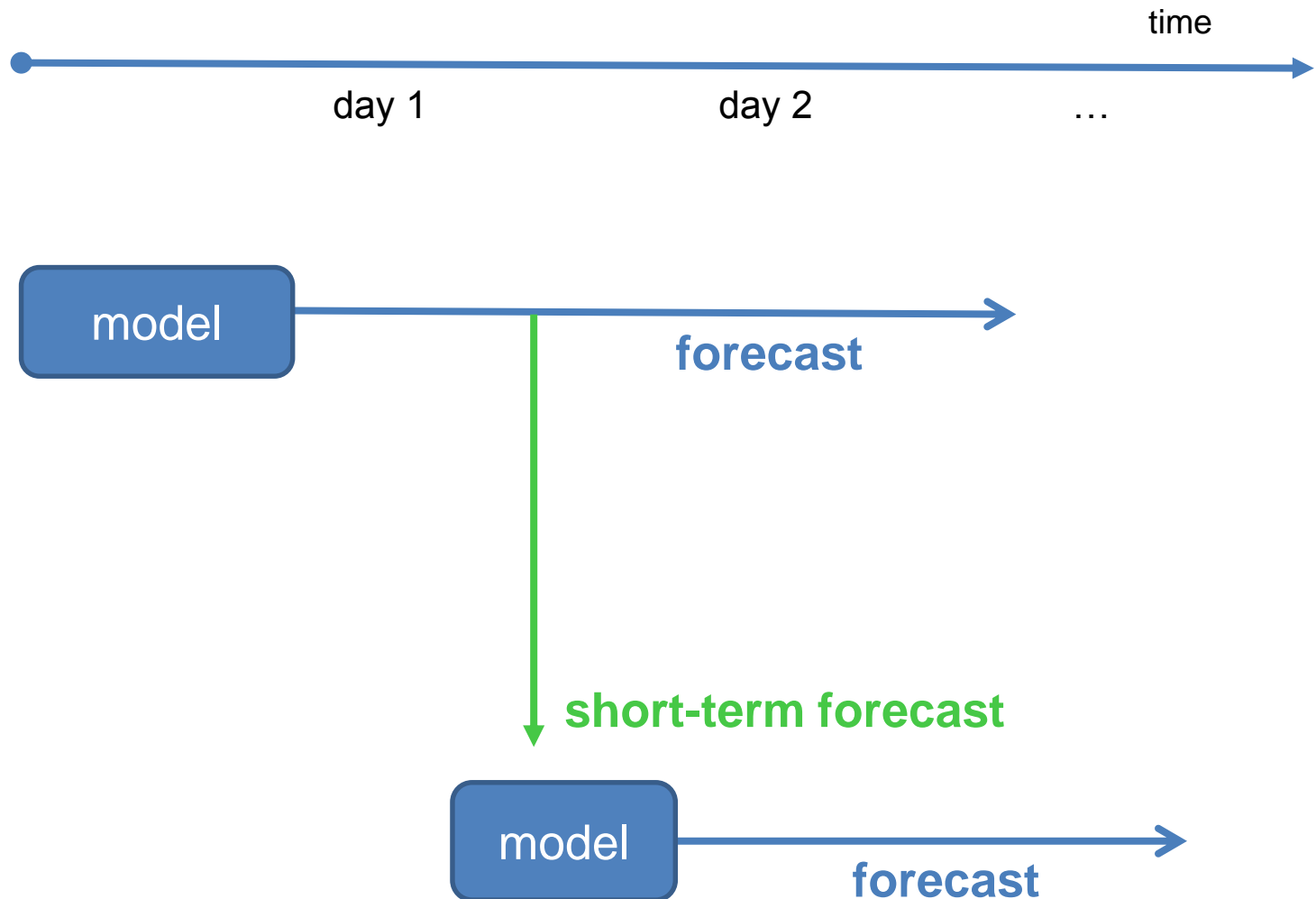
- **spatial spreading of information from observations**
- **statistically consistent increments between neighbouring grid points**

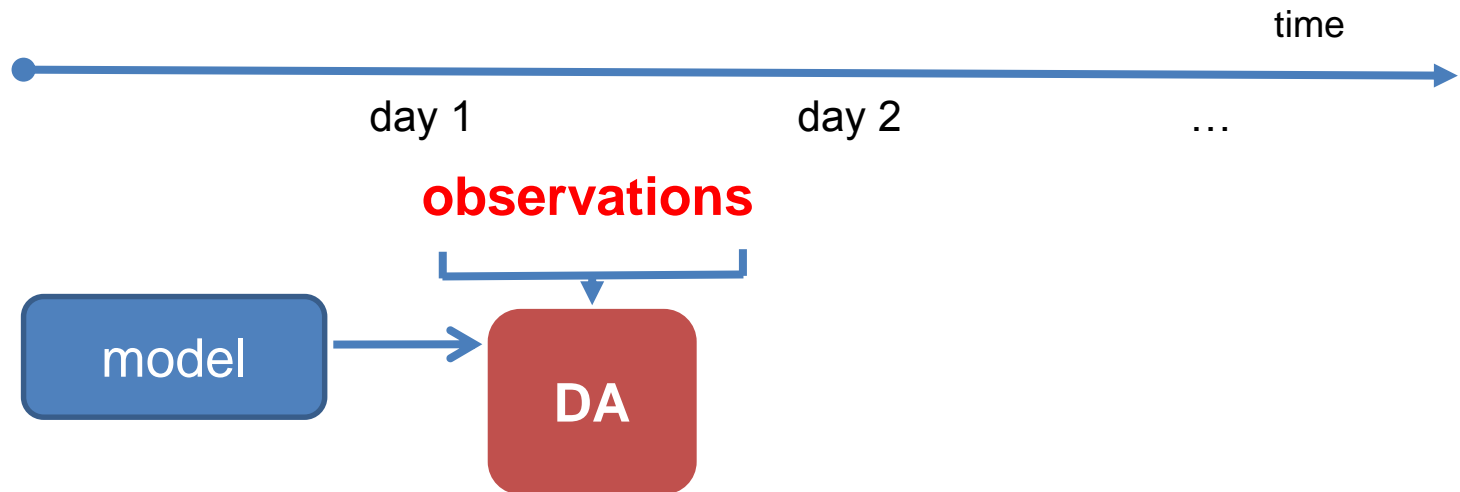
- **spatial spreading of information from observations**
- **statistically consistent increments between neighbouring grid points**
- **multivariate analysis**

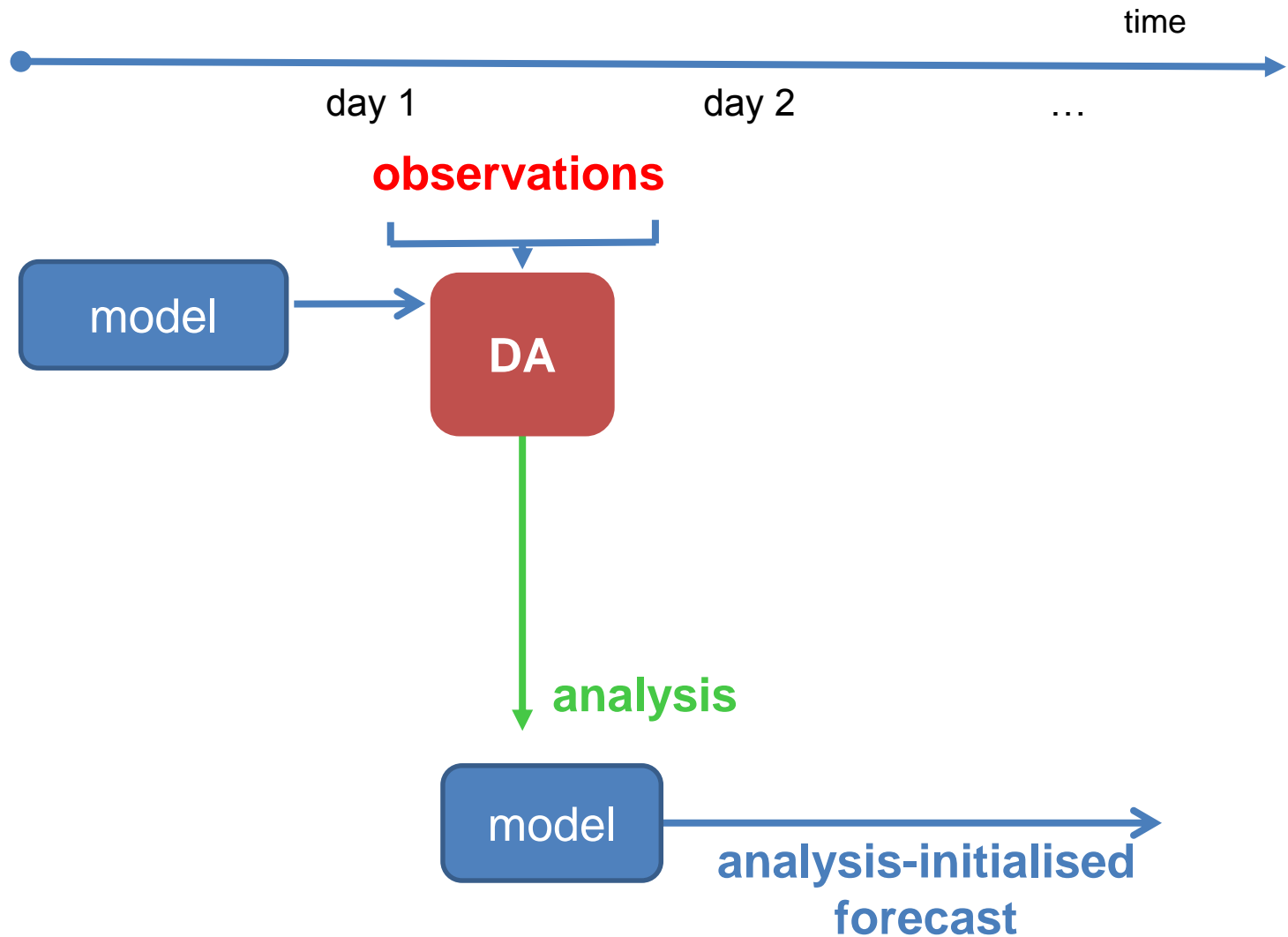


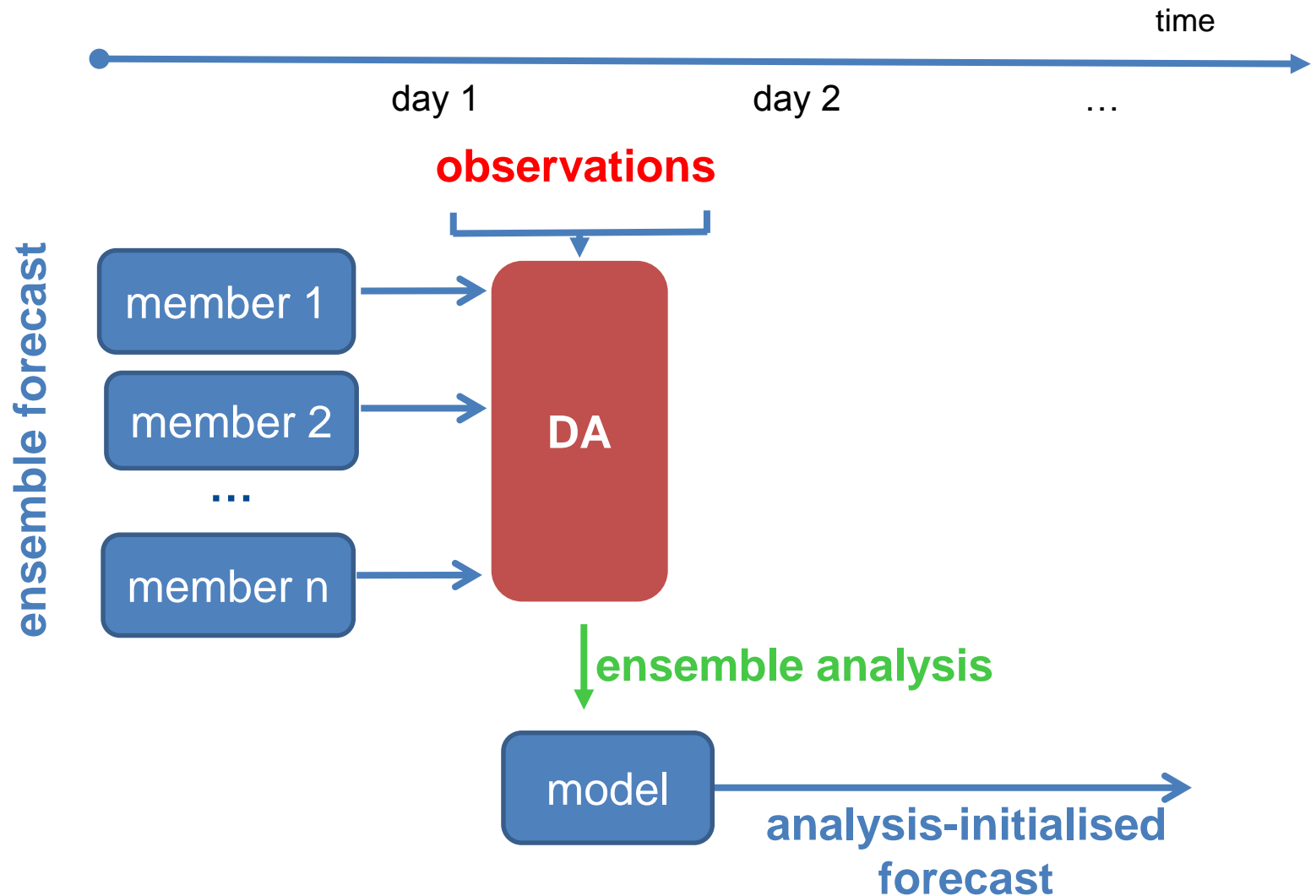
DA flow



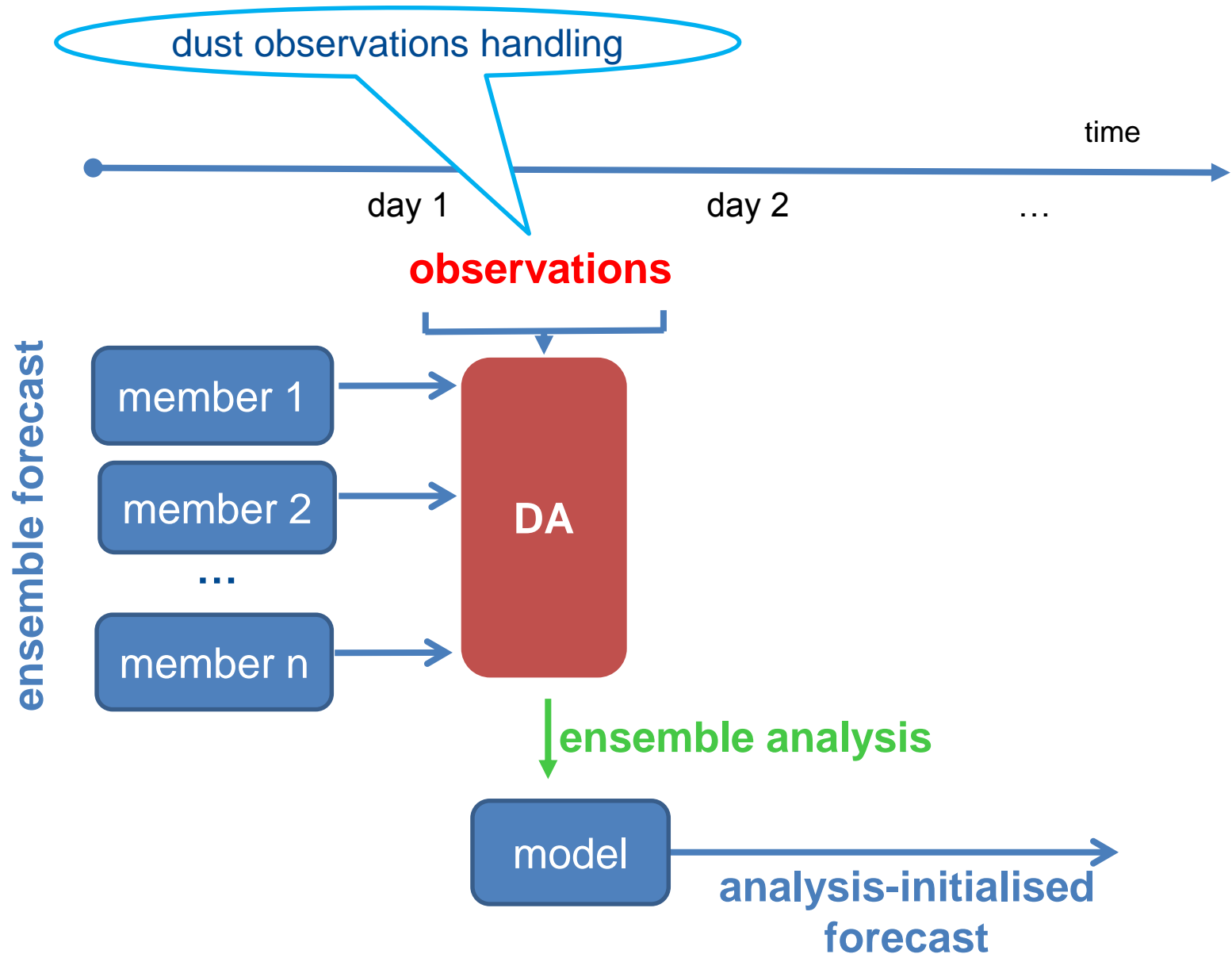


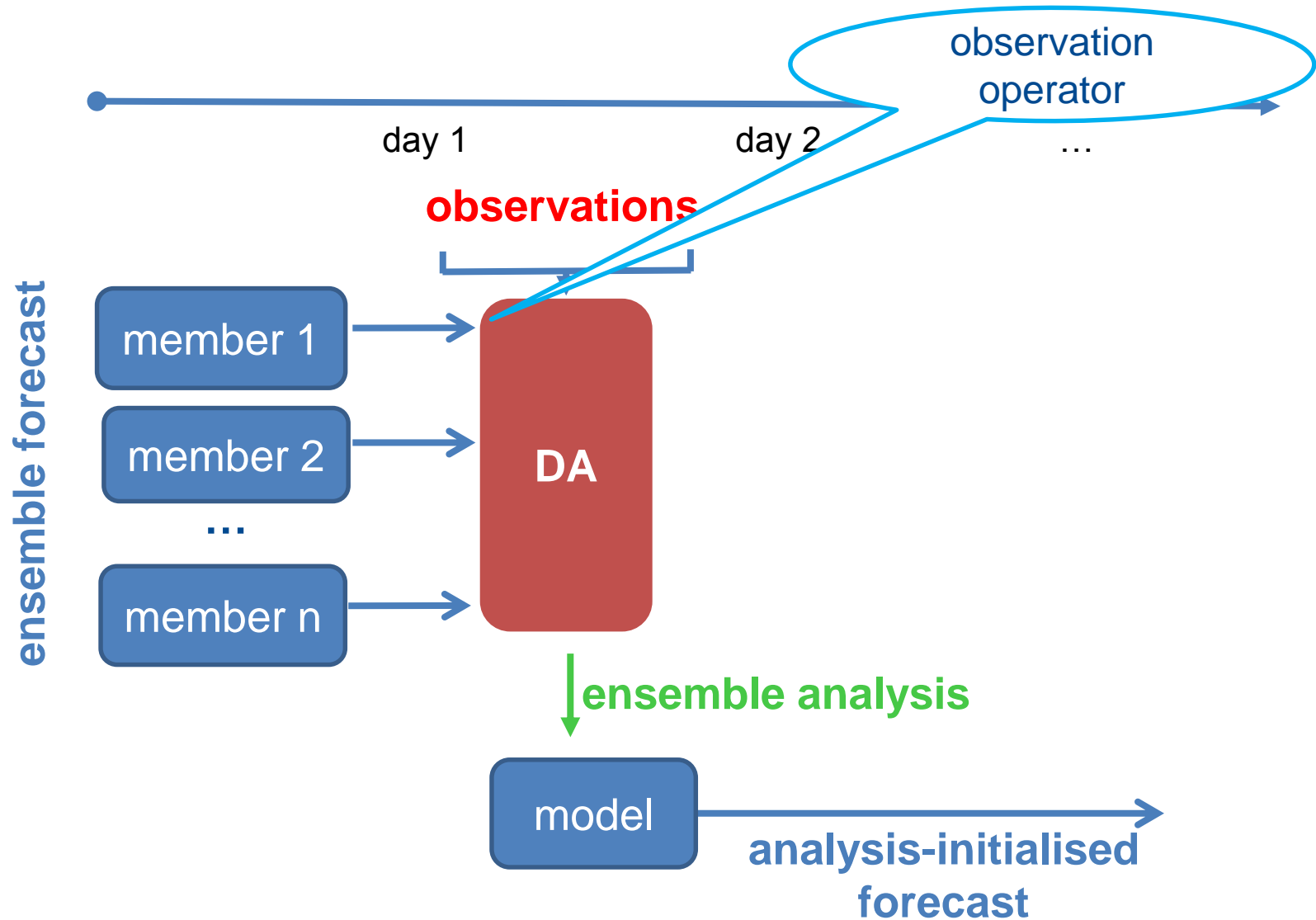


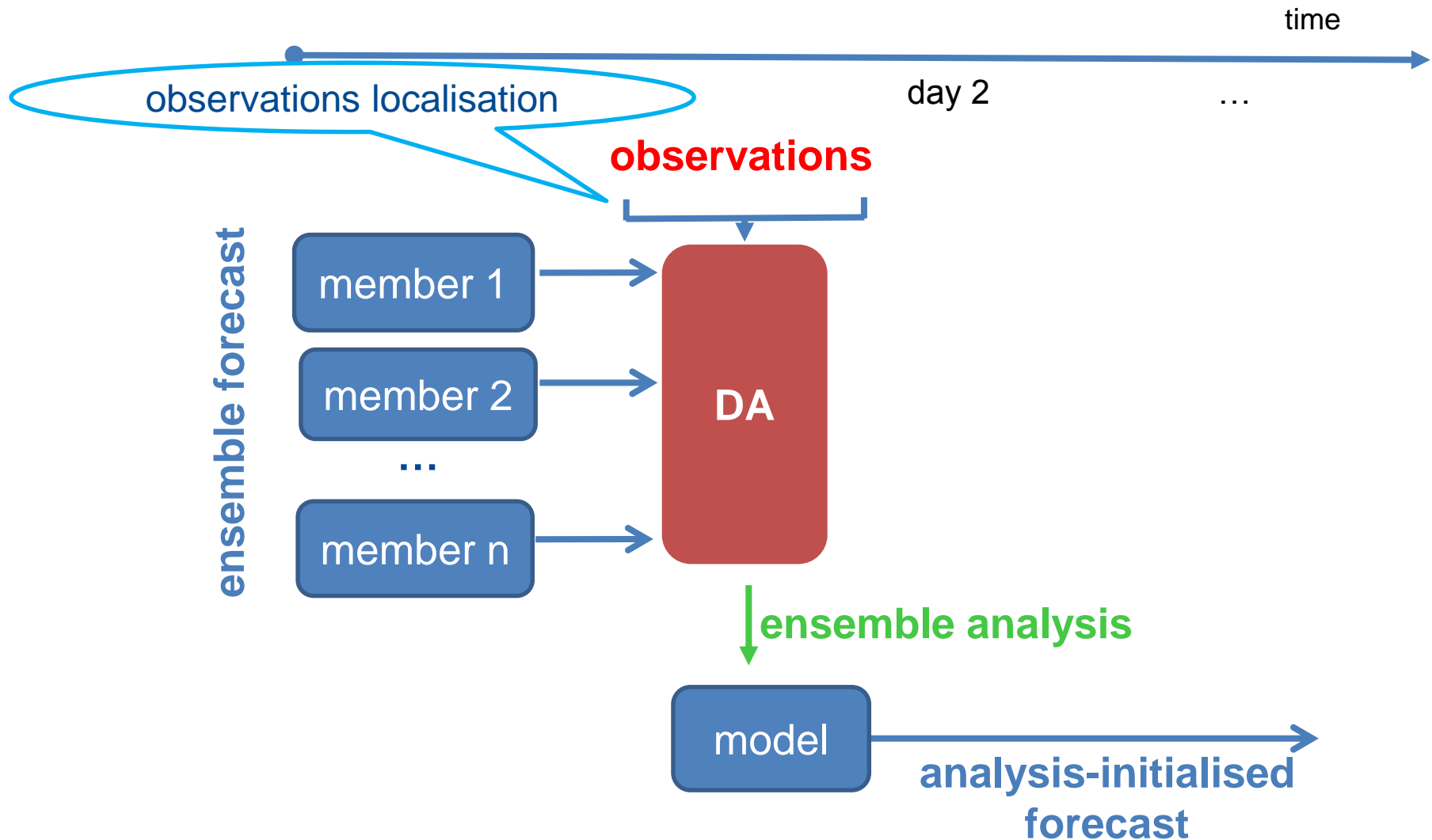


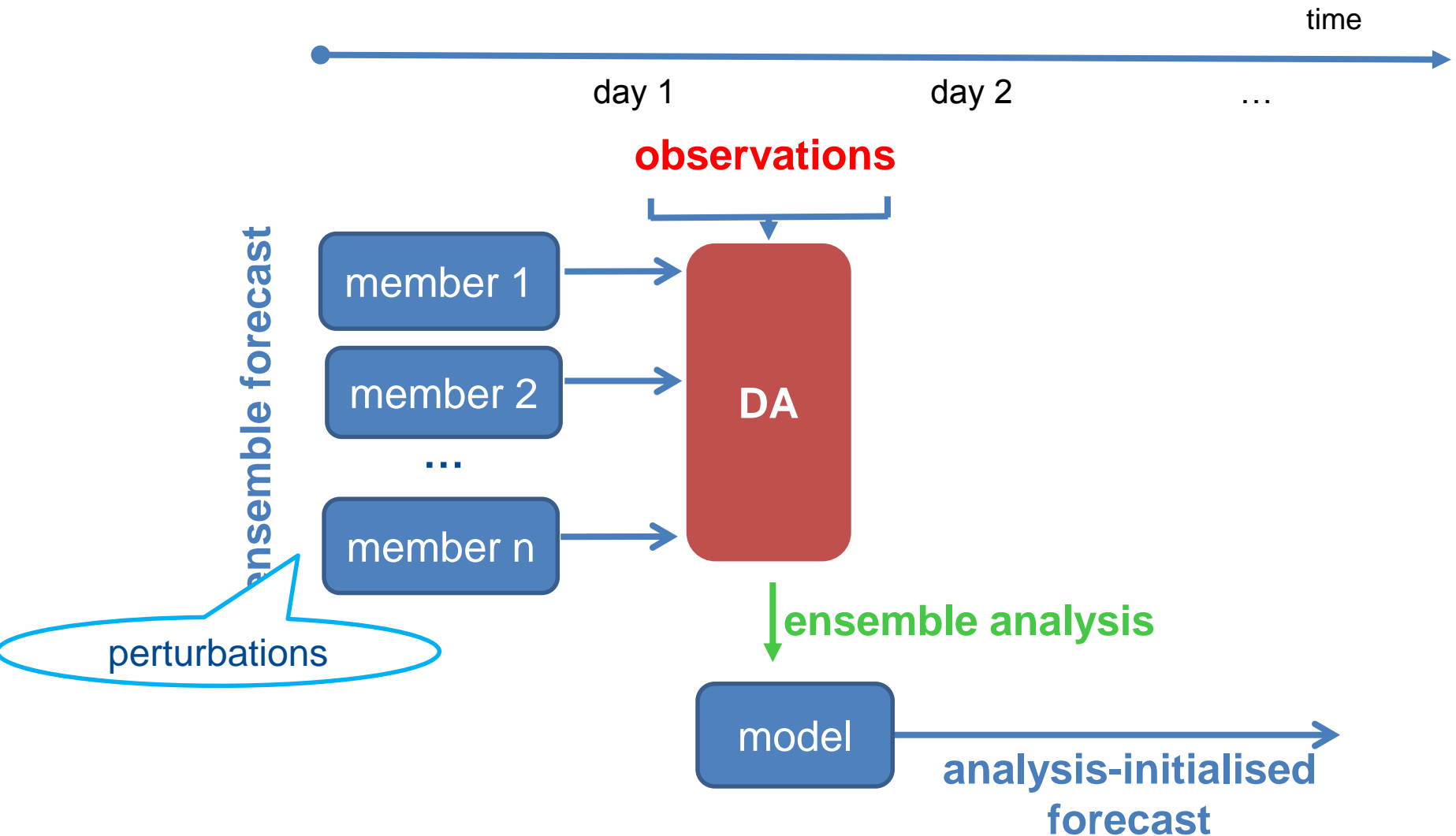


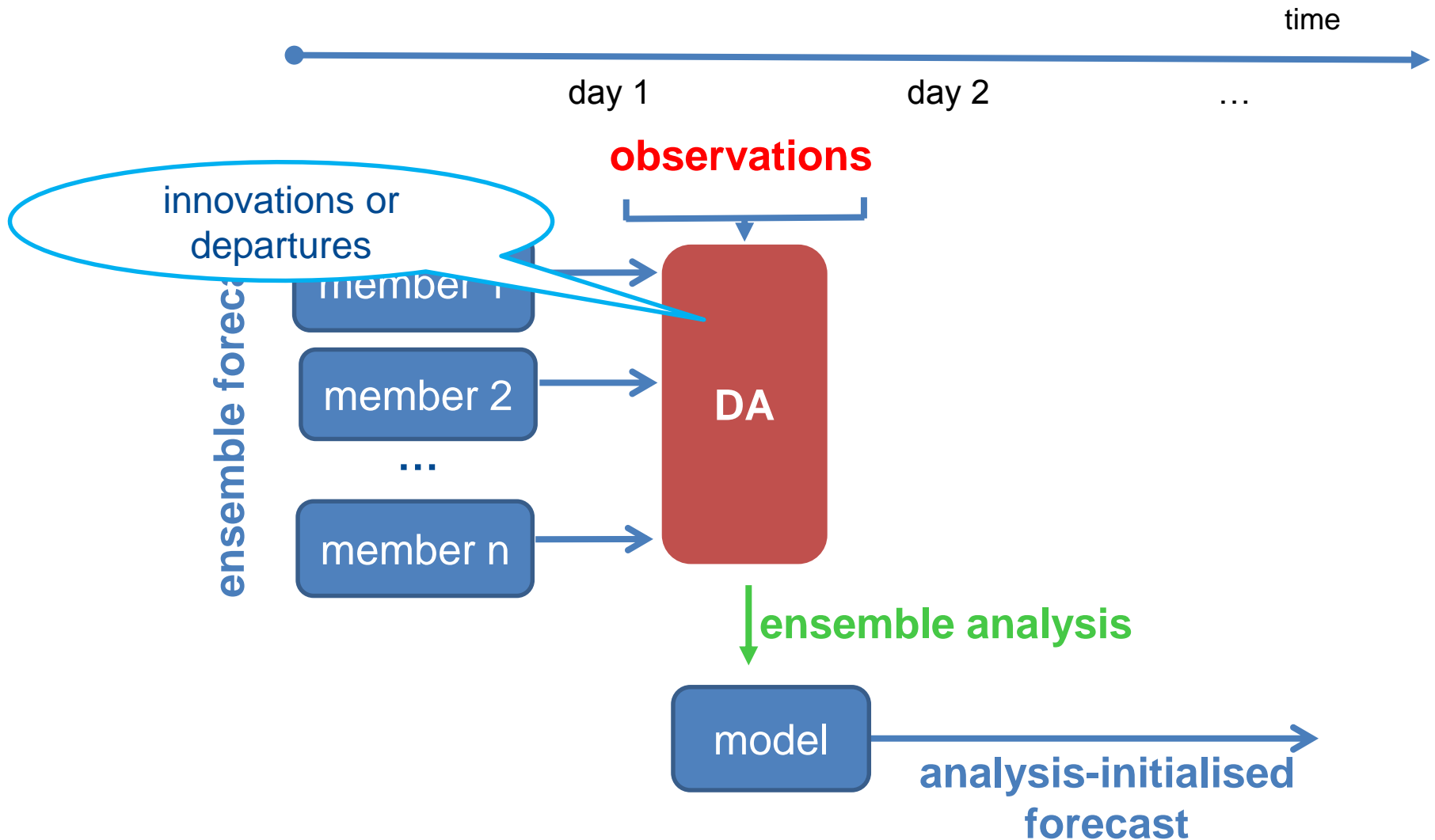


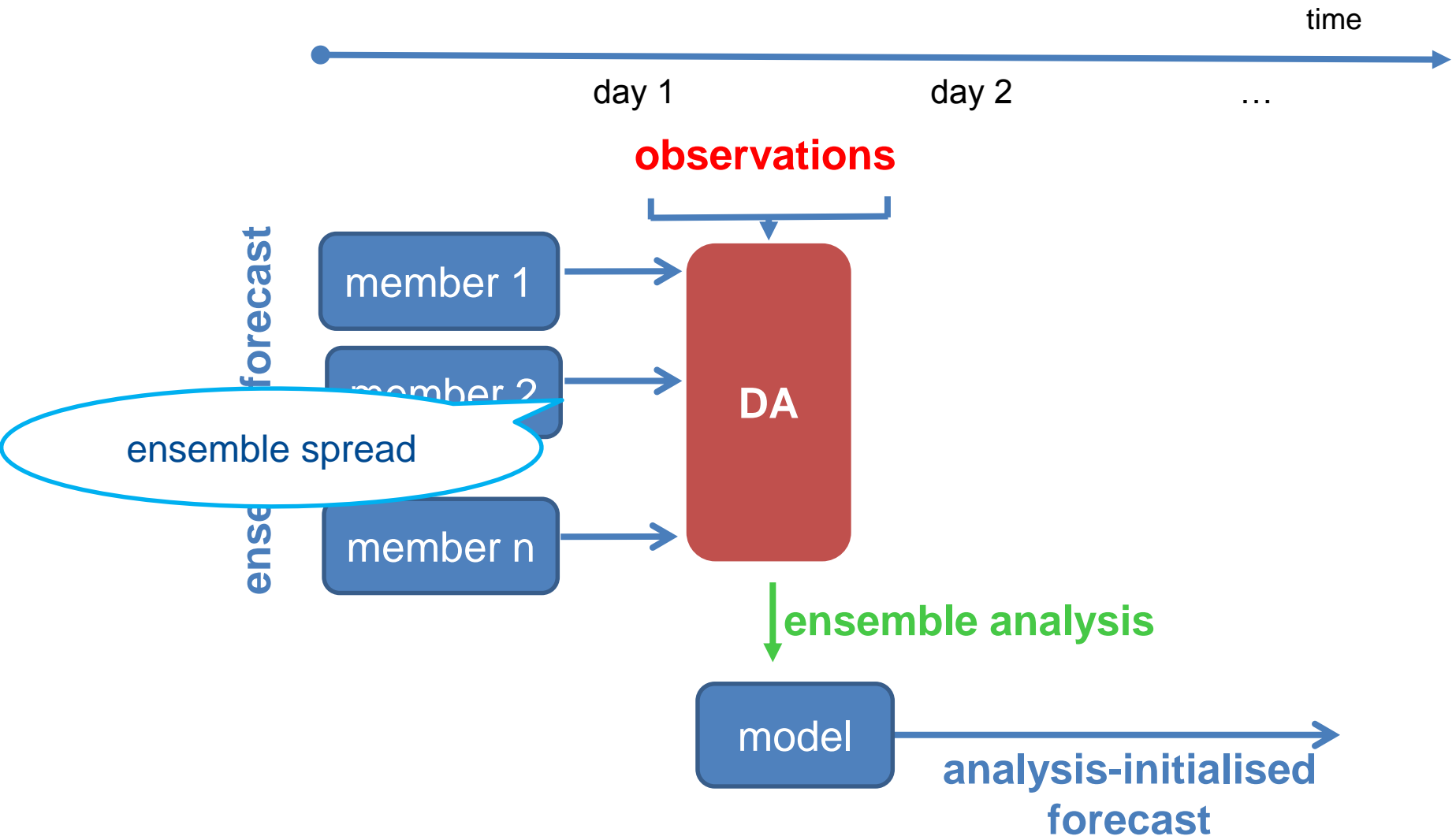


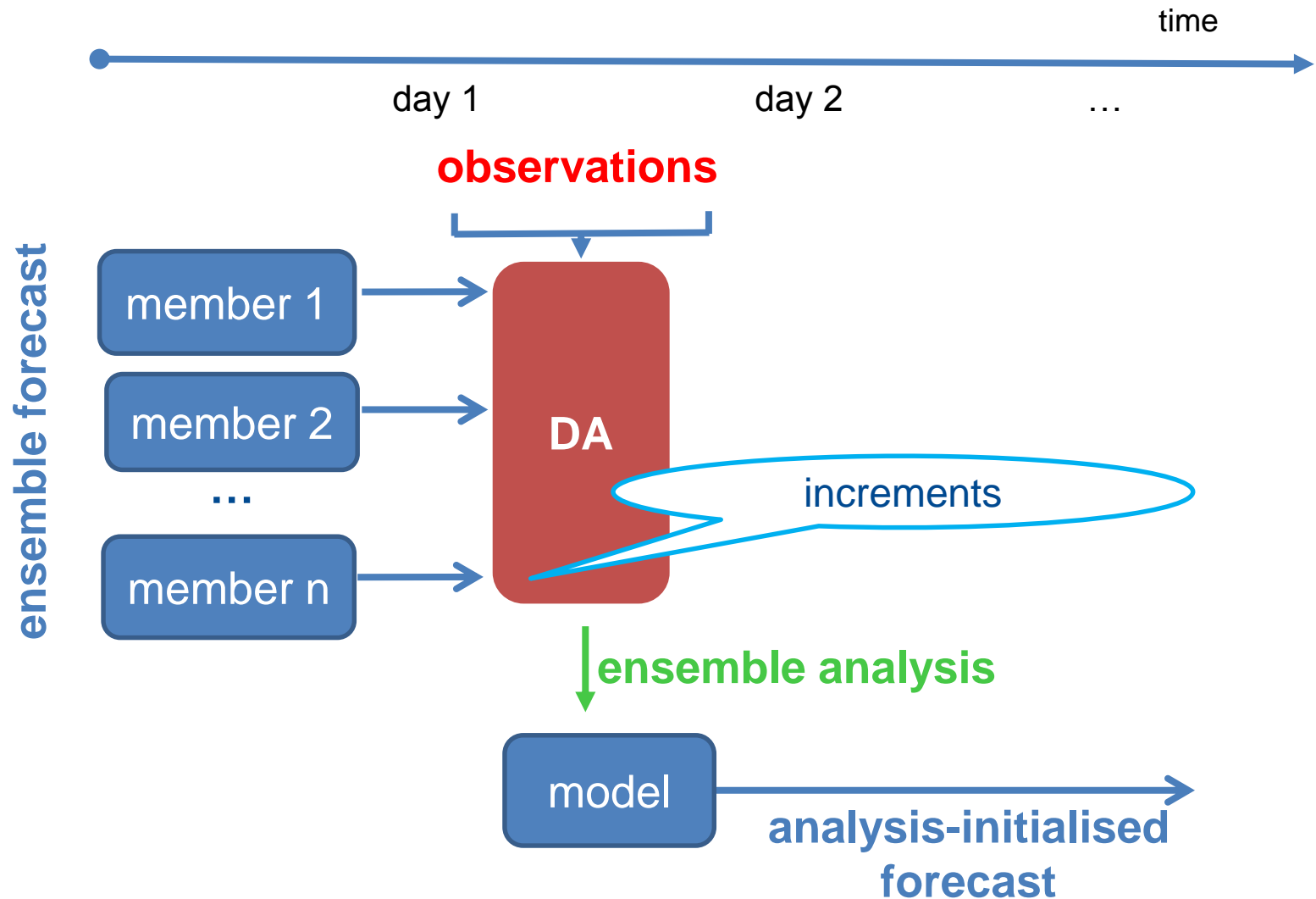


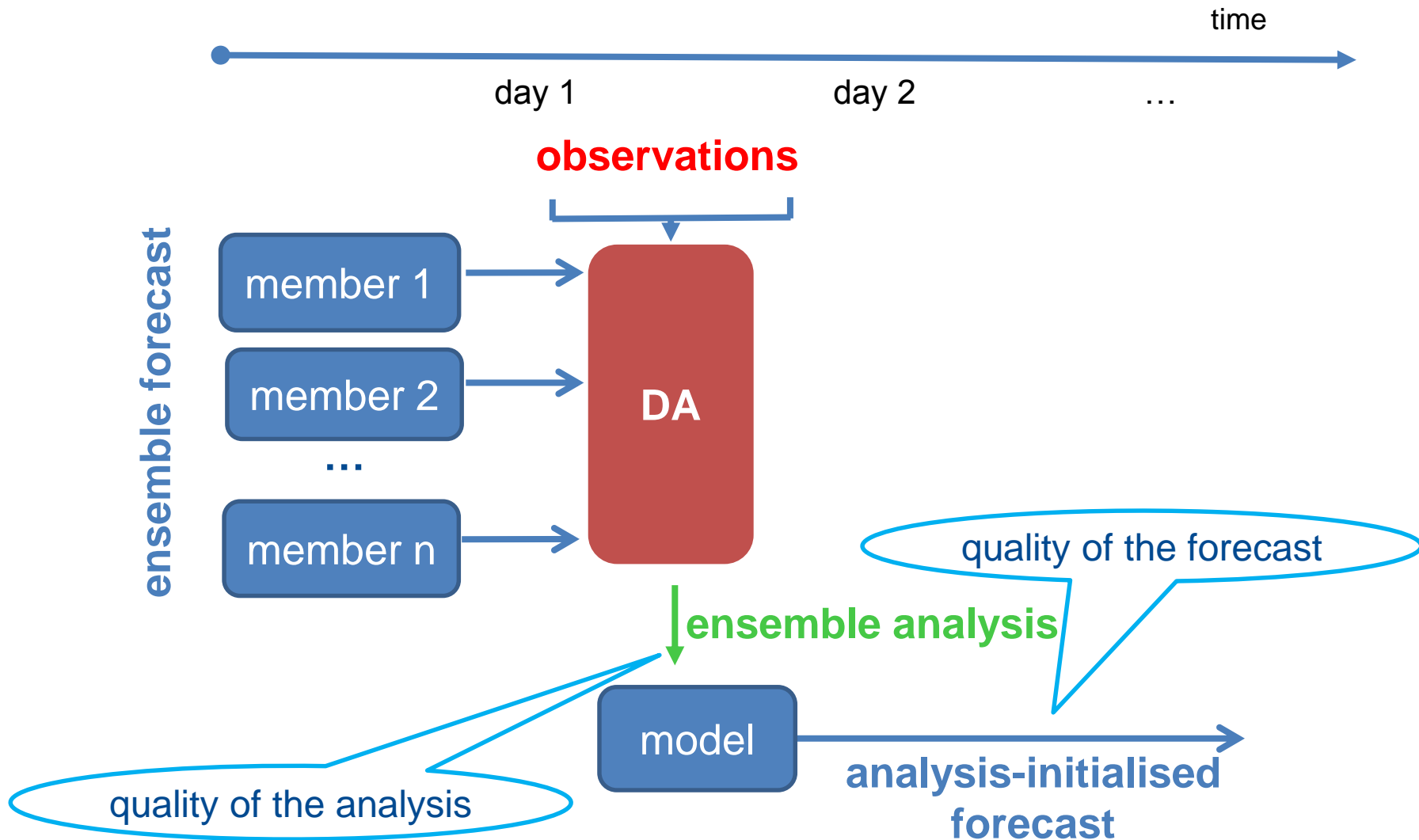
















**Barcelona  
Supercomputing  
Center**

*Centro Nacional de Supercomputación*



EXCELENCIA  
SEVERO  
OCHOA

# Thank you!

For further information please contact  
[enza.ditomaso@bsc.es](mailto:enza.ditomaso@bsc.es)