The detection and assimilation of cloud-affected infrared satellite radiances

ECMWF/EUMETSAT NWP-SAF Satellite Data Assimilation Training Course

23-27 March 2026

Tony McNally, Reima Eresmaa, **Chris Burrows**, Ethel Villeneuve

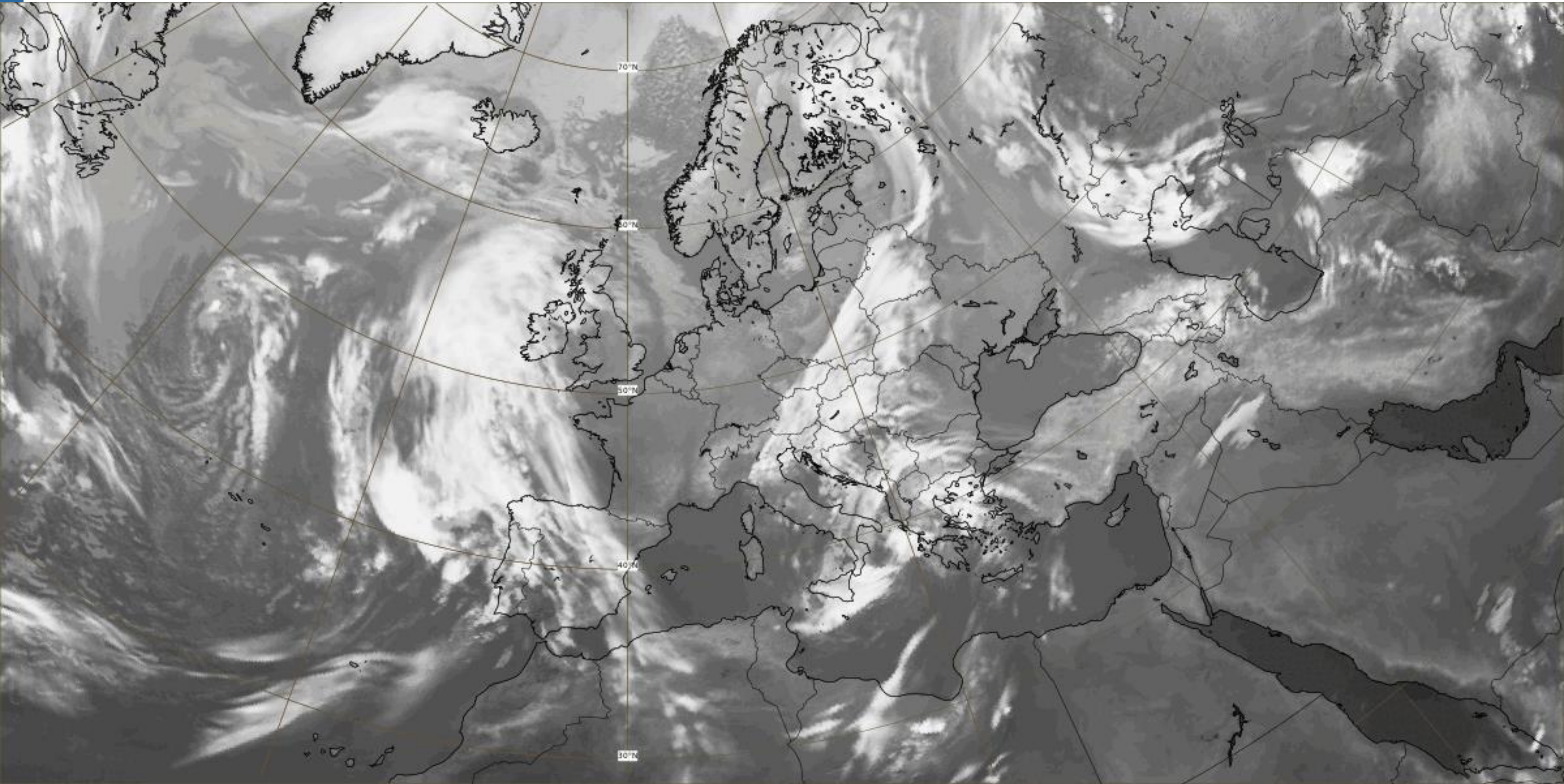
chris.burrows@ecmwf.int

Outline

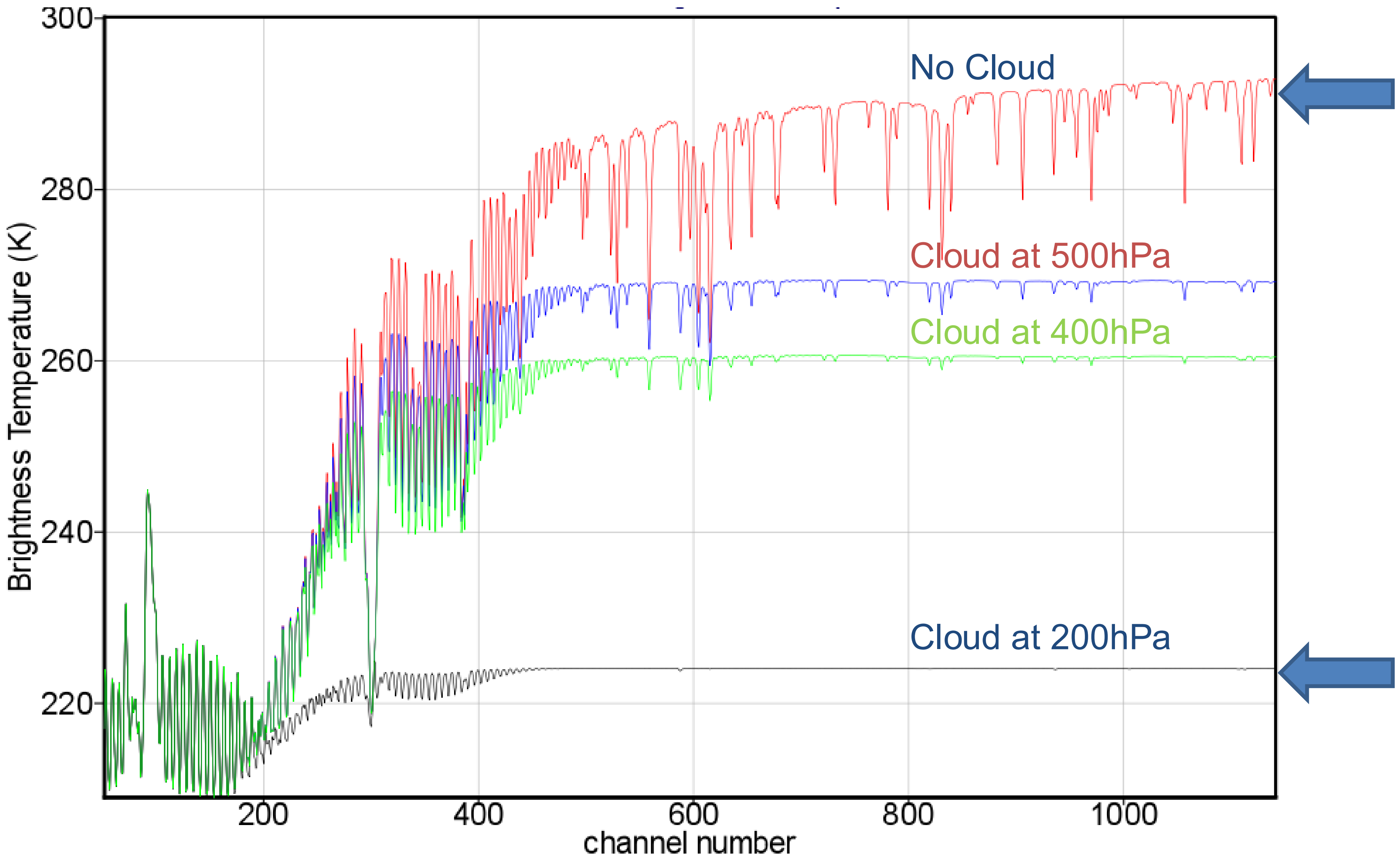
- **The effect of cloud on infrared radiances**
- **Cloud detection/rejection methods:**
  - simple departure checks
  - co-located imager information
  - pattern recognition
  - hybrid approach
  - machine learning
- **“All-sky” infrared assimilation:**
  - simplified approach
  - recent progress

ECMWF

The detection and assimilation of clouds in IR radiances



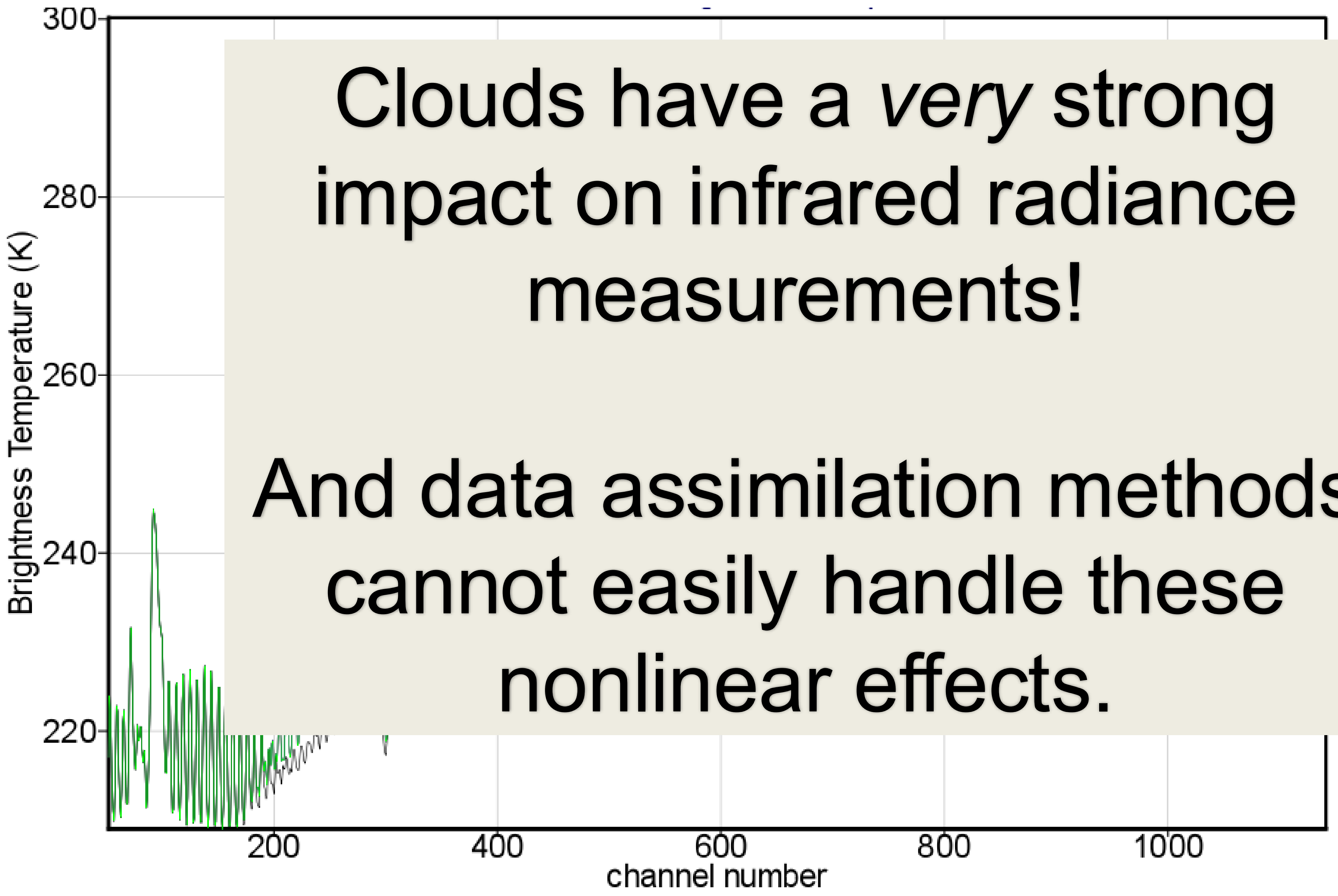
IR spectra with (and without) clouds



Channels sensitive to the surface in clear scenes.

Same channels are strongly sensitive to cold high clouds!

IR spectra with (and without) clouds



Clouds have a *very* strong impact on infrared radiance measurements!

And data assimilation methods cannot easily handle these nonlinear effects.

Channels sensitive to the surface in clear scenes.

Same channels are strongly sensitive to cold high clouds!

Big question: How should we handle clouds when assimilating infrared radiance observations??

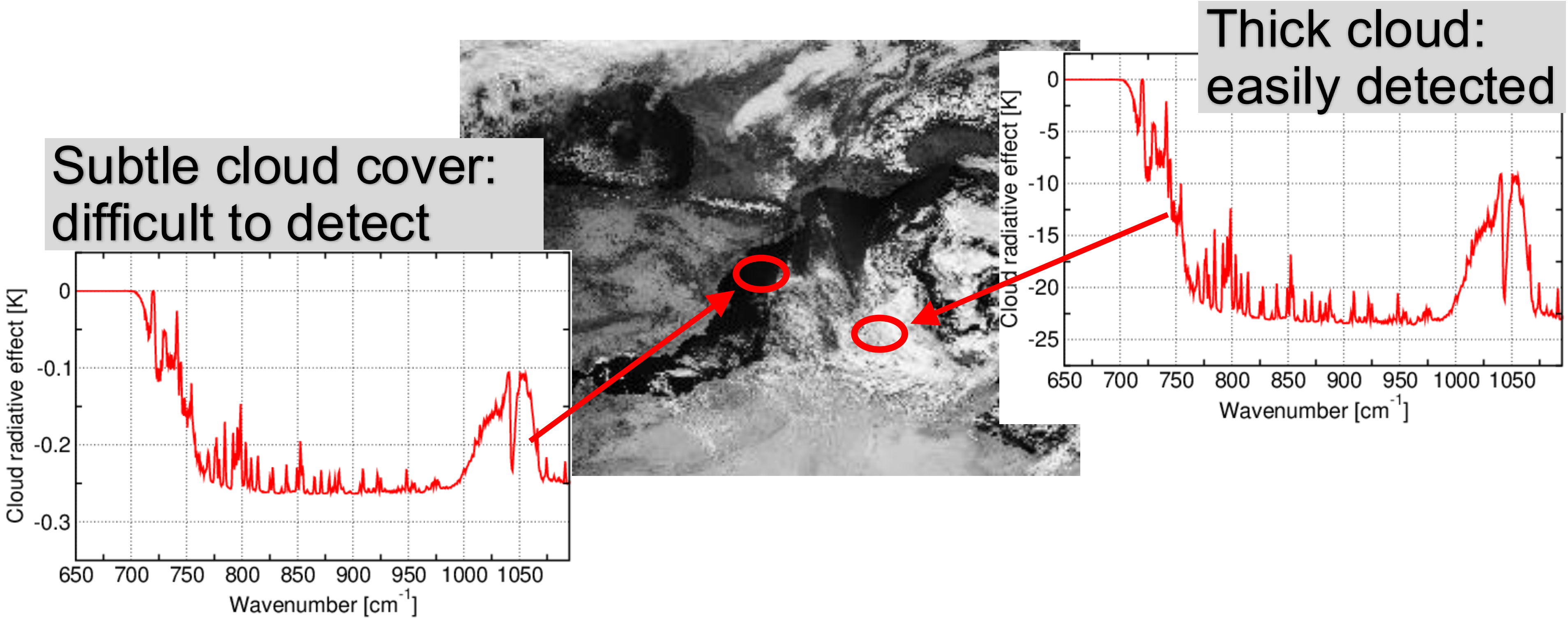
- **Option 1**: detect and reject cloud-contaminated observations.
- **Option 2**: Explicitly estimate cloud parameters from the radiances within the data assimilation (T, Q, O3 etc)

Big question: How should we handle clouds when assimilating infrared radiance observations??

- **Option 1**: detect and reject cloud-contaminated observations.
- **Option 2**: Explicitly estimate cloud parameters from the radiances within the data assimilation (T, Q, O3 etc)

ECMWF

It is not trivial to get the cloud detection right in some IR sounder footprints



There is a "grey zone" where radiative effect of cloud is comparable with meteorological signals we're looking for. Care is needed here, or the analysis can be degraded!

ECMWF

Note!

For this section, most simulated observations use cloud-free model information.

So, cloudy scenes in the observations will (in general) look colder than the model equivalents.

Cloud detection methods

Simple

- Window channel departure (O-B) checks
- Co-located imager checks
- Pattern recognition algorithms
- Hybrid systems

Complex

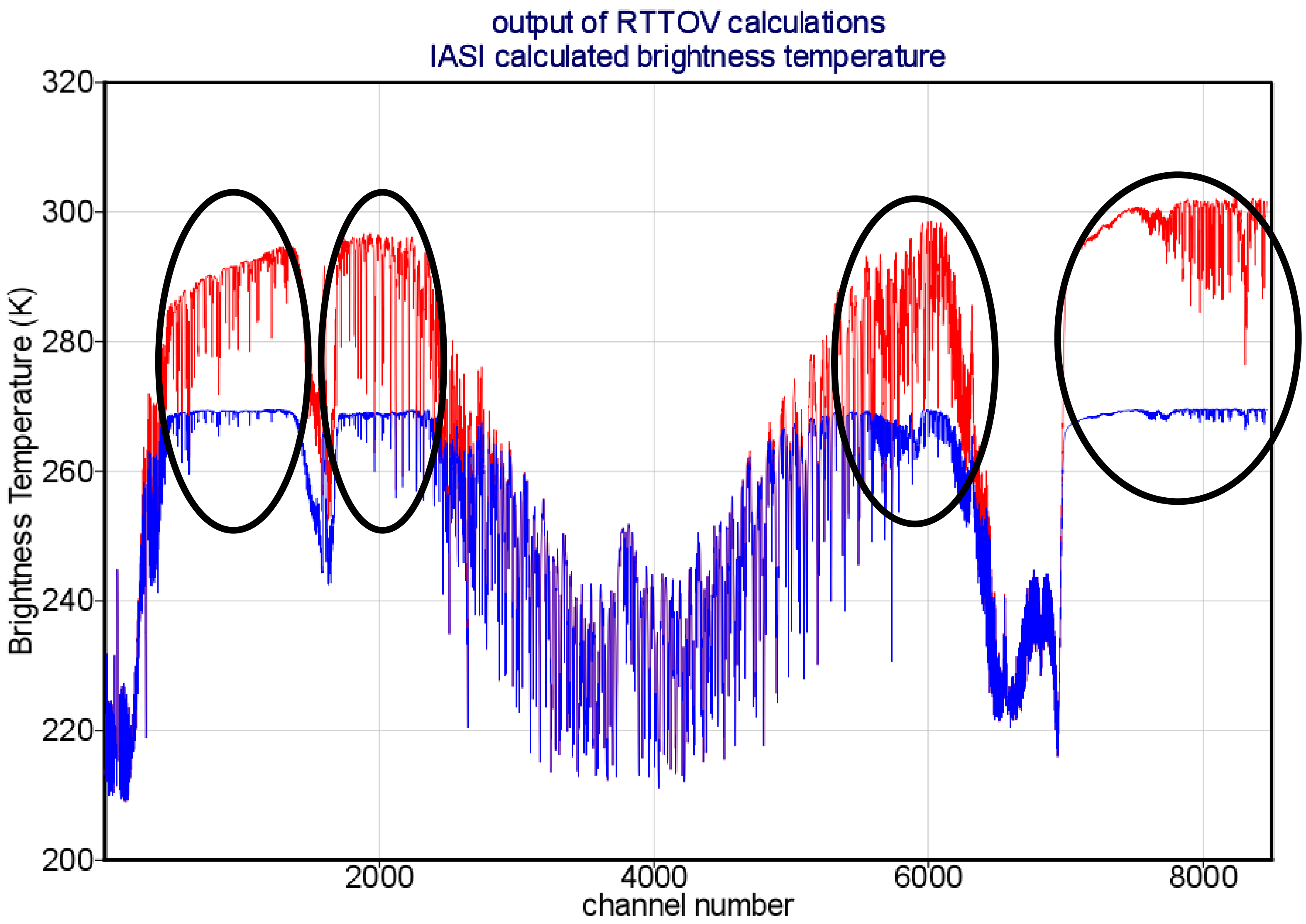
Cloud detection methods

Simple

- Window channel departure (O-B) checks
- Co-located imager checks
- Pattern recognition algorithms
- Hybrid systems

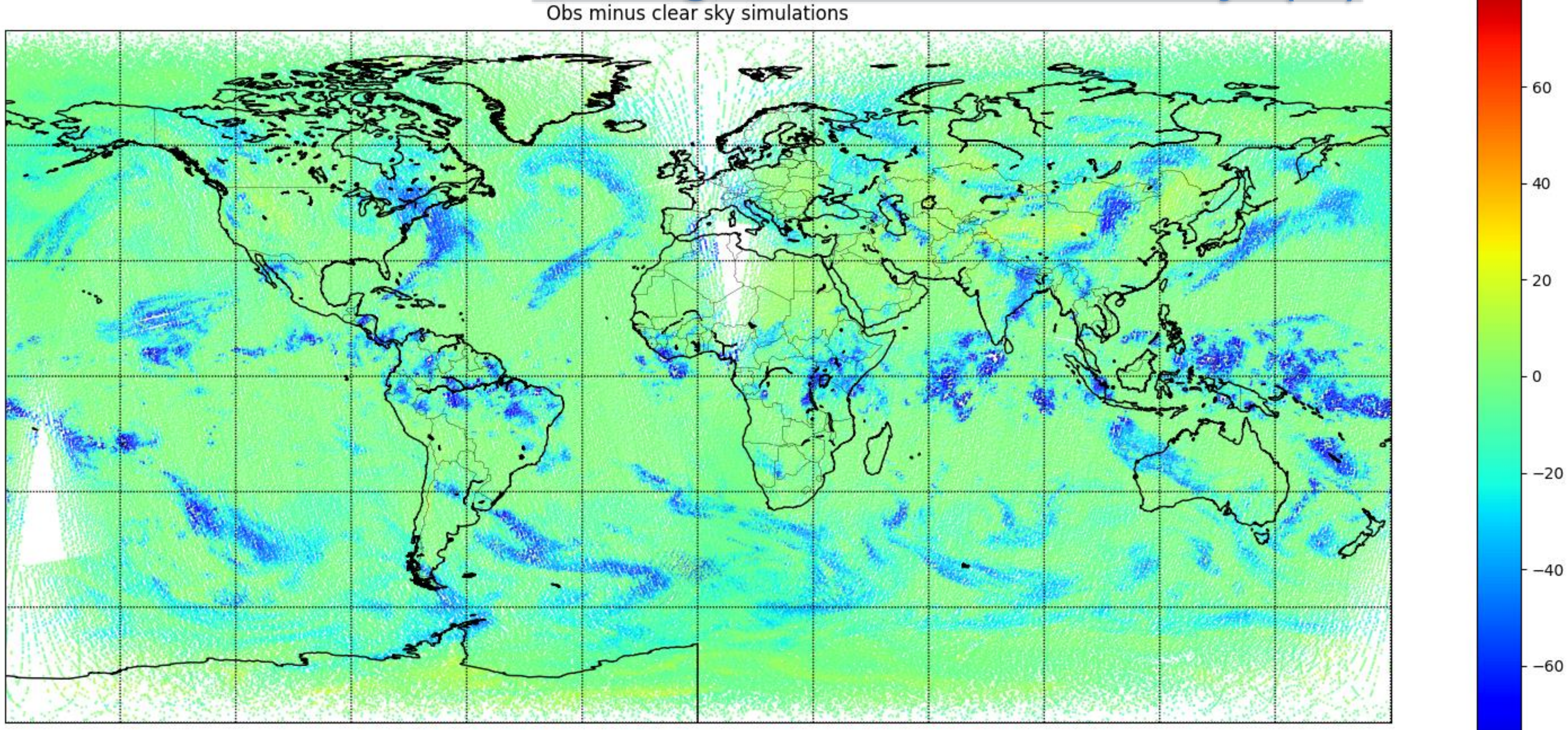
Complex

Window channels have the highest sensitivity to cloud

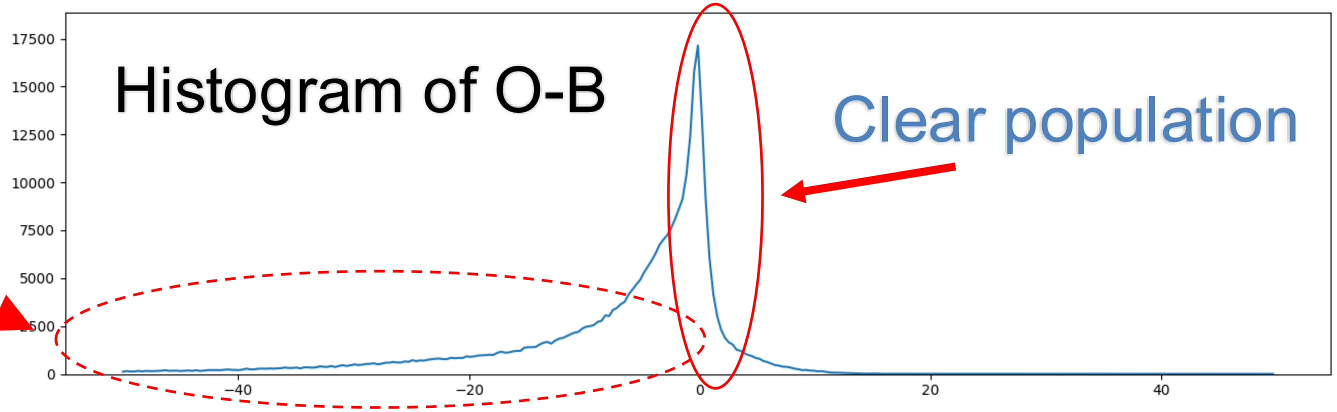


This is the classic spectral signature of the presence of cloud.

Observed radiance at 11 μm minus radiance calculated from background in *clear sky* (K)



Cold departures indicating cloud contamination in obs. The "**cold tail**".

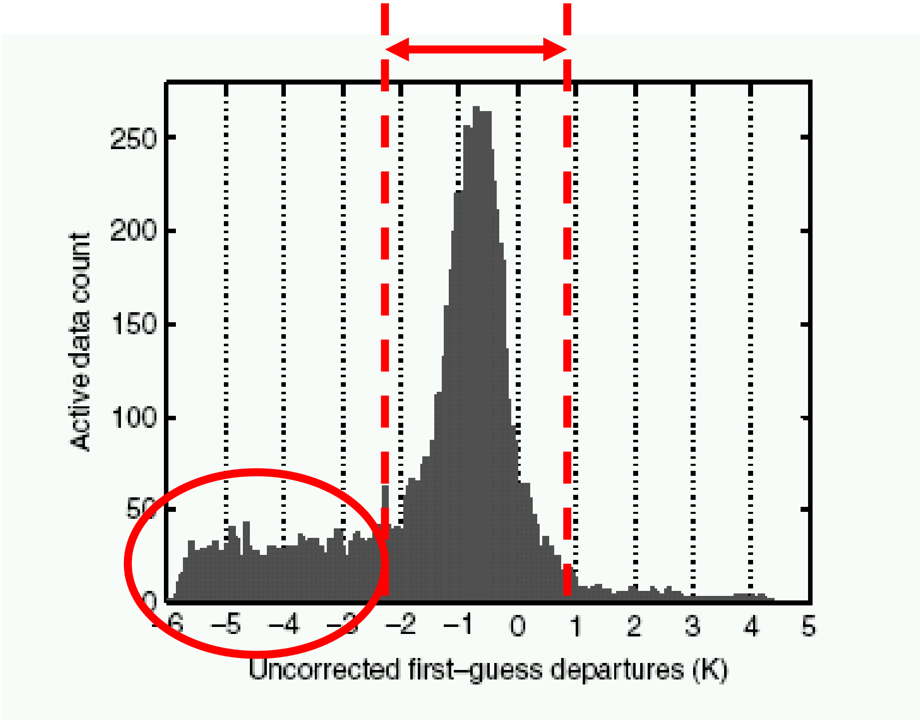


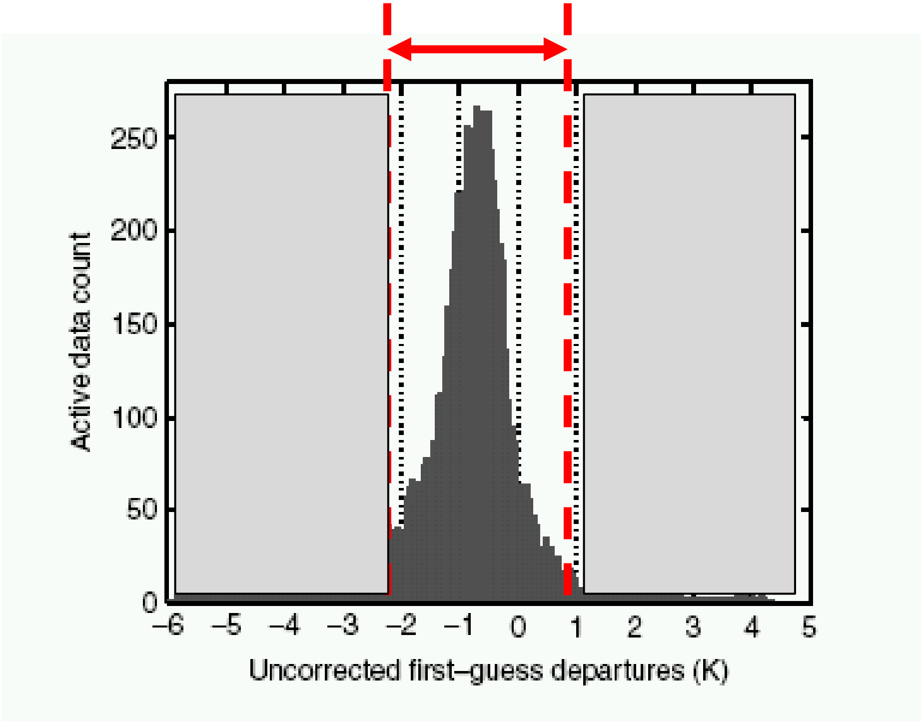
ECMWF

Simple window channel departure check

$$\Delta BT_{thresh1} < \left(y_{obs} - H(x_{clear})\right) < \Delta BT_{thresh2}$$

RT operator



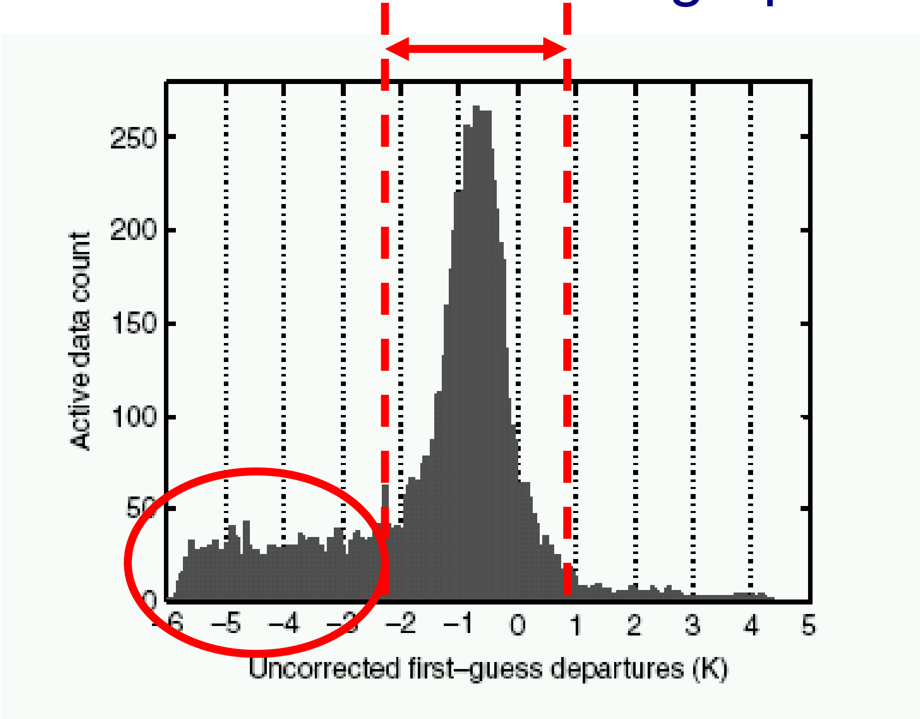


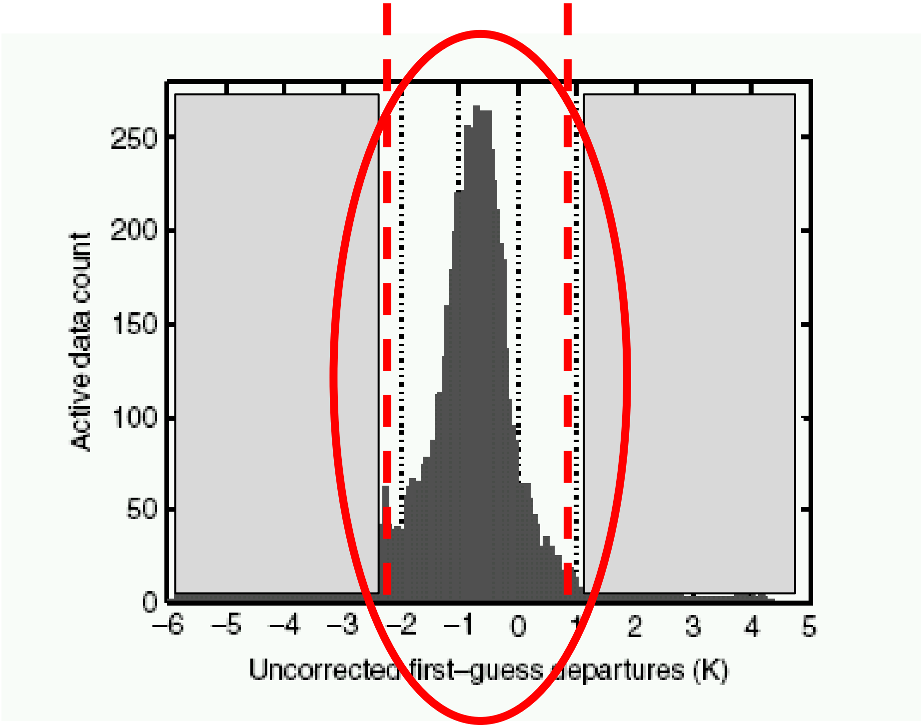
ECMWF

Simple window channel departure check

Removes most cloud-affected obs. However.....

- Some cloud contamination remains in the "**grey zone**".
- The resulting histogram is very **non-Gaussian**.
- Clear data from high-peaking channels are rejected.





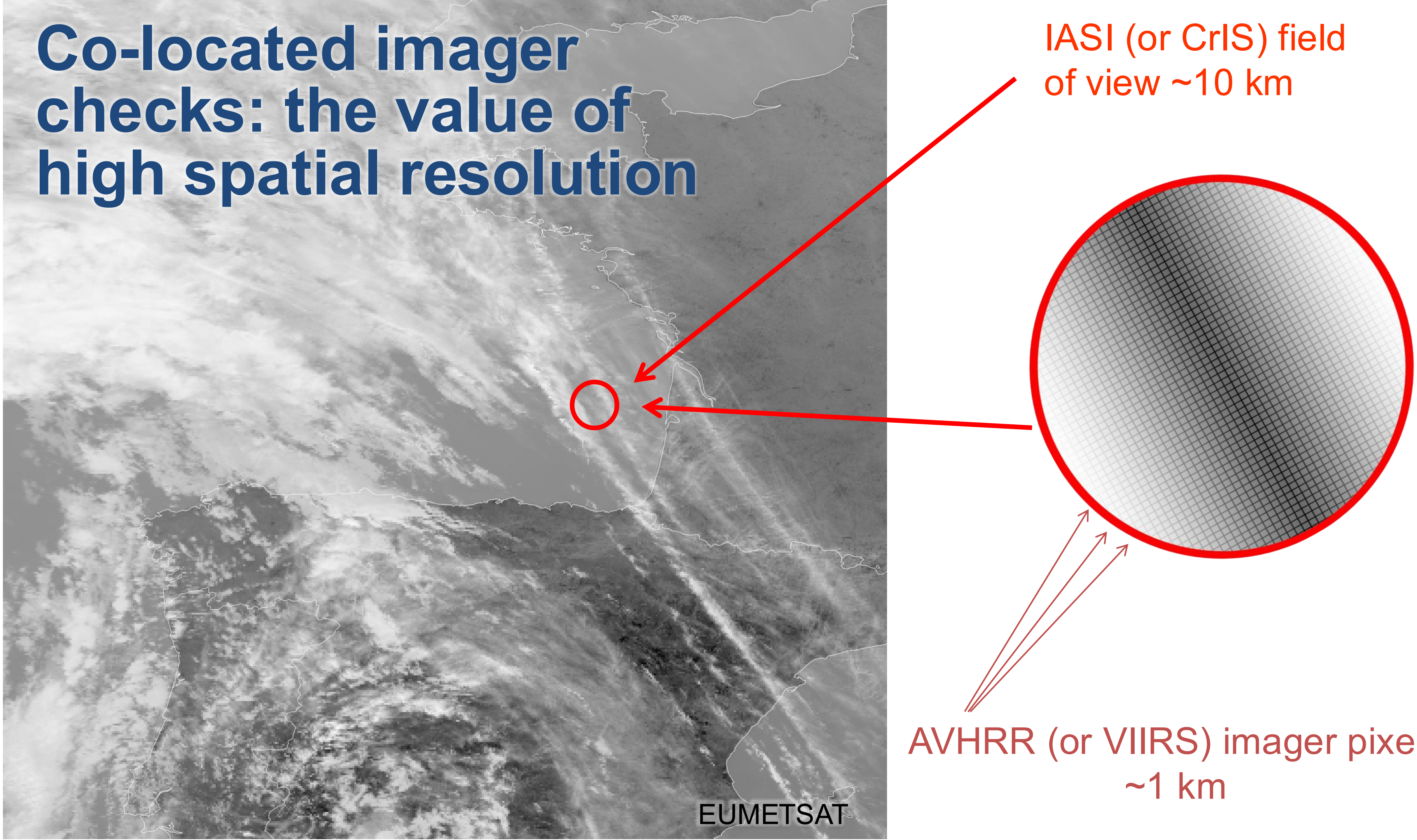
Cloud detection methods

Simple

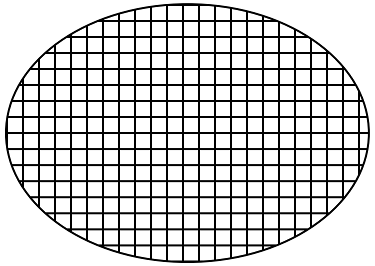
- Window channel departure (O-B) checks
- Co-located imager checks
- Pattern recognition algorithms
- Hybrid systems

Complex

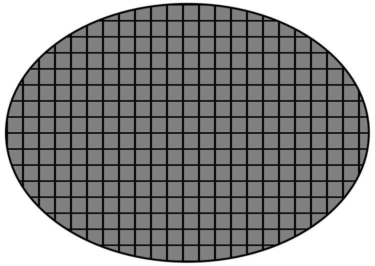
Co-located imager checks: the value of high spatial resolution



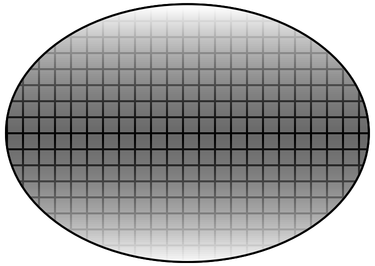
We can evaluate the <u>mean</u> and <u>variance</u> of Tb imager values inside the sounder field of view



Homogenous cloudy
(low mean Tb, low variance)

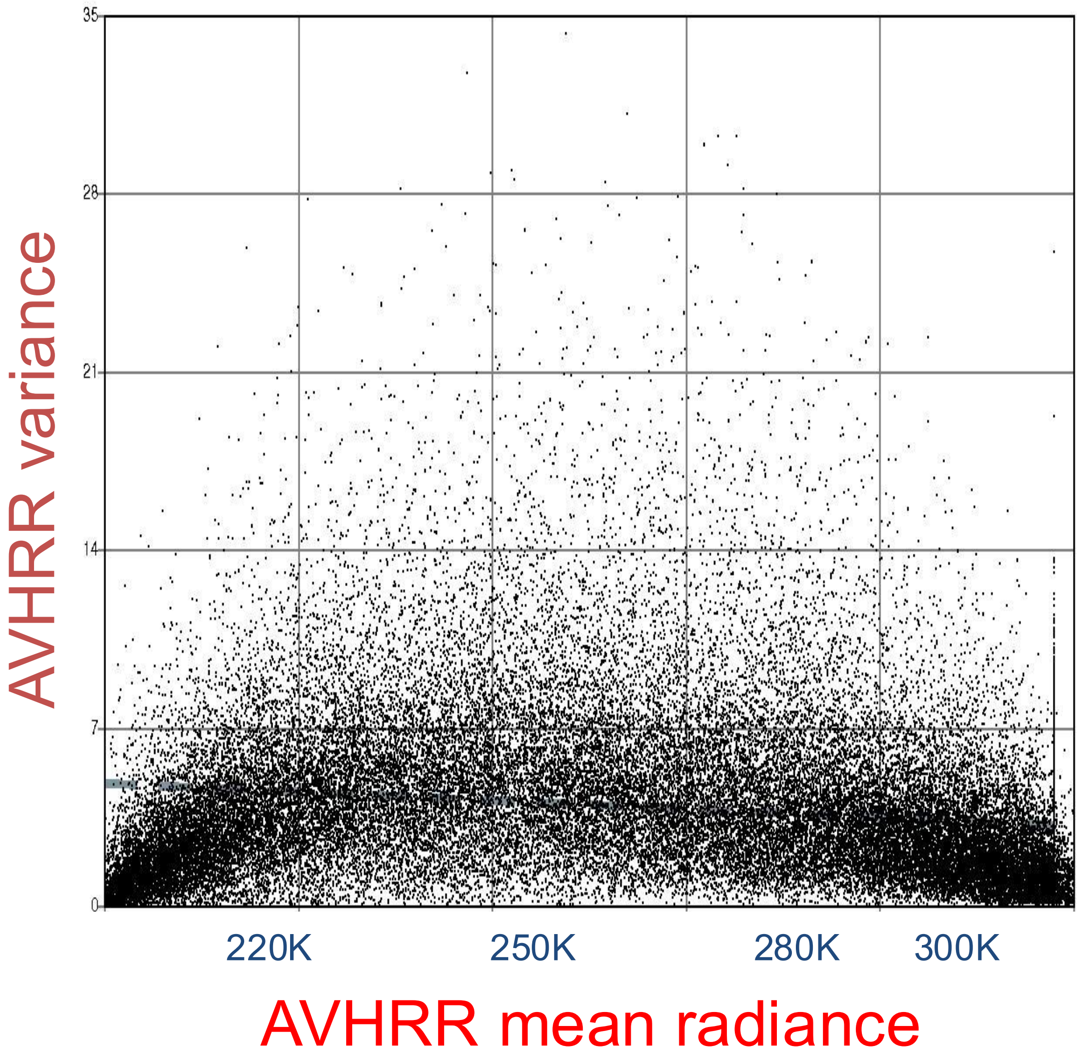


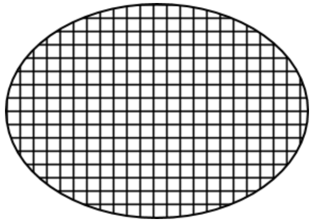
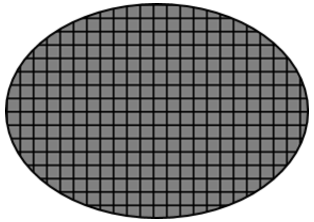
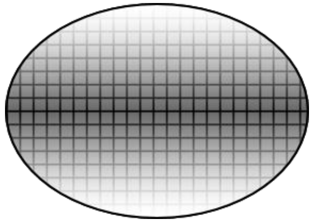
Homogenous clear
(high mean, low variance)

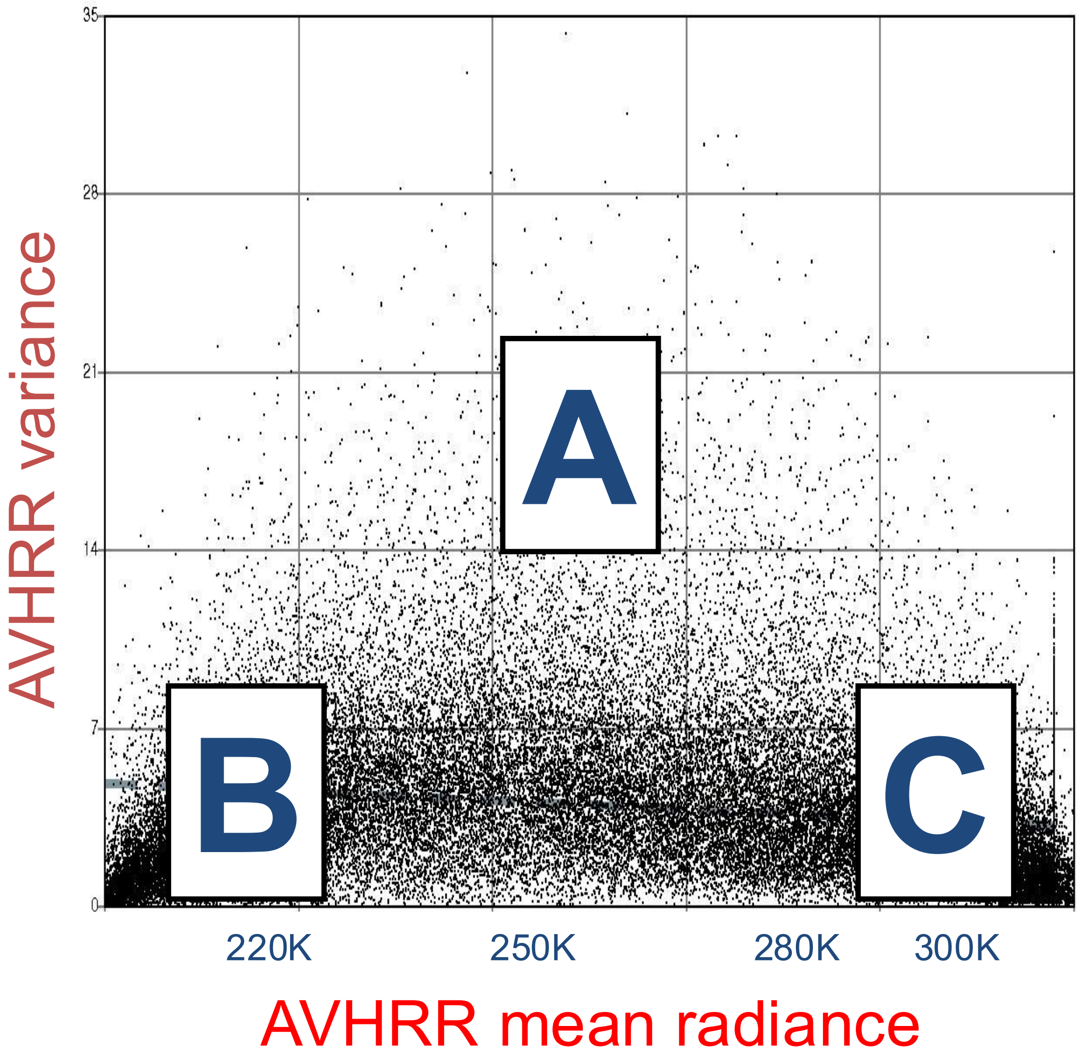


Mixed cloud scene
(any mean, high variance)

Scatter plot of variance of AVHRR imager pixels within the IASI footprint versus mean brightness temperature

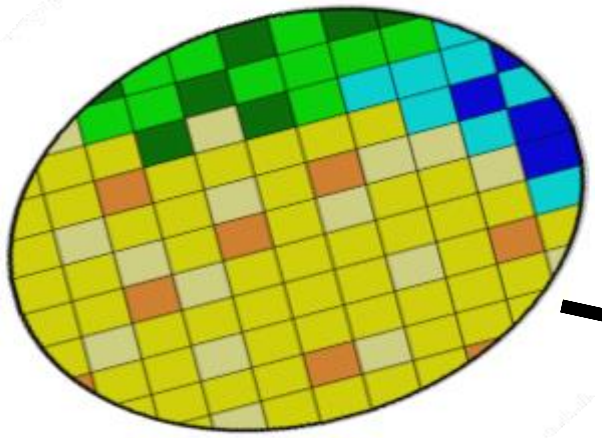


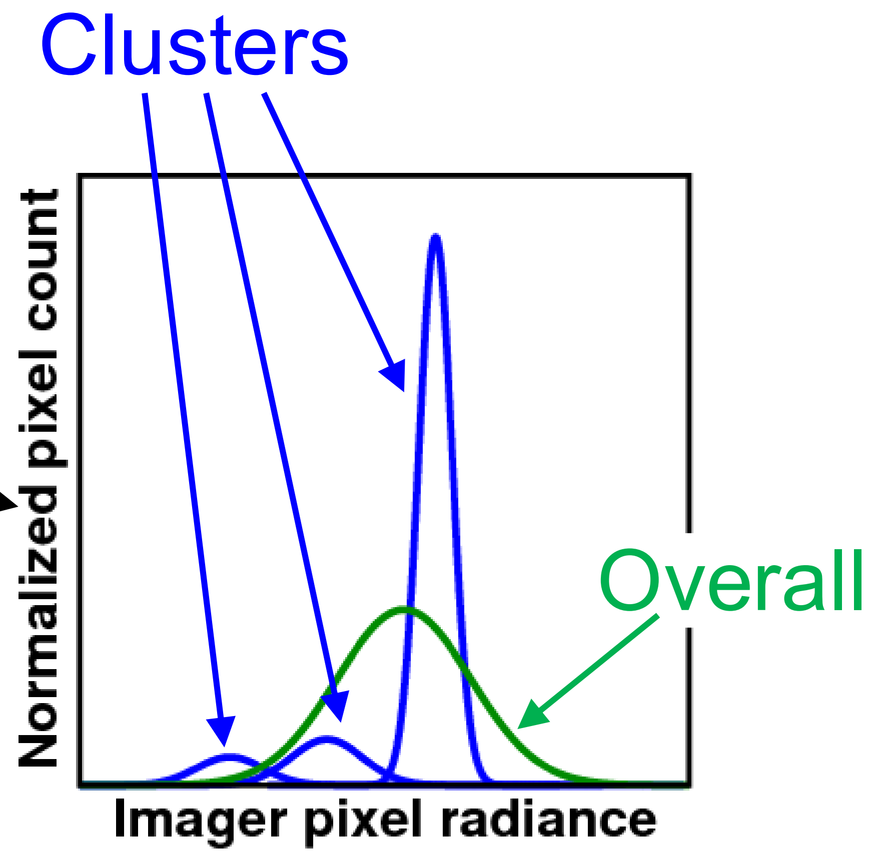
Where would , , and  be in this scatter plot ?



More sophisticated: image *clustering*

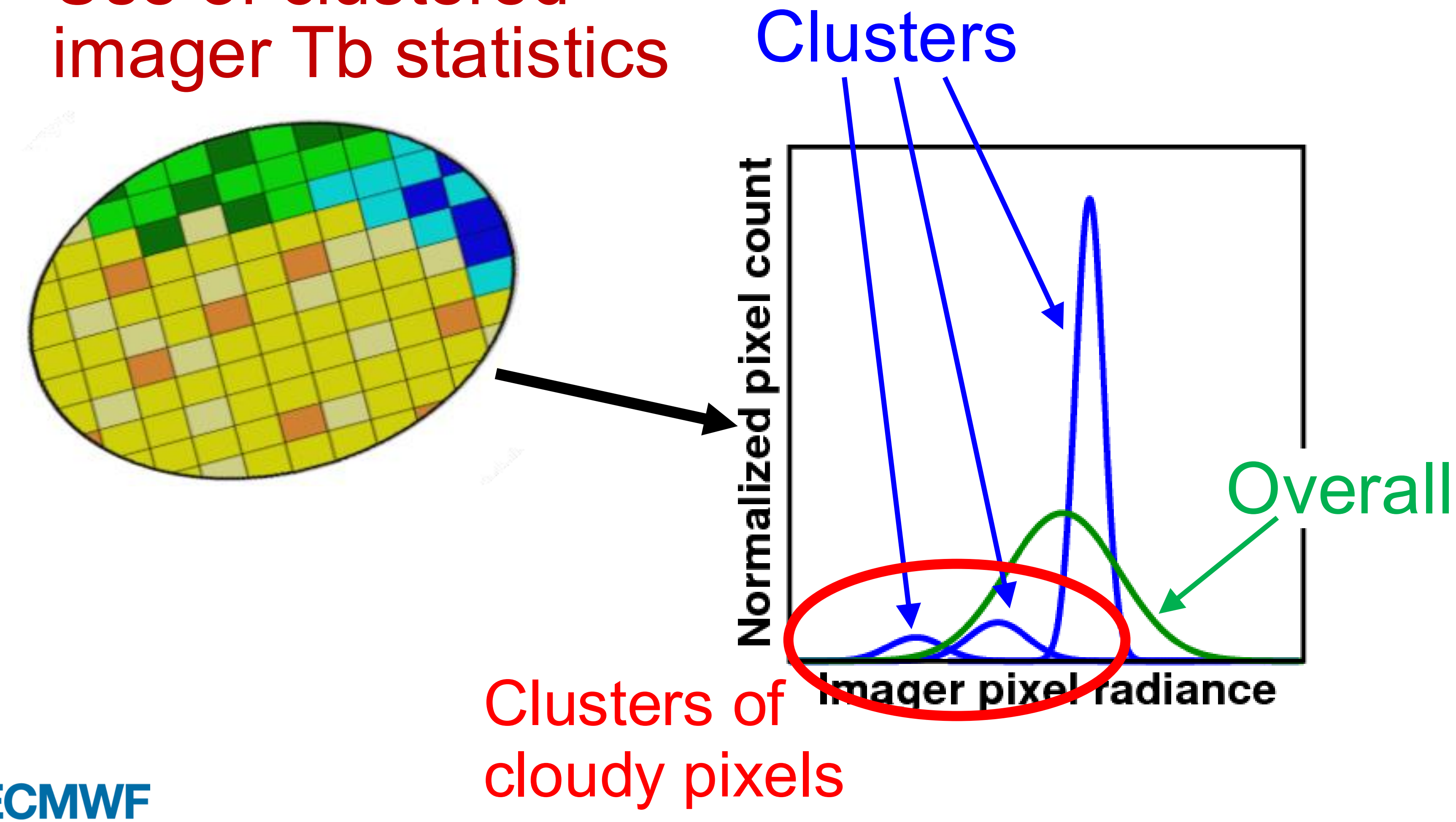
Use of clustered imager Tb statistics





More sophisticated: image *clustering*

Use of clustered
imager Tb statistics



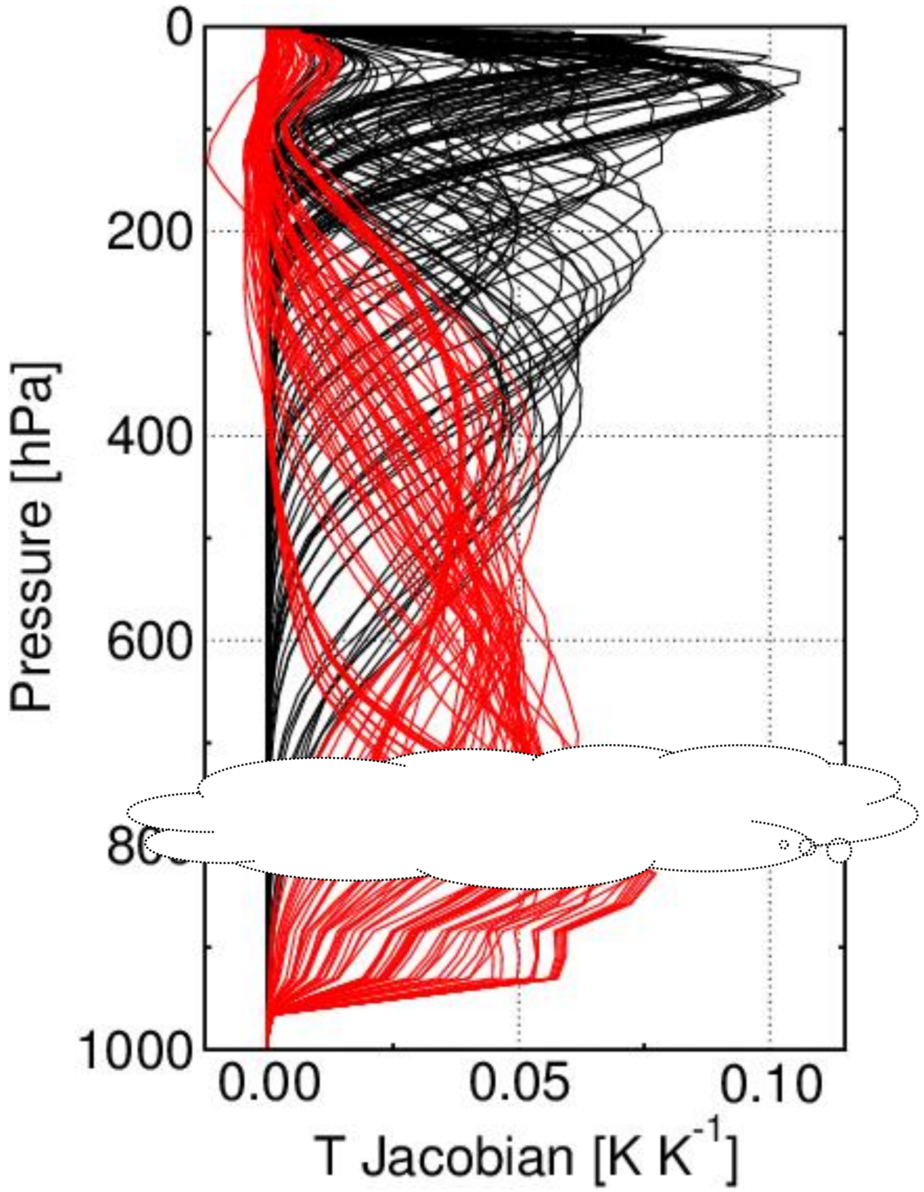
Cloud detection methods

Simple

Complex

- Window channel departure (O-B) checks
- Co-located imager checks
- Pattern recognition algorithms
- Hybrid systems

Pattern recognition algorithms

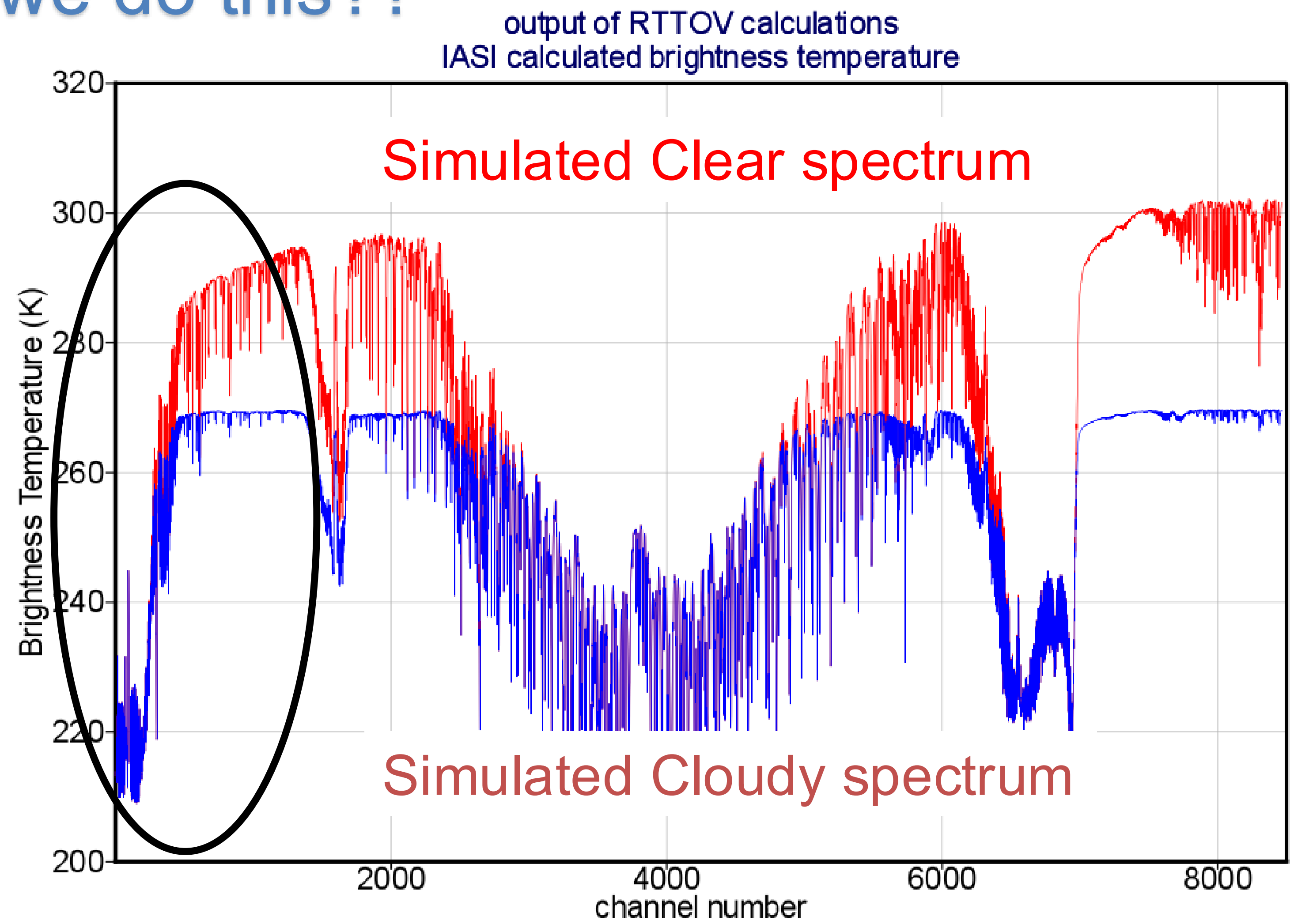


If channels are sensitive only to the atmosphere above a cloud (**black** lines), we would like to keep this data.

So here, we only reject cloud-contaminated channels (**red** lines), and keep the rest!

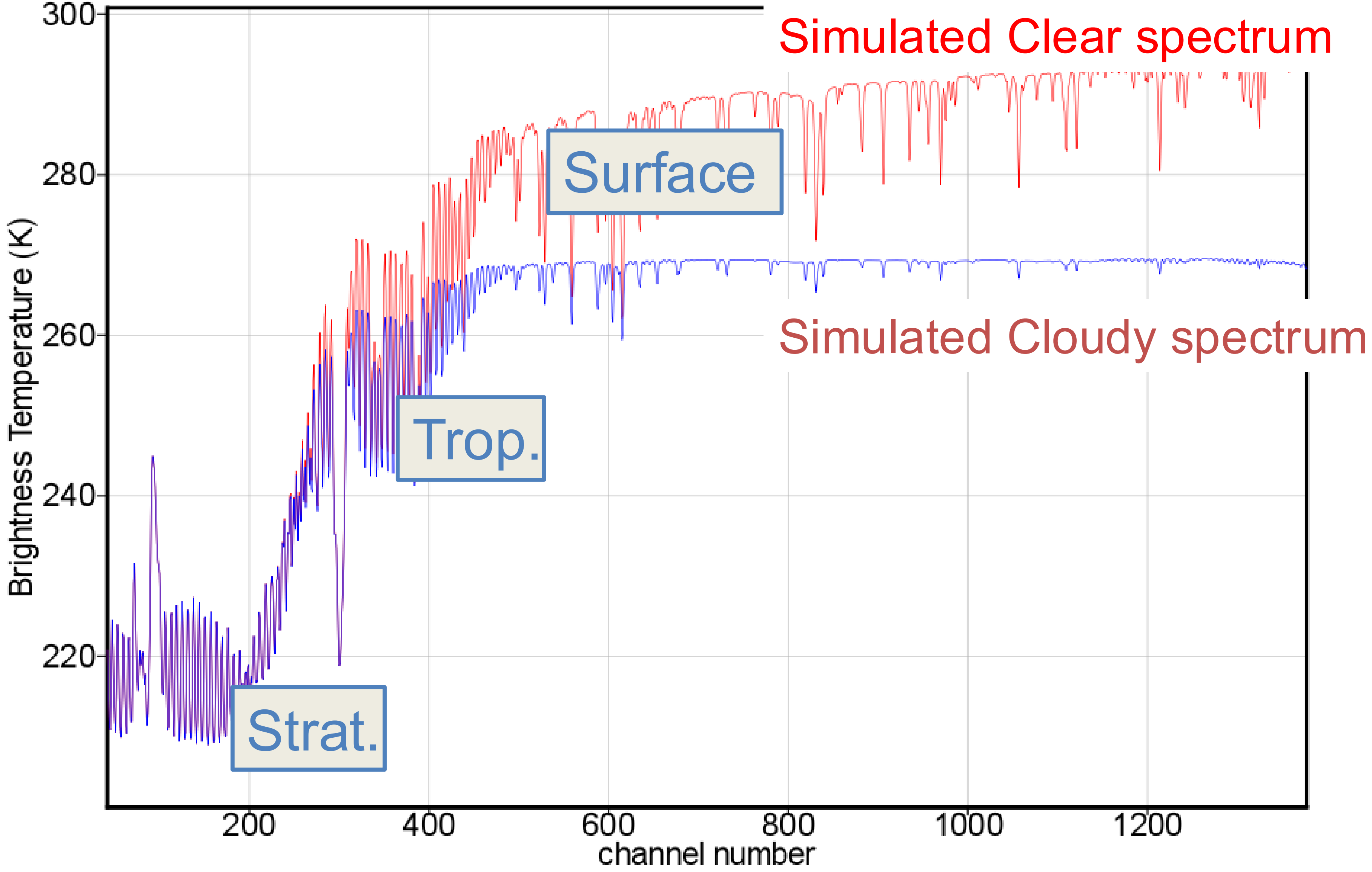
McNally & Watts 2003, https://doi.org/10.1256/qj.02.208

How do we do this??



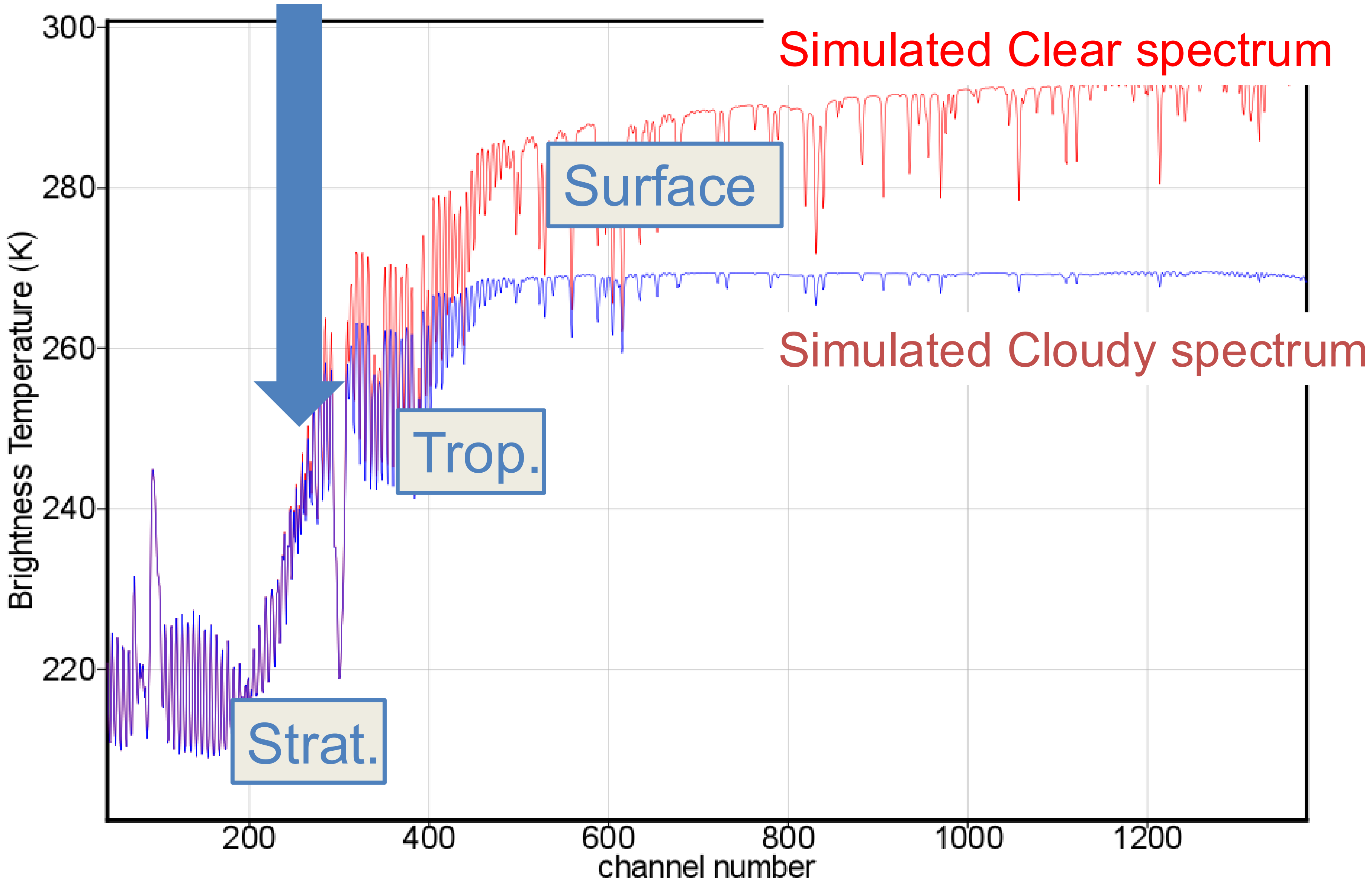
Let's zoom into the long-wave part of the spectrum which has good vertical resolution

ECMWF

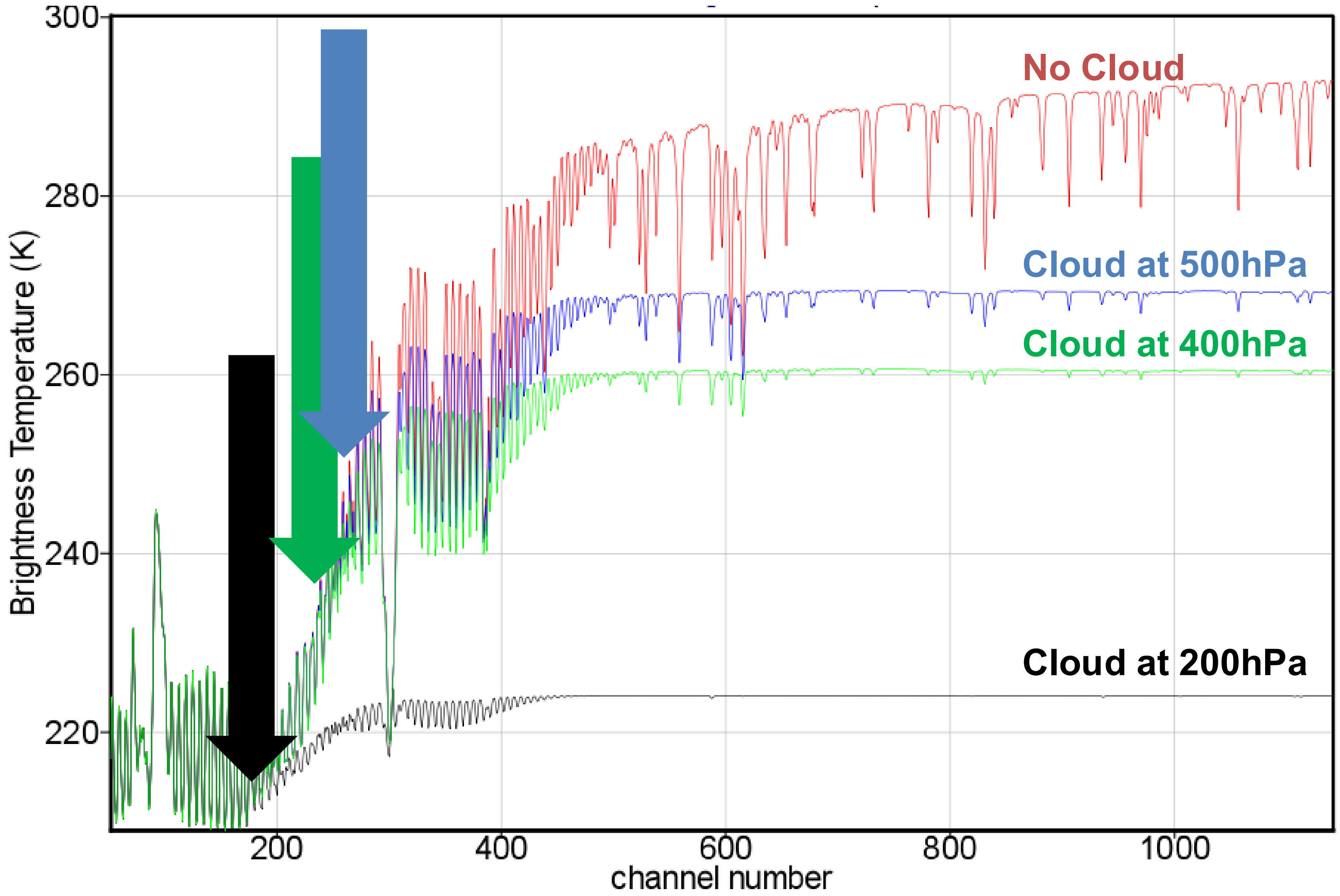


Simulated Clear spectrum
Simulated Cloudy spectrum
Surface
Trop.
Strat.
Brightness Temperature (K)
300
280
260
240
220
200
400
600
800
1000
1200
channel number

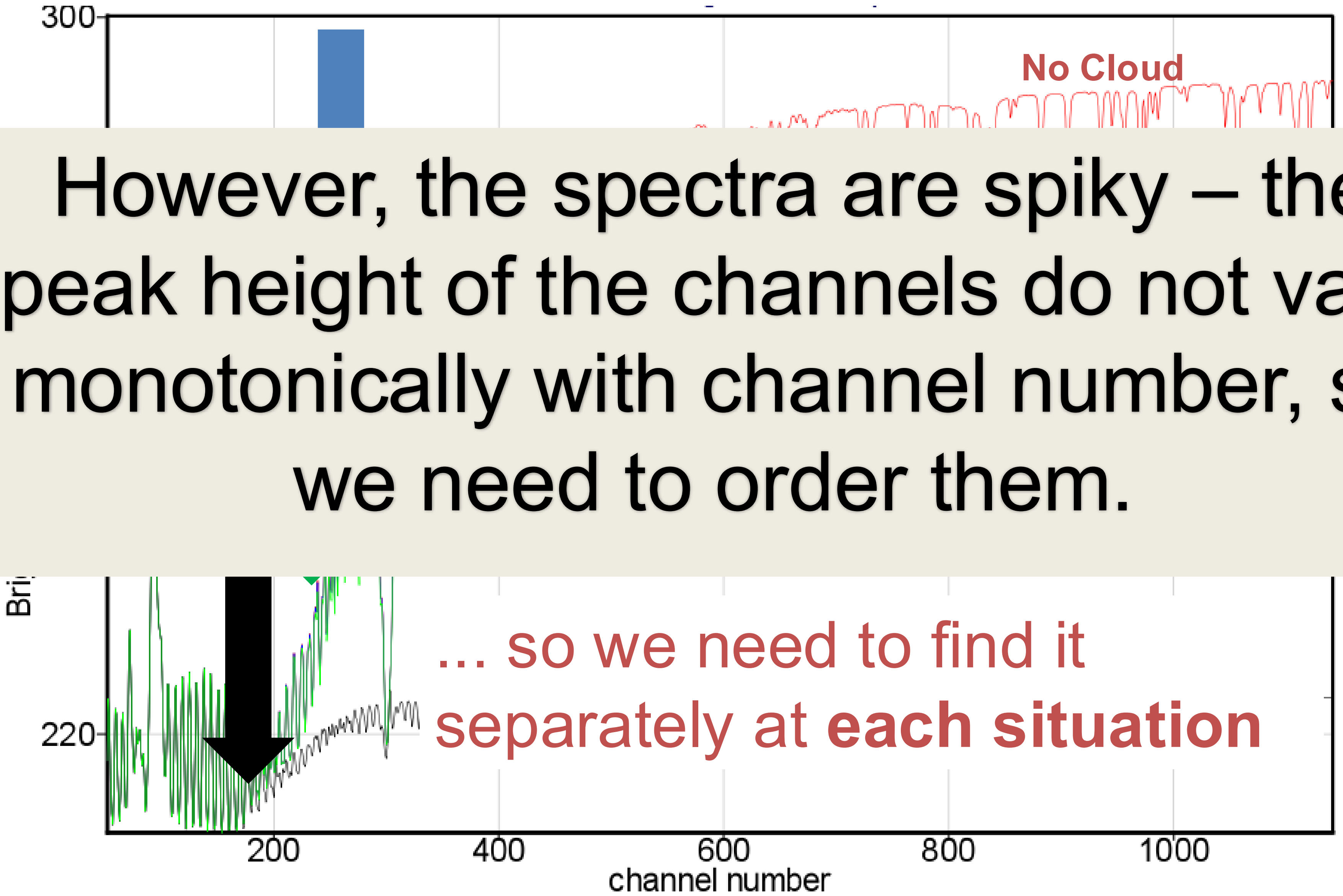
Break point


Simulated Clear spectrum
Simulated Cloudy spectrum
Surface
Trop.
Strat.
Brightness Temperature (K)
300
280
260
240
220
200
400
600
800
1000
1200
channel number

Break point depends on cloud height …



Break point depends on cloud height …

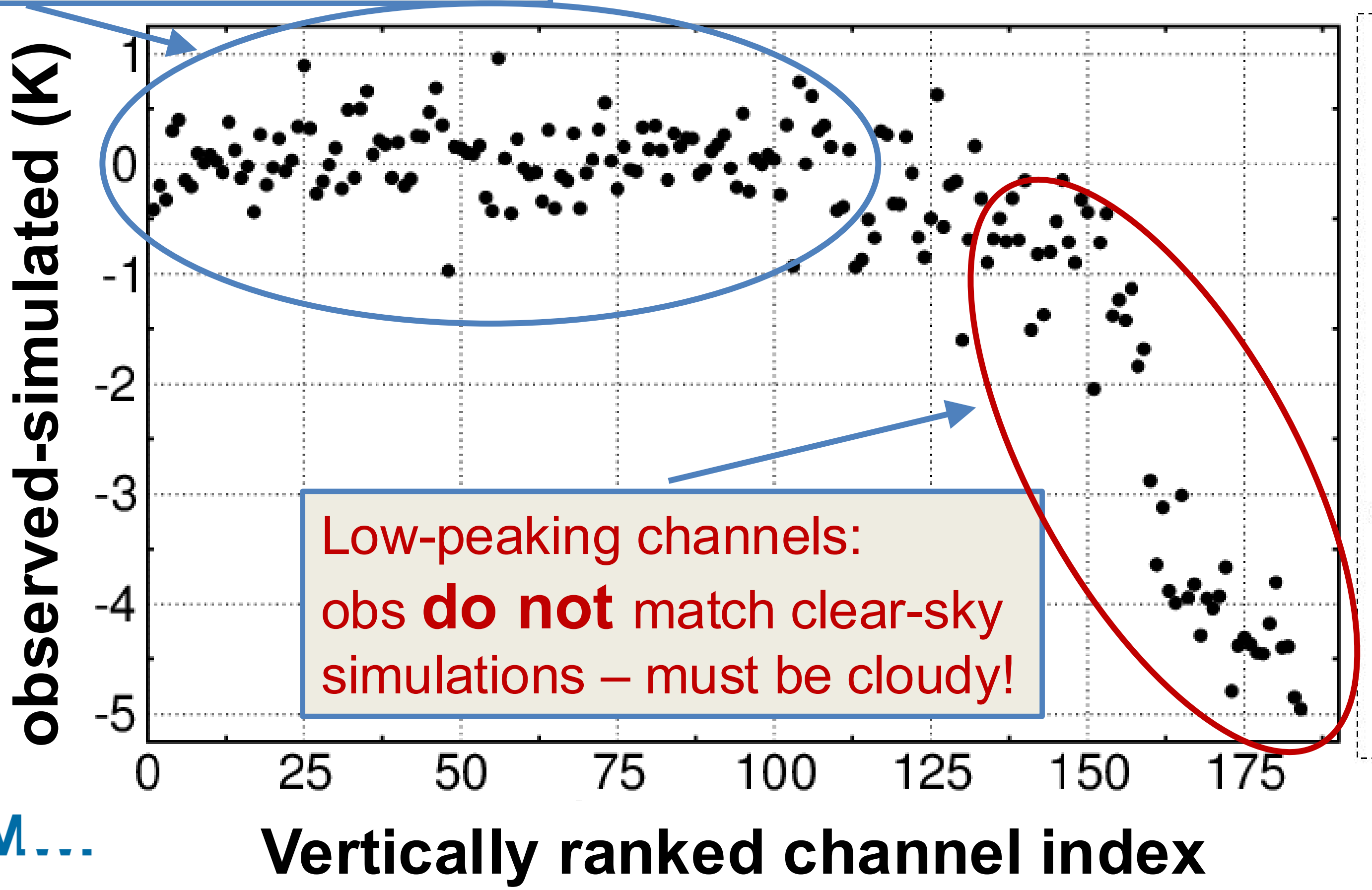


However, the spectra are spiky – the peak height of the channels do not vary monotonically with channel number, so we need to order them.

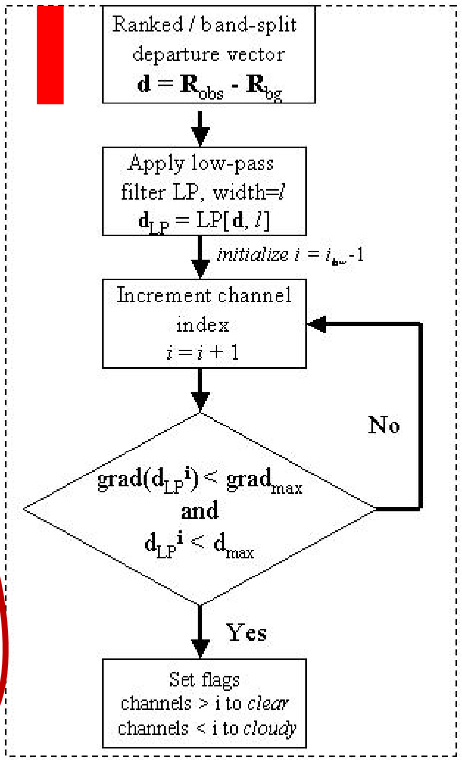
... so we need to find it separately at **each situation**

- First we compute the observations minus (clear-sky) simulations.
- Then re-order (rank) the channels according to their height of cloud sensitivity.

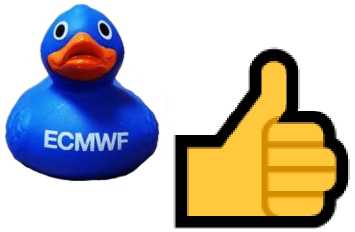
High-peaking channels:
obs match clear-sky simulations.

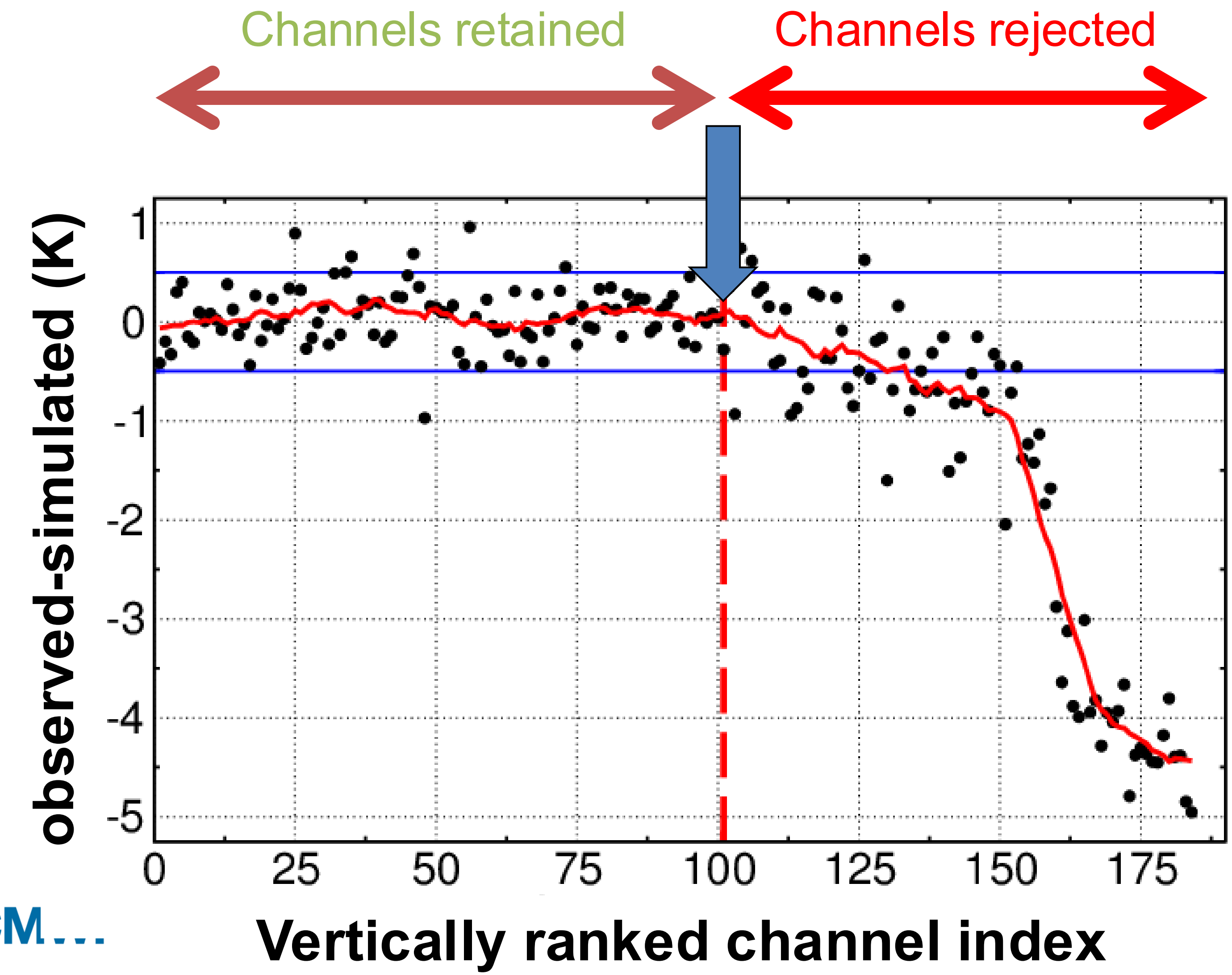


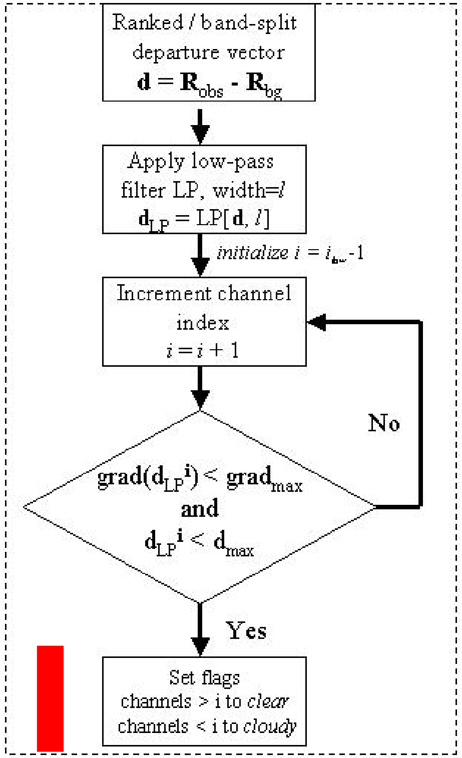
Low-peaking channels:
obs **do not** match clear-sky simulations – must be cloudy!

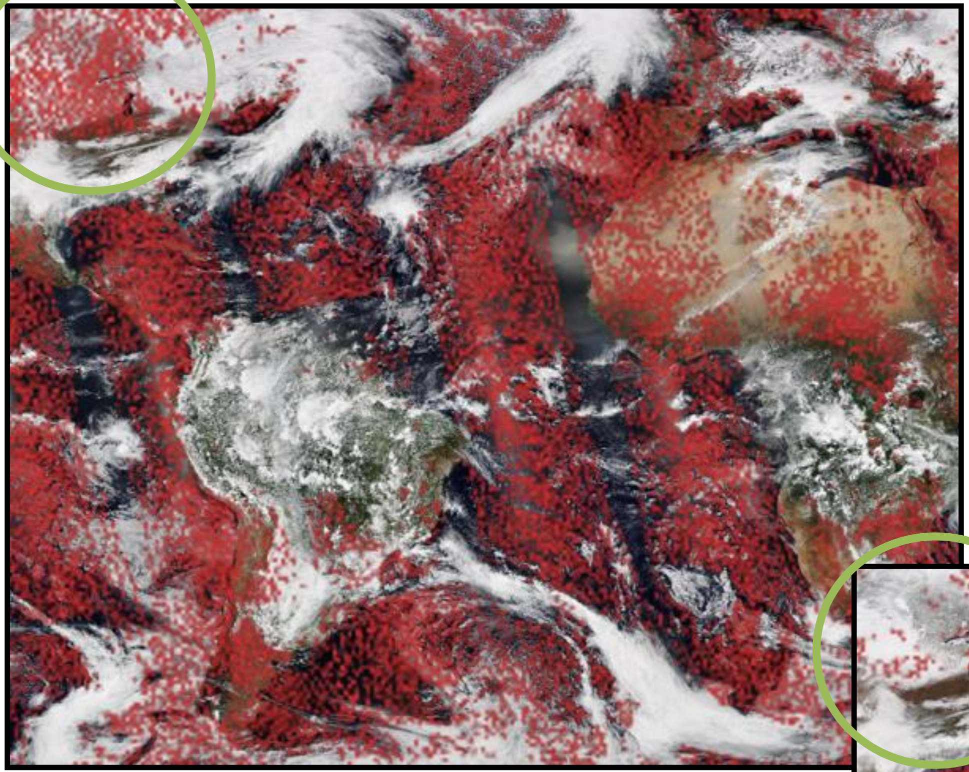


… and finally we set flags to indicate cloud-affected channels

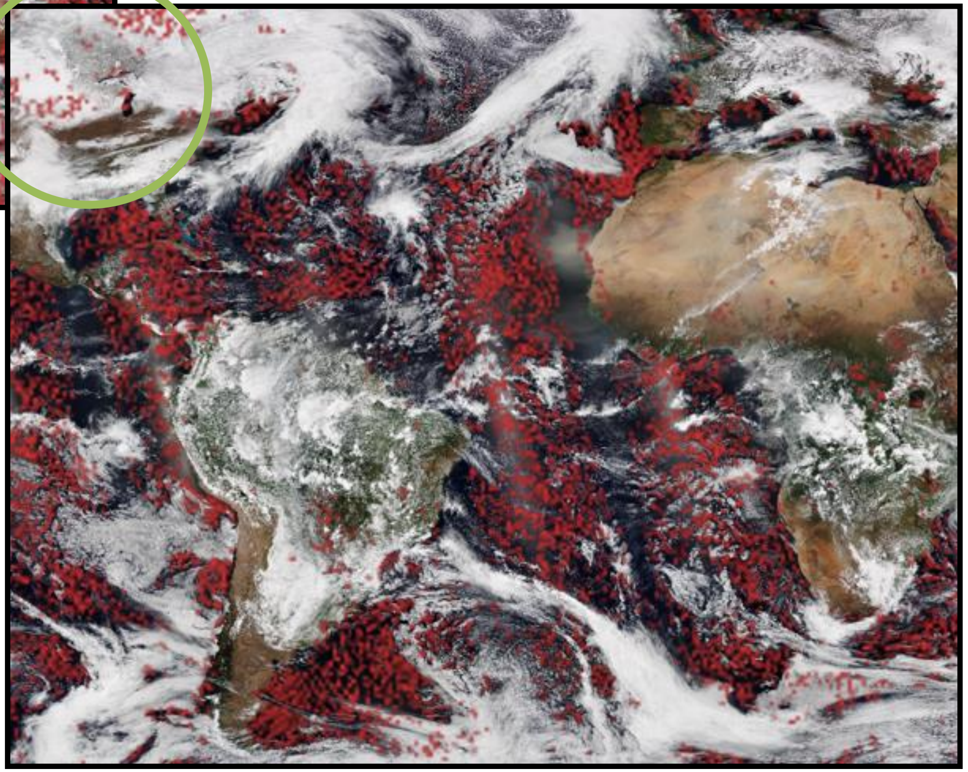








CrIS channel at 14.2 μm
(peak pressure 350 hPa)



CrIS channel at 13.6 μm
(peak pressure 600 hPa)

Cloud detection methods

Simple

Complex

- Window channel departure (O-B) checks
- Co-located imager checks
- Pattern recognition algorithms
- Hybrid systems

The ECMWF implementation is a **hybrid** scheme that combines the use of co-located imager data with departure-based pattern recognition.

This is complementary and helps to prevent the misidentification of cloud.

*See Eresmaa (2014) QJRMS **140**, 2342-2352 for details*

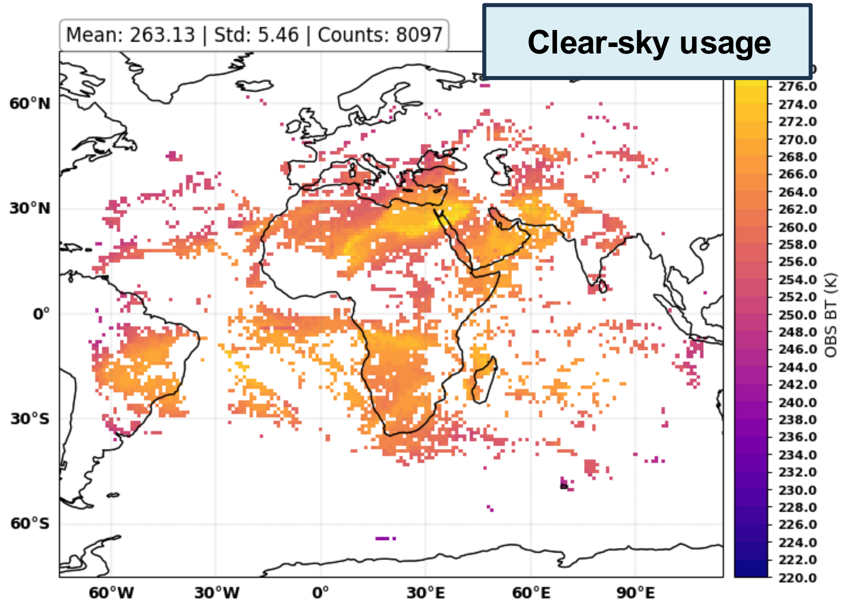
Big question: How should we handle clouds when assimilating infrared radiance observations??

- **Option 1**: detect and reject cloud-contaminated observations.
- **Option 2**: Explicitly estimate cloud parameters from the radiances within the data assimilation (T, Q, O3 etc)

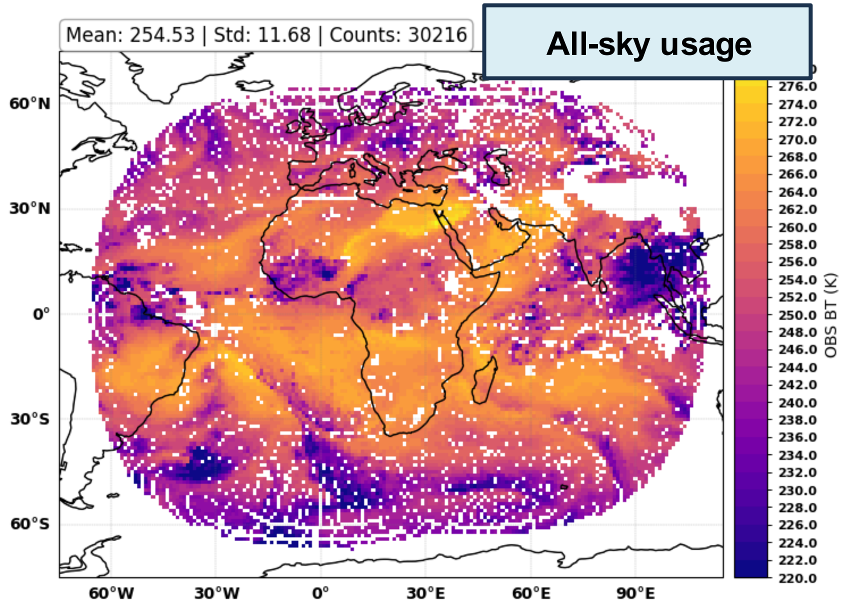
Allsky assimilation of IR radiances

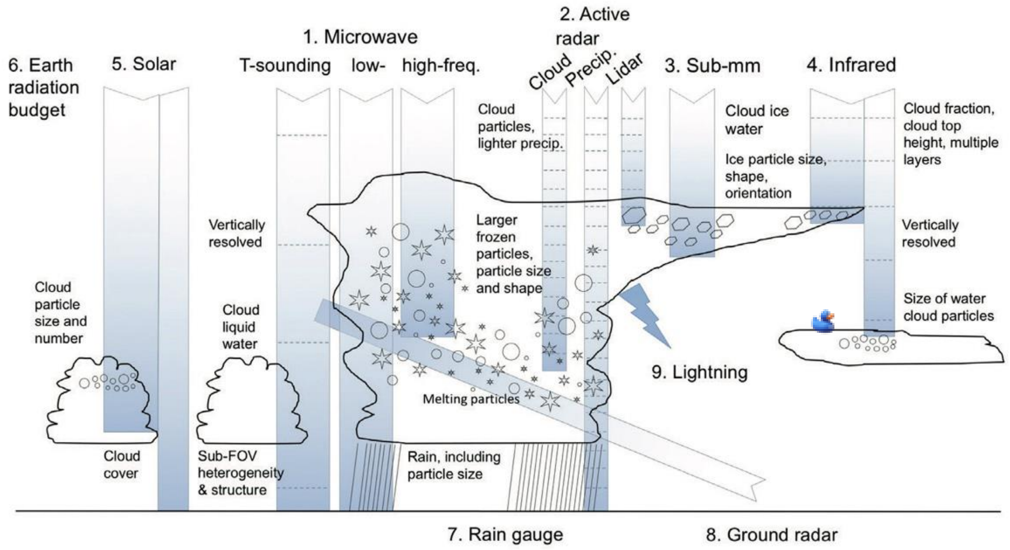
Motivation:

- To further exploit the potential of IR sensors. Currently GEOS data are used over clear-sky only leading to massive rejections. By assimilating GEOS radiances in allsky we can expect to assimilate at least 60% more data for upper-troposphere sensitive channels and 3x more data for lower-peaking channels (mid- and low-troposphere).
- Complementary with other types of observations. MW allsky observations operationally assimilated in the IFS.
- Unified assimilation process with other satellite observations. Free from complex cloud-detection algorithm or external cloud masks.



MET-9 and MET-10 SEVIRI WV7.3 observation BT on 2024-06-10 at 09:45 UTC : use in clear-sky (top) and all-sky (bottom)



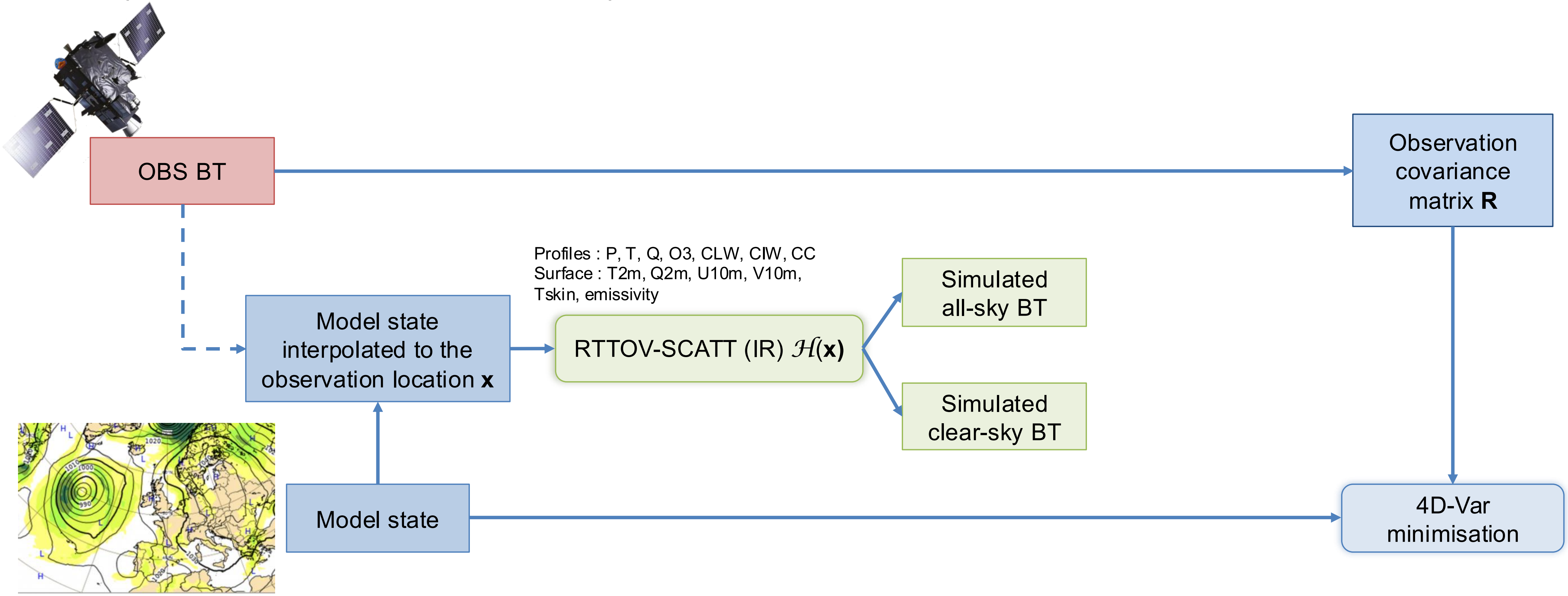


"The main measurement techniques for cloud and precipitation available in the global observing system and the micro- and macrophysical aspects to which they are most sensitive" -- from Geer et al. 2019 (ECMWF Newsletter)

Allsky assimilation of IR radiances

Challenges:

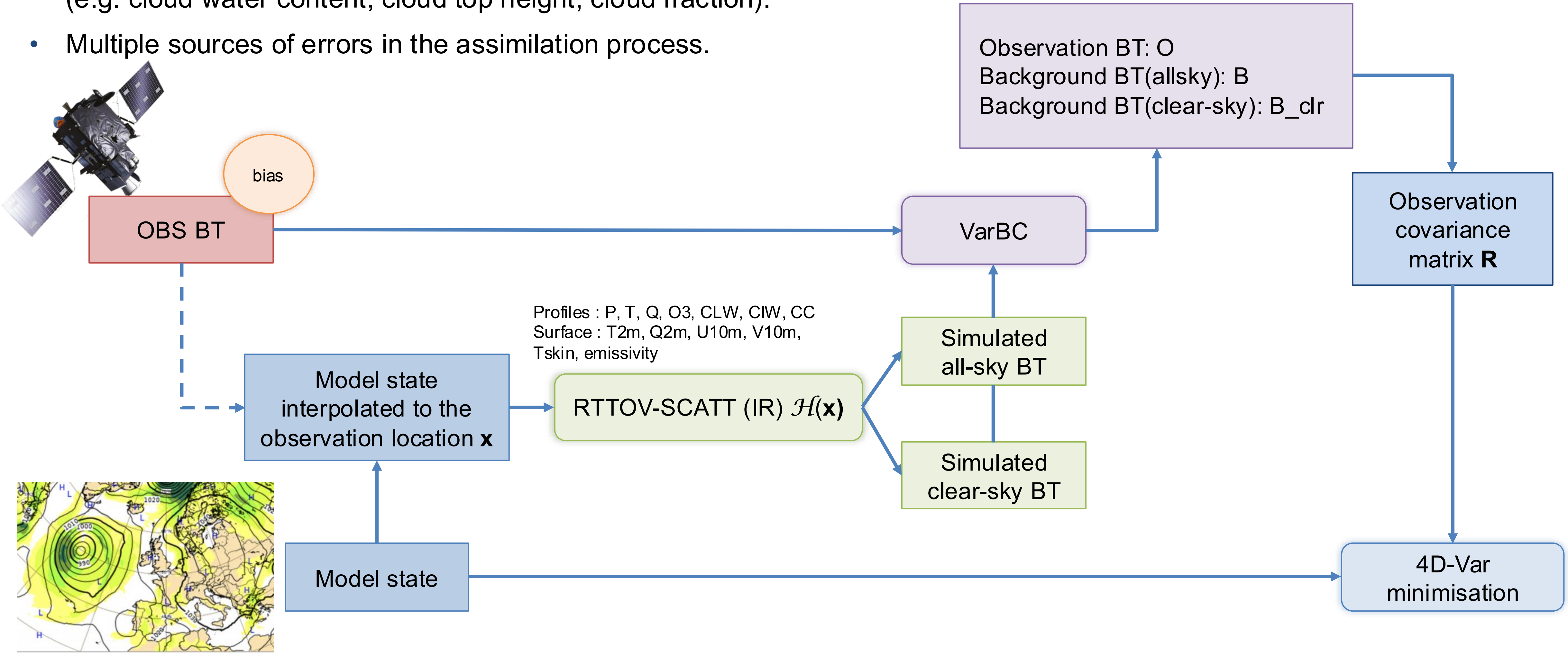
- Small errors in the representation of hydrometeors can induce large differences in the simulated BT (e.g. cloud water content, cloud top height, cloud fraction).
- Multiple sources of errors in the assimilation process.



Allsky assimilation of IR radiances

Challenges:

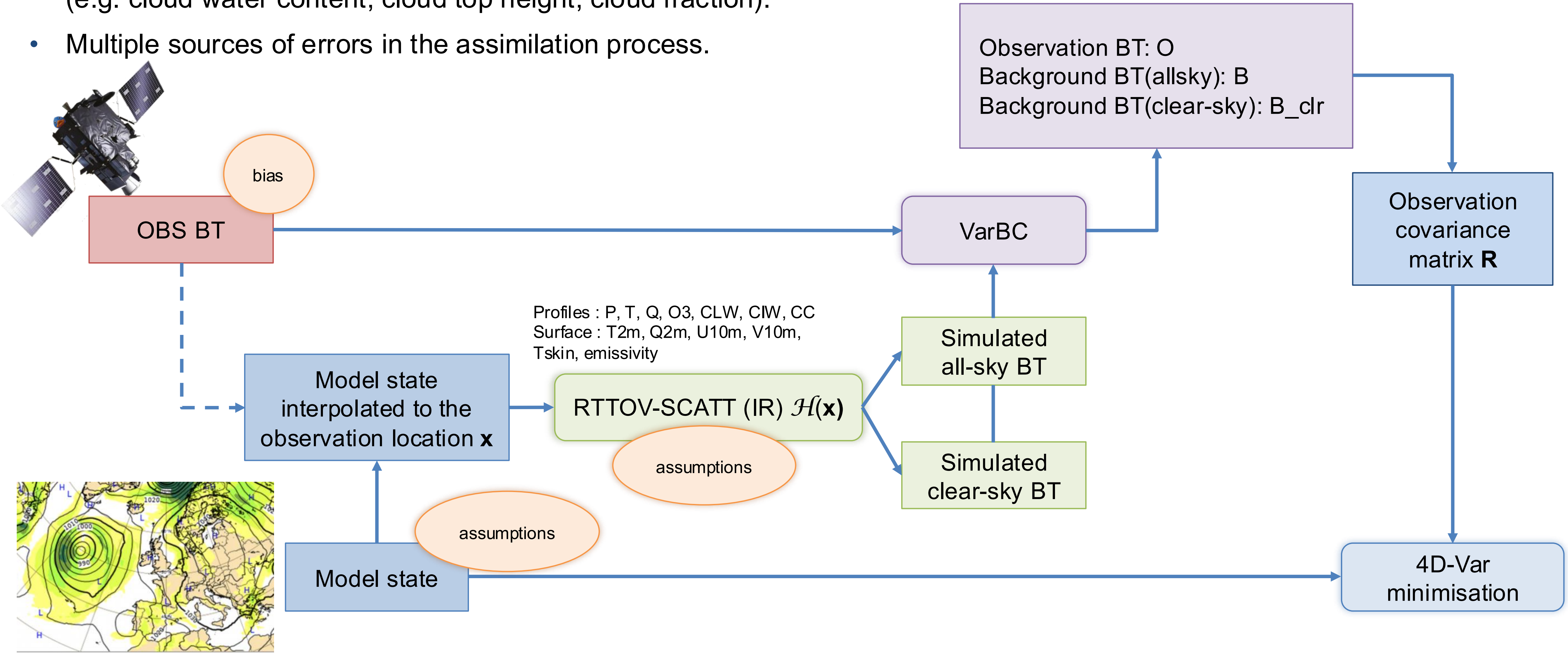
- Small errors in the representation of hydrometeors can induce large differences in the simulated BT (e.g. cloud water content, cloud top height, cloud fraction).
- Multiple sources of errors in the assimilation process.



Allsky assimilation of IR radiances

Challenges:

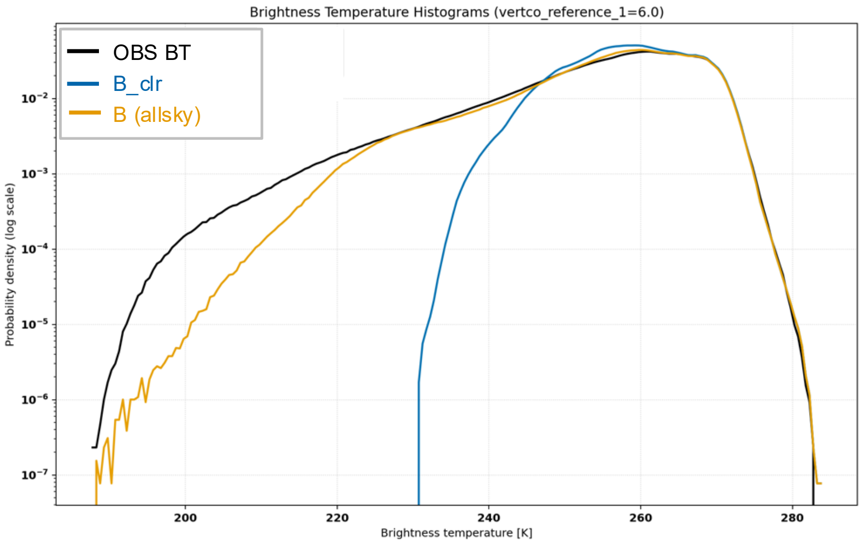
- Small errors in the representation of hydrometeors can induce large differences in the simulated BT (e.g. cloud water content, cloud top height, cloud fraction).
- Multiple sources of errors in the assimilation process.



Allsky assimilation of IR radiances: observation operator

Radiative transfer model RTTOV-SCATT (v14):

- Input: model state at observation location.
- Output: simulated brightness temperature :
  - B_clr: clear-sky simulation (neglects cloud scattering and absorption).
  - B: all-sky simulation (includes cloud properties).



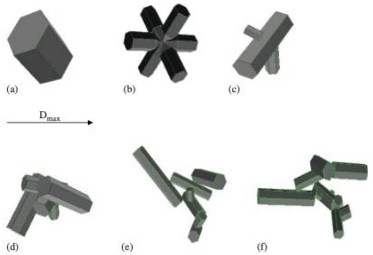
Observed (black) and simulated brightness temperatures in clear-sky (blue) and all-sky (orange) by RTTOV for SEVIRI's WV channel 7.3 µm.

- Assumptions and parametrisations for IR simulations:
  - Microphysics

Ice optical properties: shape and size distribution (***Baran et al. 2014, 2018***) parametrised in terms of temperature and ice content.

Cloud liquid water effective parameter (***Martin et al. 1994***).

Scattering parametrisation: Delta-Eddington solver (***Bauer et al. 2006***).



Ice shapes used in Baran et al. (2014)

- Macrophysics

Cloud overlap scheme: maximum overlap assumption.

Sub-grid variability, representativeness errors.

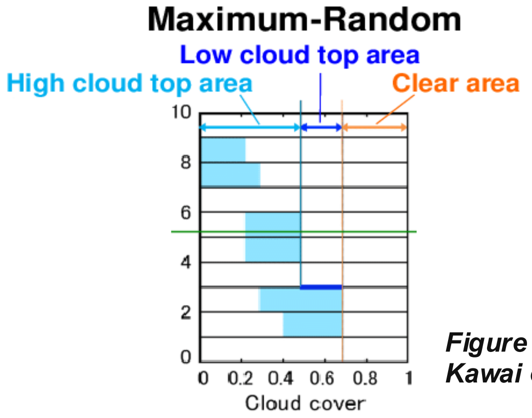


Figure from Kawai et al. (2014)

ECMWF

Allsky assimilation of IR radiances

Challenges:

- Small errors in the representation of hydrometeors can induce large differences in the simulated BT (e.g. cloud water content, cloud top height, cloud fraction).
- Multiple sources of errors in the assimilation process.

Observation BT: O
Background BT(allsky): B
Background BT(clear-sky): B_clr

bias

OBS BT

VarBC

Observation covariance matrix **R**

Profiles : P, T, Q, O3, CLW, CIW, CC
Surface : T2m, Q2m, U10m, V10m, Tskin, emissivity

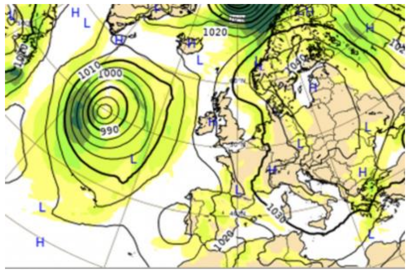
Model state interpolated to the observation location **x**

RTTOV-SCATT (IR) $\mathcal{H}(\mathbf{x})$

assumptions

Simulated all-sky BT

Simulated clear-sky BT



assumptions

Model state

4D-Var minimisation

Allsky assimilation of IR radiances: R matrix

Specification of observation error covariance matrix for all-sky assimilation:

- Observation errors are larger over cloudy scenes than over clear sky. For allsky assimilation, the **R** matrix needs to account for RT uncertainties, cloud representation errors and representativeness errors.

Allsky assimilation of IR radiances: R matrix

Specification of observation error covariance matrix for all-sky assimilation:

- Observation errors are larger over cloudy scenes than over clear sky. For allsky assimilation, the **R** matrix needs to account for RT uncertainties, cloud representation errors and representativeness errors.
- Following the operational microwave allsky assimilation approach at ECMWF (Geer and Bauer, 2010):
  1. Compute a cloud proxy to characterise the scene.

  Quantify the impact of clouds on the *observed* BT: O_cloudprox = B_clr – O

  Quantify the impact of clouds on the *simulated* BT: B_cloudprox = B_clr – B

  The symmetric cloud proxy CSYM is defined as the average of the cloud impact on the background and on the observation (***Okamoto et al. 2014, Geer et al. 2019***).

$$\text{CSYM} = \frac{\overbrace{(\text{B\_clr} - \text{B})}^{\textit{Cloud impact on background}} + \overbrace{(\text{B\_clr} - \text{O})}^{\textit{Cloud impact on observation}}}{2}$$

  Higher CSYM values indicate scenes where clouds have a stronger impact.

Allsky assimilation of IR radiances: R matrix

Specification of observation error covariance matrix for all-sky assimilation:

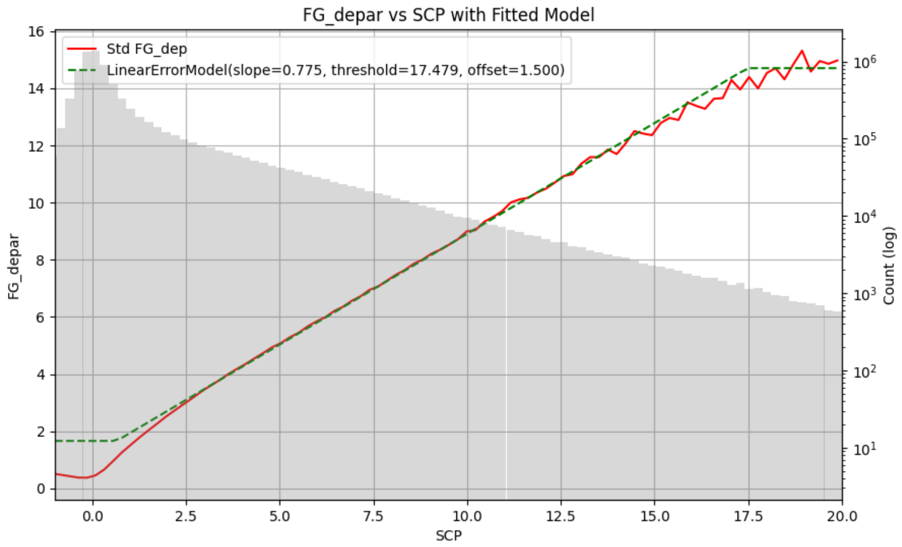
- Observation errors are larger over cloudy scenes than over clear sky. For allsky assimilation, the **R** matrix needs to account for RT uncertainties, cloud representation errors and representativeness errors.
- Following the operational microwave allsky assimilation approach at ECMWF (Geer and Bauer, 2010):
  1. Compute a cloud proxy to characterise the scene.

$$\mathrm{CSYM} = \frac{(B\_clr - B) + (B\_clr - O)}{2}$$

  2. Scene-dependent observation error model fit to first guess standard deviation error.

  Almost linear-relation between the first guess departure standard deviation error and CSYM.

  Inflated fixed value for low CSYM values (clear-sky). The error value grows linearly until a diagnosed maximum value (saturated scenes).



First guess departure standard deviation (red) as function of CSYM - SEVIRI WV6.3. Linear fit in dashed green.

ECMWF

Allsky assimilation of IR radiances: R matrix

Specification of observation error covariance matrix for all-sky assimilation:

- Observation errors are larger over cloudy scenes than over clear sky. For allsky assimilation, the **R** matrix needs to account for RT uncertainties, cloud representation errors and representativeness errors.
- Following the operational microwave allsky assimilation approach at ECMWF (Geer and Bauer, 2010):

  1. Compute a cloud proxy to characterise the scene.

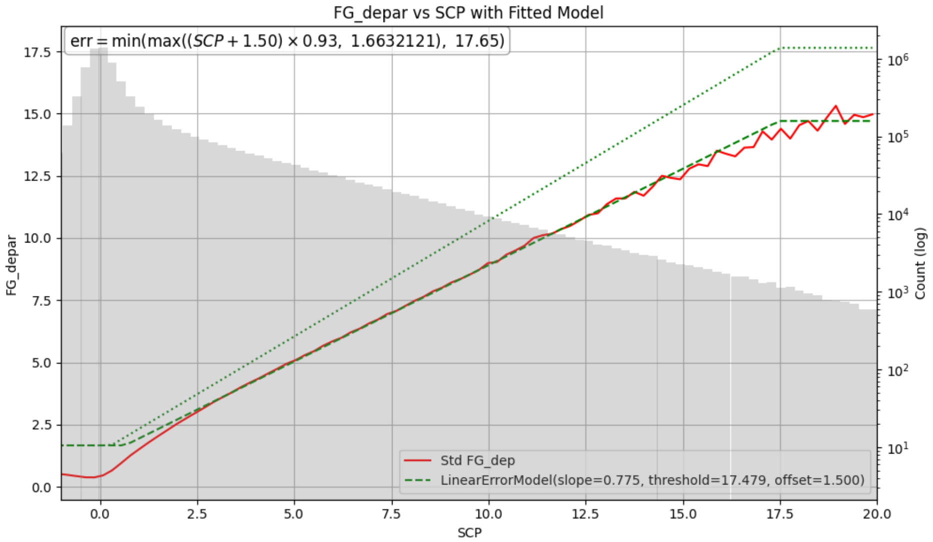
  $$\text{CSYM} = \frac{(B\_clr - B) + (B\_clr - O)}{2}$$

  2. Scene-dependent observation error model fit to first guess standard deviation error.

  Almost linear-relation between the first guess departure standard deviation error and CSYM.

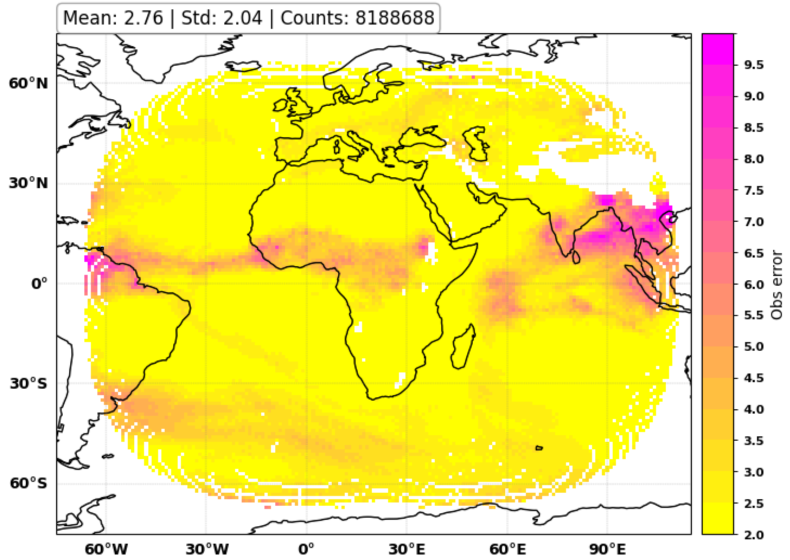
  Inflated fixed value for low CSYM values (clear-sky). The error value grows linearly until a diagnosed maximum value (saturated scenes).

  3. Inflation factor applied.



Left: First guess departure standard deviation (red) as function of CSYM - SEVIRI WV6.3. Linear fit in dashed green. Inflated model by a factor of 1.5 in dotted green

Right: Prescribed observation error



ECMWF

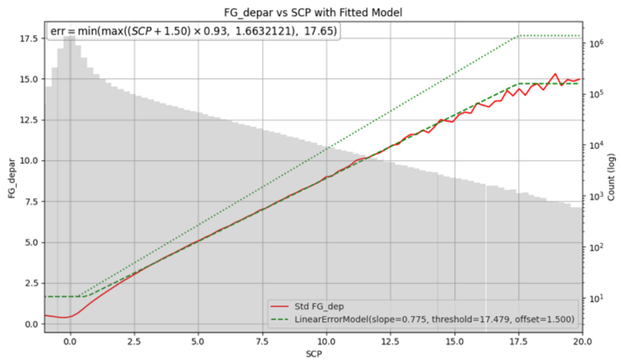
Allsky assimilation of IR radiances: R matrix

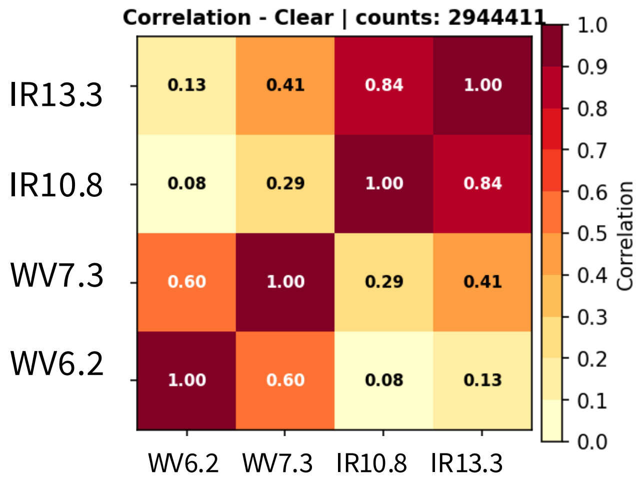
Specification of observation error covariance matrix for all-sky assimilation:

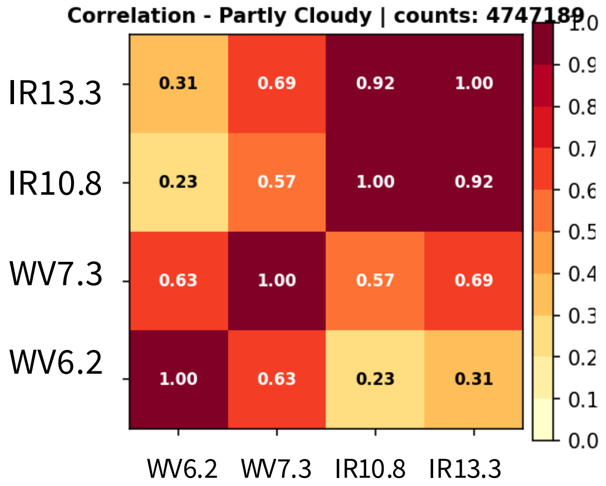
- Observation errors are larger over cloudy scenes than over clear sky. For allsky assimilation, the **R** matrix needs to account for RT uncertainties, cloud representation errors and representativeness errors.
- Following the operational microwave allsky assimilation approach at ECMWF (Geer and Bauer, 2010):
  1. Compute a cloud proxy to characterise the scene.

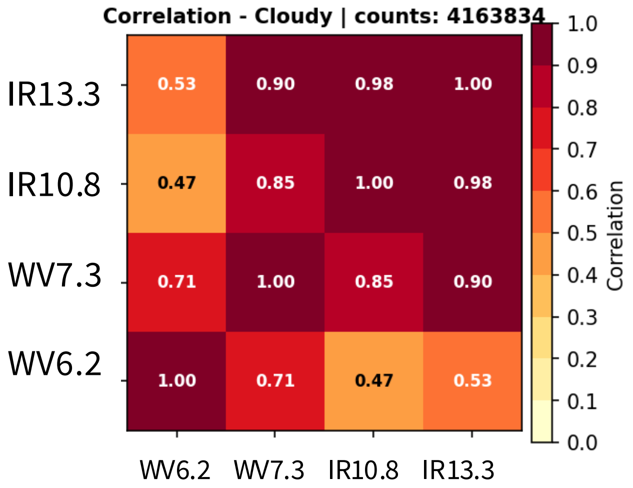
$$\text{CSYM} = \frac{(B\_clr - B) + (B\_clr - O)}{2}$$

  2. Scene-dependent observation error model fit to first guess standard deviation error.
  3. Inflation factor applied.
- Interchannel correlation error varies with the cloudiness of the scene.









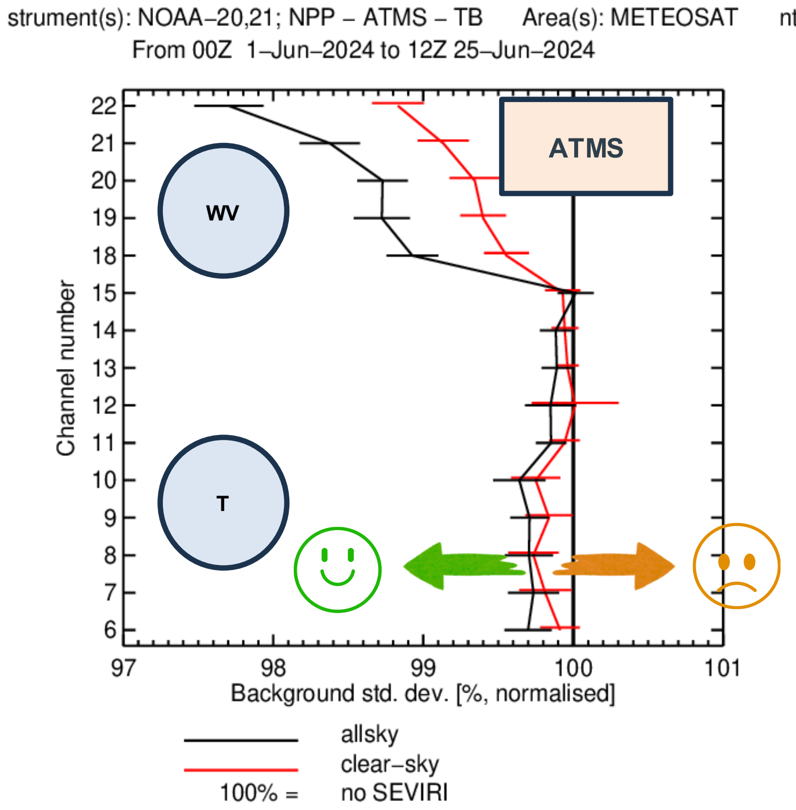
Allsky assimilation of IR radiances: impacts

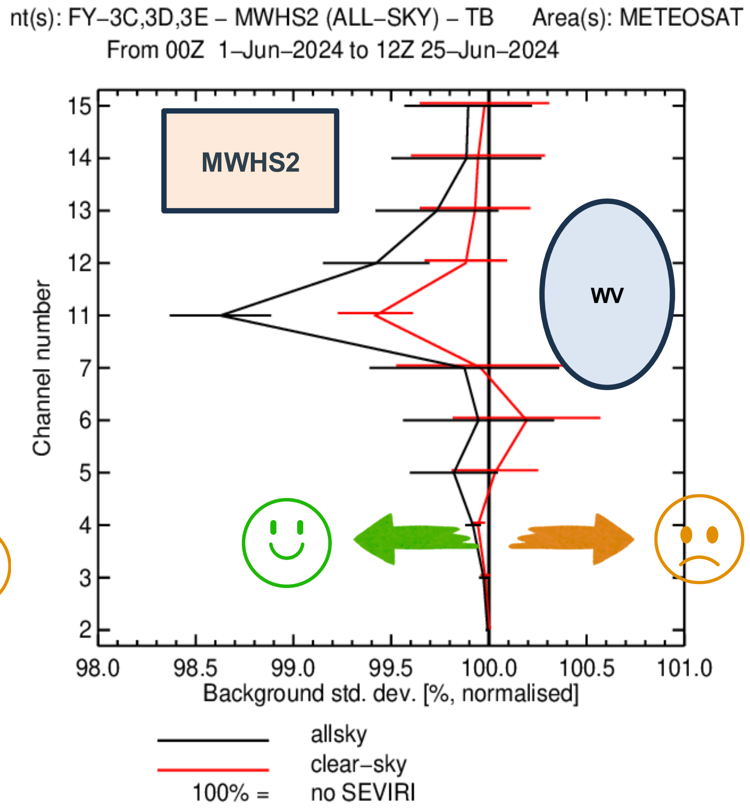
Experiment configuration:

Meteosat-SEVIRI water vapour channels (6.3 μm and 7.3 μm) assimilated in clear-sky (current operational) or allsky.

Improved fit to clear-sky (CrIS, ATMS) and allsky (MWHS2, SSMIS) observations.

- Control (100%) is an experiment where SEVIRI is fully rejected
- Red line: Control + SEVIRI assimilated in clear-sky
- Black line: Control + SEVIRI assimilated in all-sky





ECMWF

Summary

- Infrared satellite observations are **very** sensitive to cloud.
- Identifying cloud in “clear-sky” assimilation can be done in several ways:
  - Window channel departure checks
  - Co-located imager checks
  - Pattern recognition
  - Machine learning

- All-sky infrared assimilation is a difficult challenge. But fast progress is being made!

- If you want to do cloud detection, the CADS package developed by the NWP SAF is available here: https://nwp-saf.eumetsat.int/site/software/cloud-and-aerosol-detection/

- Feel free to speak to me later, or contact me by email!

Questions ?