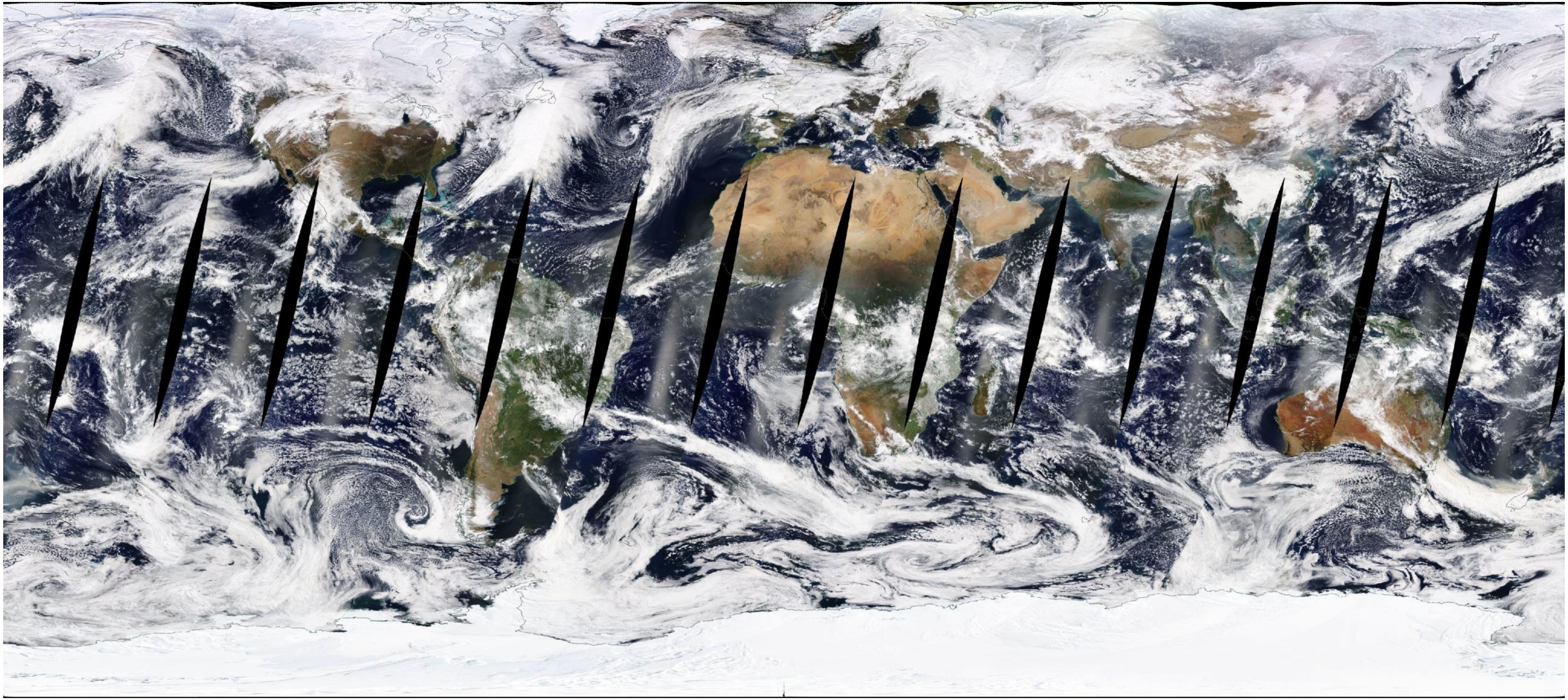
Biases in all-sky data assimilation: ignore, screen, correct?

Katrin Lonitz

Thanks to Alan Geer, Philippe Chambon, Yanqiu Zhu, Stefano Migliorini, Kozo Okamoto, Masahiro Kazumori, Min-Jeong Kim, Jason Otkin, David Duncan, Tony McNally

ECMWF

All-sky conditions



MODIS visible image on 2 February 2020

Which Level 1 data do we use in all-sky conditions?

Microwave

- AMSU-A
- ATMS
- MHS
- GMI
- SSMIS
- MWHS-2
- AMSR2
- Cloudradar
- Precipitation Radar (GPR)

Infrared

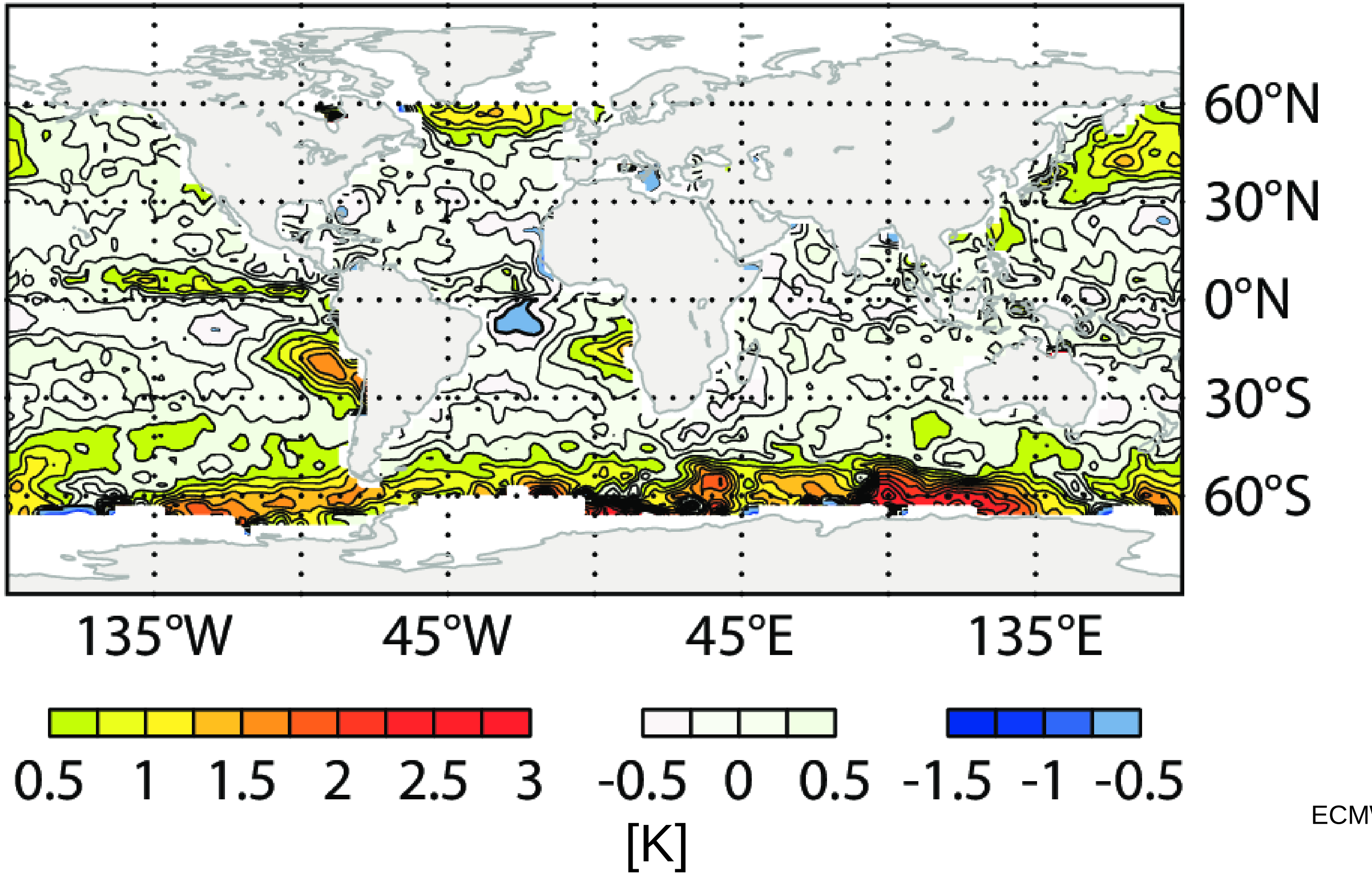
- Himawari-8
- IASI

VIS

- SEVIRI
- Lidar

Which biases do we face in all-sky conditions?

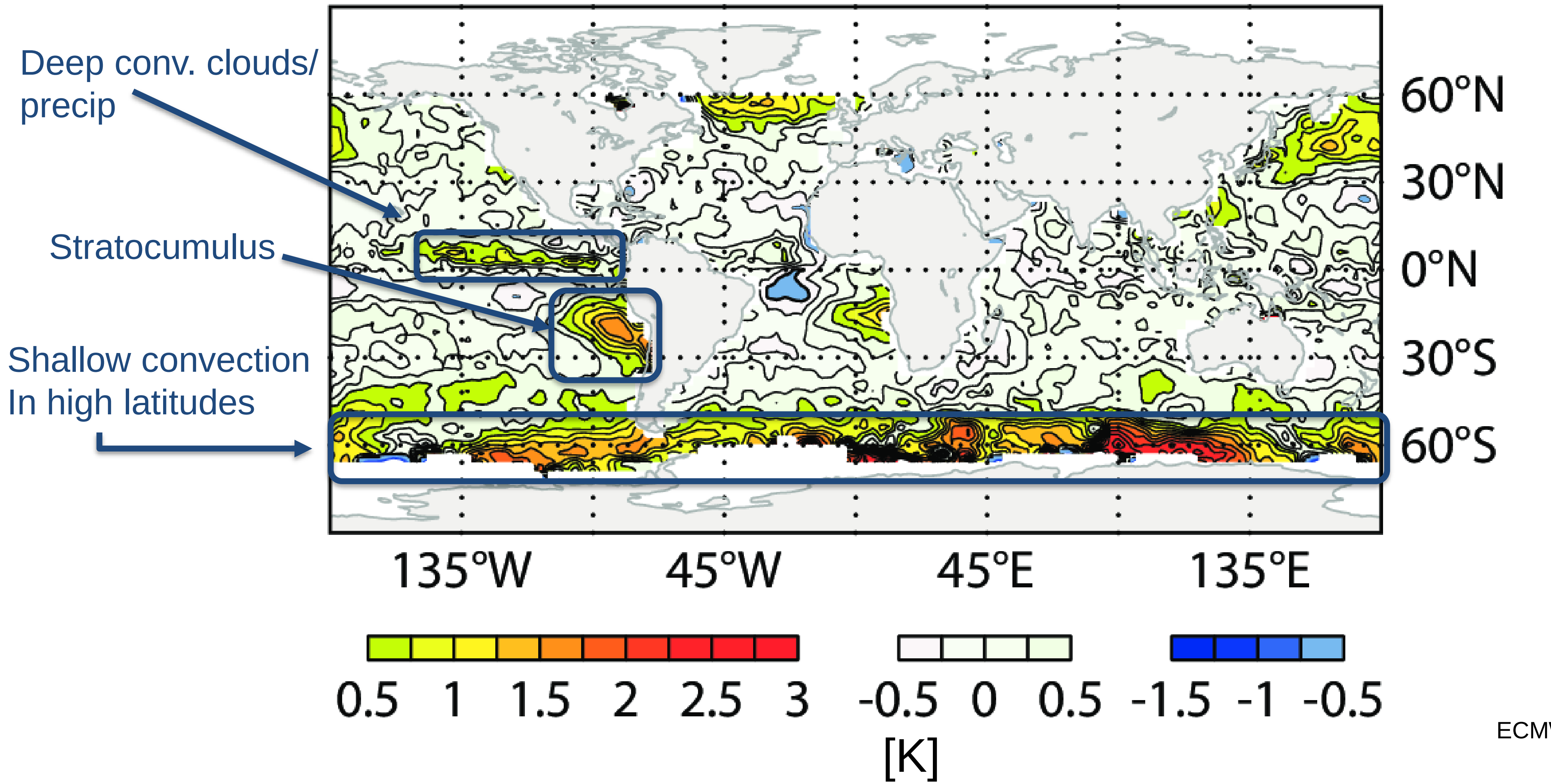
- FG departures from SSMIS, 92 GHz, IFS Ops (HRES), May 2014 – April 2015



ECMWF Newsletter No. 146

Which biases do we face in all-sky conditions?

- FG departures from SSMIS, 92 GHz, IFS Ops (HRES), May 2014 – April 2015

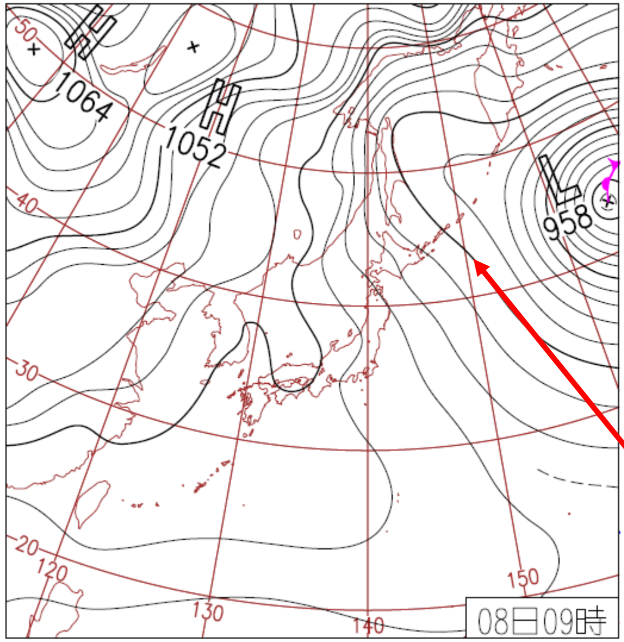


ECMWF Newsletter No. 146

Masahiro Kazumori

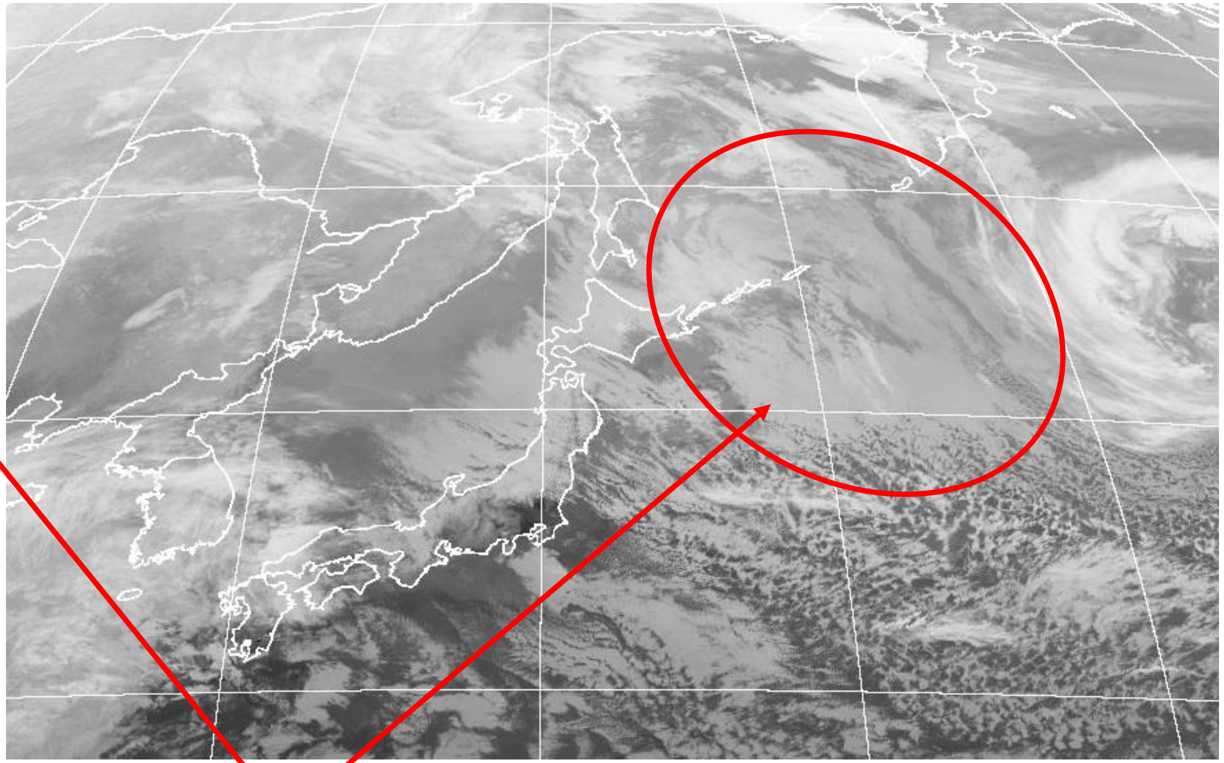
Cold-air outbreak area in mid- and high latitudes

Surface Weather Chart
February 8, 2017



8日(水)西から天気下り坂
冬型の気圧配置は西から次第に緩む。夜には日本海西部と四国の南で低気圧が発生し、西日本で雨や雪。太平洋側は東日本を中心に晴れ。那覇市でヒカンザクラが満開。

2017年2月8日03UTC Channel 13 Himawari-8

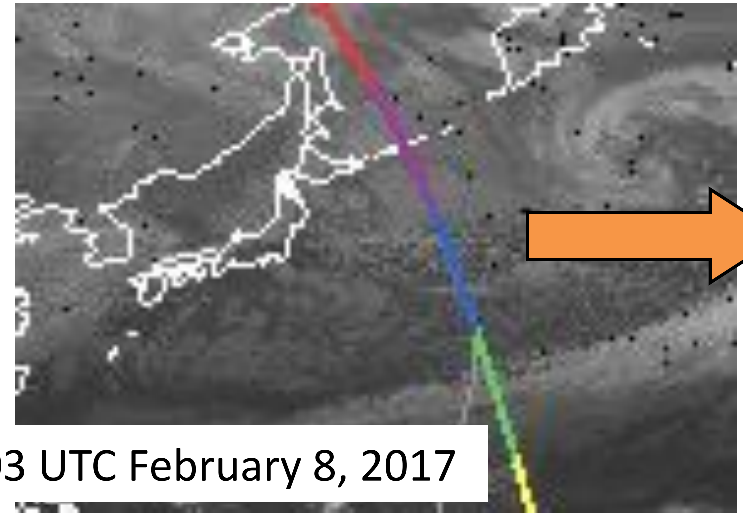


Occurrence of low-level clouds

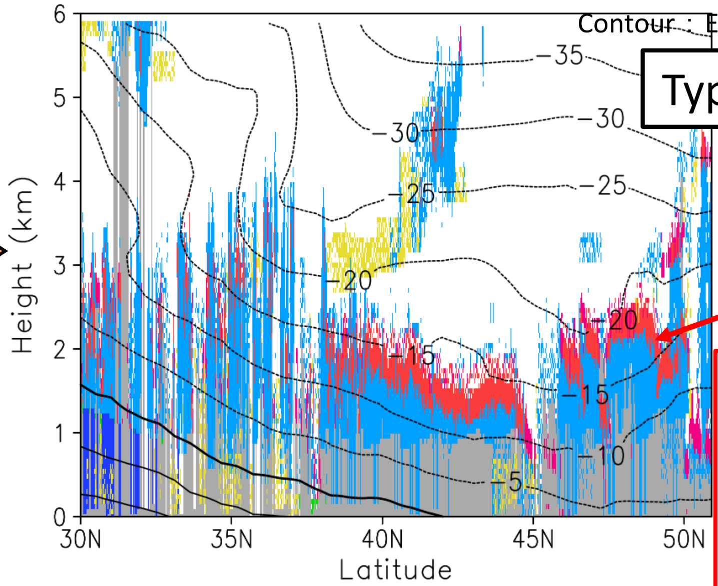
Example near Japan

Vertical cross section observed by CloudSat CALIPSO

Masahiro Kazumori



03 UTC February 8, 2017



Type of hydrometeor

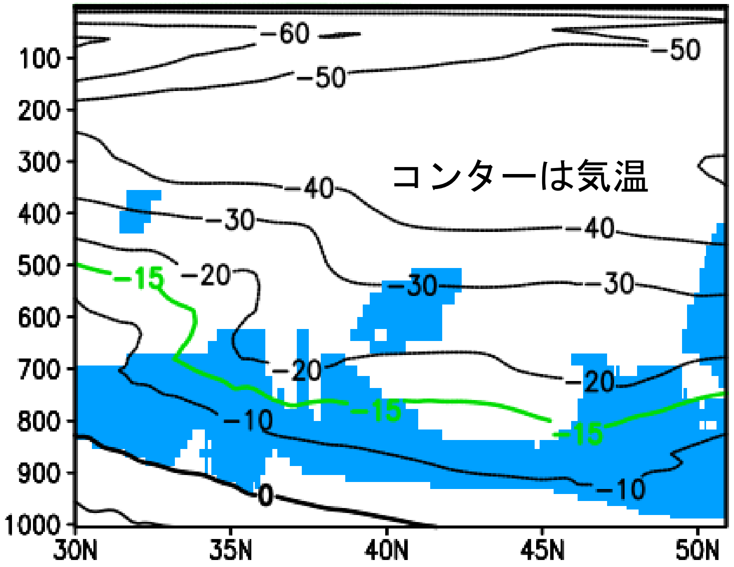
Rain

Cloud liquid
（Super-cooled）

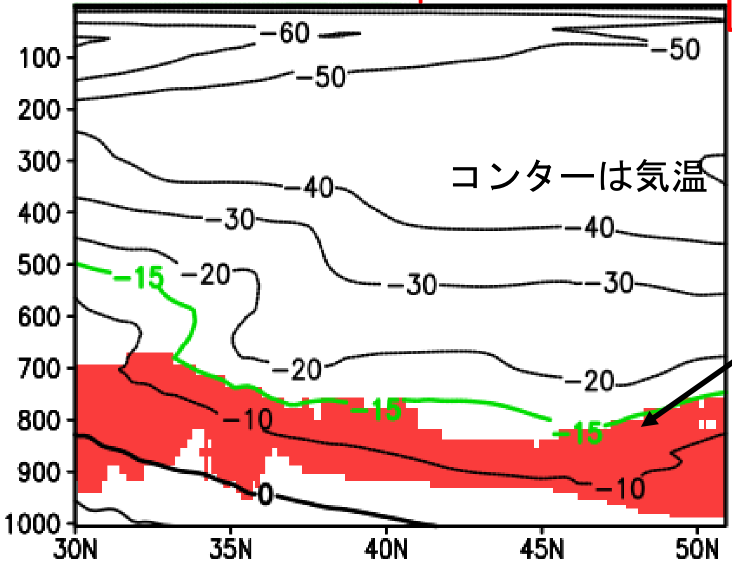
Cloud ice

Cloud liquid over the top of low-level cloud （supercooled）

Model cloud ice area



Model cloud liquid area



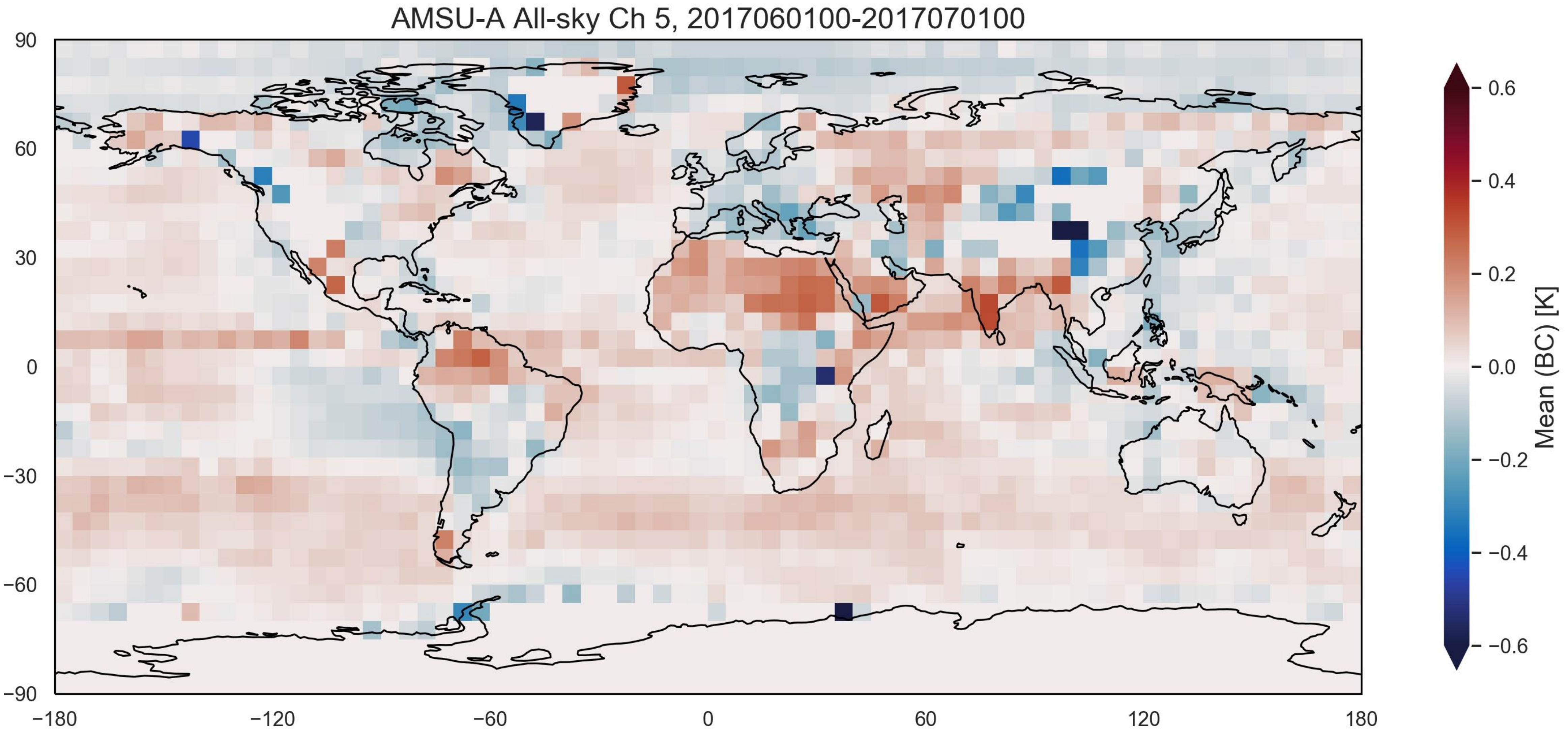
No cloud liquid water below -15℃

This value is a threshold of model cloud physics in JMA global model

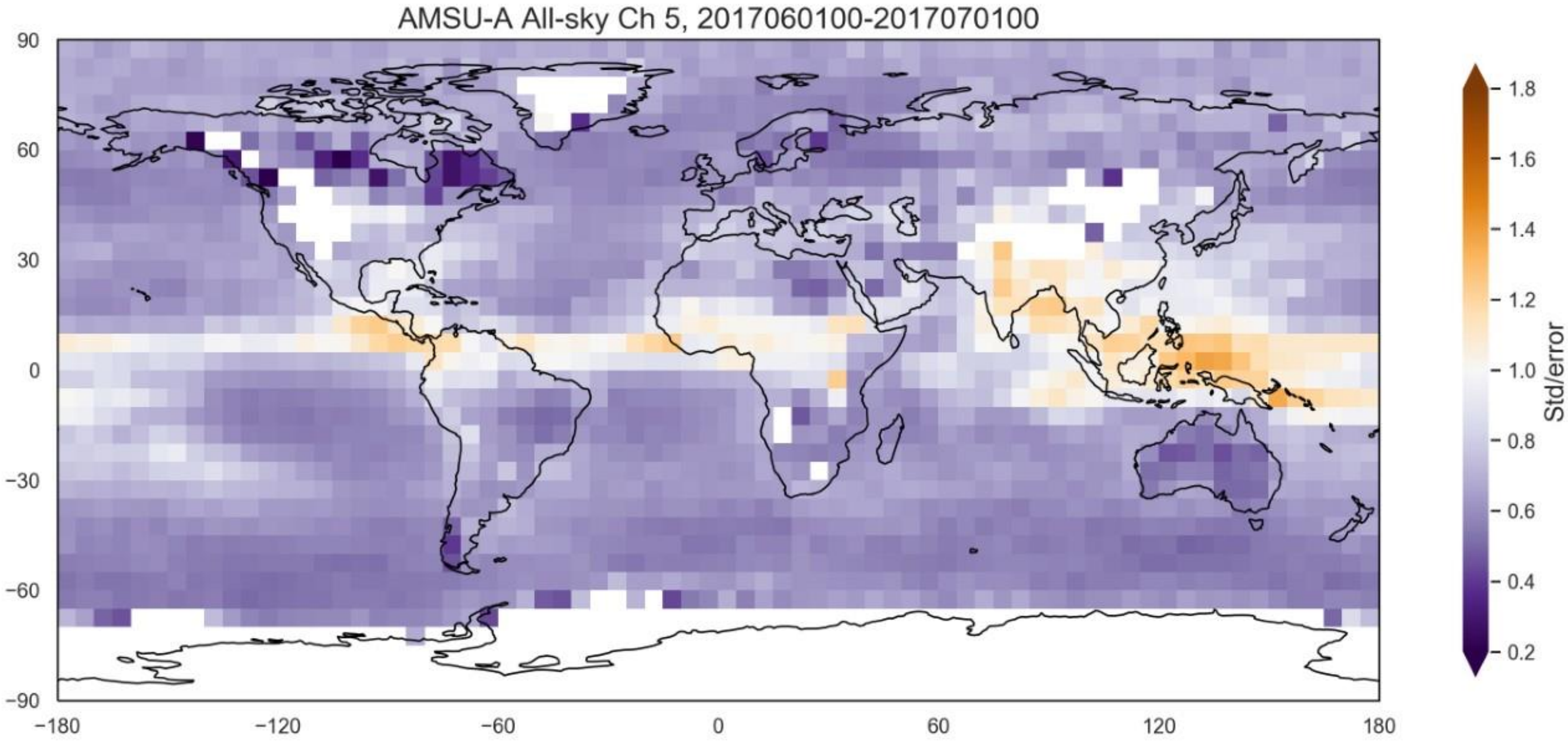
No cloud liquid water due to no super-coolded liquid water over the top of clouds

David Duncan

Mean FG departure for AMSU-A channel 5 (53.6hGHz) - ECMWF



Std. dev. In FG dep norm. by error for AMSU-A channel 5 (53.6hGHz)

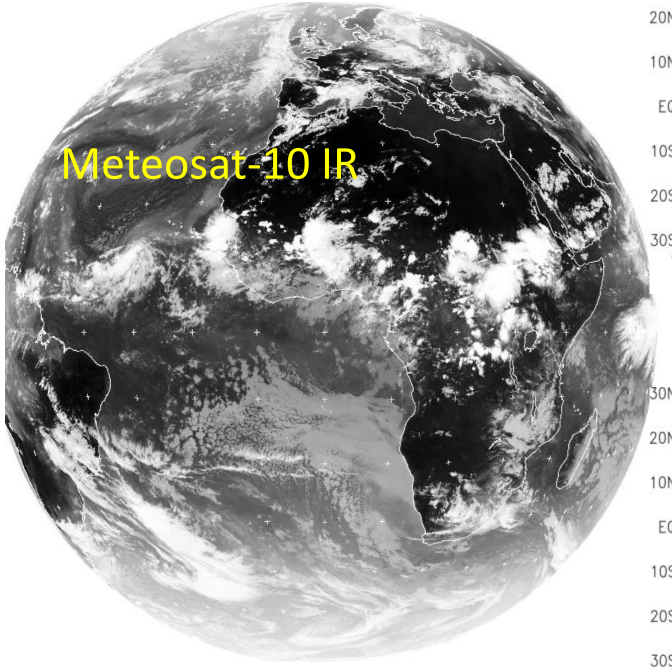


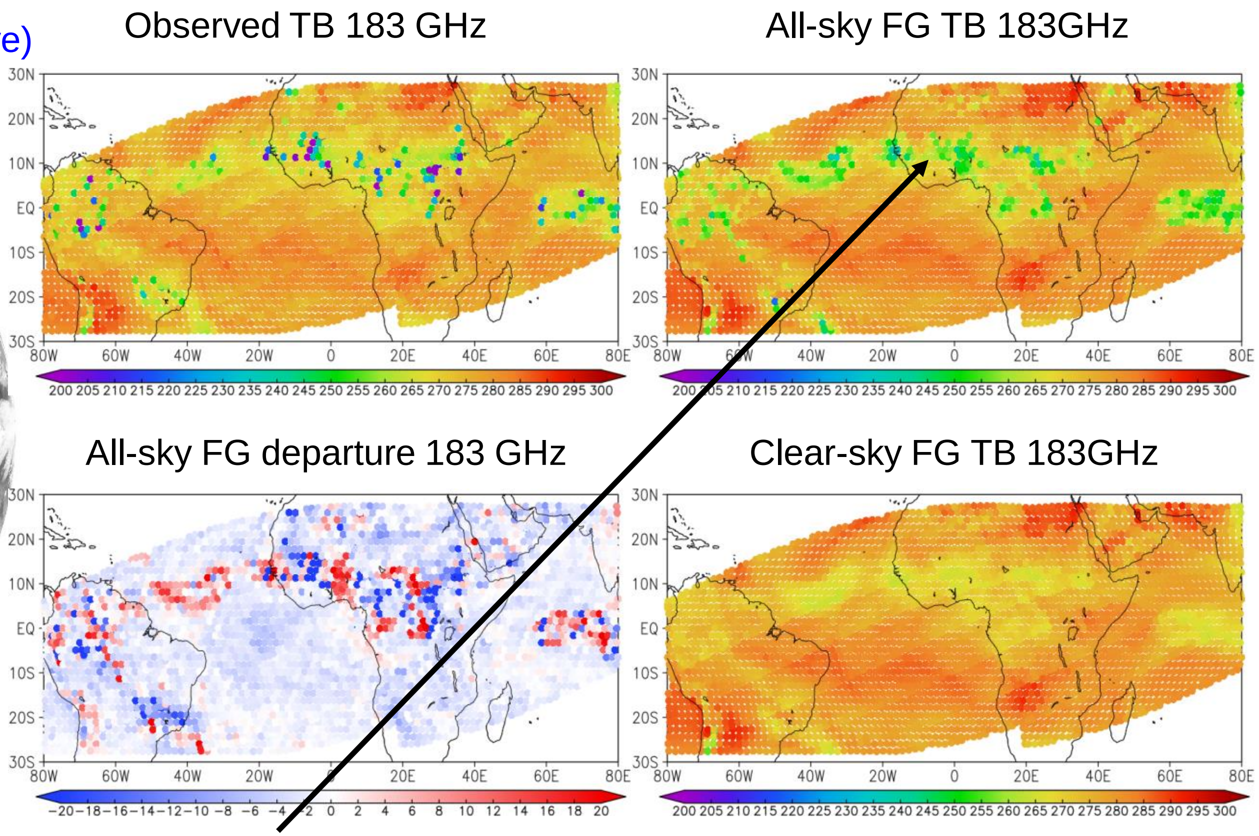
Masahiro Kazumori

Forecast model bias issues in all-sky MW DA @JMA

- **Underestimation of strong convective clouds in the JMA global model?**

SAPHIR 183.31+6.8 GHz
August 14, 2016 (all-sky passive)

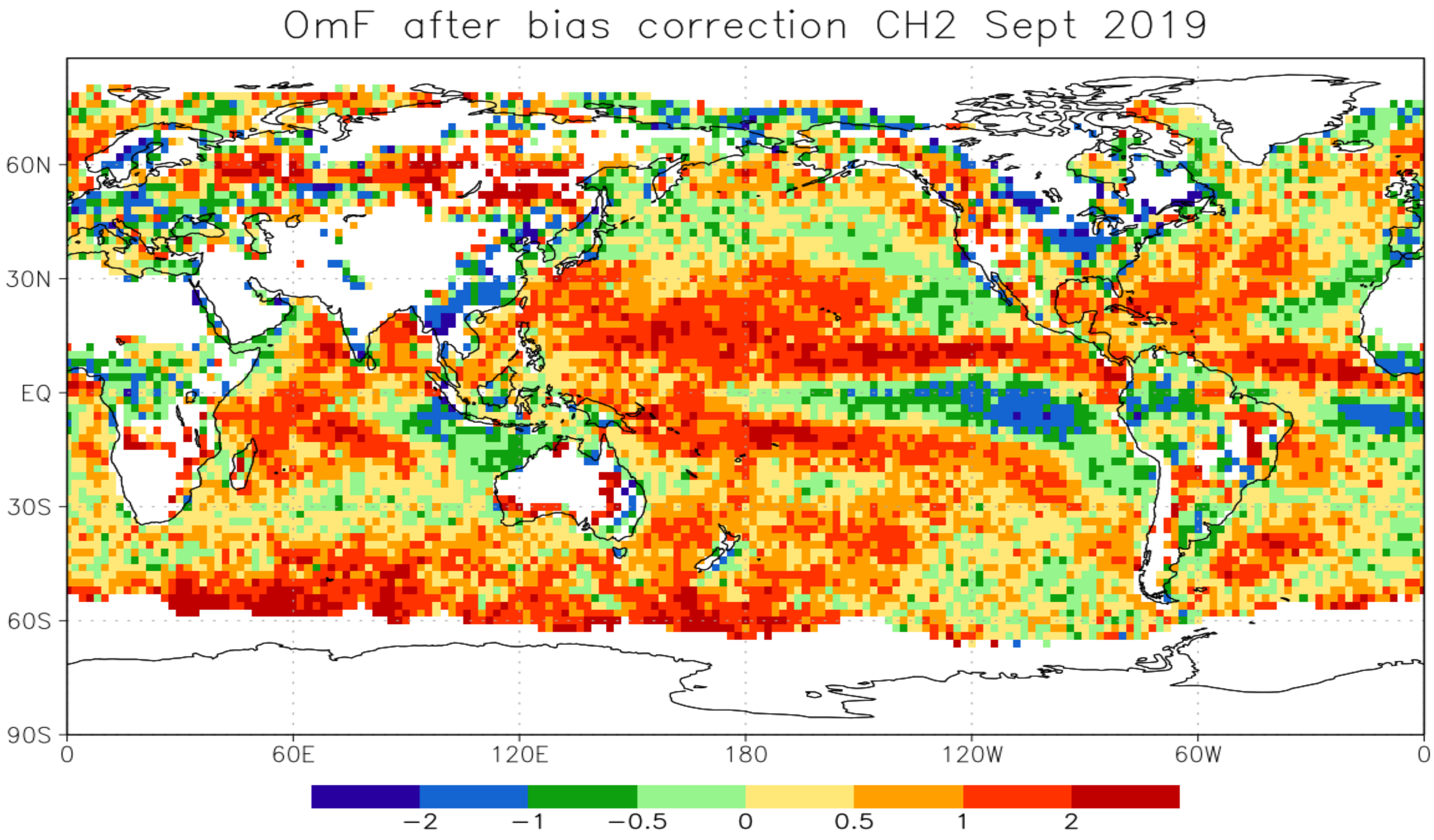




Model's convective clouds are weak and broadly spread.
Model's precipitation representation is crucial for all-sky 183 GHz humidity sounding radiance assimilation.

Yanqiu Zhu

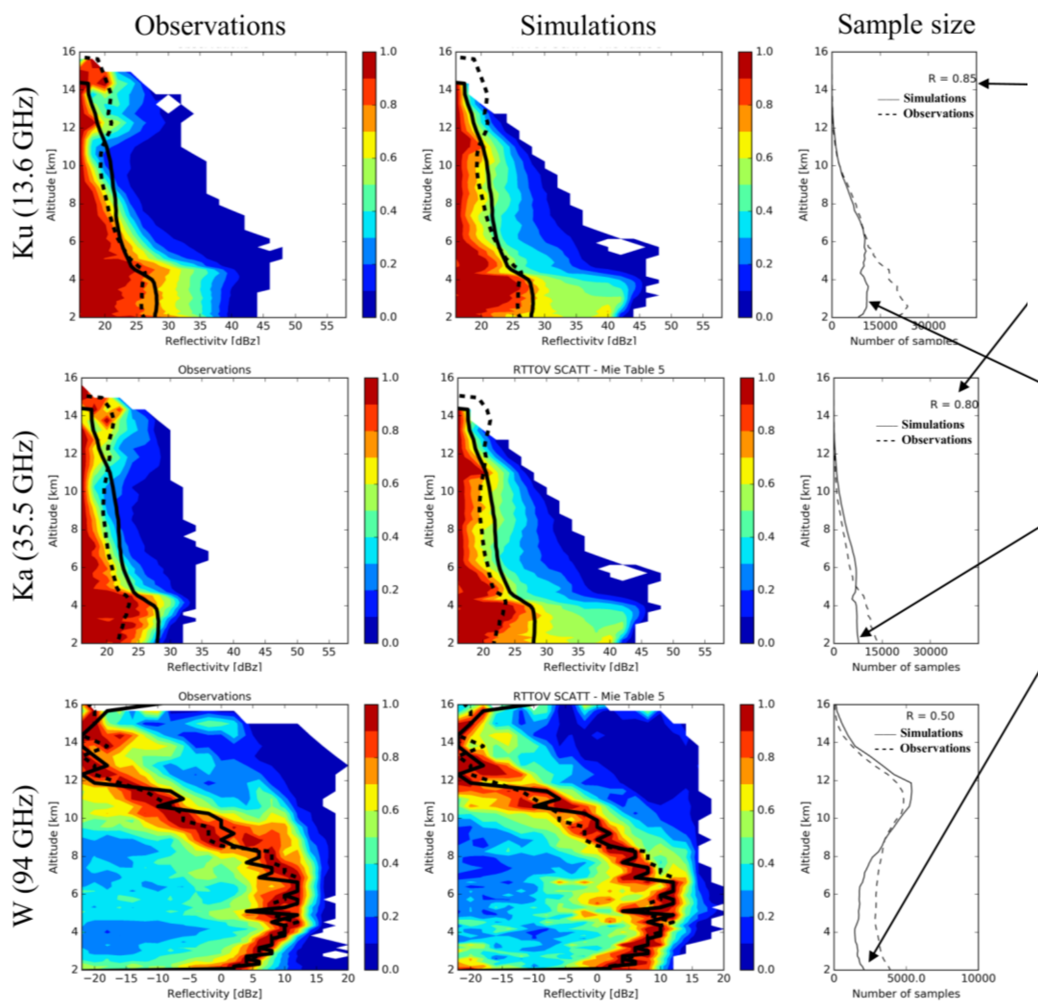
Mean OmF for ATMS NPP in operational FV3GFS at NCEP



ATMS NPP Channel 2 (31.4vGHz) , Sept 2019

GPM/DPR & Cloudsat/CPR in combination with AROME in Tropics

Fabrice Duruisseau,
Philippe Chambon,
Ghislain Faure,
Alan Geer



Solid line = simulations
Dashed line = observations

Less profiles with simulated reflectivites >16dBz for DPR and >-24dBz for CPR simulations below the freezing level.

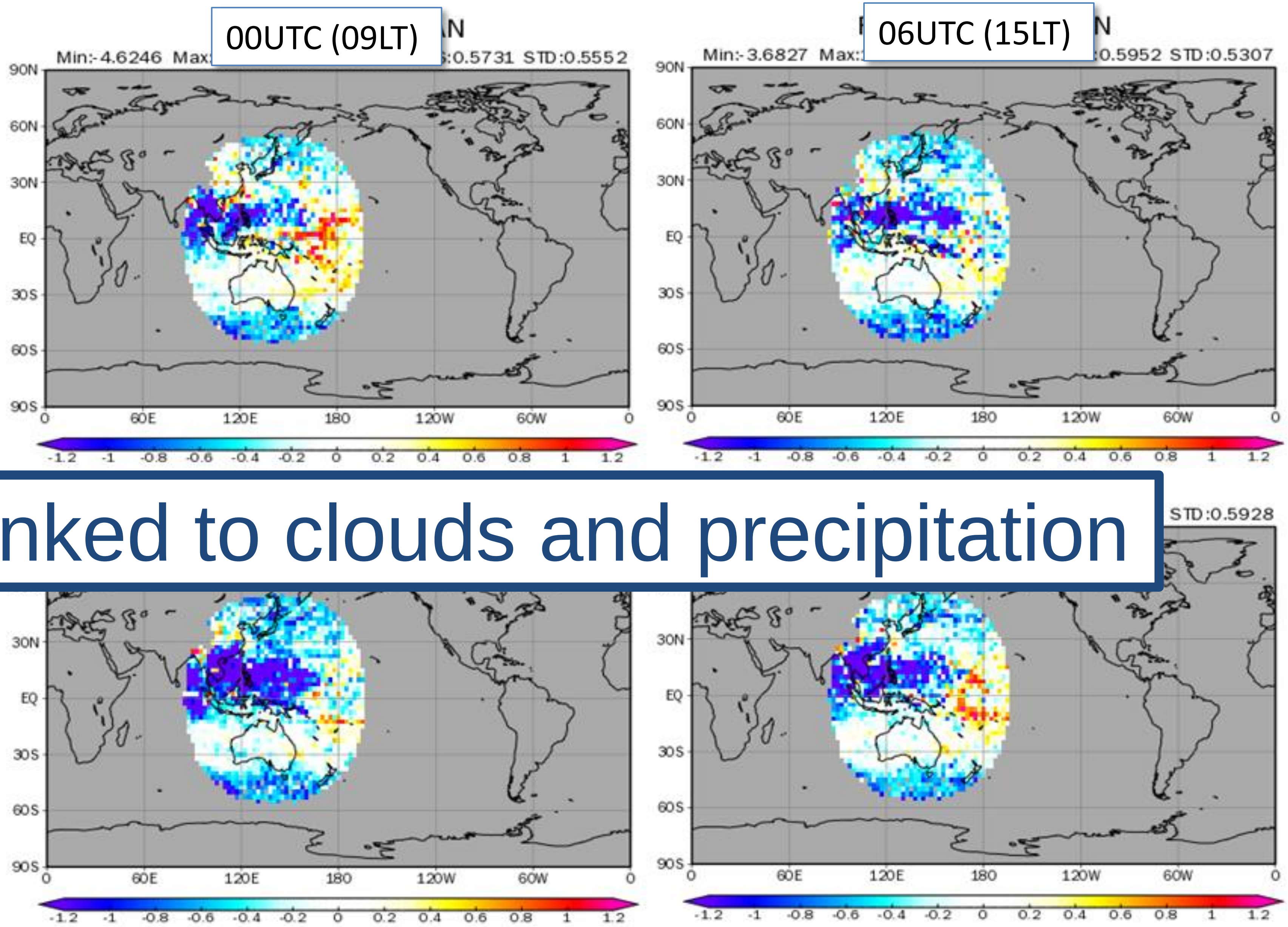
- AROME micro physics tend to keep not enough ice for the cloud shield of the convective systems but to produce too much snow
- Modifying the auto-conversion function within the micro-physics scheme seems to slightly improve it

Norm. histogram of attenuated reflectivities. Normalisation done by maximum number of occurrence at each elevation. August 2017over Antilles.

Kozo Okamoto

Forecast model bias issues in all-sky IR DA @JMA

- O-B at Himwari8/AHI band 08 (6.2µm) in August 2018
- Model does not well represent diurnal variation of high clouds
  - Wider and thicker around the conve
    the aftern



Biases linked to clouds and precipitation

Which biases do we face in all-sky conditions?

Additional biases due to the presence of cloud and precipitation.

- These biases are:
  - representativity
  - biases in RT transfer: assumptions and approximations of hydrometeor properties can cause biases in simulated radiances (Petty and Huang, 2010)
  - model biases
  - sampling bias (Geer and Bauer, 2010)

Which techniques do we have to deal with biases?

- (Variational) bias correction methods, e.g. Dee, 2004; Auligne *et al.*, 2007
  - does not work for situation dependent biases (CAO outbreaks)
- Screen
- Ignore
- Non-linear bias correction
- Tools using Machine Learning and/or Topological Data Analysis (TDA)

Shall we correct all (all-sky) biases?

We should correct instrument biases!

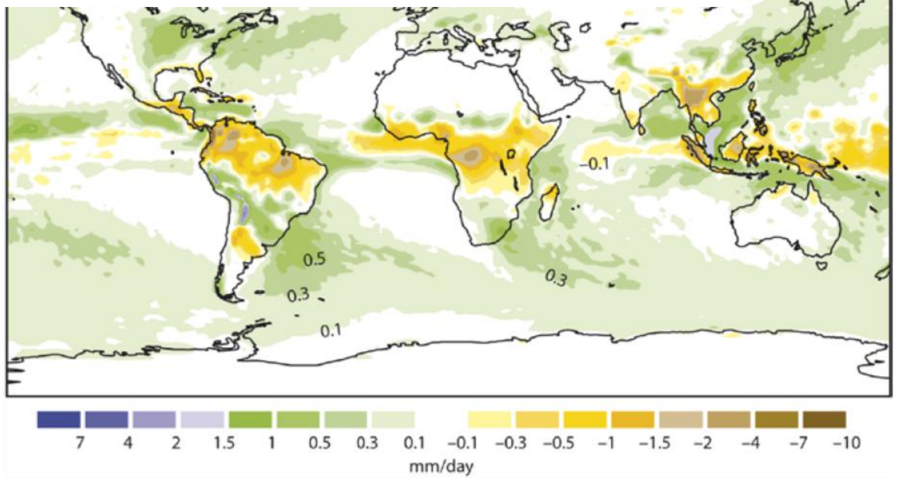
We should correct biases in RT transfer!

But should we correct biases of the model??

- Depends on time scale in which bias is adjusted by the model when corrected
  - If it adjust quickly, correct!
  - If it adjust slowly, do not correct

From ECMWF ERA Report 10, 2011

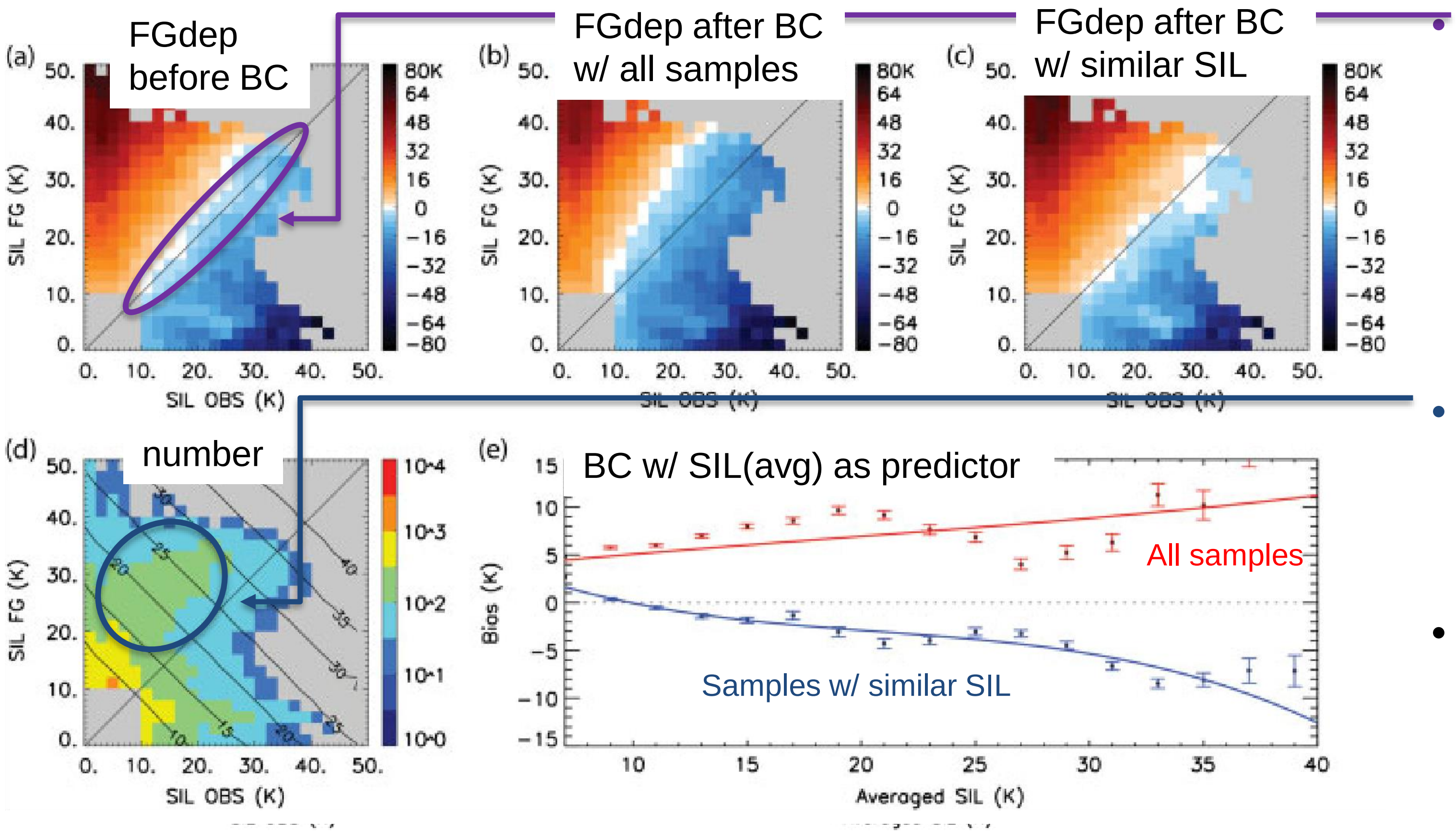
Diff. in Precip between T+24-36h and T+0-12h of ERA Interim



However, we should avoid asymmetric sampling while doing corrections!

Construct bias correction with symmetric sample

Chambon et al. (2014; QJRMS)



Bias of FG departure for 150 GHz SSMIS as a function of Scattering Index over land (SIL) for Sept-Dec 2010/11.

- SIL(obs) = SIL(FG): neg. biases, increasing with increasing SIL (bias due to RTM?)
- Excessive frequency of high scattering index samples in FG (bias due to model?)
- Bias correction (BC)
  - Use averaged scattering Index SIL(avg) as predictor
  - BC based on all samples fails to reduce the bias effectively
  - **BC based on samples with similar SIL is better**

(Variational) Bias correction + symmetric sampling

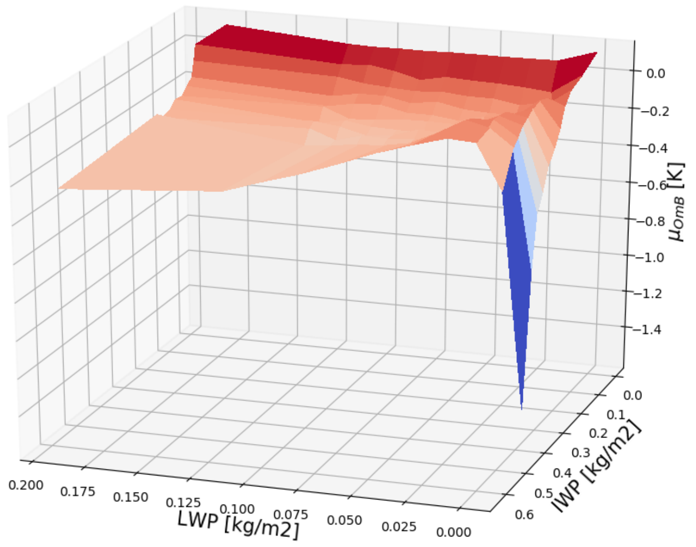
- Use selected subset to train bias correction coefficients for clear-sky and cloudy sky,
  - MetOffice for MHS in clear skies
  - Nasa/GMAO for GMI in clear and cloudy skies,
  - NOAA/NWS AMSUA &ATMS in clear and cloudy skies,
- BC coefficients trained for using only subset but BC is then applied to all-data

Stefano Migliorini,
Brett Candy

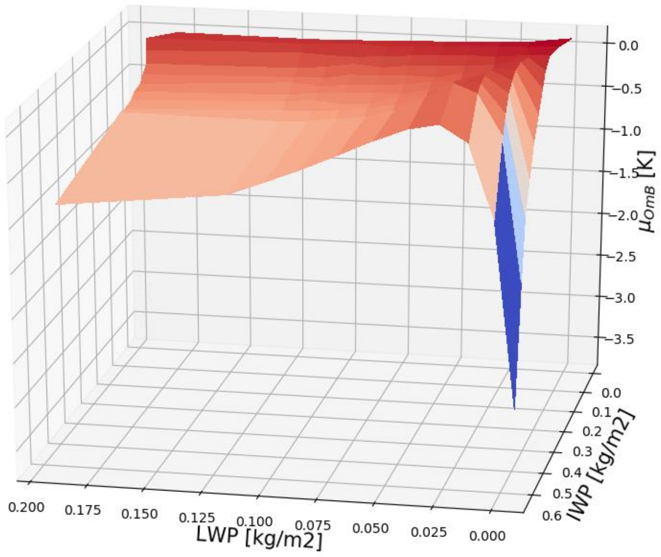
MetOffice: Ongoing work for all-sky MHS

- FG departures from all-sky MHS trials show presence of bias for scenes affected by frozen cloud (i.e. high-level cloud with high IWP and low LWP)
- This motivates retraining of initial bias correction for MHS channel 3, 4 and 5 (i.e. the assimilated channels)

MHS ch 4



MHS ch 5



MetOffice: Ongoing work for all-sky MHS

Stefano Miglorini,
Brett Candy

- Current VarBC includes all radiances for bias correction, including cloud-affected radiances
- No specific cloud or precipitation bias predictors are used.
- Currently developing selective VarBC where we can choose subset of radiances for bias correction, e.g**. clear-sky only radiances**
- Plan to perform assimilation trials without biased data and compare results with full (non-precipitating) radiance dataset

NASA GMAO – VarBC work for GMI

- All-sky MW: VarBC pred = no retrieved cloud water path as pred., **only near-clear sky observations with near-clear sky background profiles** are used in updating the bias correction coefficients:
  - the observed cloud index, CIo, is less than 0.05,
  - the simulated cloud index, CIg, is less than 0.05, and
  - the absolute difference between CIo and CIg is less than 0.005.
- Roughly 35% of the assimilated observations for each channel are used to update the bias correction coefficients

Min-Jeong Kim

NASA GMAO – VarBC work for GMI

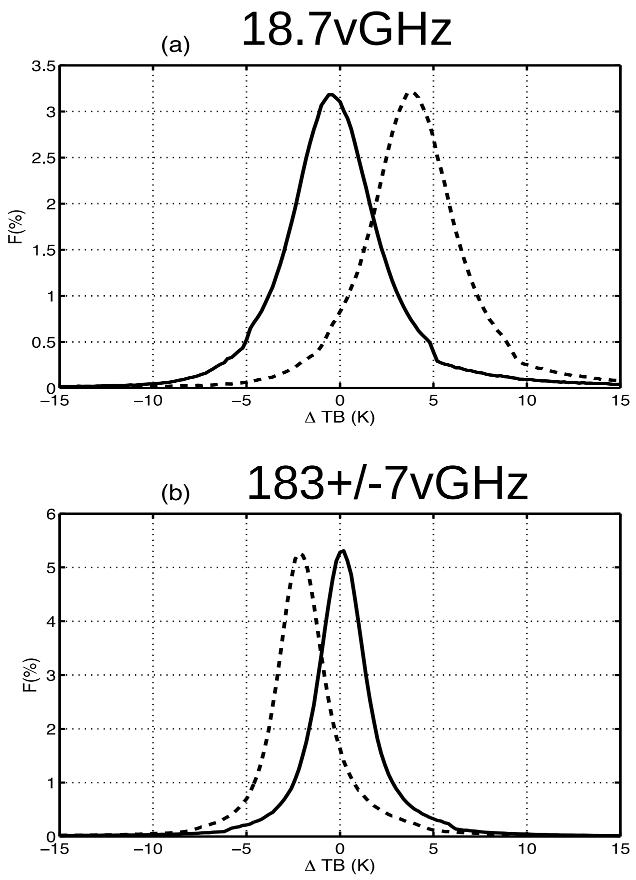


Figure 1: Histograms of all-sky first-guess departures in GMI (a) channel 3 and (b) channel 13 before (dashed line) and after (solid line) bias correction is applied.

- Remaining biases due to thick cloud and heavy precipitation
- Additional bias correction to correct cloud-amount FG bias:
  - CI_{avg} and CI_{avg}^2 = additional BC predictors
  - BC coefficients for these predictors are updated each cycle **using only data where both the observed and simulated cloud indices are greater than 0.05 and their absolute difference is less than 0.005**

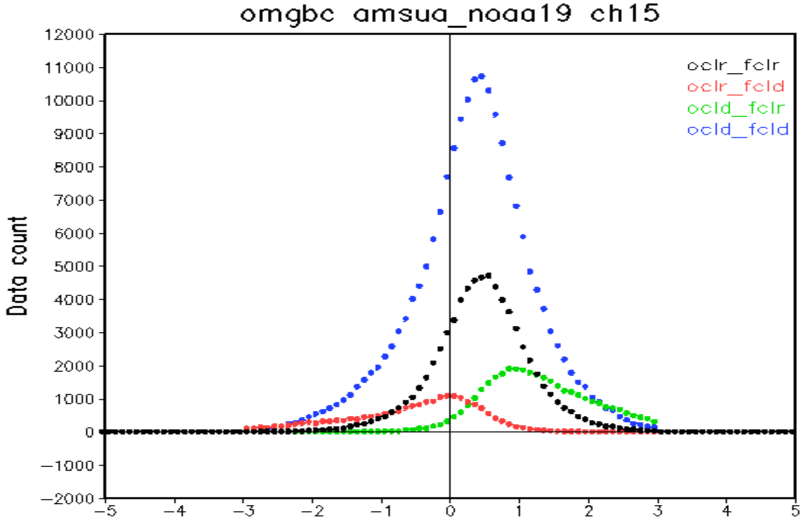
Yanqiu Zhu

NOAA/NWS: VarBC for AMSU-A and ATMS

- Based on cloud liquid water calculated from radiance observation (O) and first guess (F),
  1) O:clear vs. F:clear
  2) O:clear vs. F:cloudy
  3) O:cloudy vs. F:clear
  4) O:cloudy vs. F:cloudy

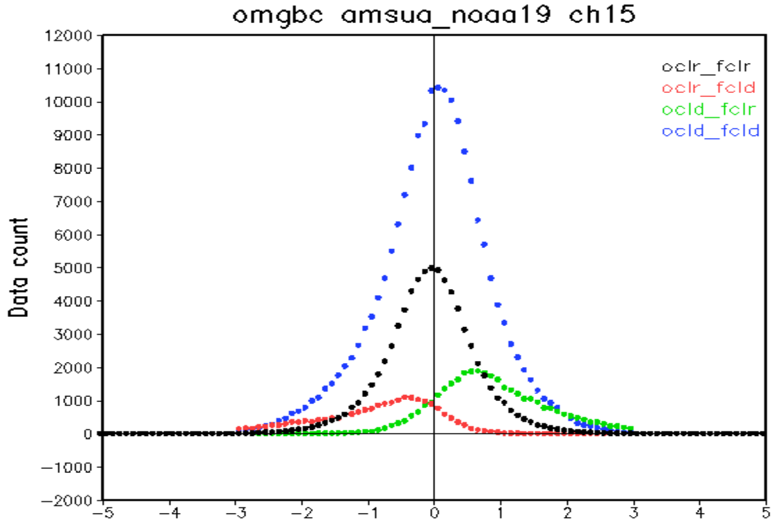
  Mismatched clouds due to model error
- **BC coefficients are obtained using only a selected data sample with consistent cloud info between the first guess and the observation**
- Use latest bias coef. available to bias correct the data with mismatched cloud info

without using the all-sky strategy



Normalized OmF w/ BC

using the all-sky strategy

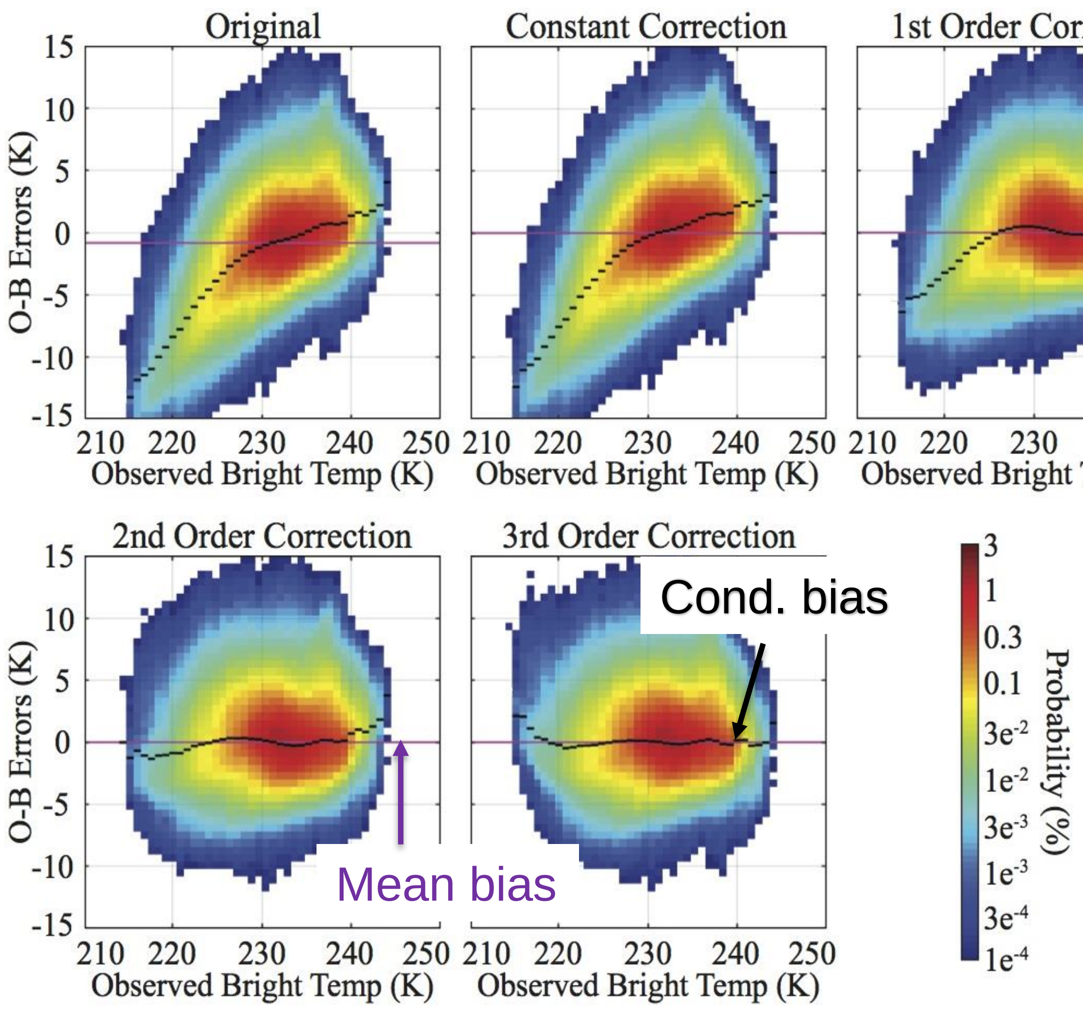


Normalized OmF w/ BC

Otkin et al. (2018; MWR)

Nonlinear Bias Correction Method

Remove linear and nonlinear conditional biases from all-sky satellite observations using a Taylor series expansion of the OMB departures

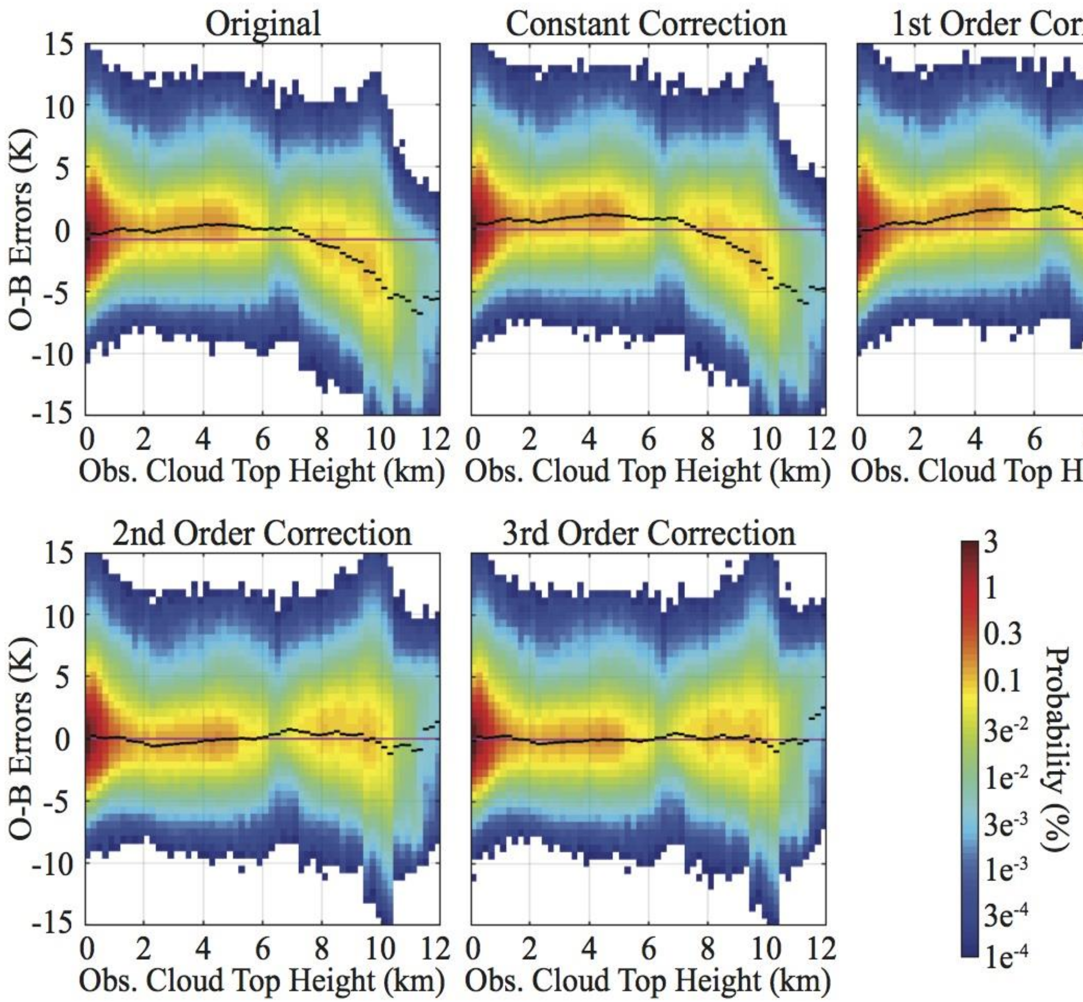


- Results evaluated for original, 0th (constant), 1st (linear), 2nd (quadratic), and 3rd (cubic) order Taylor series expansions
- Assessed to how bias varies as a function of the predictor value

Otkin et al. (2018; MWR)

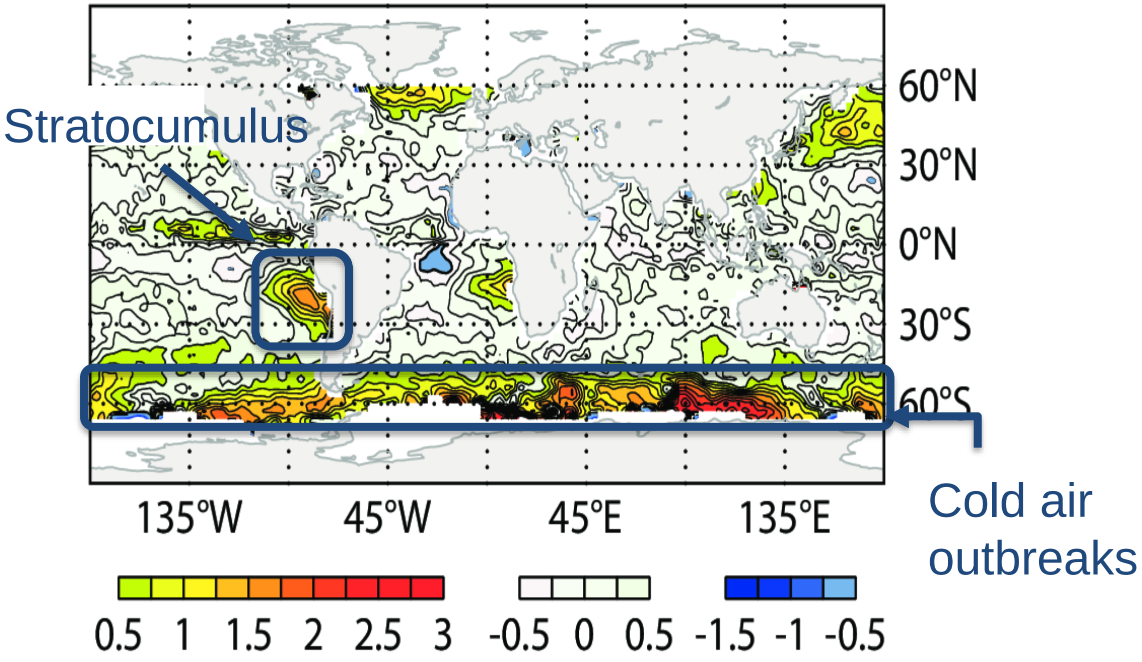
Retrieved Cloud Top Height Predictor

Cloud top height can serve as an effective bias predictor for IR Tbs when higher order Taylor series terms are used

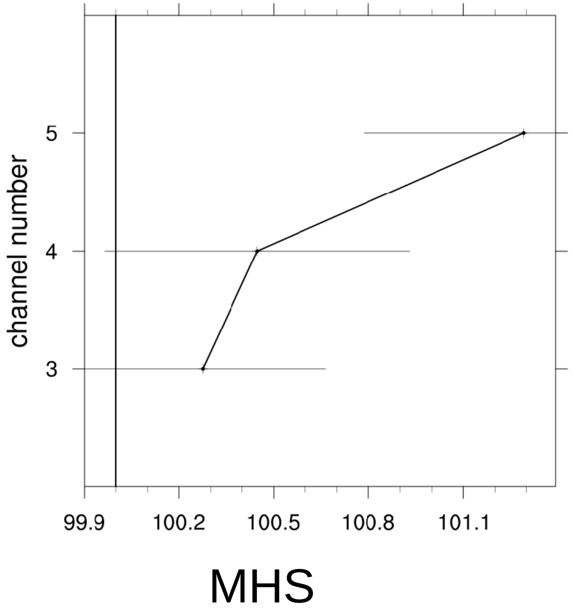


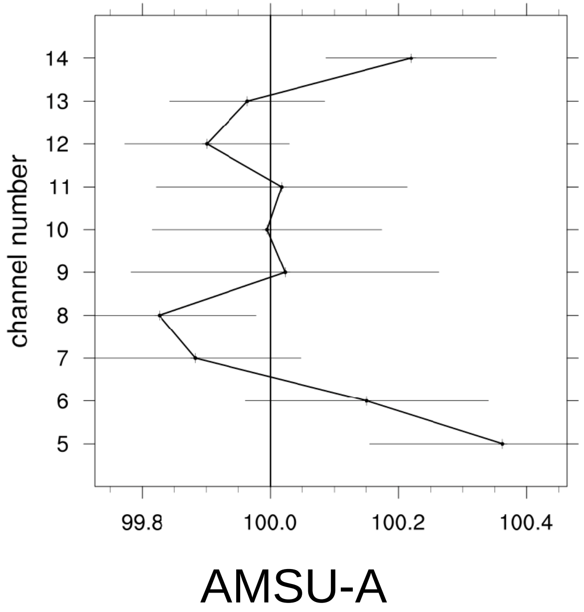
- Nonlinear conditional bias error pattern in the original distribution
- Constant and linear BC terms unable to remove the conditional biases
- Arch pattern in the 1st order conditional biases removed when using the 2nd order quadratic term
- Some additional small reductions in the biases after using 3rd order term

What to do if (Var)BC cannot be applied?



- **Screening:** cold-air outbreaks (Lonitz and Geer, 2015):
  lower tropospheric stability (LTS) <12 K and mostly liquid clouds
- **Ignoring:** Stratocumulus areas (Lonitz and Geer, 2017)



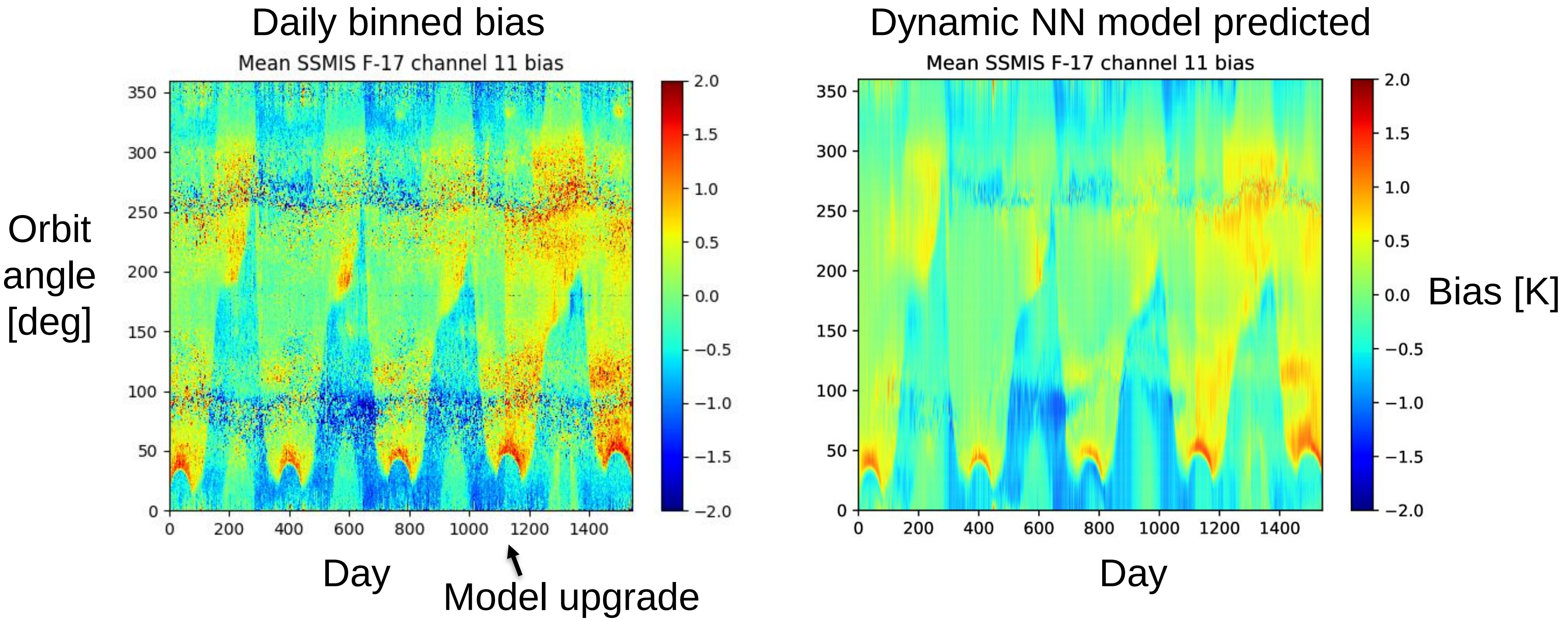


Normalised difference in standard deviation of first-guess departures between Sc off and control for different instruments between 20◦S - equator and 80◦W - 100◦W. Results cover the time period from 1st July 00 UTC to 31st December 2013 12 UTC.

Alan Geer

New methods…Machine Learning

- Learning about ML: satellite bias correction applications
- Example of 183GHz bias for SSMIS-F17

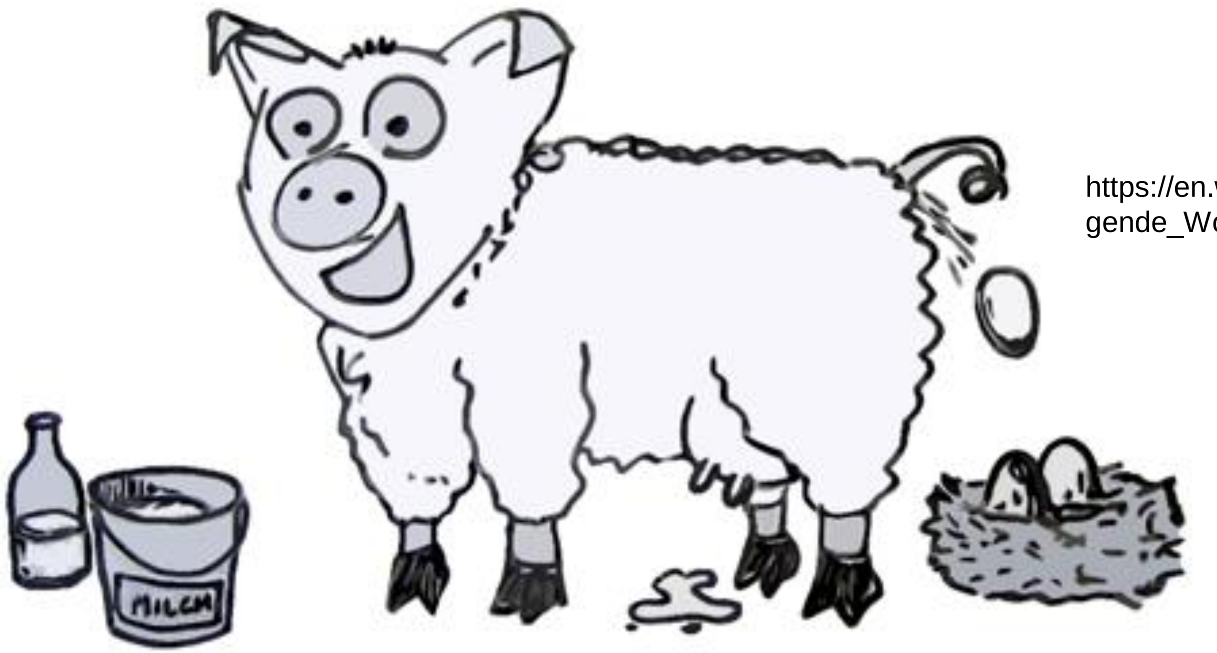


Topological Data Analysis (TDA)

- TDA is an interdisciplinary field spanning topology and data analysis
- Tool to uncover patterns in data.
- Based on the philosophy that data has shape, and that shape has meaning.
- Persistence homology (PH) is a technique from TDA that can identify clusters, holes, and voids within a set of points.
- More persistent topological features are detected over a wide range of spatial scales and are deemed more likely to represent true topological features of the underlying space rather than artefacts of sampling, noise, or other factors.
- Persistent features could be biases in all-sky conditions?

Things to talk about and discuss

- There is not just **one** tool to detect and correct all biases under all-sky conditions

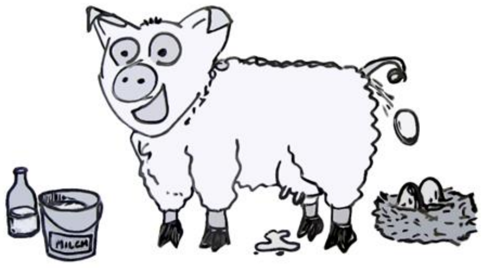


https://en.wiktionary.org/wiki/eierlegende_Wollmilchsau

“eierlegende Wollmilchsau” (egg-laying wool milk sow) ~ Swiss army knife

Things to talk about and discuss

- There is not just **one** tool to detect and correct all biases under all-sky conditions



https://en.wiktionary.org/wiki/eierlegende_Wollmilchsau

- Current techniques on detecting/correcting bias are based often on looking into PDF of FG departures, with the aim of mean FG departures being zero when all biases are zero
  - This might lead to oversee compensating biases or filter out “signals”
- Should we look into new techniques to find biases?
- Best: Use various metrics to detect bias!
- Come up with new metric to evaluate success of...also fits to (independent) obs, etc.

Call for EUMETSAT Research Fellowships at ECMWF

- All-sky assimilation of radiances from microwave instruments in Numerical Weather Prediction
  - 5-year contract
  - Deadline 16 March 2020

- Assimilation of geostationary radiances in Numerical Weather Prediction
  - 3-year contract
  - Deadline 16 March 2020

For more information see: https://www.eumetsat.int/website/home/AboutUs/Jobs/Vacancies/index.html