The assimilation of cloud and rain-affected observations at ECMWF

Peter Bauer

Tony McNally, Alan Geer, Philippe Lopez, Angela Benedetti, Carla Cardinali, Niels Bormann, Andrew Collard, Dick Dee, Marta Janisková and Graeme Kelly, Sabatino Di Michele.

ECMWF

History at ECMWF

Early 1990s	Experiments on diabatic forcing through normal mode initialization (Puri and Miller, Heckley et al.)
1998-2002	Development of 1D+4D-Var assimilation approach using TRMM surface rain rate retrieval products (Marécal & Mahfouf)
1999-2003	Cloudy radiance calculations and 1D-Var retrieval studies with HIRS, AMSU data (Chevallier et al.)
2001-2003	Experimental 1D+4D-Var with SSM/I radiances (Moreau et al.)
2001-2003	Tests with ARM ground-based cloud-radar data (Janisková et al.)
2002-2004	Tests with TRMM PR active data (Benedetti et al.)
2004-2006	Tests with SEVIRI cloudy radiance assimilation (Szyndel et al.)
2003-2005	Operational 1D+4D-Var (Bauer et al.)
2006-present	Revision and extension of operational 1D+4D-Var to other instruments (Geer et al.)
2006-present	Experimental 4D-Var radiance assimilation (Bauer et al.)
2006 present	Tests with surface radar network data (Lopez et al.)
2006-present	Tests with MODIS optical depth retrievals (Benedetti & Janisková)

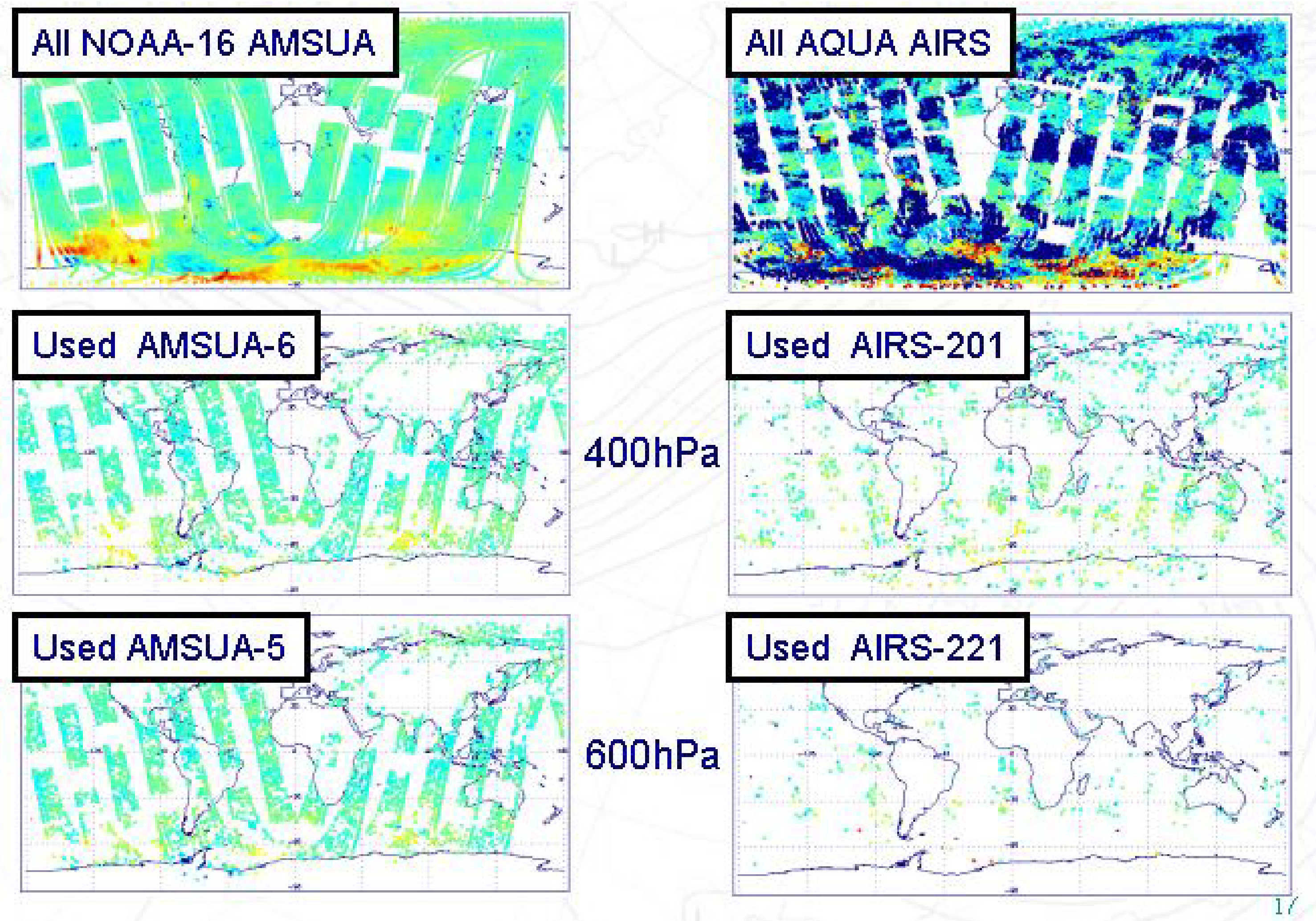
… and continuing developments of linearized moist physical parameterizations (Lopez & Janisková) and data assimilation aspects (Hólm et al.).

Data loss due to clouds/precipitation/surface effects

All data

Upper tropospheric

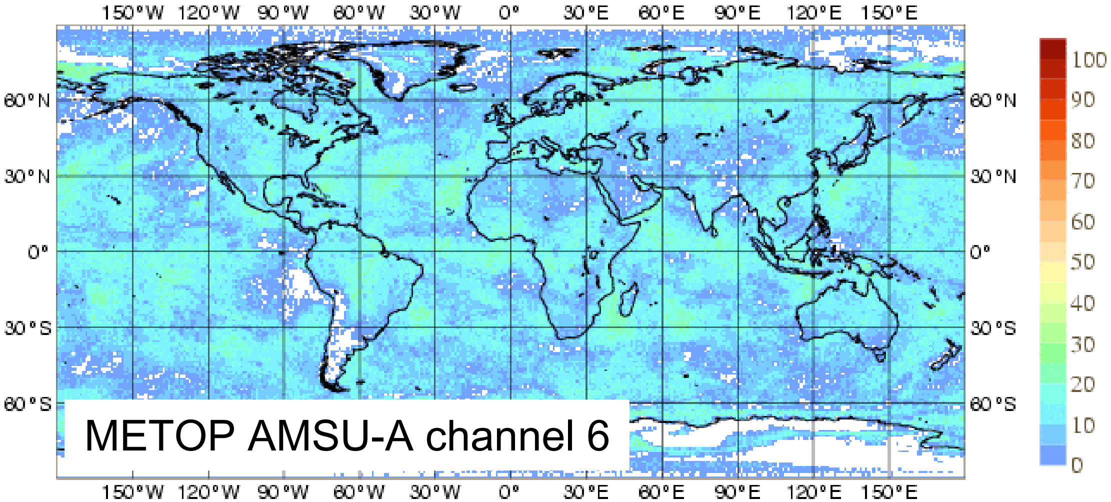
Lower tropospheric

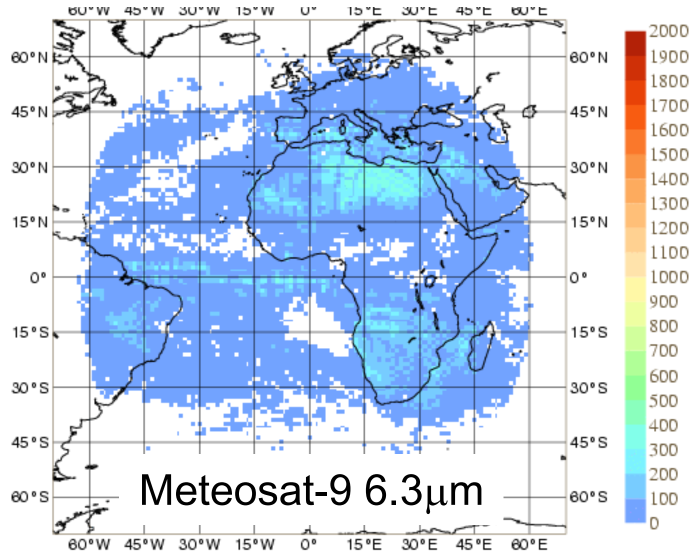


(Courtesy T. McNally)

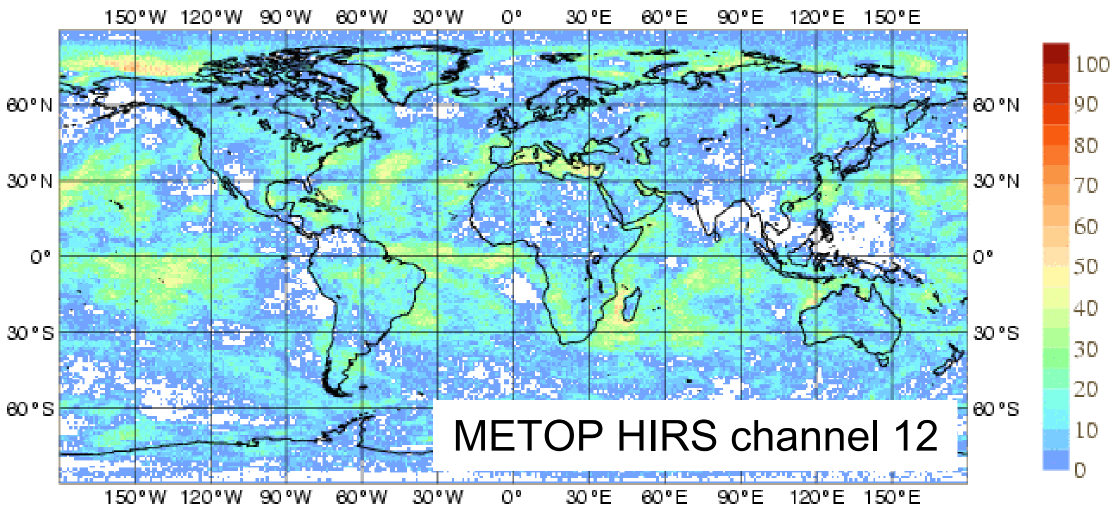
Why assimilate data in cloud-affected areas?

01-04/08/2007 Number of used observations



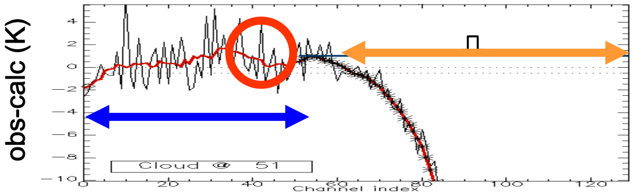


Only 5% of the data volume passes the screening/thinning



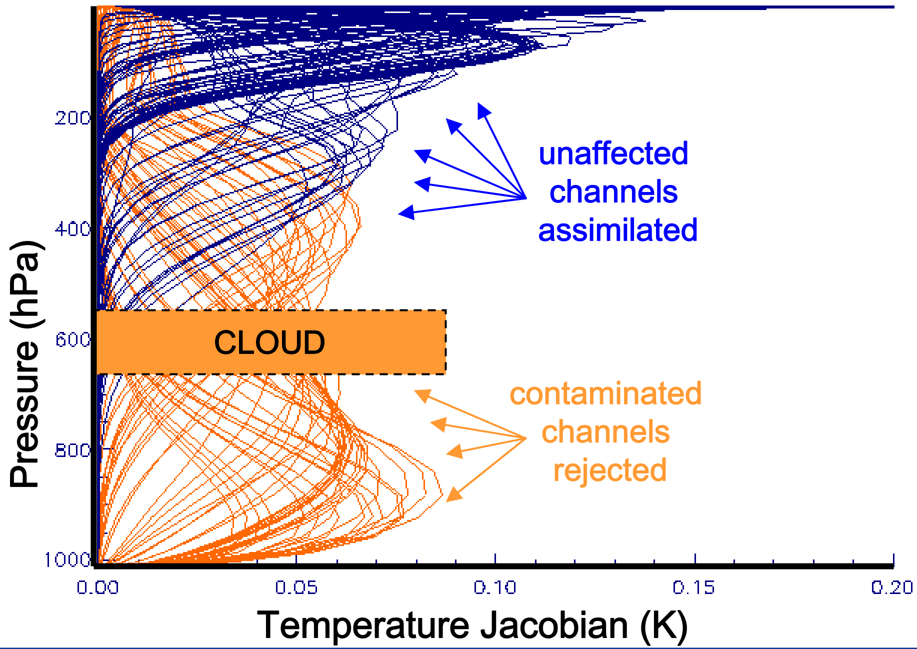
Retaining useful information above clouds

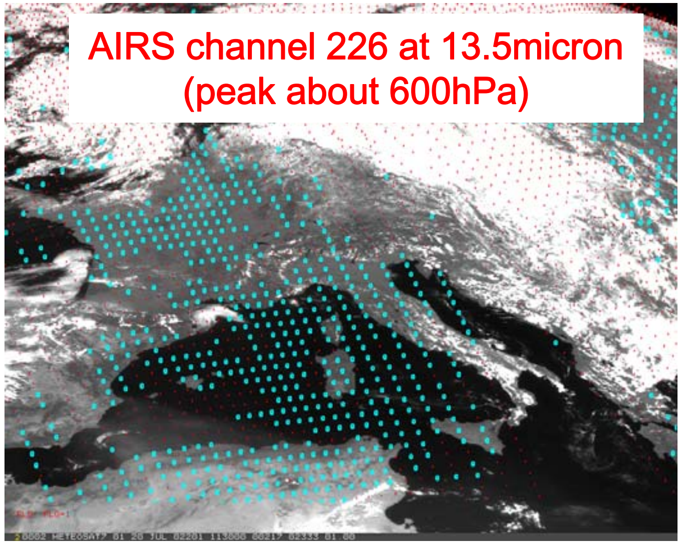
A non-linear pattern recognition algorithm is applied to departures of the observed radiance spectra from a computed clear-sky background spectra.

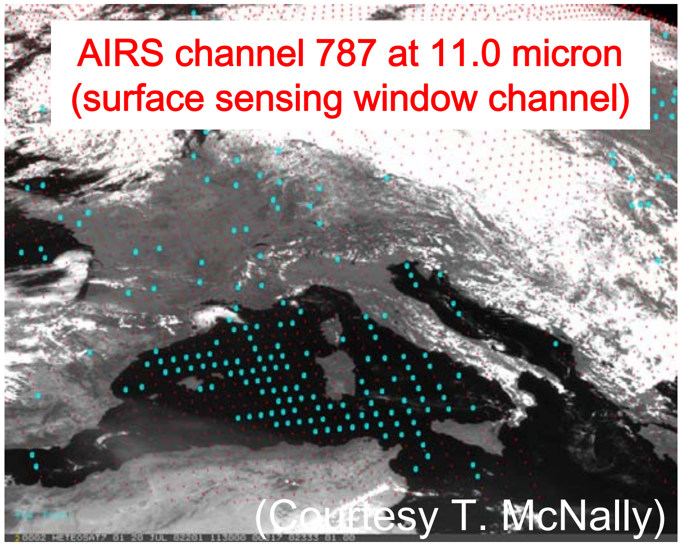


Vertically ranked channel index

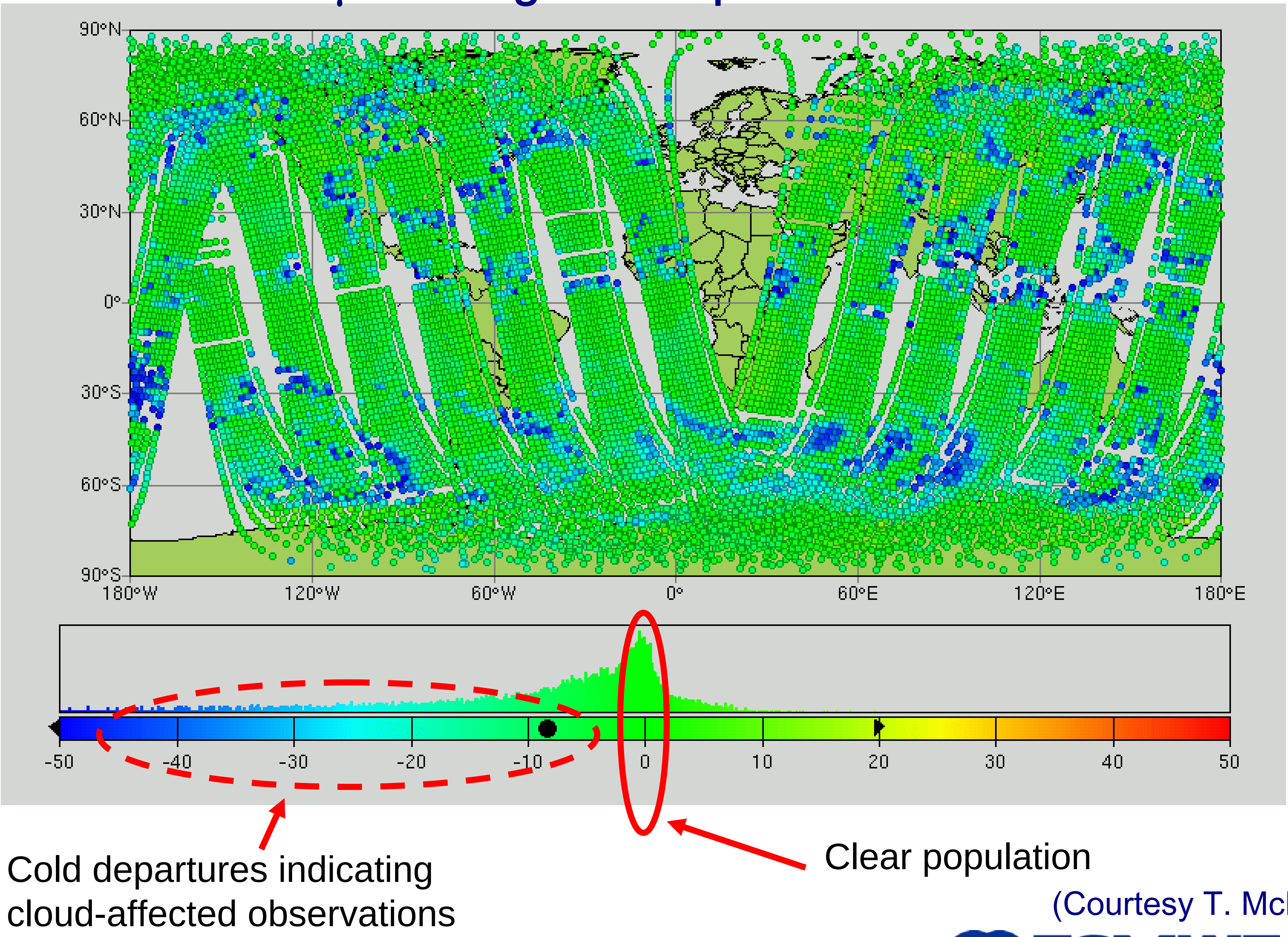
This identifies the characteristic signal of cloud in the data and allows contaminated channels to be rejected.





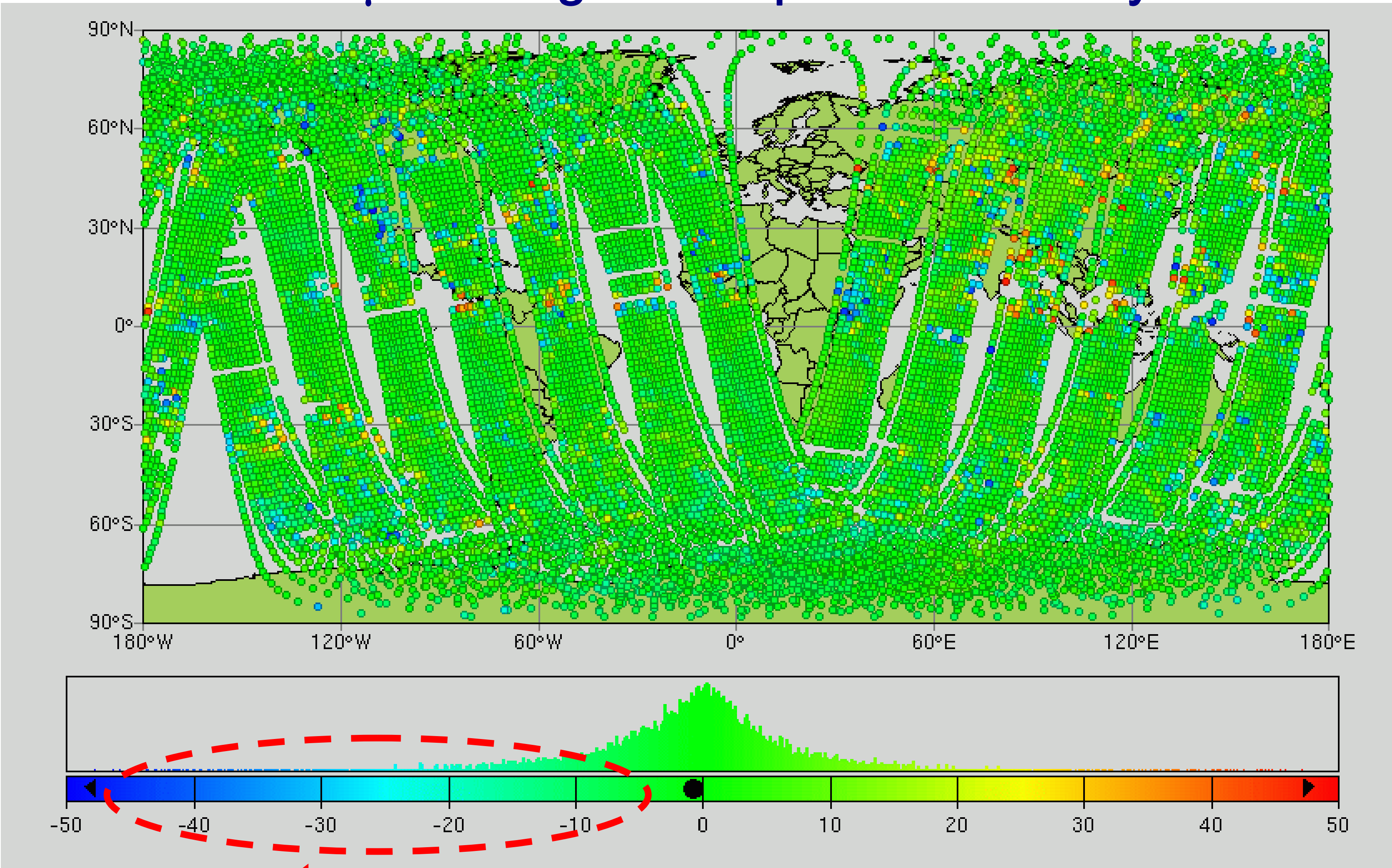


HIRS 11μm first-guess departures: clear skies



(Courtesy T. McNally)

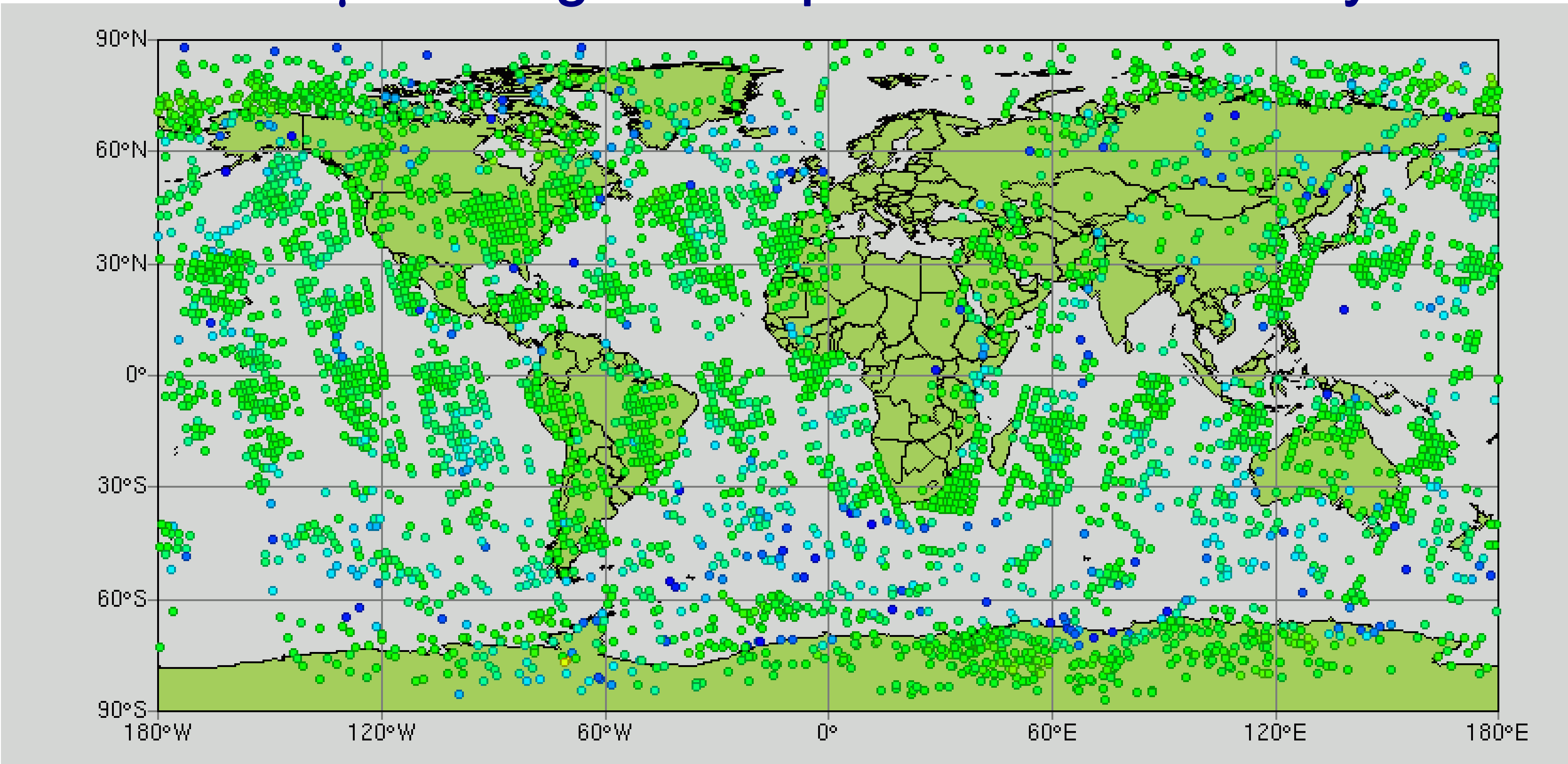
HIRS 11μm first-guess departures: cloudy skies



Many clouds with significant radiance signals are accurately represented by the NWP model and RT modelled !

(Courtesy T. McNally)

HIRS 11μm first-guess departures < 1K: cloudy skies

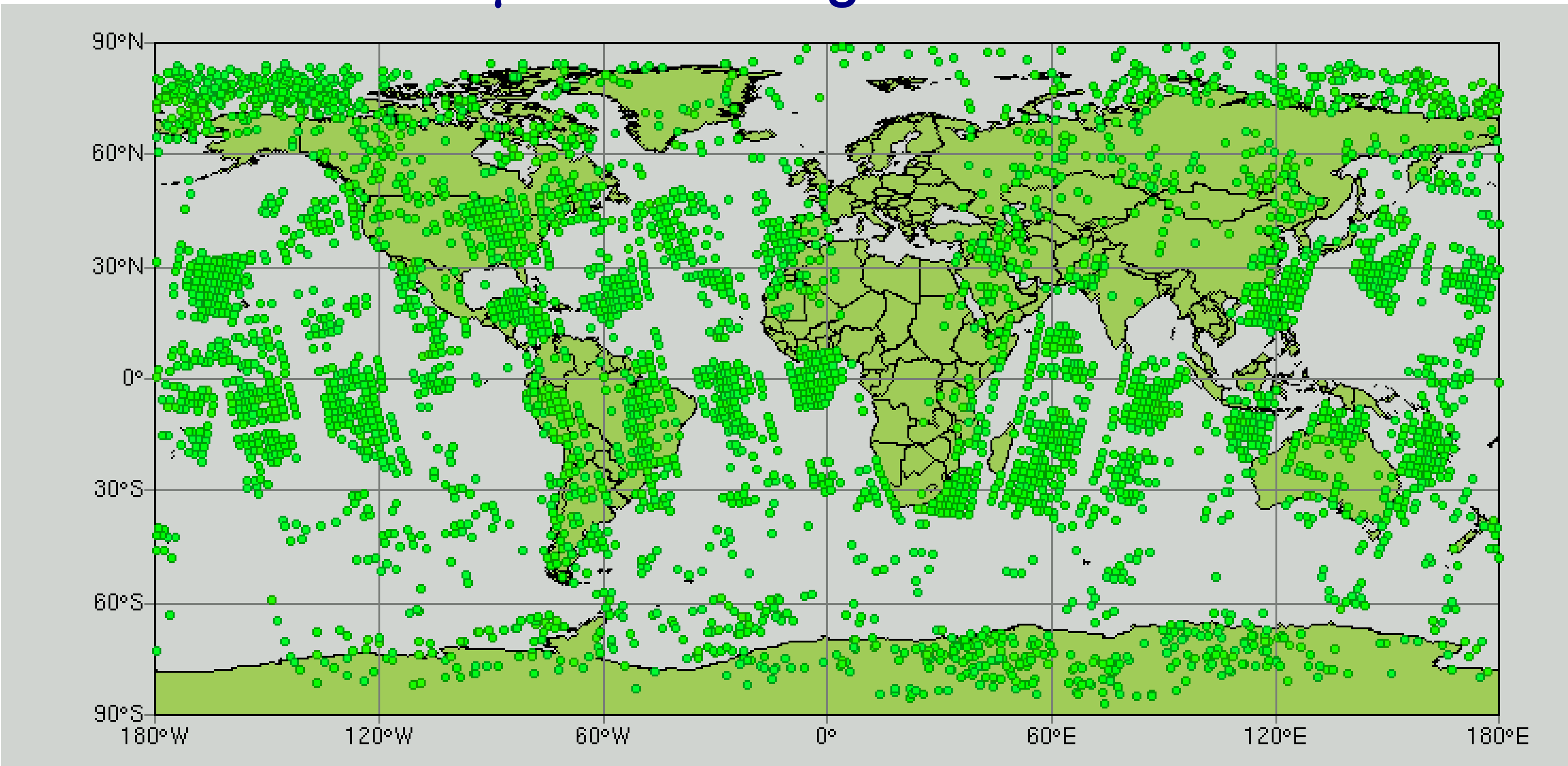


At these locations the cloud signal may still be large but we know *a priori* the NWP model clouds and the RT calculation are accurate (agreeing with the observation better than 1K):

1) Tangent linear approximation is likely to be valid
2) Background errors are not extreme
3) These locations out number completely clear locations

(Courtesy T. McNally)

HIRS 11μm data usage: clear skies



(Courtesy T. McNally)

Issues

- Predictability.
- Sensitivity of:
  - forecast to initial conditions in areas with clouds/precipitation;
  - precipitation forecast to initial conditions.
- Potential violation of Gaussian pdf (i.e. pdf-shape important; 0-value issue); radiance vs. product (radar rain rates, cloud optical depths) assimilation.
- Inversion problem may be underconstrained (adding clouds introduces too many degrees of freedom) and operators to inaccurate.
- Potential violation of linearity (sensitivity has strong dependence on state; multiple cost-function maxima).
- Potential issues related to definition of observation (+modelling) errors (e.g. cloud modelling uncertainties much larger than (T,q)-signal); representativeness i.e. model fields vs. observations. Bias definition (correction) more complex.
- More room for discrepancy between non-linear and linearized moist physics.

- 4D-Var system optimized for clear-sky observations:
  - choice of control variable and its behaviour near saturation (T, q, ξ, η, p_s);
  - forecast error calculation;
  - inner loop resolution and activation of physics (T95-no phys./159/255 vs. T799);
  - data sampling/representativeness (static thinning, ΔL_{FG}-driven screening).

Issues

Predictability

Sensitivity of forecast to initial conditions near clouds

Sensitivity of cloud/precipitation forecast to initial conditions

Sensitivity of forecast to observations in cloud-affected areas

Accuracy of observation operator

Similarity of linearized and non-linear parameterizations

Linearity of observation operator

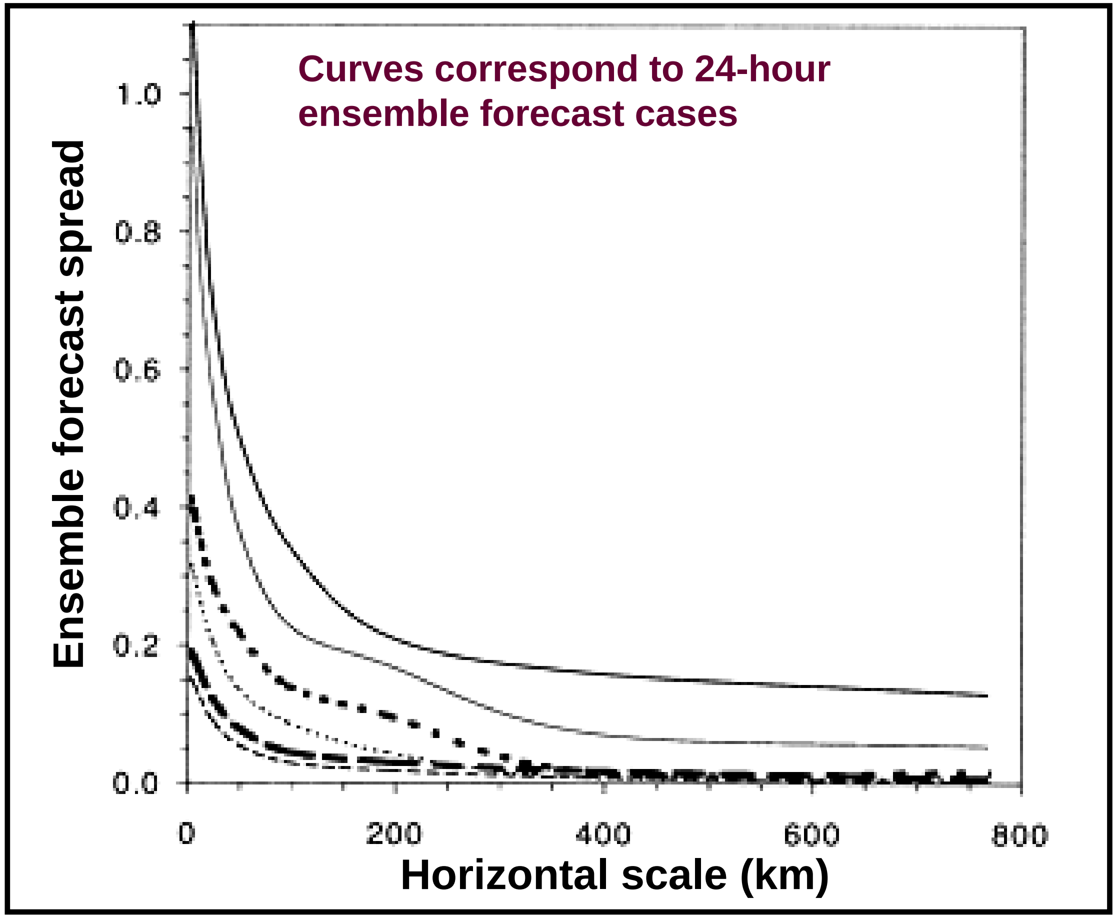
Definition of observation/background errors, biases

Choice of control variable

Predictability as a Function of Scale

Results from ensemble runs with the MC2 model (3 km resolution) over the Alps

$\sigma_{RR}/\langle RR \rangle$



→ Predictability of precipitation decreases dramatically for horizontal scales smaller than a few tens of kilometers.

(Walser et al. 2004; Courtesy P. Lopez)

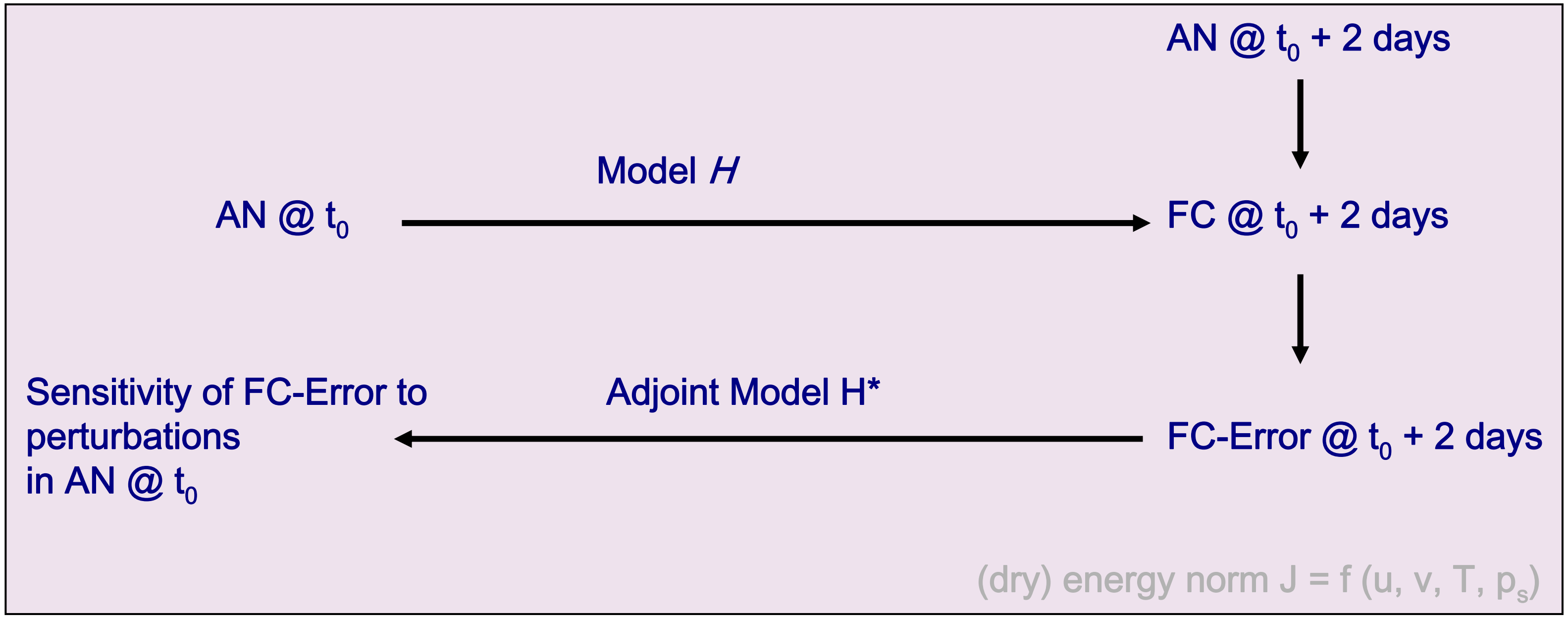
Issues

Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

Potential forecast sensitivity to initial conditions

Current systems produce rather good cloud/precipitation forecasts without assimilating any (!) direct precipitation or cloud observation.

However:
There are indications that key analysis errors occur in areas that are influenced by clouds and precipitation ⟶ Localization of error sensitivity to initial conditions.



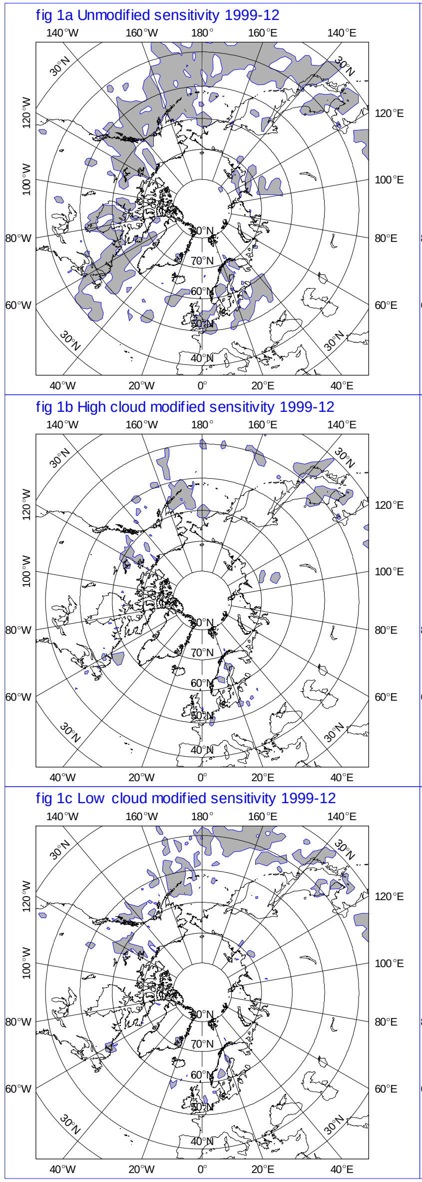
(Rabier et al. 1996, Klinker et al. 1998)

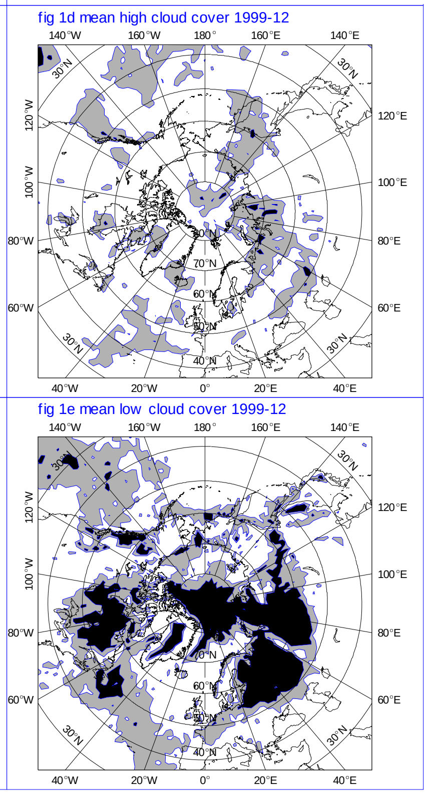
Potential forecast sensitivity to initial conditions

December 1999
600 hPa mean
T-perturbations
= sensitive areas

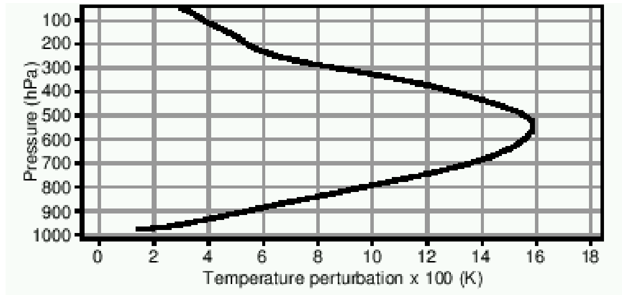
Perturbations
modified by
high cloud cover
= sensitivity scaled
with cloud cover
and height

Perturbations
modified by
low cloud cover





Mean vertical profile of T-perturbations



Mean *high* cloud cover
(clouds above 450 hPa)

Mean *low* cloud cover
(clouds below 800 hPa)

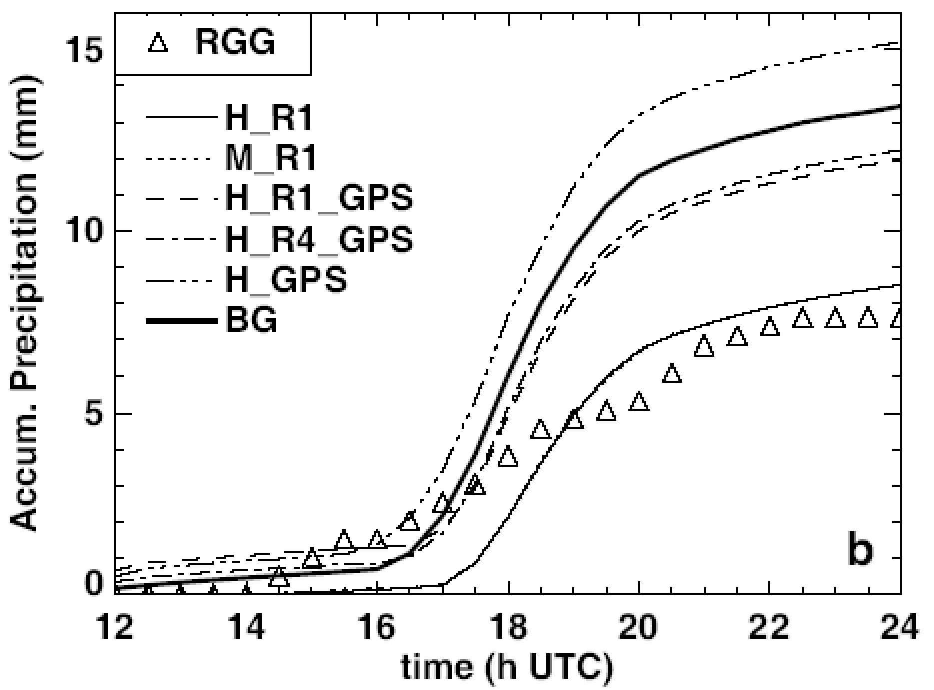
(McNally 2002)

Issues

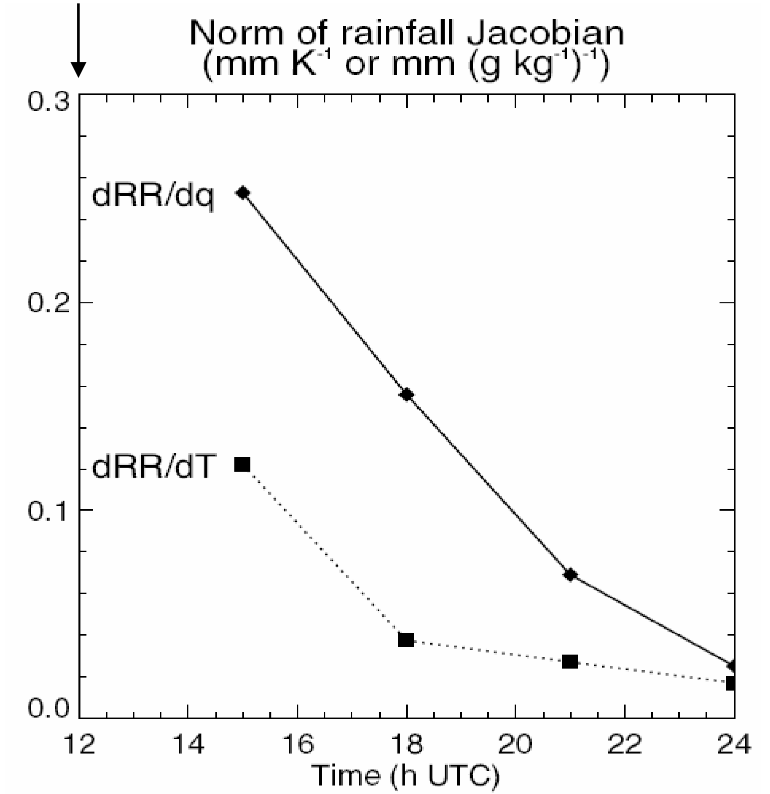
Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

How sensitive is precipitation forecast to initial conditions?

2D-Var assimilation of 3 or 12-hourly accumulated rainfall and ground-based GPS observations 24/10/2002, ARM-site.



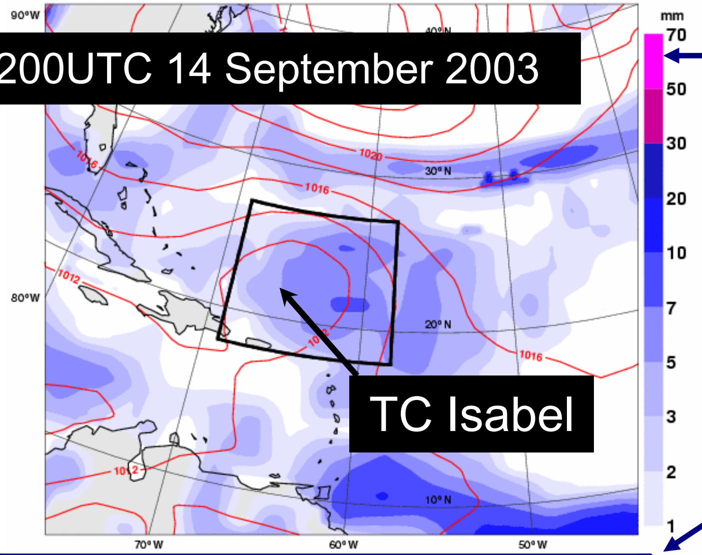
12-hour evolution of rainfall from background and various 2D-Var experiments



Sensitivity norm evolution of 3-hourly rain accumulation

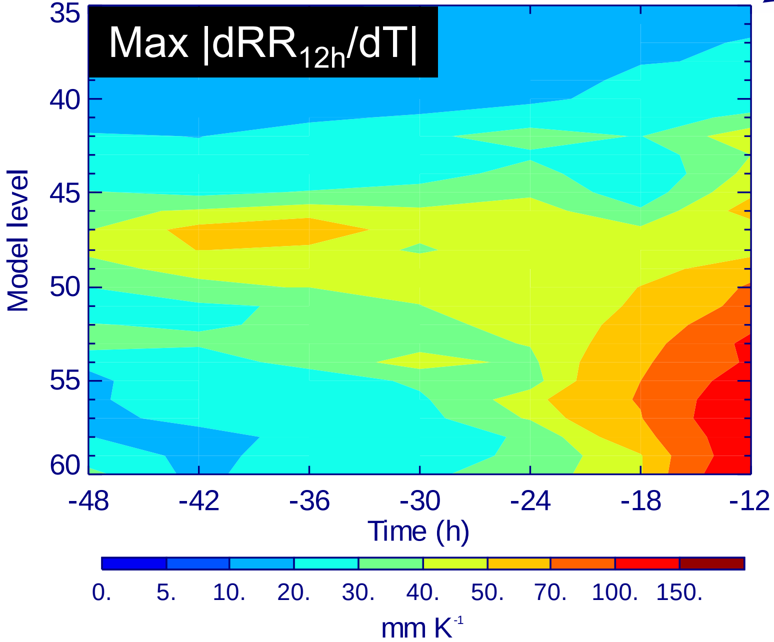
(Lopez et al. 2006)

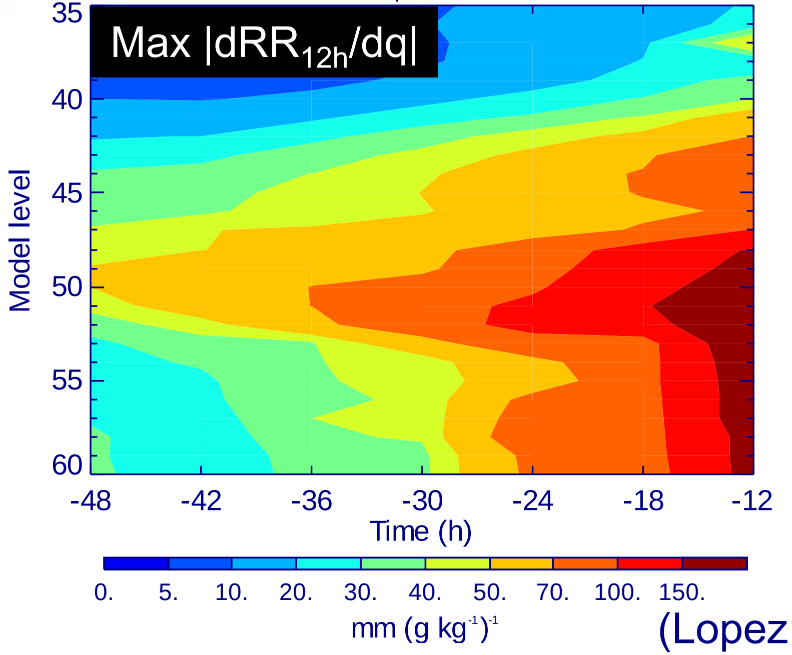
How sensitive is precipitation forecast to initial conditions?



12-hour surface rainfall (RR_{12h}) and MSLP
48-hour T95L60 forecast started at 1200UTC 12 September 2003.

Time evolution of the maximum 3D adjoint sensitivities of RR_{12h} to T (left) and q (right) for lead times of up to 36h prior to precipitation accumulation.

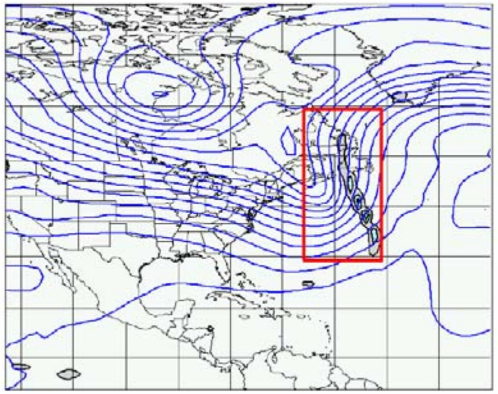




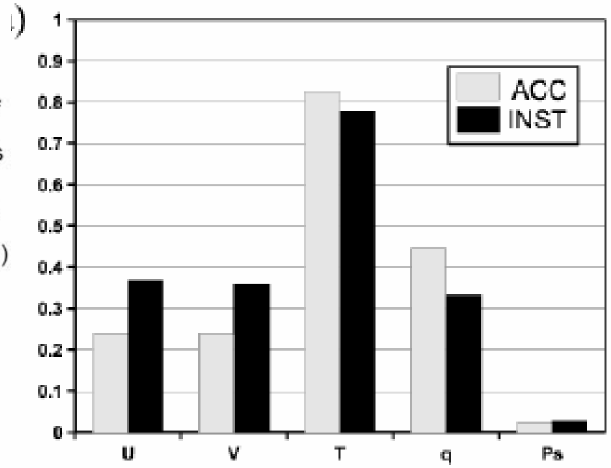
(Lopez et al. 2006)

How sensitive is precipitation forecast to initial conditions?

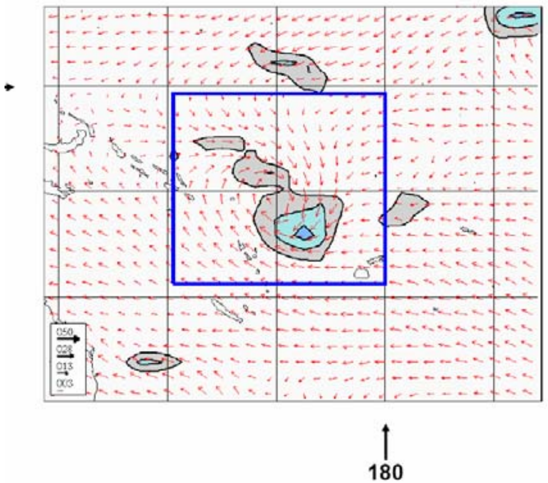
Mid-latitude cyclone
27 January 2003 12 UTC



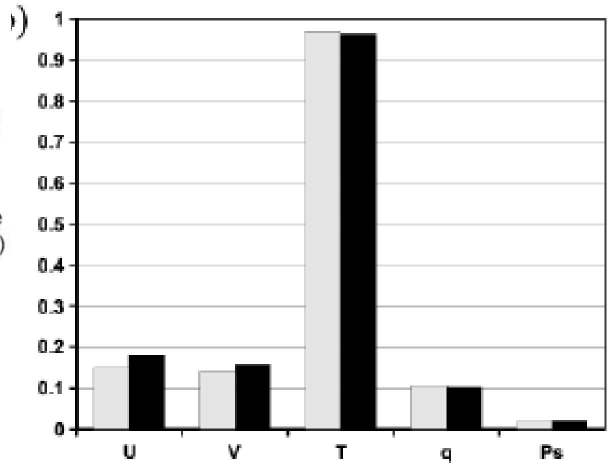
24-h forecast of
500 hPa heights
+
total pcpn rate
(contour 2 mm/h)



ZOE Typhoon
26 december 2002 00 UTC



24-h forecast of
winds at 10 m
+
precipitation rate
(contour 2 mm/h)



GEM model, 4D-Var:

24-hour adjoint sensitivity for deriving optimal perturbations (u, v, T,q) divided by total perturbations that maximize instantaneous (INST) or accumulated (ACC) precipitation

(Mahfouf and Bilodeau 2006, 2007)

Issues

Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

Forecast sensitivity to observations in analysis

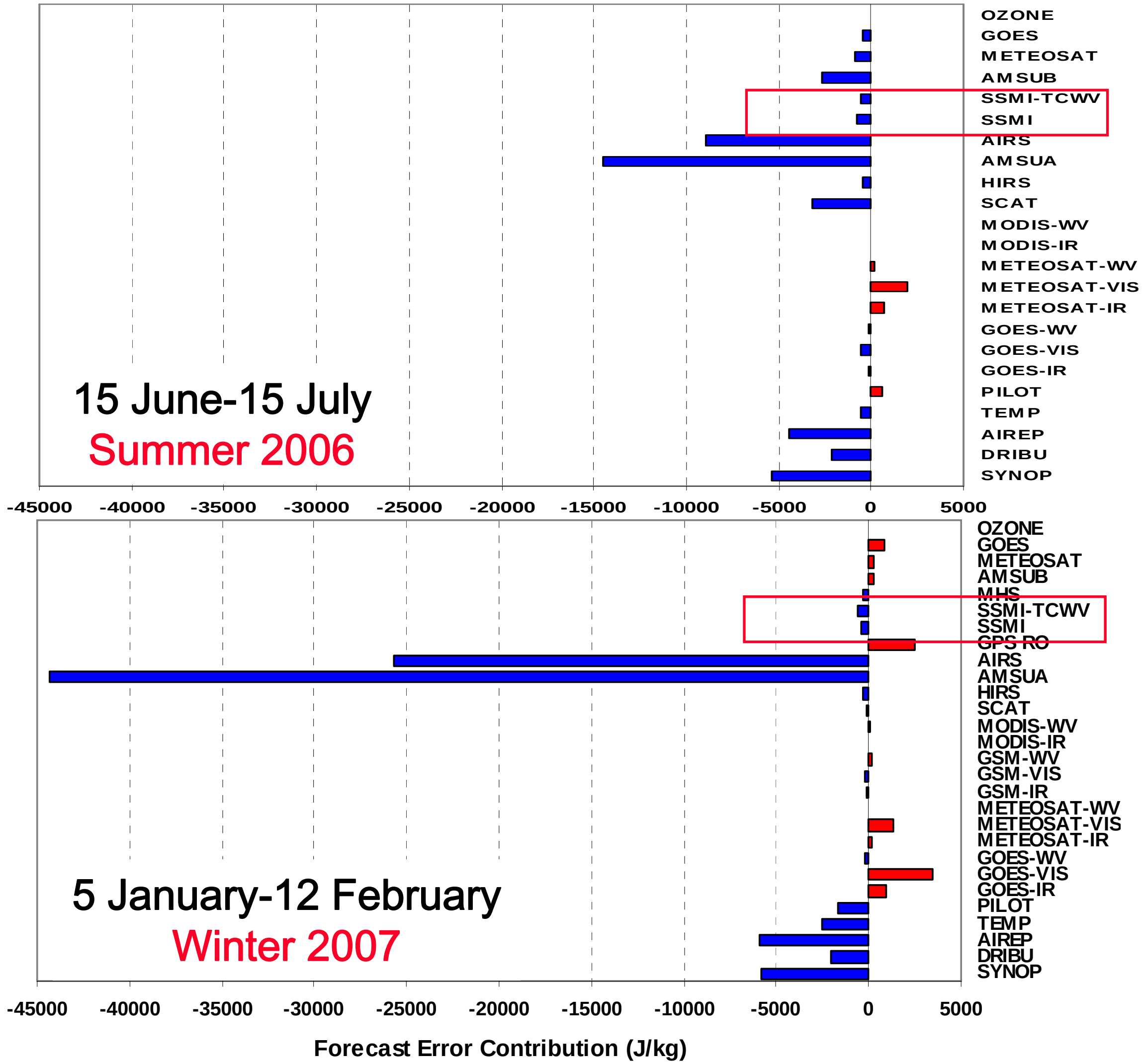
$$\delta J = \frac{\partial J}{\partial \mathbf{y}}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

(Baker and Daley)

Here:

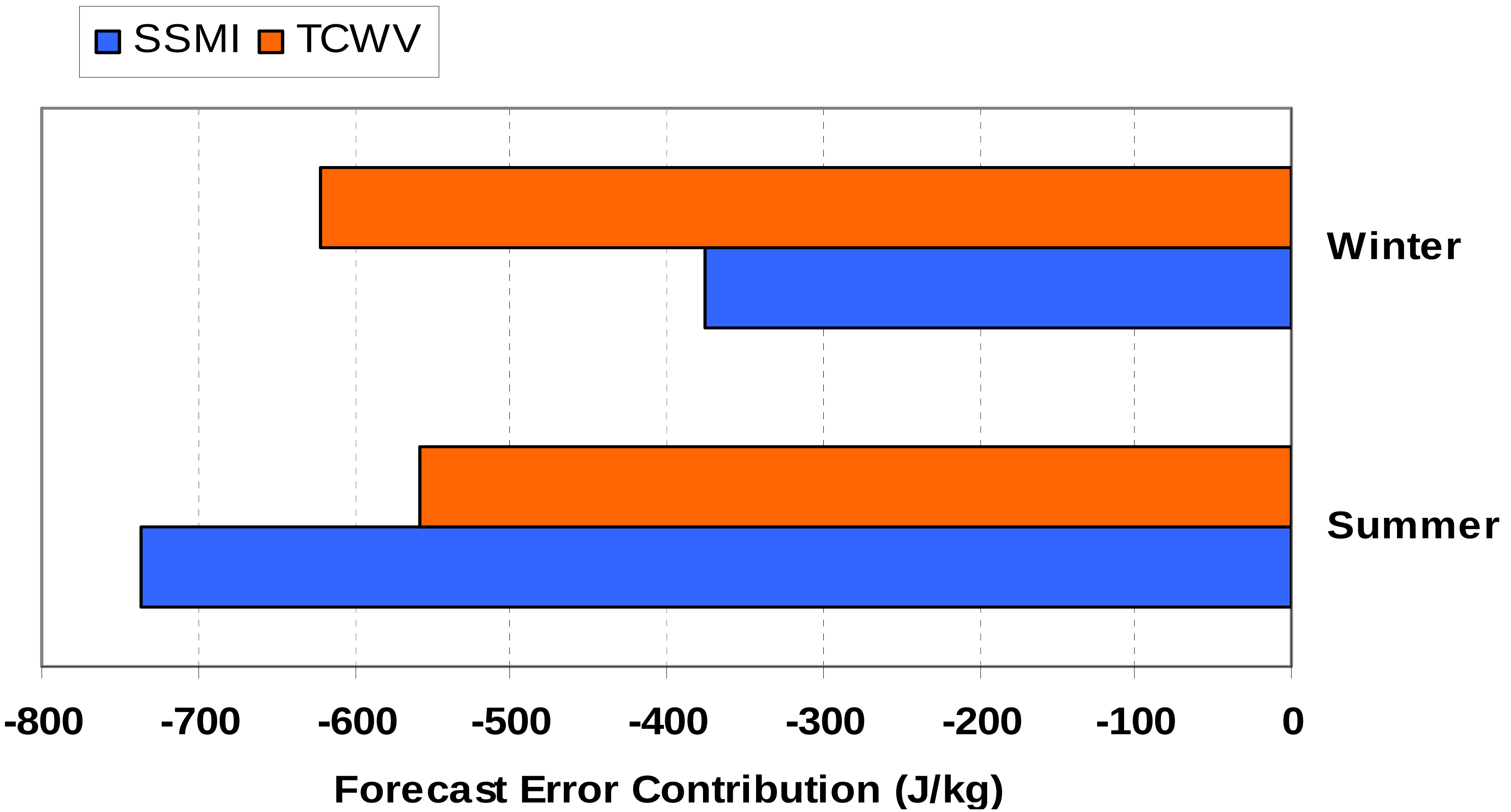
- applied to OSE experiments for winter and summer period;
- T511L60, IFS cycle 31R2;
- 24h FC-AN;
- dry energy norm but moist model AD.

(Courtesy C. Cardinali)



Forecast sensitivity to observations in analysis

Impact of 'similar' observation type in clear-skies vs. clouds/rain



(Courtesy C. Cardinali)

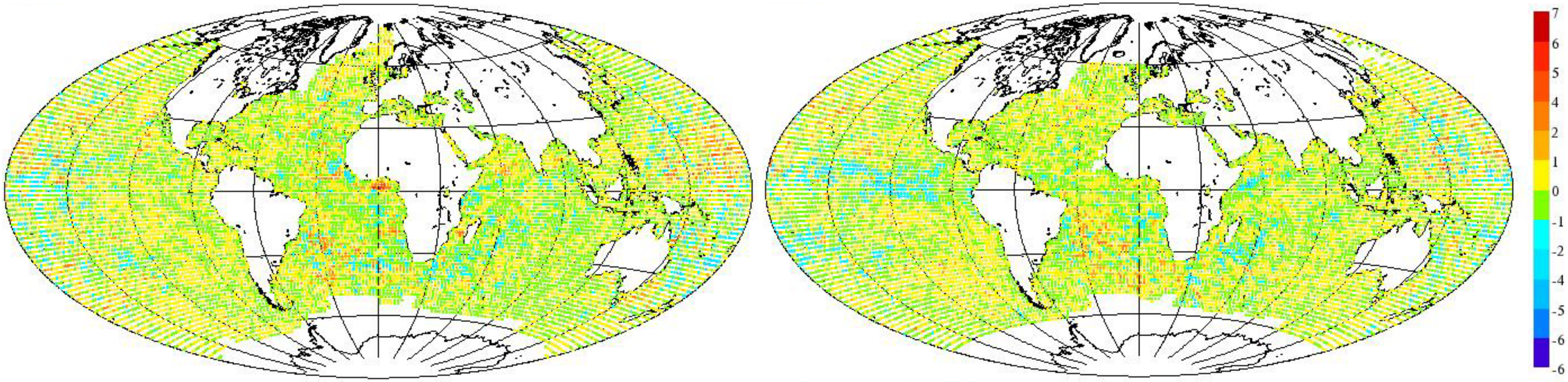
ECMWF

Forecast sensitivity to observations in analysis

CY32R1, T511L60, 20070105-20070212

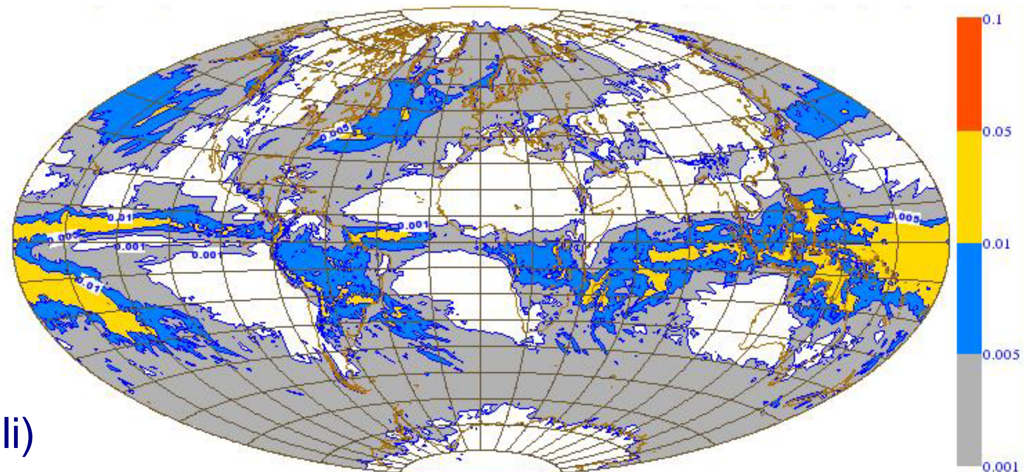
SSM/I clear-sky, winter

SSM/I clouds/rain, winter



Mean 36-12h precipitation forecast initialized at 12 UTC

[J/kg]



(Courtesy C. Cardinali)

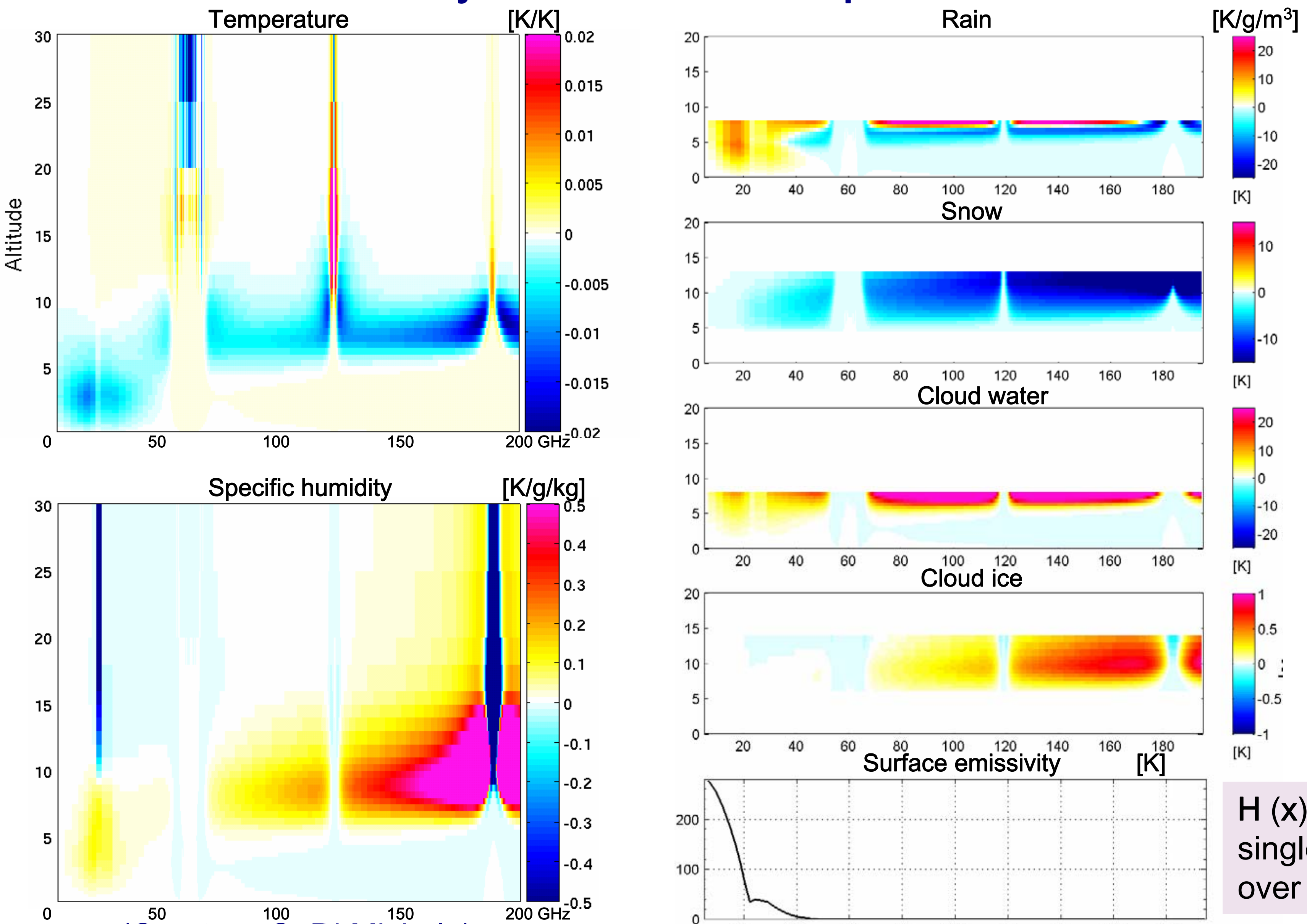
[10^{-3} mm]

ECMWF

Issues

Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

Sensitivity of observation operator – MW

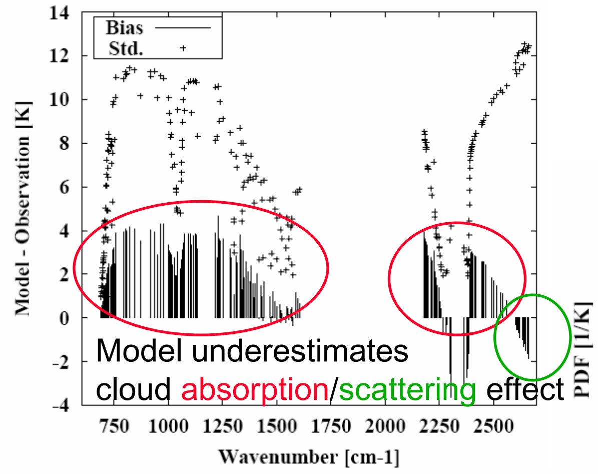


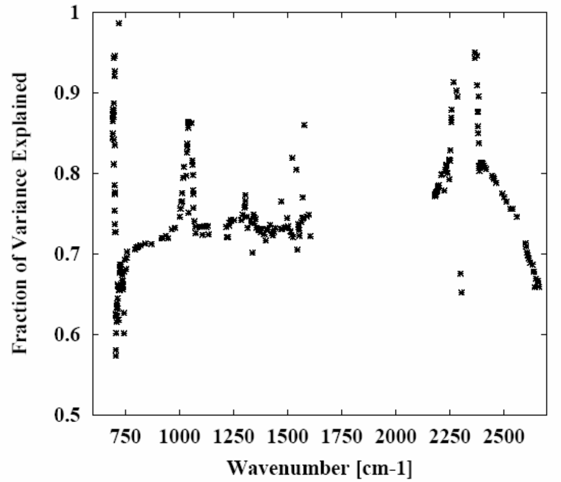
H (x)
single profile
over ocean

(Courtesy S. Di Michele)

Accuracy of observation operator – IR

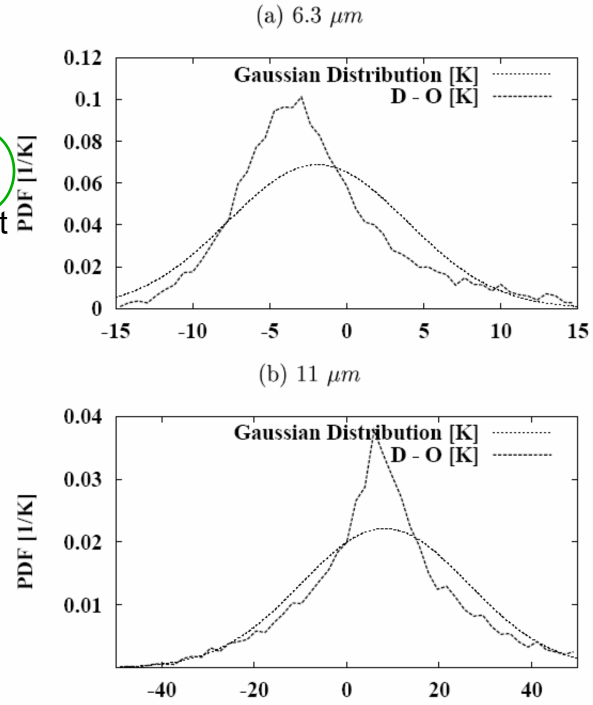
AQUA AIRS 30/11/2002





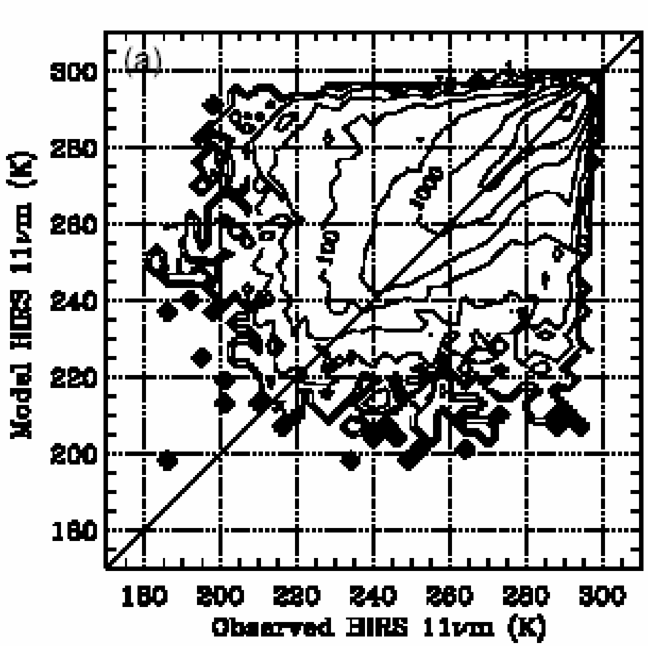
f = (Var[y]-Var[*H*(x)-y])/Var[y]
i.e. f = 1 for perfect correlation

Meteosat-7 MVIRI 30/11/2002



11μm channel produces larger departures due to less water vapour absorption

NOAA-15 HIRS channel 12 15/03/2001



(Chevallier et al. 2002, 2003)

ECMWF

Accuracy of observation operator – MW

DMSP F-13/14 SSM/I channels 1-5 09/2004

clouds+rain | clear-skies

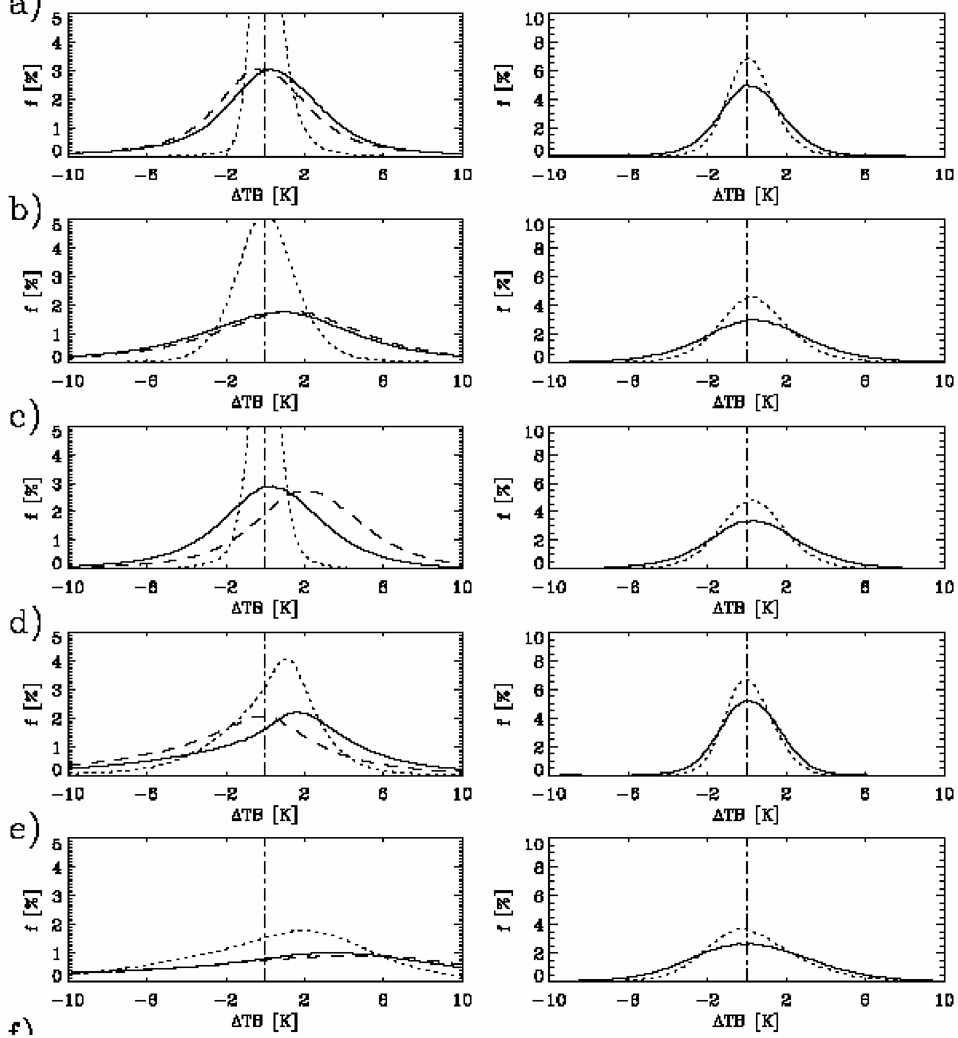
19v

19h

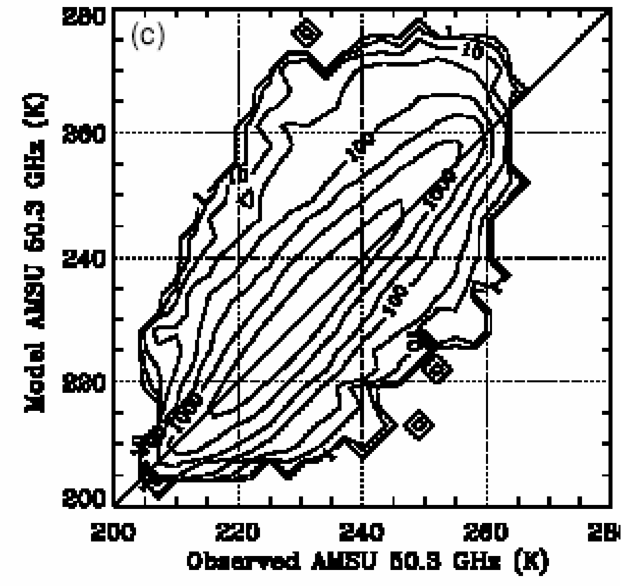
22v

37v

37h



NOAA-15 AMSU-A channel 3 15/03/2001

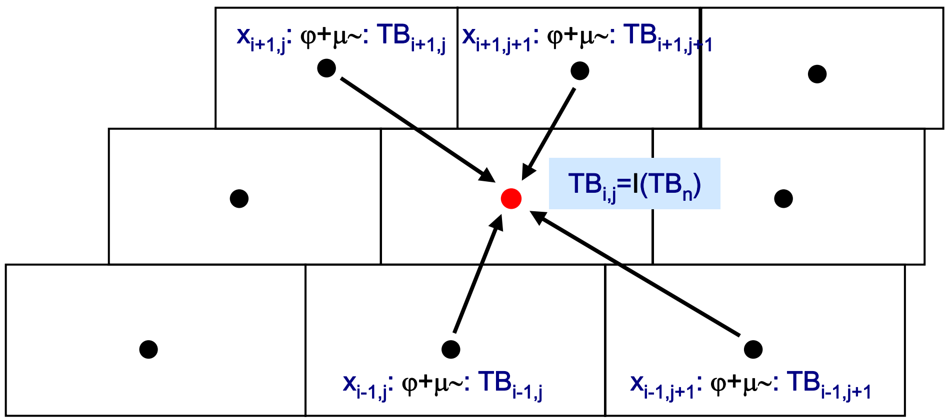


(Bauer et al. 2006,
Chevallier et al. 2002)

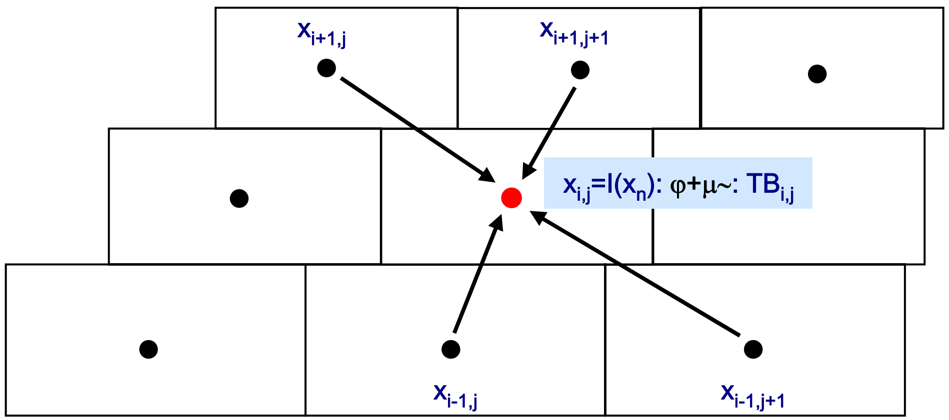
ECMWF

Options for grid-point – observation treatment

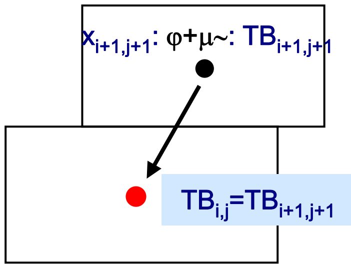
Option 1: apply operator and interpolate output = IO



Option 2: interpolate input and apply operator = II



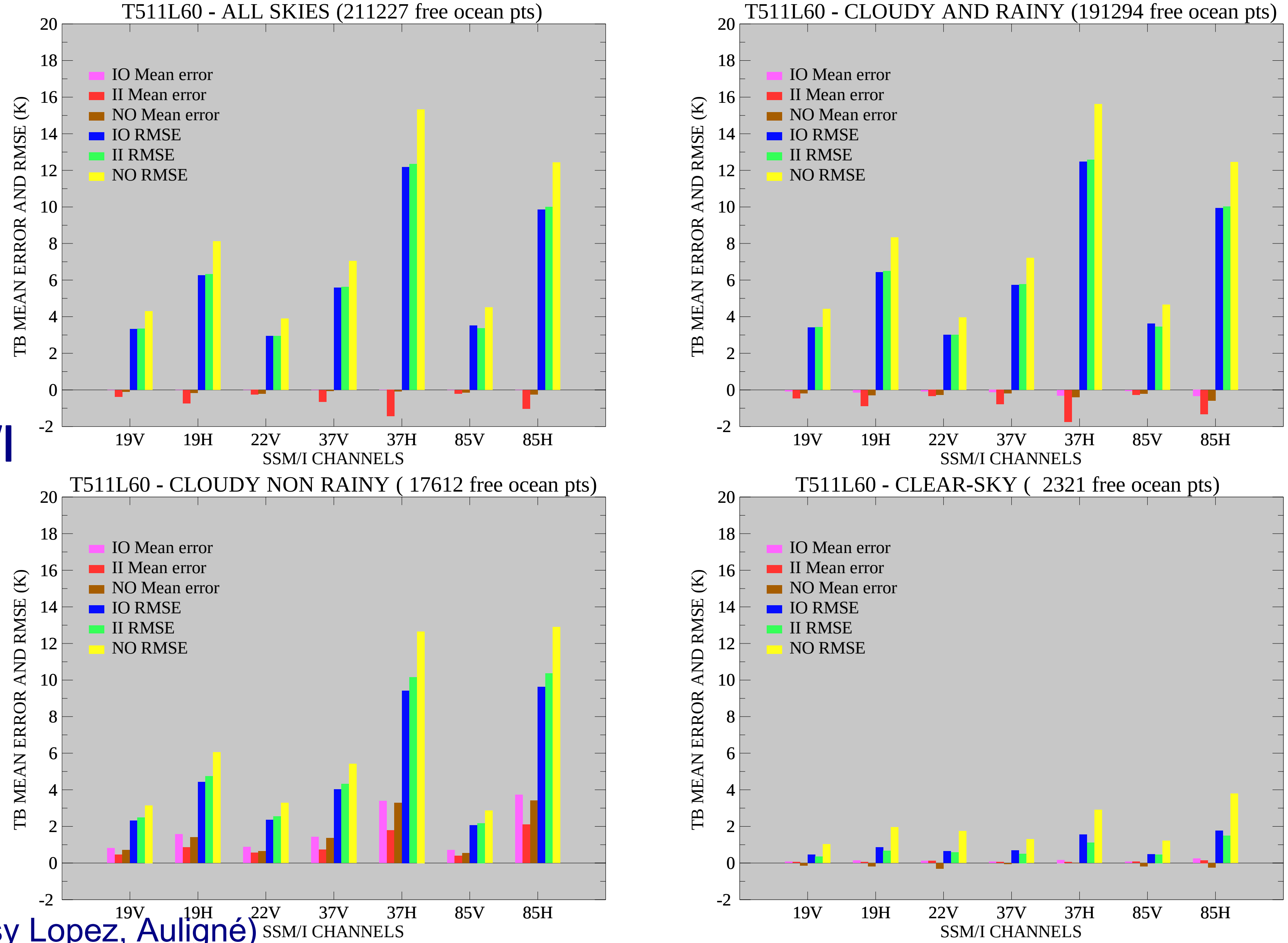
Option 3: take nearest output = NO



Operator: φ+μ~ = moist physics + microwave RT

4D-Var rain/cloud treatment of grid-point – observation problem

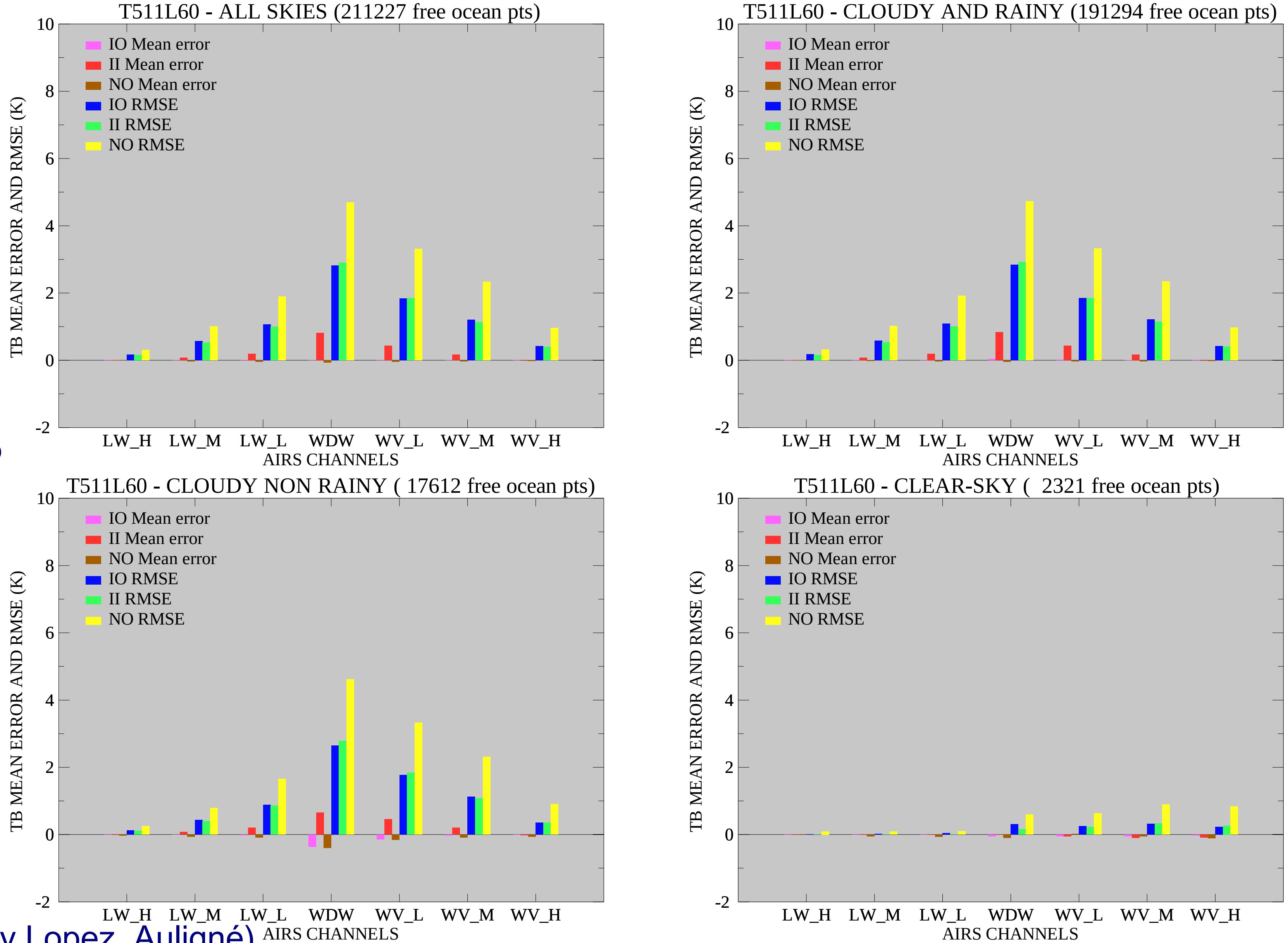
SSM/I T511



(Courtesy Lopez, Aulígné)

4D-Var rain/cloud treatment of grid-point – observation problem

AIRS
T511

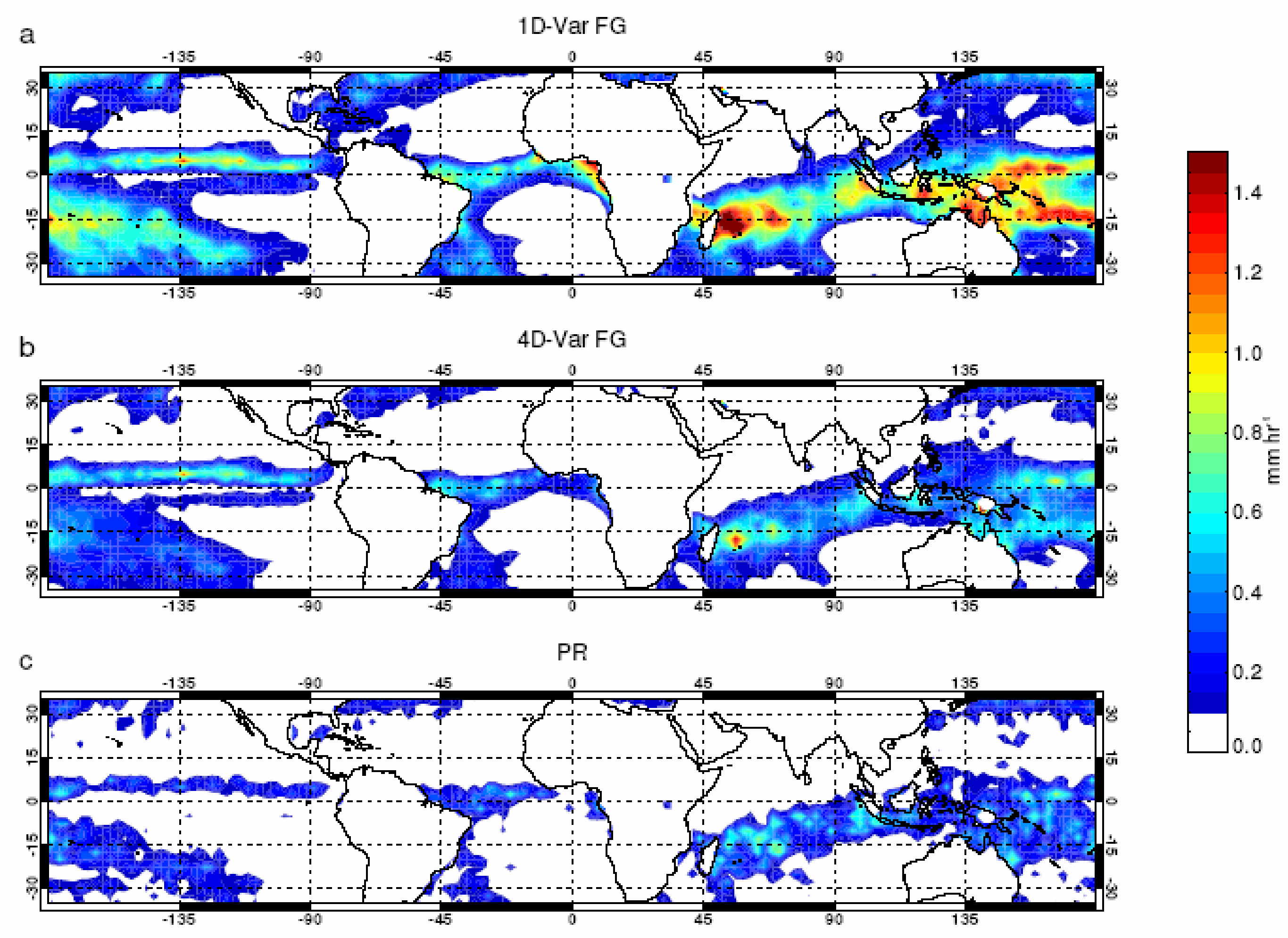


(Courtesy Lopez, Auligné)

Issues

Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

Non-linear vs. linearized moist physics



Mean surface rain rate (mm hr^{-1}) for February 2007

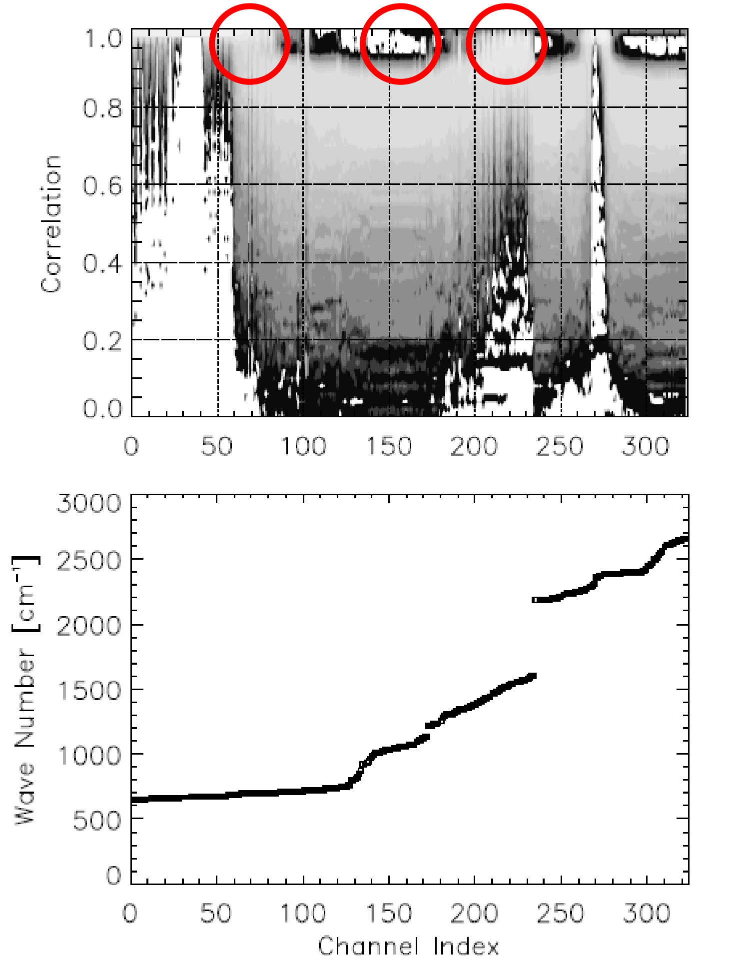
(Geer et al. 2007)

Issues

Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

Accuracy/linearity of observation operator - IR

$$M(\mathbf{x}_b + \delta\mathbf{x}) - M(\mathbf{x}_b) \text{ vs. } \mathbf{M}\delta\mathbf{x}, \quad \delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$$



Observation operator:
diagnostic cloud scheme + non-scattering RTTOV-cloud

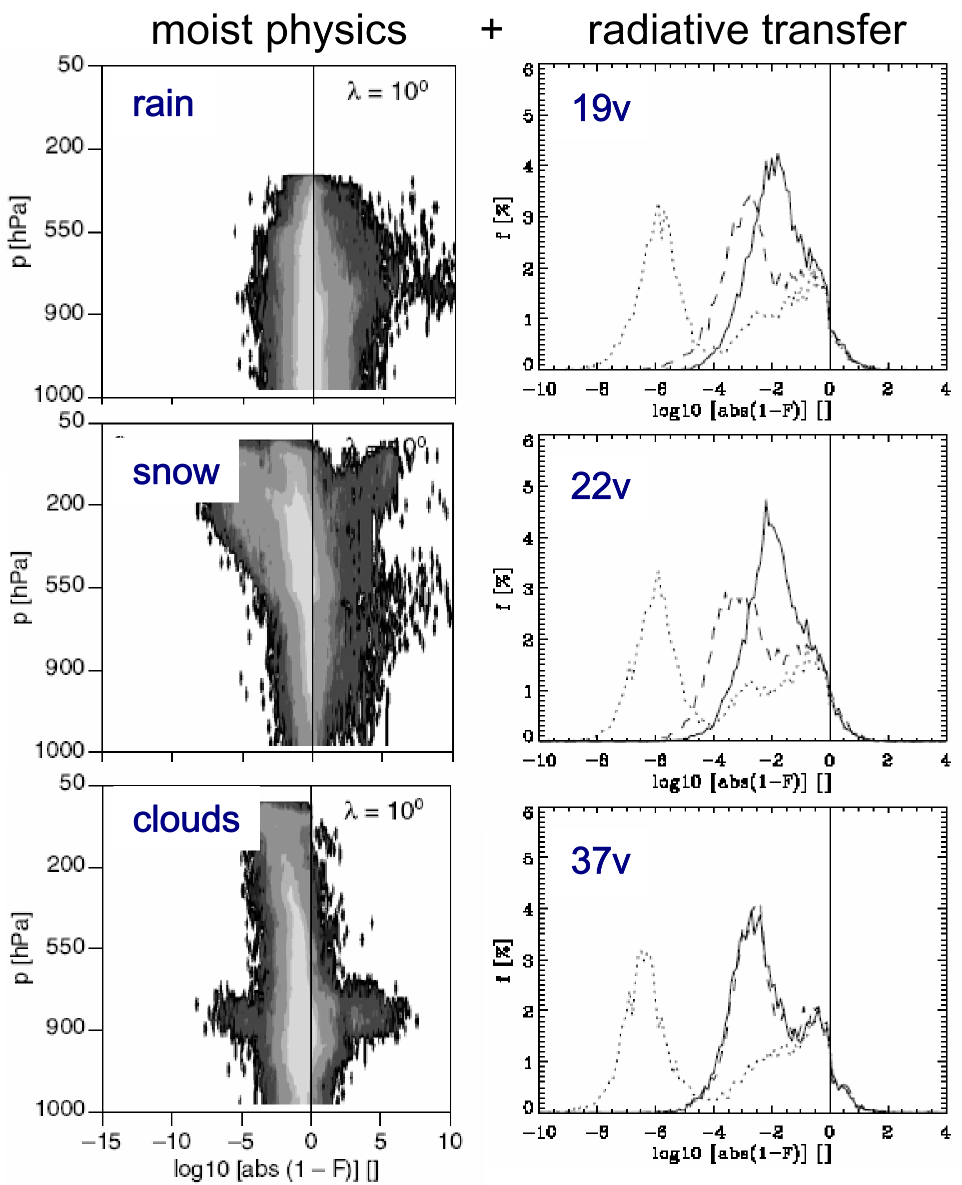
Experiment:
1 analysis cycle worth of profiles over Meteosat-7 disk
100 perturbations, $\delta\mathbf{x}$, per profile (perturbations represent **B**)
Correlations between finite difference and tangent-linear

For selection criteria:
Sensitivity i.e. cloud impact > 0.5 K
Accuracy i.e. FG-departures < 6 K
Linearity i.e. correlation > 0.85

42 of 324 AIRS channels seem suited for assimilation:
13 channels near 14.3μm (see also HIRS, IASI)
22 channels near 6.3 μm (see also SEVIRI, HIRS, IASI)
7 channels near 4.5 μm (see also HIRS, IASI)

(Chevallier et al. 2003)

Accuracy/linearity of observation operator - MW



Observation operator:
diagnostic cloud+convection scheme + scattering RTTOV-cloud

Experiment:
1 analysis cycle worth of global profiles
Perturbations: $\delta \mathbf{x} = \mathbf{x}_a - \mathbf{x}_b$

Deviation factor: $F = \dfrac{H(\mathbf{x} + \lambda\delta\mathbf{x}) - H(\mathbf{x})}{\lambda \mathbf{H}(\delta\mathbf{x})}$

$\log_{10}(|1-F|) = -1$: 10% deviation from linearity
$\log_{10}(|1-F|) = -2$: 1% deviation from linearity

3 of 7 SSM/I channels seem suited for assimilation:
19.35 (v, h) and 22.235 (h) GHz

(Bauer et al. 2006)

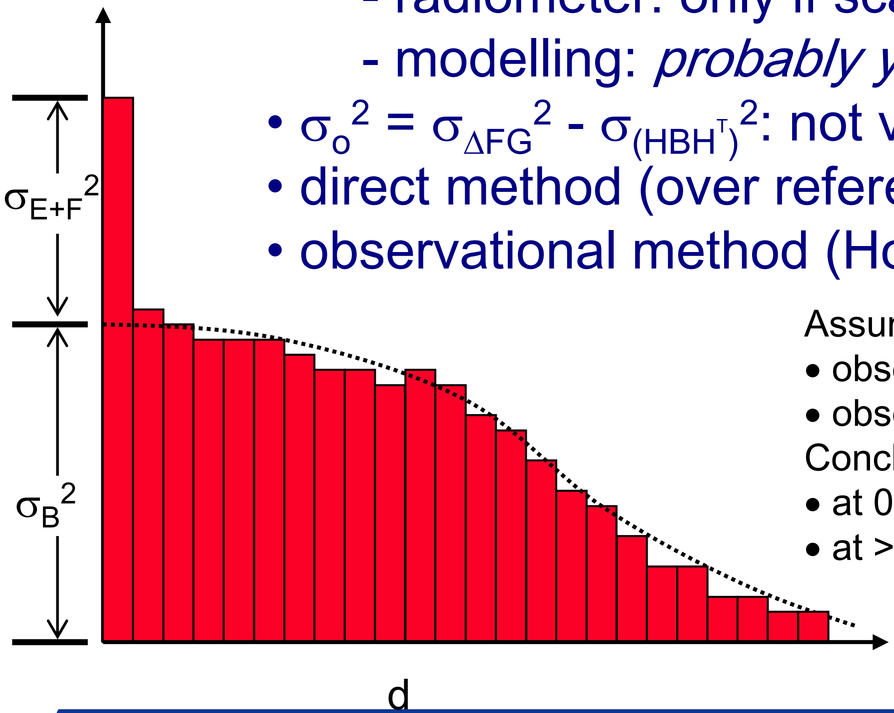
Issues

Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

Instrument + observation operator errors

Var observation errors in TB-space:

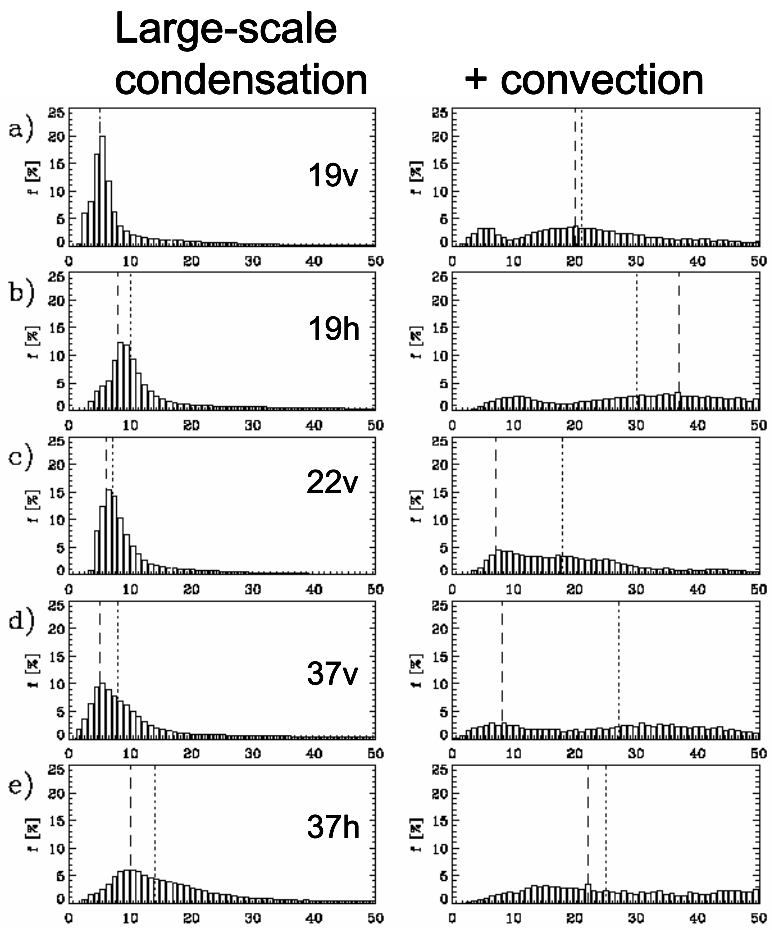
- radiometric noise (~ 1K) << observation operator error
- error correlation between channels:
  - radiometer: probably no
  - modelling: *probably yes*
- spatial correlation:
  - radiometer: only if scanwise calibration erroneous
  - modelling: *probably yes*
- $\sigma_o^2 = \sigma_{\Delta FG}^2 - \sigma_{(HBH^T)}^2$: not valid (often negative b/c **B** inappropriate)
- direct method (over reference sites): not representative
- observational method (Hollingsworth-Lönnberg):



Assumptions:
- observation network is rather dense
- observations are spatially uncorrelated (and discrete)

Conclusions:
- at 0-separation distance, the variance is $\sigma_B^2 + \sigma_{E+F}^2$
- at >0-separation distance, the variance is covariance σ_B^2 (d)

Background errors in radiance space - Observation errors?



σ_{HBH^T}

	19v	19h	22v	37v	37h	85v	85h
σ_B	2.2	4.1	2.0	3.5	7.4	2.8	7.7
σ_R	2.8	5.2	2.5	4.4	9.0	4.0	8.4
Median: σ_{HBH^T}, LS	5	10	7	8	14	7	16
Median: σ_{HBH^T}, LS+CV	21	30	18	27	25	12	22
Mode: σ_{HBH^T}, LS	5	8	6	5	10	5	12
Mode: σ_{HBH^T}, LS+CV	20	37	7	8	22	4	16

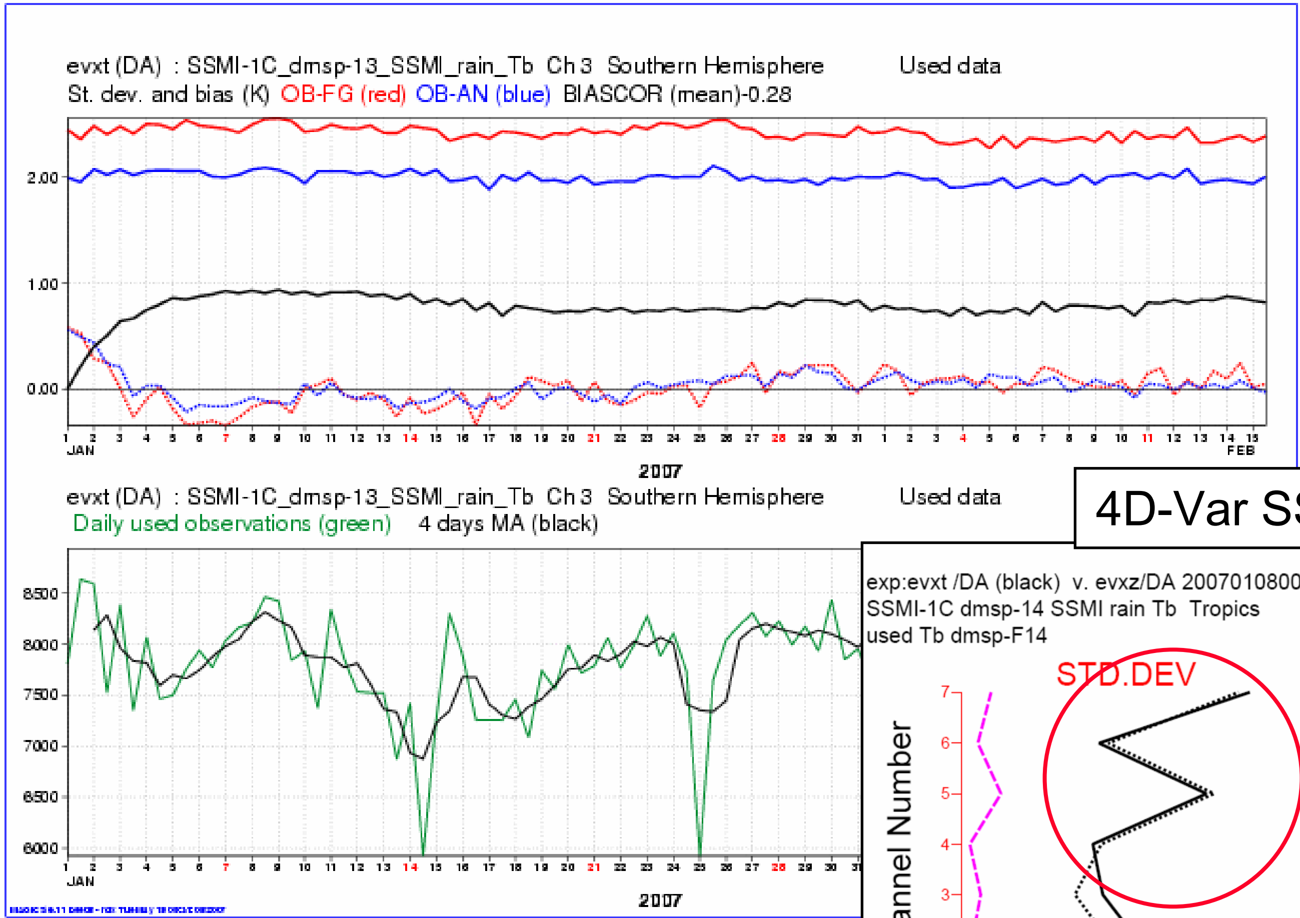
(Bauer et al. 2006)

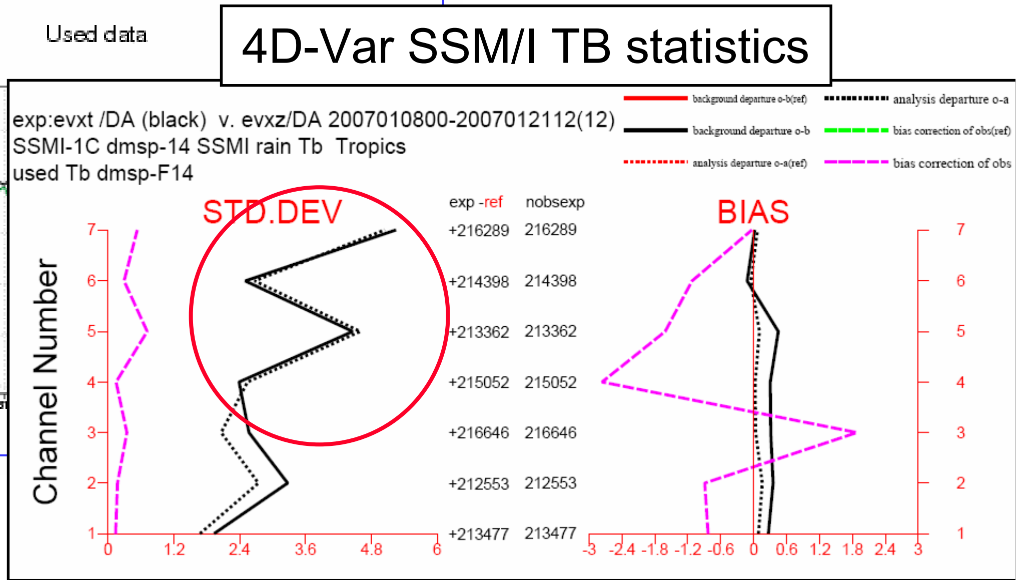
SSM/I channel frequency (GHz)	Three-channel dataset equivalent error (K) No. = 1414		Five-channel dataset equivalent error (K): Rainy profiles No. = 2137		Five-channel dataset equivalent error (K): Nonrainy profiles No. = 961		Obs error (K) specified for 1DVAR T_b LOE expt
	Mean	Median	Mean	Median	Mean	Median	
19 V	24.01	22.44	24.27	23.11	10.68	8.62	20.
19 H	44.90	41.66	45.97	44.07	29.60	24.69	40.
22 V	10.21	8.29	14.89	15.03	7.81	6.18	9.
37 V			37.73	35.46	24.42	20.30	30.
37 H			76.05	70.03	57.63	49.73	60.

(Deblonde et al. 2007)

Biases: 4D-Var SSM/I rain-affected radiance assimilation

22v TB-departure and bias correction evolution in variational bias-correction system
(skin temperature, total column water vapour, 10m wind speed, zenith angle)





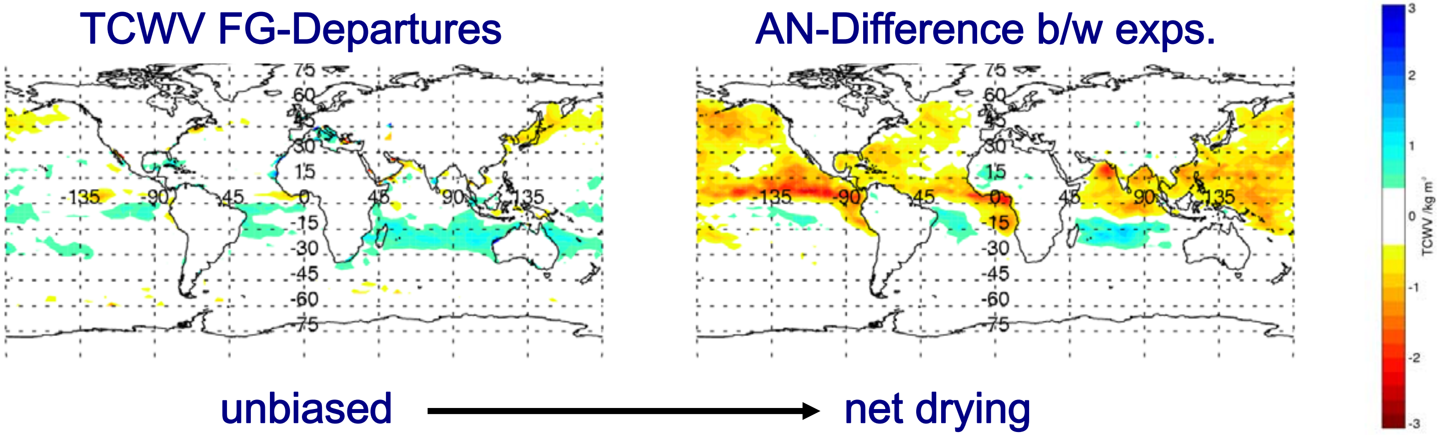
ECMWF

Issues

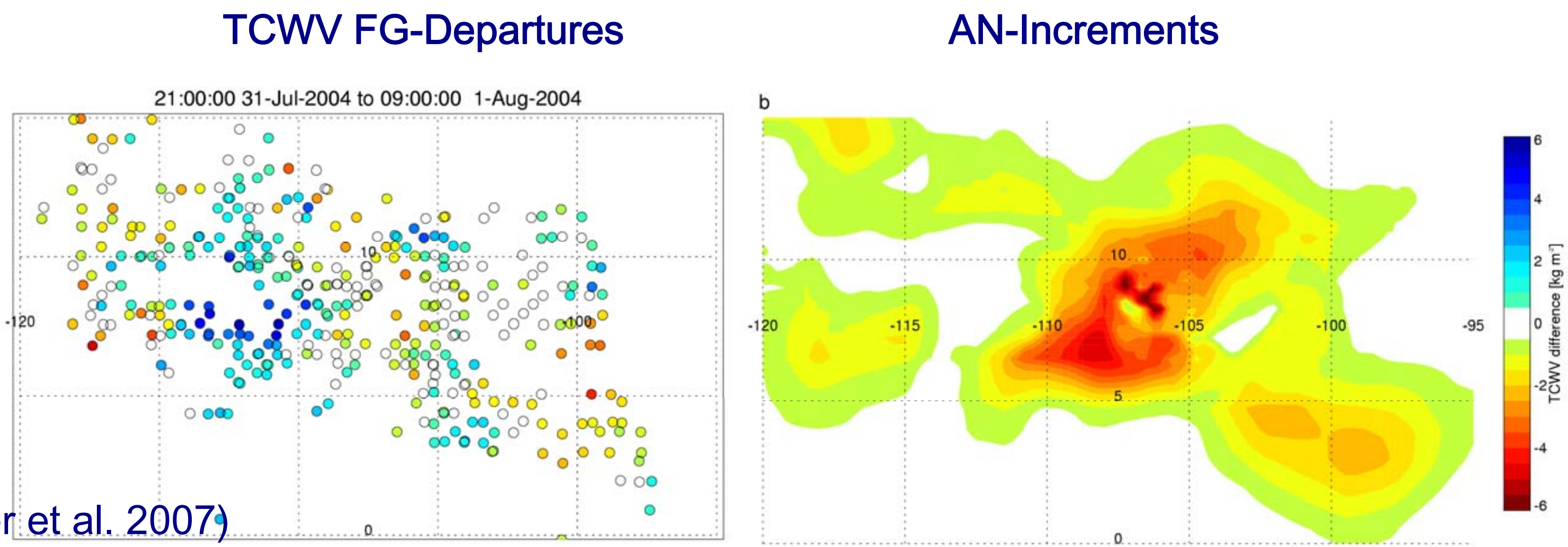
Predictability
Sensitivity of forecast to initial conditions near clouds
Sensitivity of cloud/precipitation forecast to initial conditions
Sensitivity of forecast to observations in cloud-affected areas
Accuracy of observation operator
Similarity of linearized and non-linear parameterizations
Linearity of observation operator
Definition of observation/background errors, biases
Choice of control variable

4D-Var analysis using 1D-Var TCWV (SSM/I radiances)

Mean August 2004 TCWV difference:



Example: 1st cycle ITCZ East Pacific:

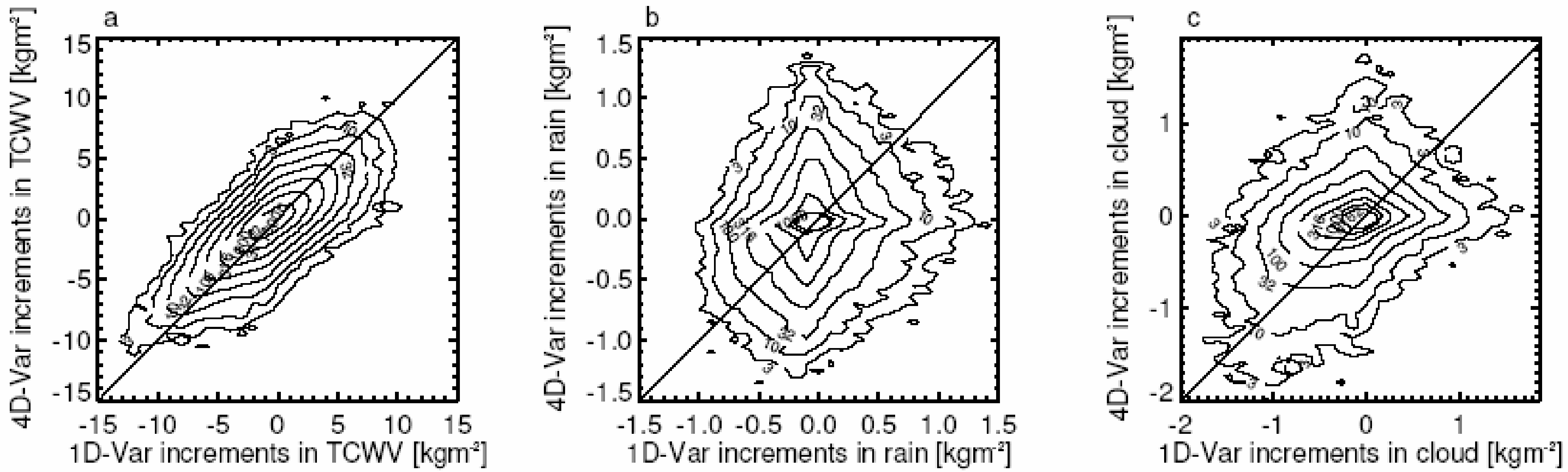


(Geer et al. 2007)

ECMWF

4D-Var analysis using 1D-Var TCWV (SSM/I radiances)

Total column water vapour | Total column rain | Total column cloud water

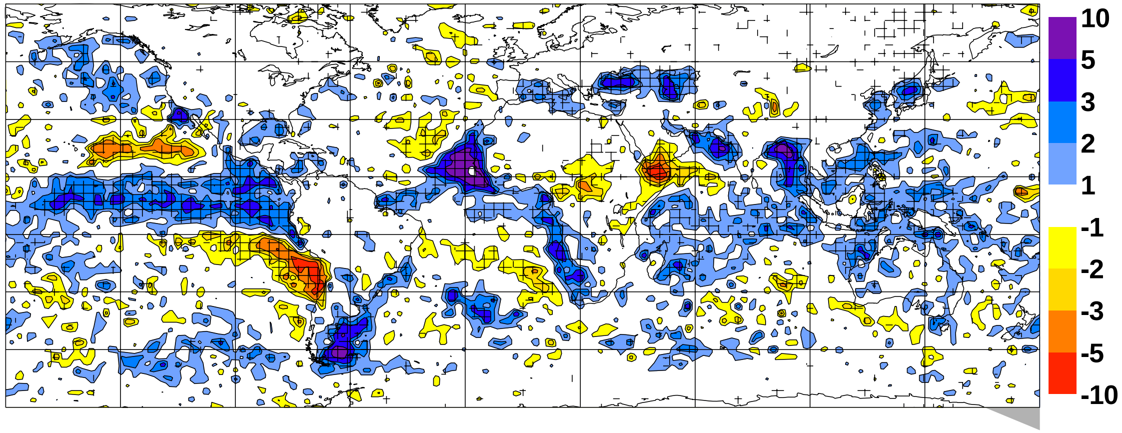


- 4D-Var integrates model in space/time.
- Other observations drive 4D-Var analysis.
- Choice of 1D+4D and moist control variable reduces impact.

(Geer et al. 2007)

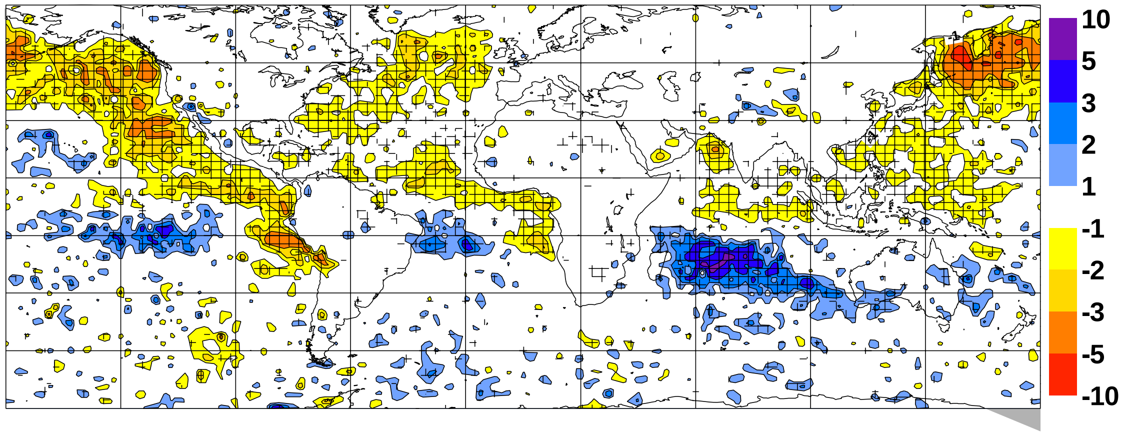
1D+4D vs. 4D-Var assimilation of rain-affected radiances: Difference in TCWV analysis

Mean TCWV difference [%]



4D-Var - 1D+4D-Var
4D-Var

(20070108-2070214)



1D+4D-Var - NORAIN
1D+4D-Var

(20060701-20060808)

Conclusions

Issues

Predictability

Reevaluate in ECMWF model context.

Sensitivity of forecast to initial conditions near clouds

Sensitivity of cloud/precipitation forecast to initial conditions

Favour 4D-Var to maximize balanced sensitivity.

Sensitivity of forecast to observations in cloud-affected areas

Continuously evaluate forecast sensitivity in evolving DA-system.

Accuracy of observation operator

RT: cloud absorption ok, scattering less good; Mφ: large-scale cond. easier than convection; interpolation problem significant.

Similarity of linearized and non-linear parameterizations

Much improved with CY32R3 also thanks to cloud/rain assimilation efforts.

Linearity of observation operator

Apart from improved models, may be evaluated in analysis.

Definition of observation/background errors, biases

Observation ≈ RT+Mφ -errors, observational method vs. error correlation;
B rather unspecific and low-resolution; VarBC suitable for linear channels.

Choice of control variable

Total water control variable, also facilitates B-estimation.

Options for using cloud/rain information

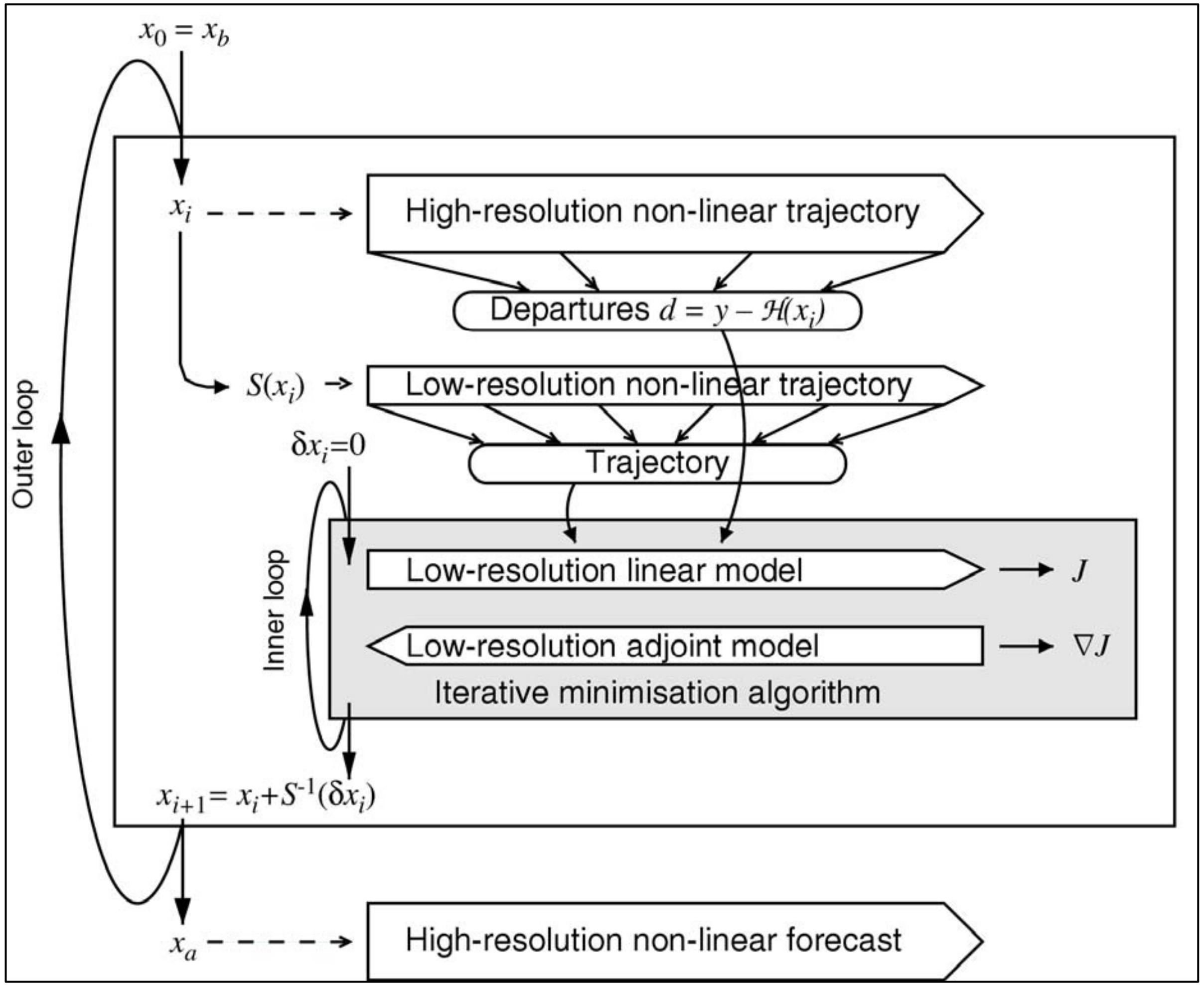
(assuming 4D-Var system oriented towards radiance assimilation)

- Screening of cloud-affected observations. Effective for all but SSM/I in clouds/rain.
- Use pre-processed data where cloud signal has been removed (cloud-cleared). Trade-off between yield and cloud contamination.
- Simulation of cloud-affected observations for diagnostics but no active assimilation. Currently only for few data and diagnostics.
- Include cloud contributions in FW/TL/AD operators but only maintain dry gradients for model update, i.e. only avoid aliasing in RT-model. Under development for most IR/MW instruments.
- 1D+4D-Var assimilation: Subset 4D-Var with 1D-Var retrieval of pseudo-observation using cloud-affected observations; assimilate pseudo-observation in (dry) 4D-Var. Operational for SSM/I, soon other MW-imagers.
- 4D-Var assimilation: Fully account for sensitivity of 4D-Var to gradients from moist observation operator. Under development for MW-imagers+.

Incremental Formulation of 4D-Var

$$J(\delta \mathbf{x}) = \delta \mathbf{x}^T \mathbf{B}^{-1} \delta \mathbf{x} + [\mathbf{H}\delta \mathbf{x} - \mathbf{d}]^T \mathbf{R}^{-1} [\mathbf{H}\delta \mathbf{x} - \mathbf{d}]$$

$$\delta \mathbf{x} = \mathbf{x} - \mathbf{x}_b, \qquad \mathbf{d} = \mathbf{y} - H(\mathbf{x})$$



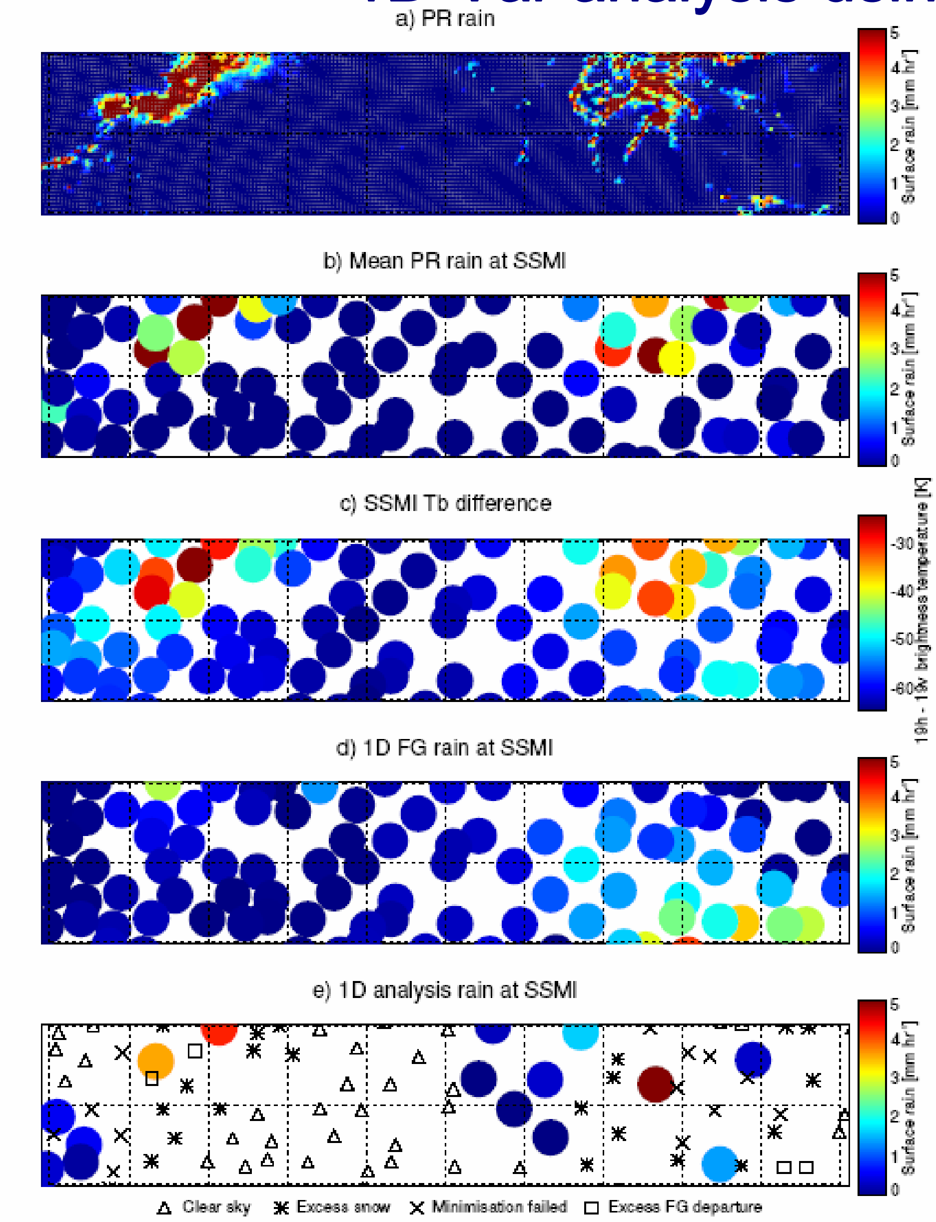
$5 \cdot 10^6$ control variables
10^6 observations/12h

(Trémolet 2004)

ECMWF

Examples

1D-Var analysis using SSM/I radiances



TRMM PR rain rates [mm hr^{-1}]

… averaged to SSM/I spatial resolution

SSM/I 19h-19v TB-difference [K]

Model FG-rain rates [mm hr^{-1}]

1D-Var analysis rain rates [mm hr^{-1}]

(Geer et al. 2007)

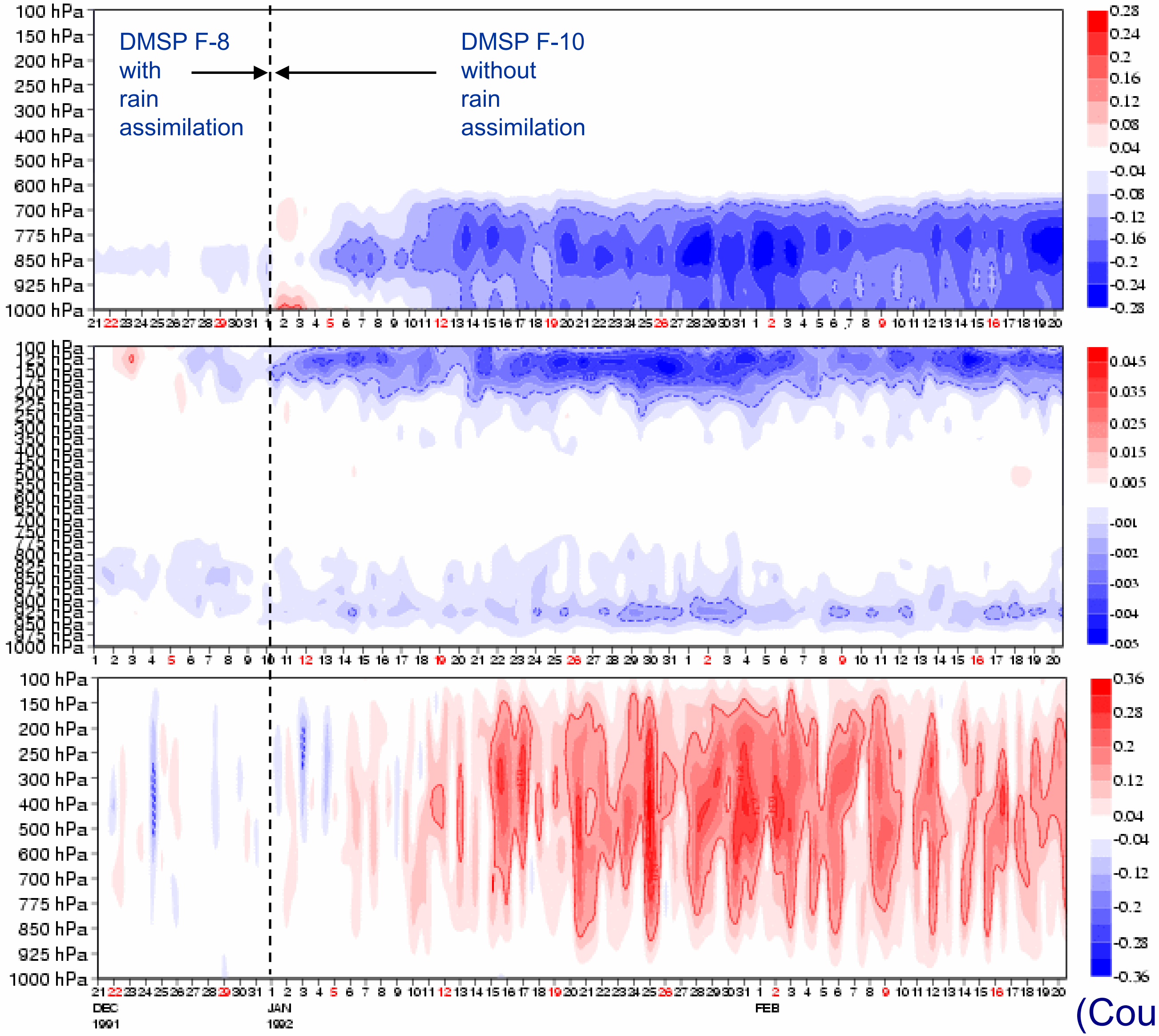
1D+4D-Var assimilation of rain-affected radiances: ERA-interim

Analysis difference b/w 2 ERA expts.

Specific humidity

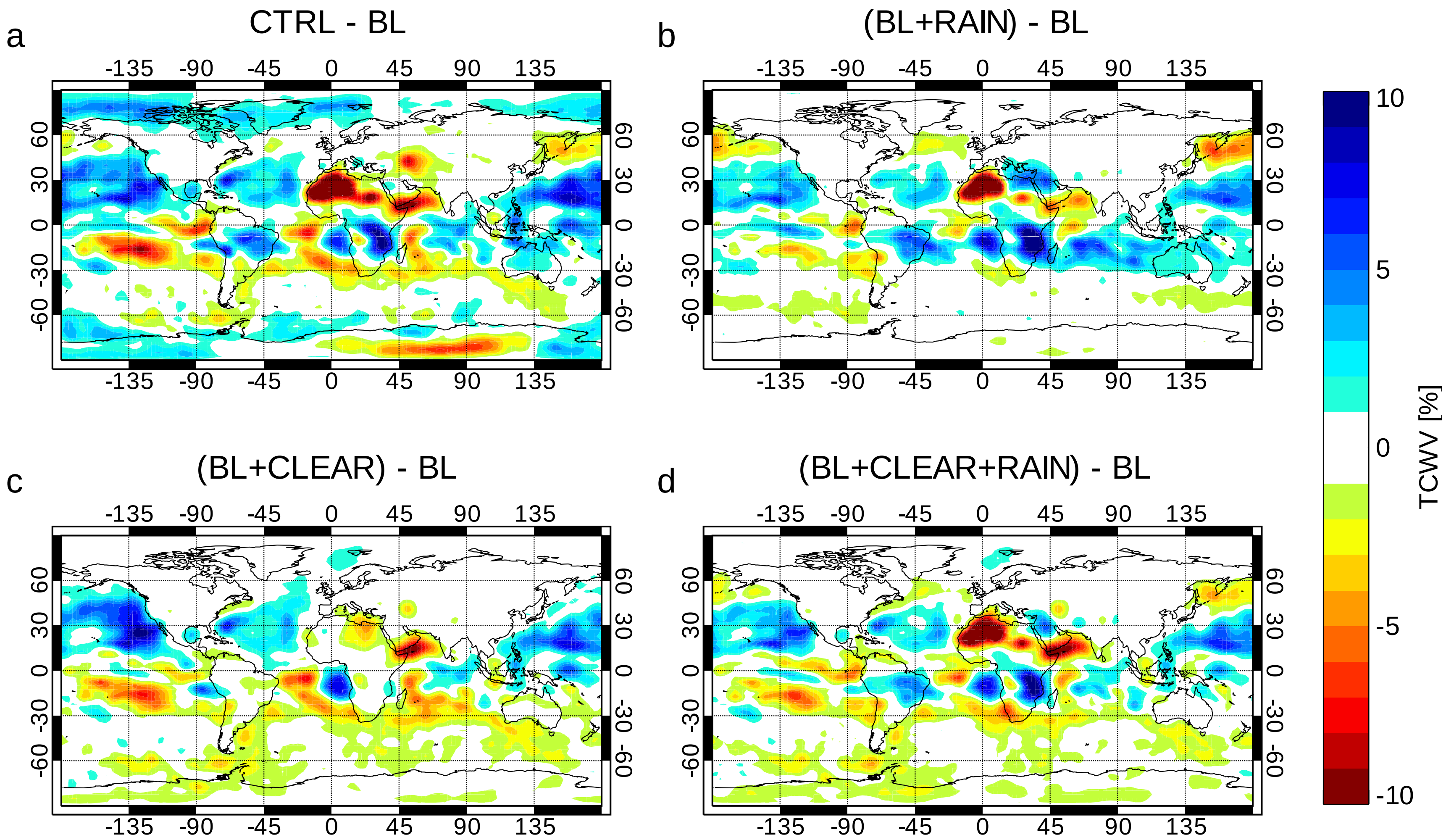
Cloud cover

Vertical wind



(Courtesy D. Dee)

1D+4D-Var assimilation of rain-affected radiances: Dependence of moisture (TCWV) climate on satellite data usage



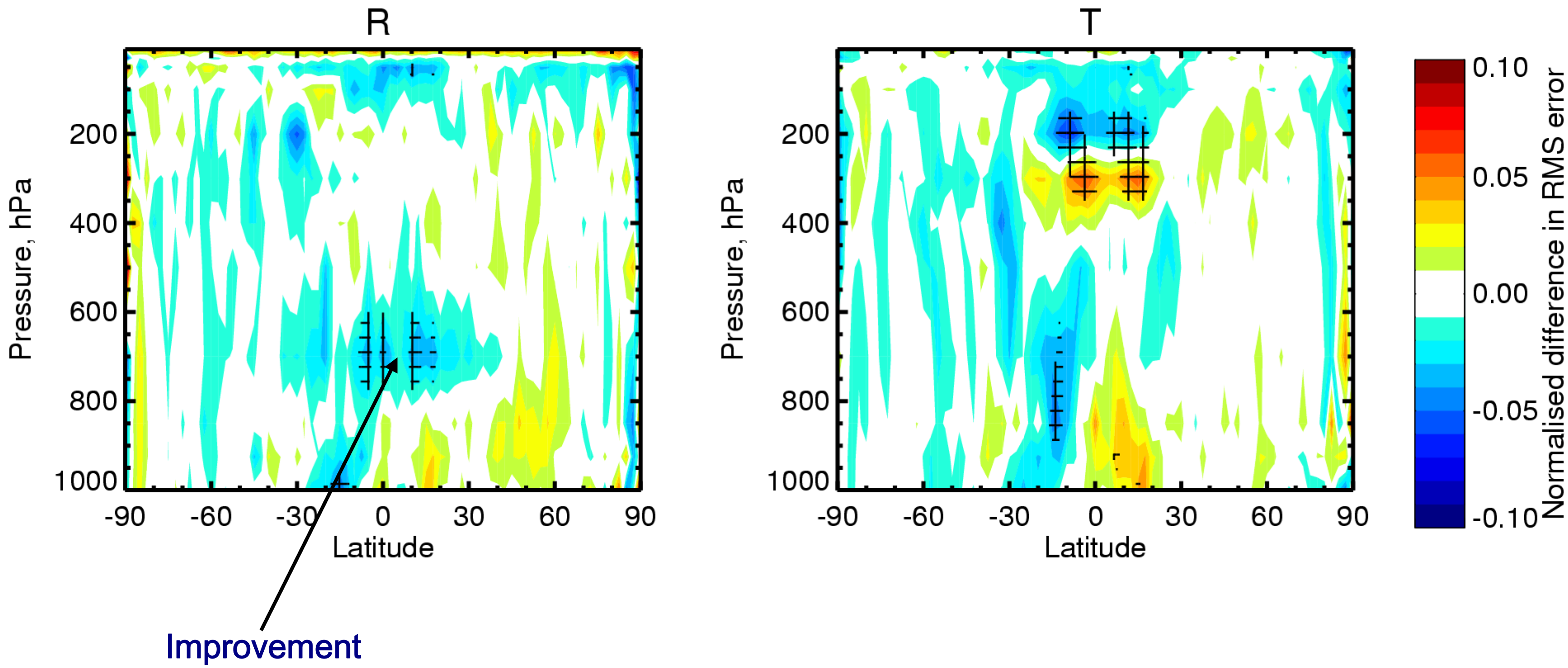
(Kelly et al. 2007)

1D+4D-Var assimilation of rain-affected radiances: Forecast performance in operations

Operational implementation: June 2005

T+48h normalised RMS forecast error difference (averaged over 60 days):

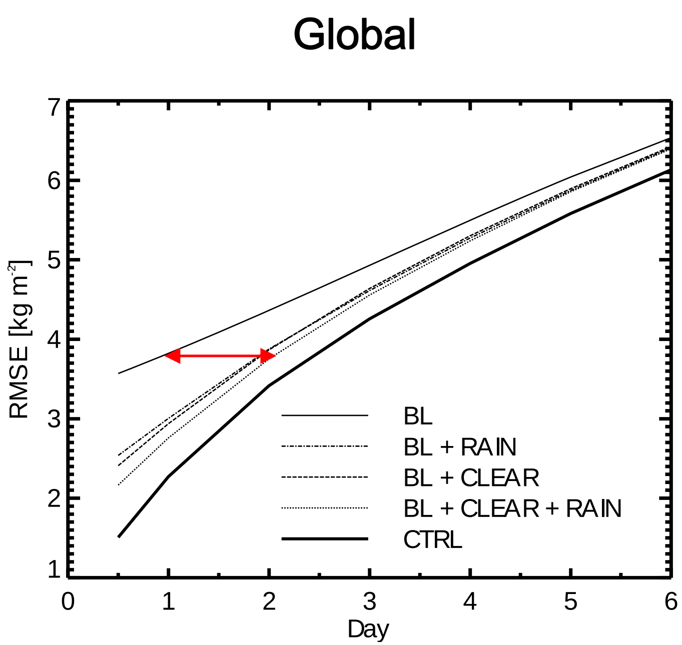
RAIN – NO RAIN

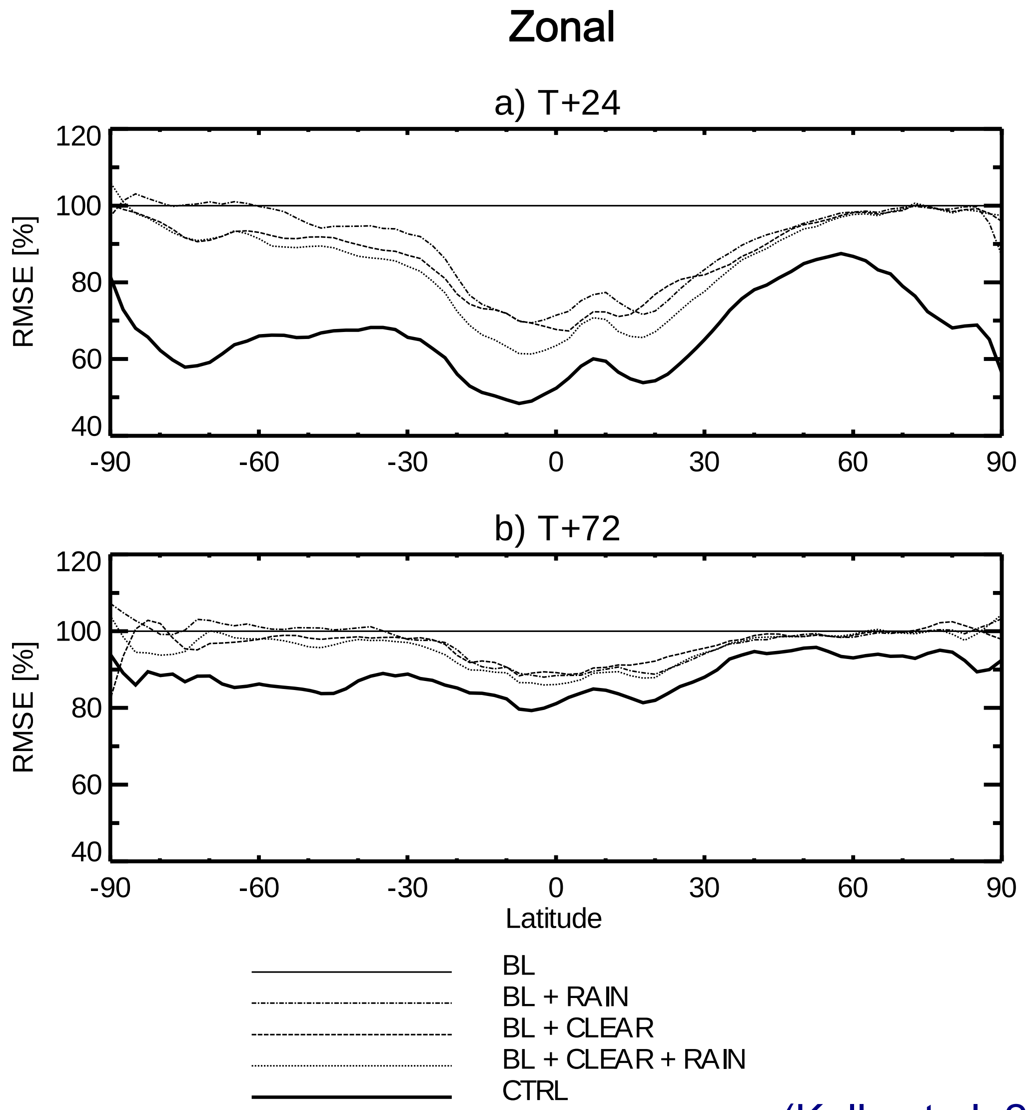


(Geer et al. 2007)

1D+4D-Var assimilation of rain-affected radiances: Forecast performance in operations

Total column water vapour





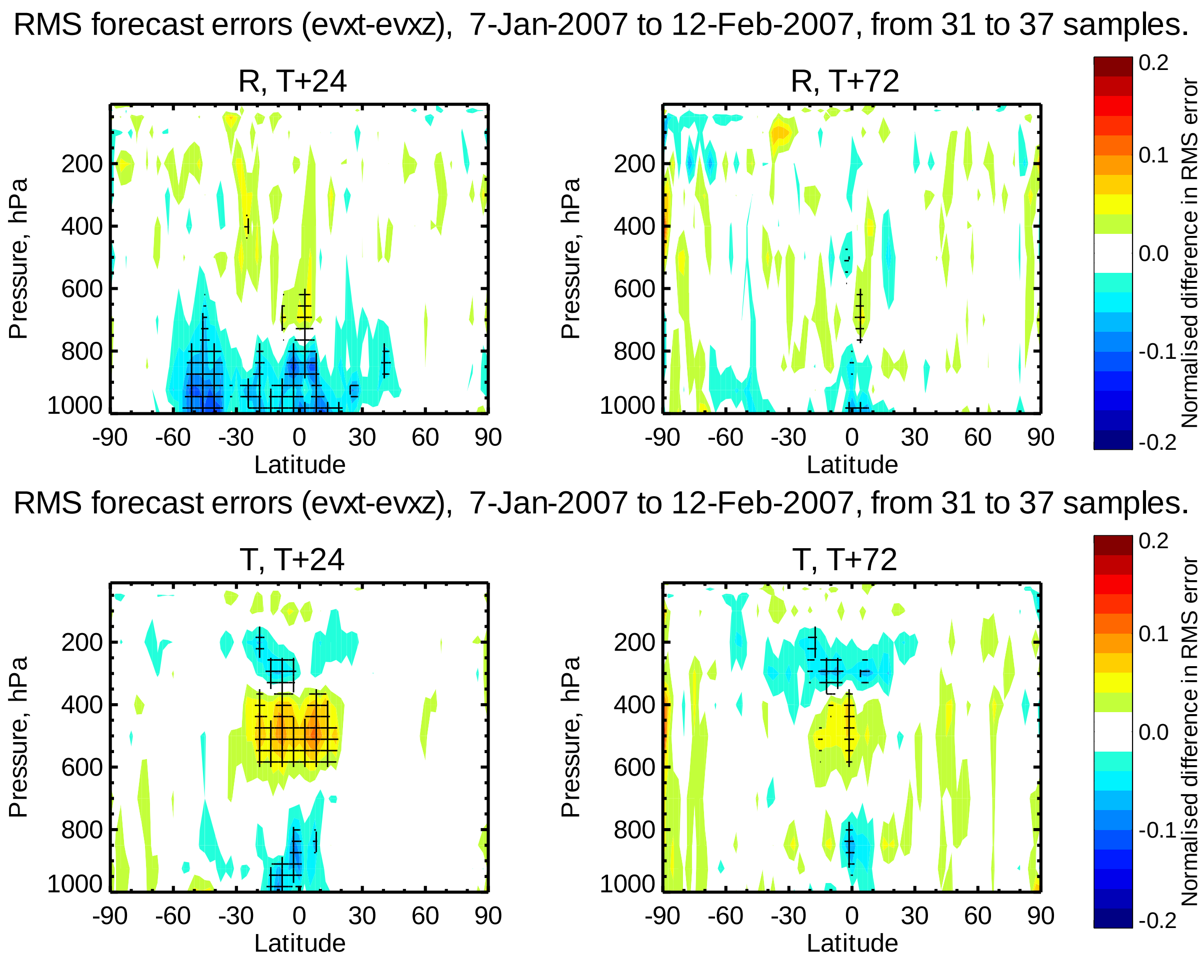
(Kelly et al. 2007)

1D+4D vs. 4D-Var assimilation of rain-affected radiances: 4D minus 1D+4D forecast RMSE

(against own analysis)

Relative humidity

Temperature



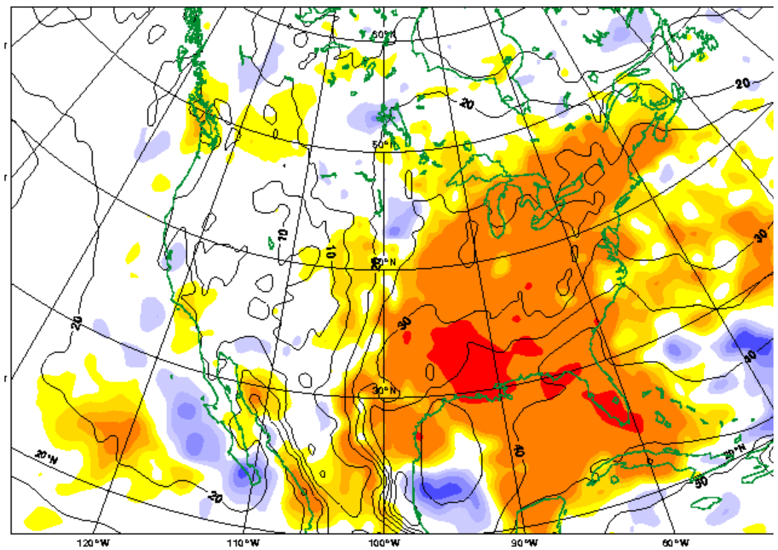
Examples: Products – Surface radar-derived rain rates

Experiments

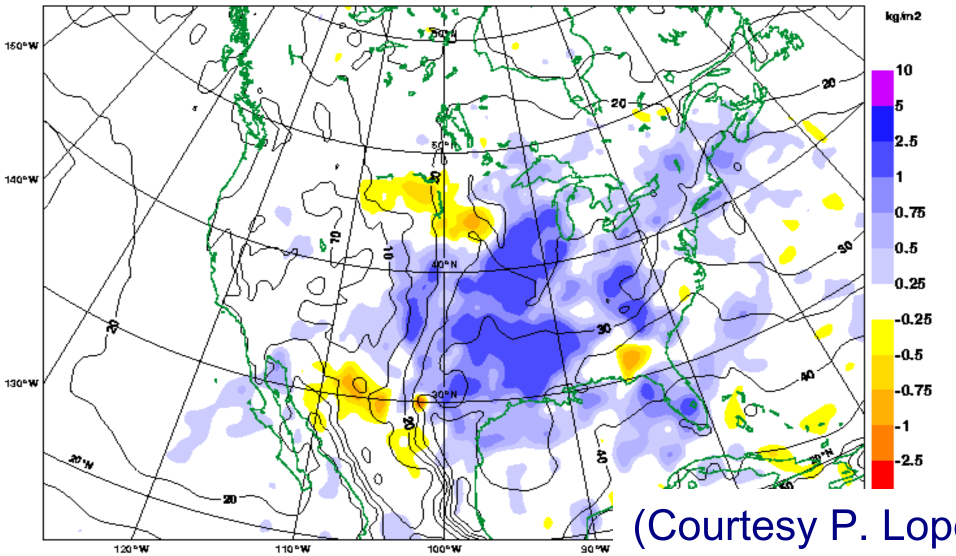
- *control*: all observations (incl. SSM/I rain assimilation over oceans);
- *noqUS*: withdraw TEMP-q, RH2m, HIRS, AMSU-B, SSM/I, AIRS, GOES-WV over US;
- *noqUS+StageIV*: as noqUS but add Stage IV data.

Mean differences of
TCWV analyses at 00UTC

noqUS – control



noqUS+StageIV – noqUS

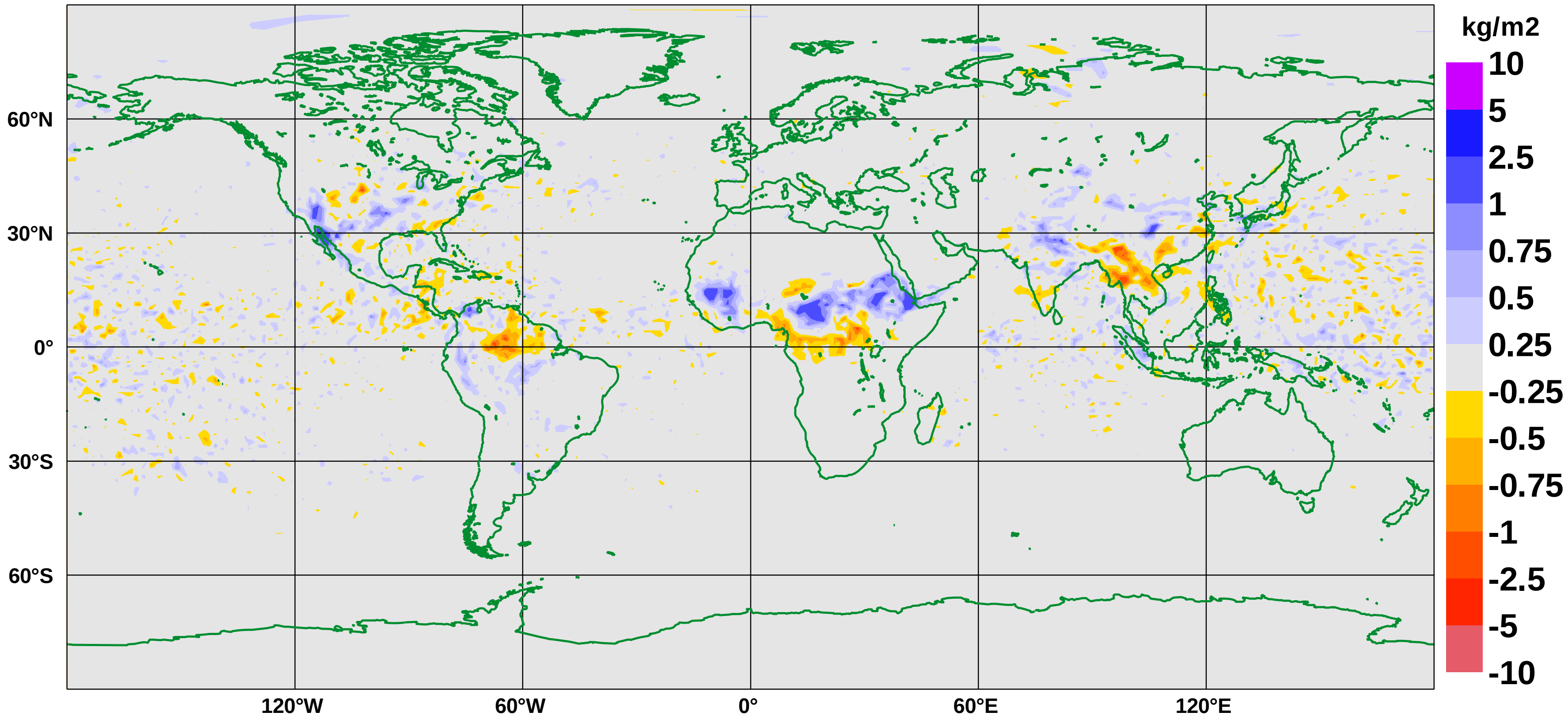


(Courtesy P. Lopez)

Examples: Products – Satellite data-derived rain rates

Assimilation over Land: TRMM 2A12 Rain Rates

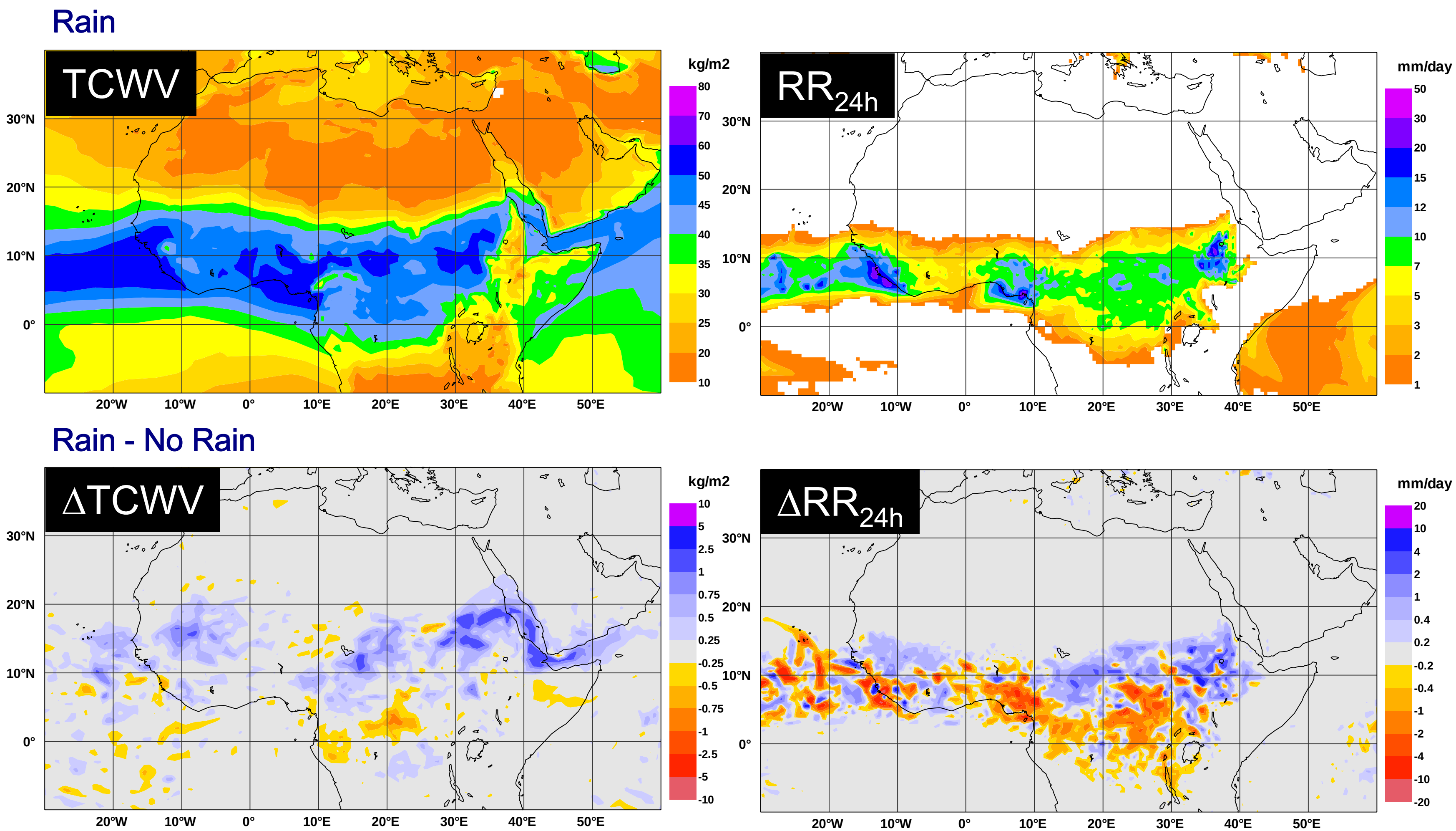
Mean TCWV analysis increments 1-25/07/2006 at 00UTC



(Courtesy P. Lopez)

Examples: Products – Satellite data-derived rain rates

24-hour Forecast Differences: Rain – No Rain



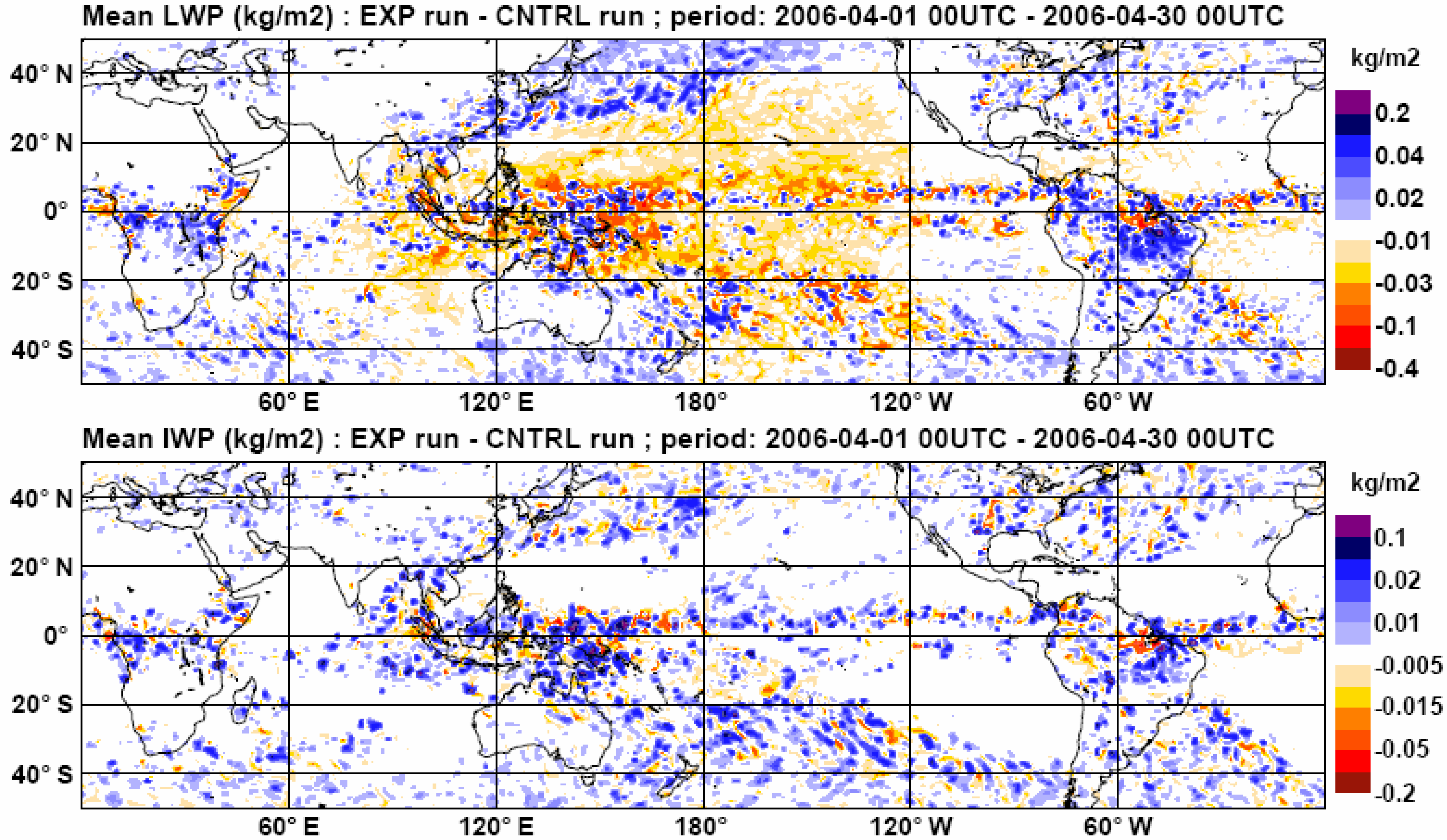
(Courtesy P. Lopez)

ECMWF

Examples: Products – MODIS optical depth

- Assimilation of 0.55 μm optical depth derived from effective particle size product (25 km).
- Observation operator combines moist physics and liquid+ice optical property parameterizations.
- T511L60 DC experiment for April 2004.

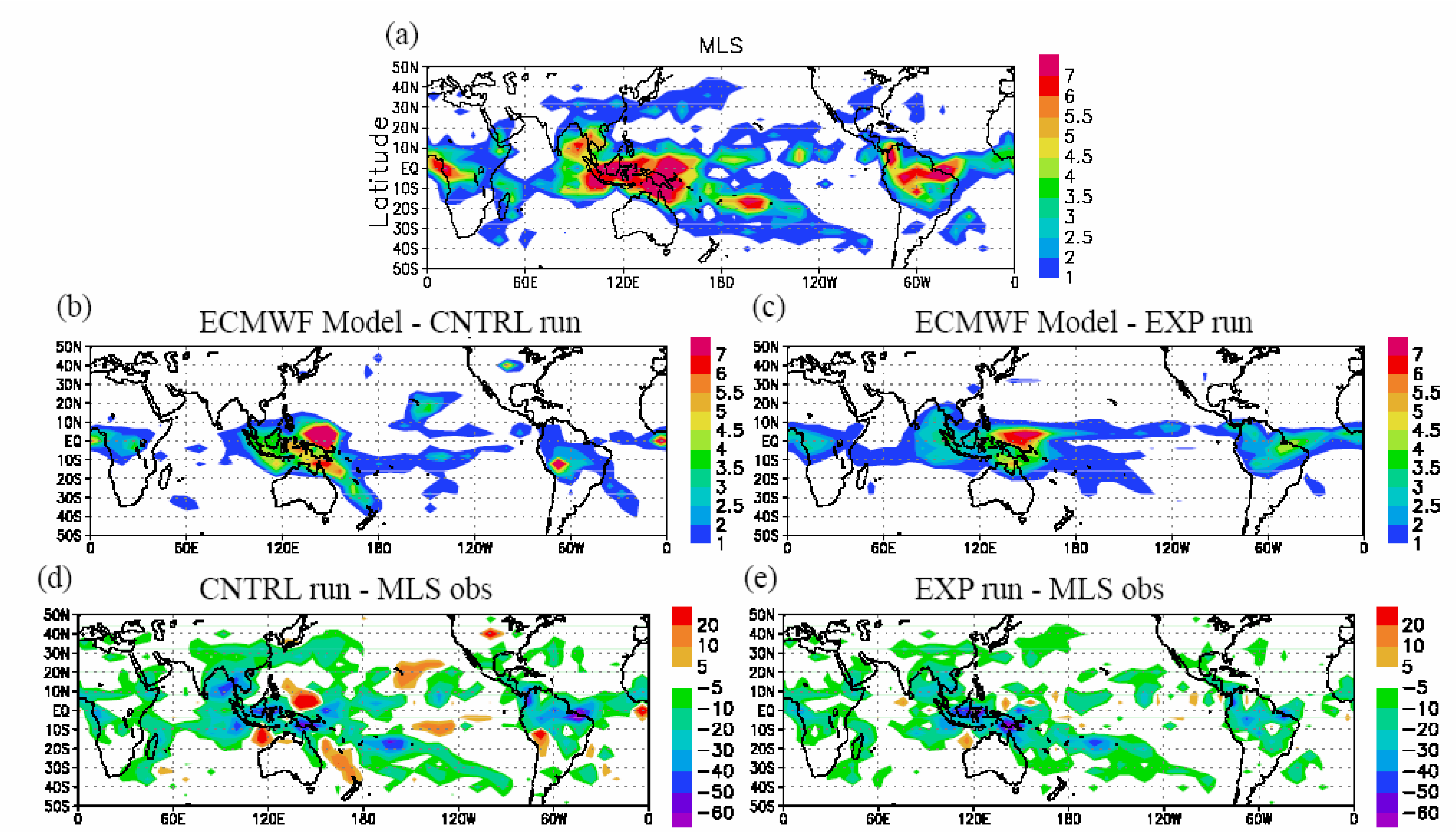
Change in liquid water and ice mean analyses



(Benedetti and Janisková 2007)

Examples: Products – MODIS optical depth

Ice water content (mg m^{-3}) at 215 hPa



(Benedetti and Janisková 2007)