COSMIC 1D-Var Update and Validation

Tae-Kwon Wee

COSMIC Program Office
UCAR

IROWG-8 Virtual Meeting, 8 April 2021

Background

- **Purpose**

  Optimal estimation of the thermodynamic states (Tpq; unobserved but needed by many data users) along the trajectory of ray tangent points.

  Based on the same principle/method with var data assimilation but differs in purpose and practice

- **Requirements**

  Accurate (meeting mission requirements and other standards), reliable (stable), and fast enough for real-time data processing

- **Information utilized**

  Observation and a priori (background) and their error statistics, and physical constraints (hydrostatic, sub-adiabatic, non-negative q) and relationships (e.g., forward models)

  Uses the background to tackle underdeterminancy, while attempting to minimize its adverse influence on the retrieval

- **Status**

  Comprehensive validation completed for C2 and other missions including COSMIC

  In use for C2 data processing and will be used for re- and post-processing of other missions

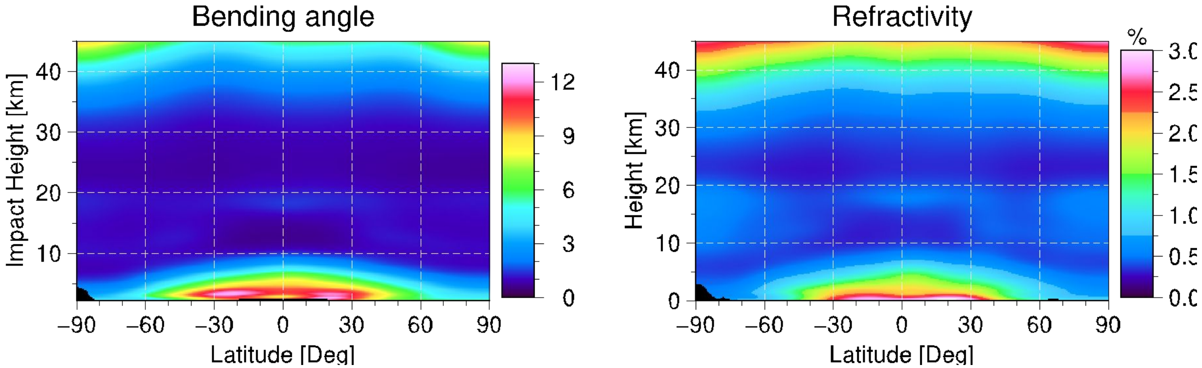
COSMIC 1D-Var Update

	OLD (v 1.0)	NEW (v 2.0)
Formulation	Incremental (Courtier et al., 1994), Control-Variable Transform (Parrish-Derber, 1992)	
Control Variables	T, p_L, RH* (other options available)	
Observation error	Extreme quality	Statistical (Hollingsworth-Lönnberg, 1986)
Background error (**B**)	Empirical	Statistical (HL86 + NMC-method)
Correlation	Univariate	Multivariate* (NMC-method)
Resolution	3 latitude zones, seasonal	5°×5° lat-lon grid, monthly
Construction	Run-time (expensive)	Precomputed $\mathbf{C}^{1/2}$ on fixed levels (“no cost”)
Observation operator	Refractivity	Variational Abel Transform (Wee, 2018)
Terrain	None	Digital Elevation Model
Ducting	Disregarded	Truncation of model grids
I/O	ascii	netcdf

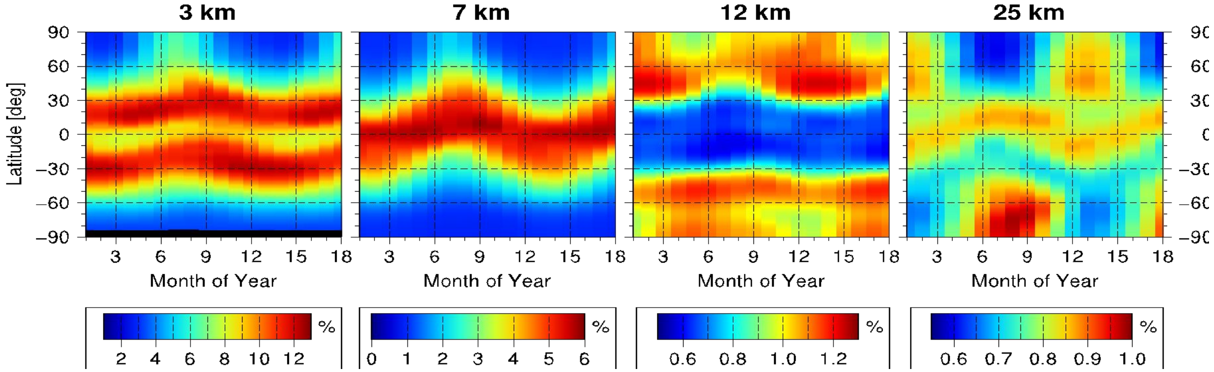
Observation error specification

- Hollingsworth-Lönnberg method (1986) to nearby RO-RO pairs (Wee, 2018)
- Locally accurate (statistically precise) OE, varies with height, latitude, and month
- Provides corresponding background errors for both bending angle refractivity

annual-mean structure



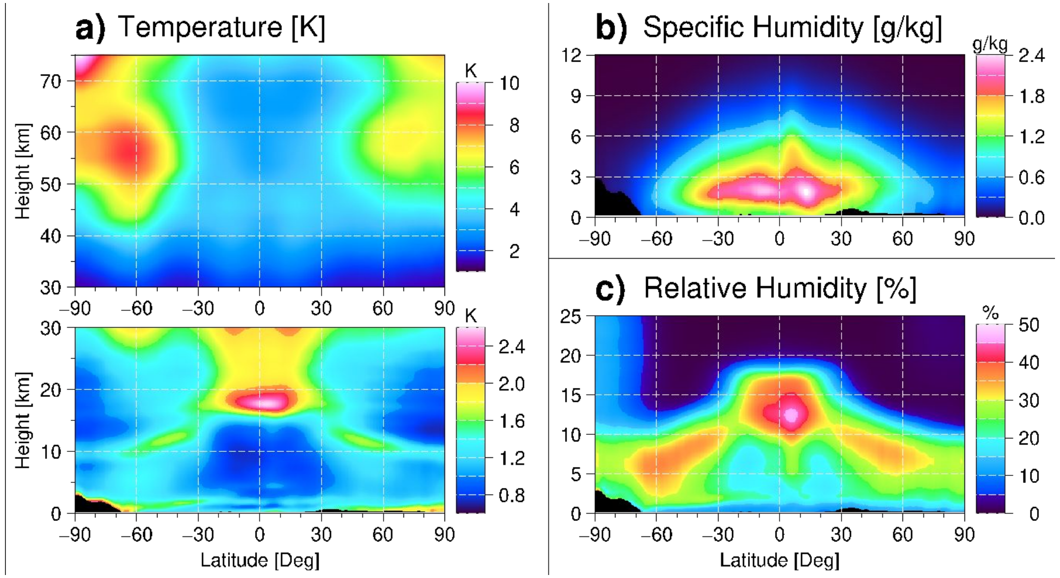
temporal variation



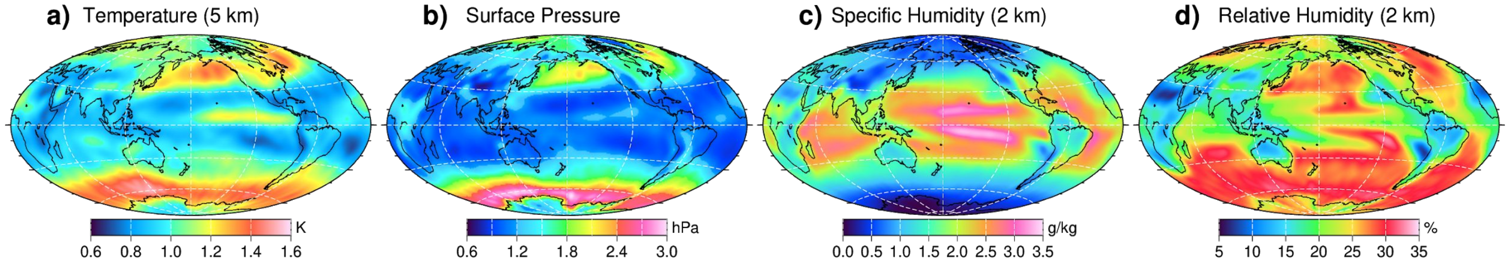
Background error spec.

- NMC method scaled towards the Hollingsworth-Lönnberg (1986) with radiosonde (RS)
- Varies with height, latitude, longitude, and month

zonal- and annual-mean structure

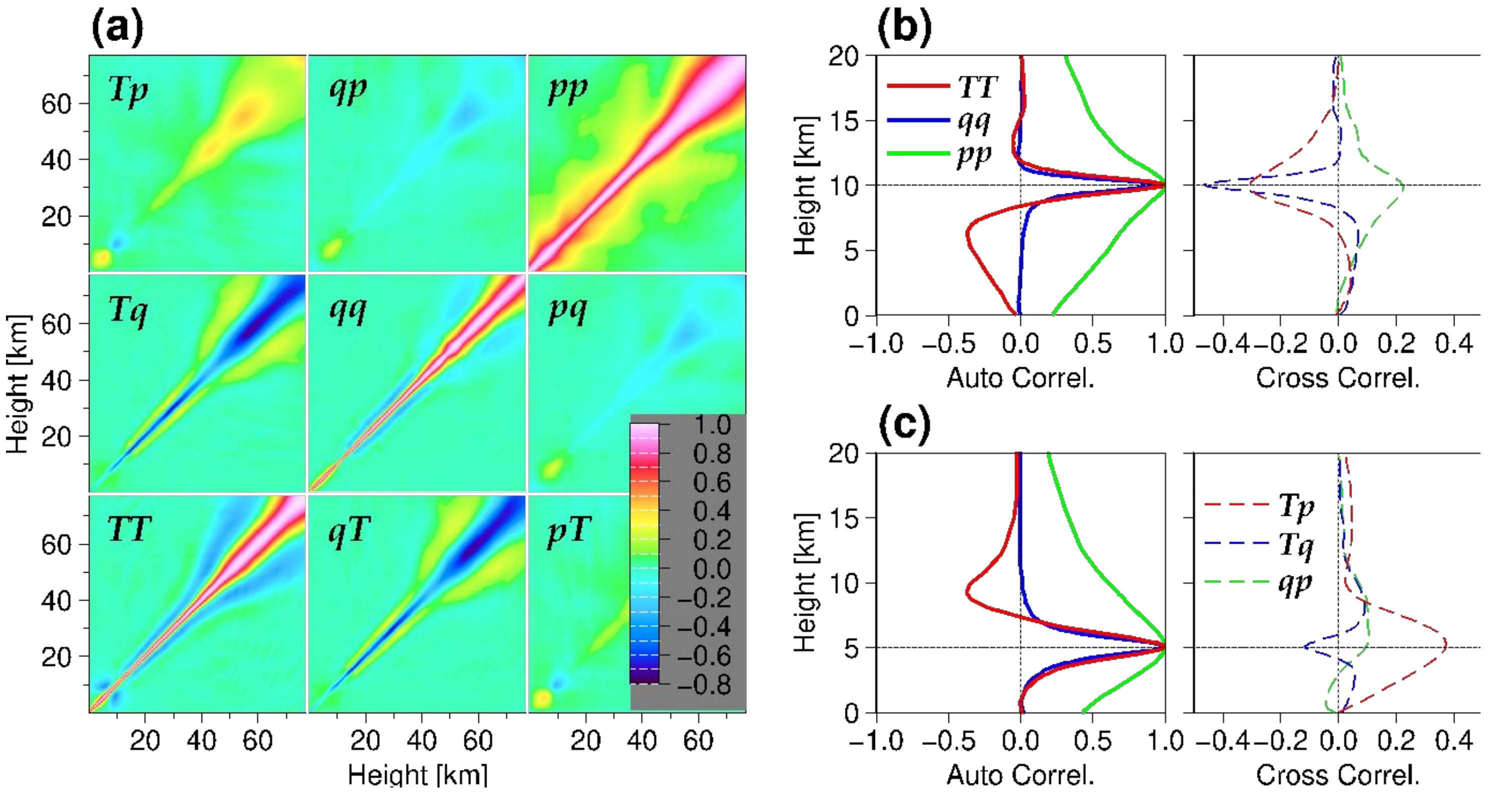


horizontal structure



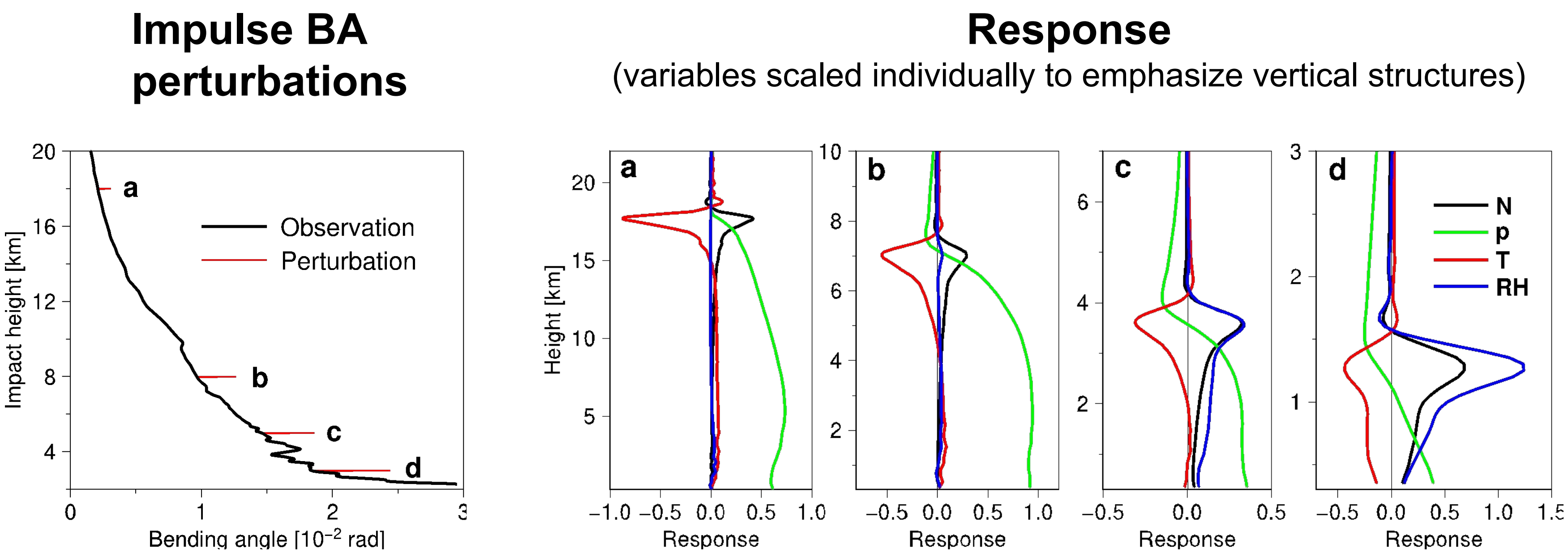
Background error correlation

- 1D-Var can utilize **full error covariance matrices** (w/o approx. & simplification)
- NMC method with ECMWF forecasts over a 9-y period
- Varies with latitude and season



1D-Var response to single-level ob pert.

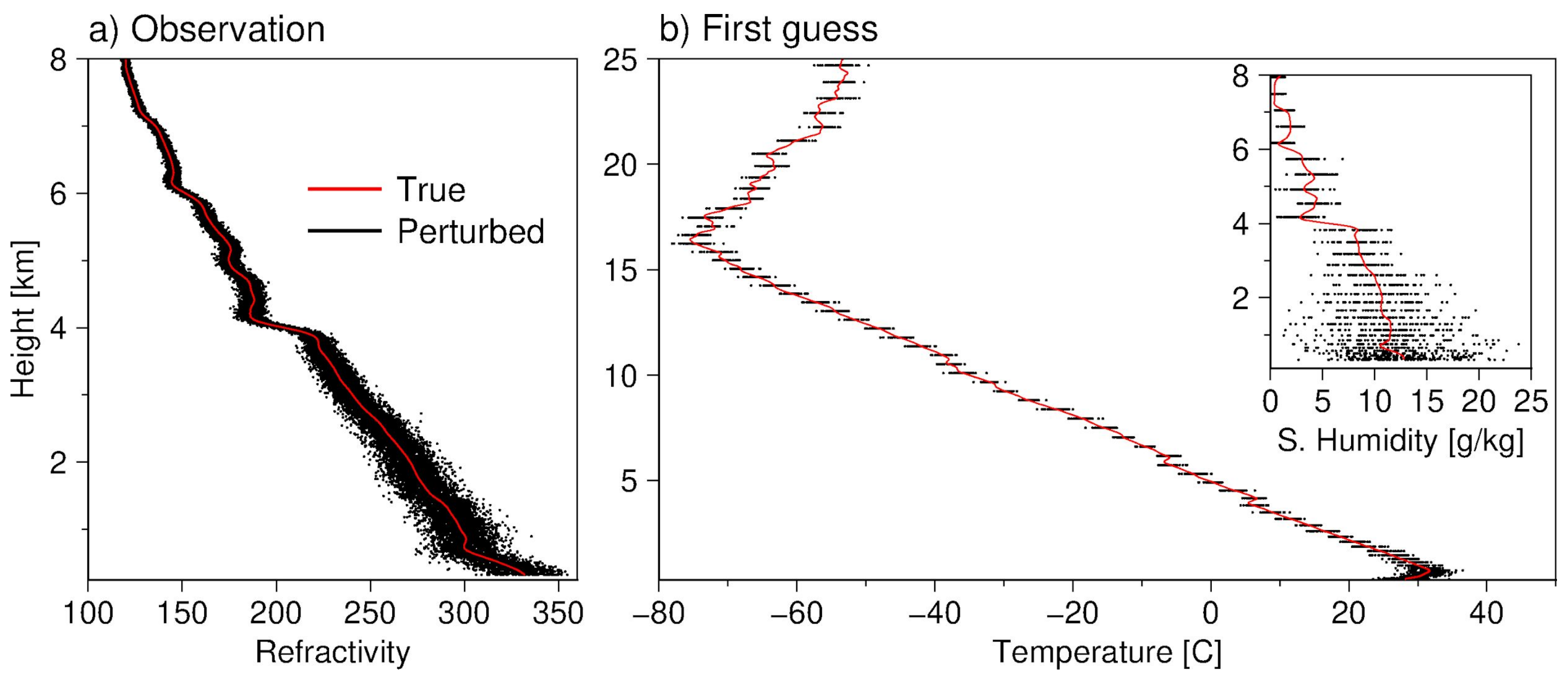
(Useful to understand how information is spread vertically)



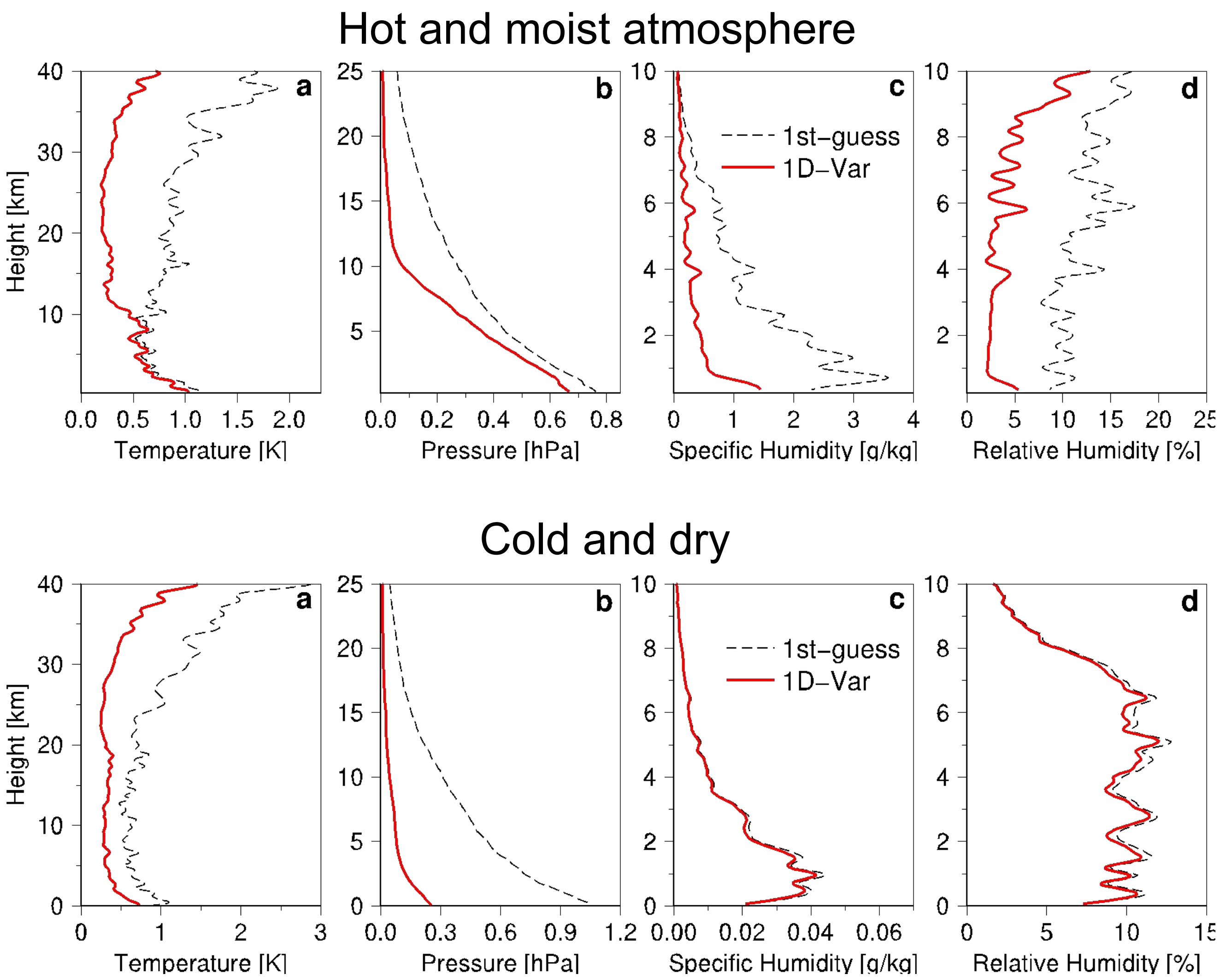
- 1D-Var response is non-local and broad, because **B** and **R**, along with Abel transform and hydrostatic balance op, spread information to wide extents
- **R** and **B** (cross error correl. in particular) are important to tackle the underdetermined retrieval problem
- If the observation holds a good number of independent pieces of information, 1D-Var might be able to reconstruct the whole atmosphere

Validation using synthetic data

- A sanity check whether 1D-Var behaves as designed and intended
- <u>A controlled experiment where the **truth** and **B** and **R** are well known</u>, opposed to real-world validation
- For a given RO event, a smoothed GRUAN sounding serves as the truth:
  - 200 observed RO soundings (made available by perturbing the true phase Doppler)
  - 200 first-guess (FG) soundings (by perturbing FG using EOFs of **B**)
  - 40,000 realizations of RO-FG pairs and 1D-Var retrievals
  - Refractivity is taken as the observation (for simplicity's sake)

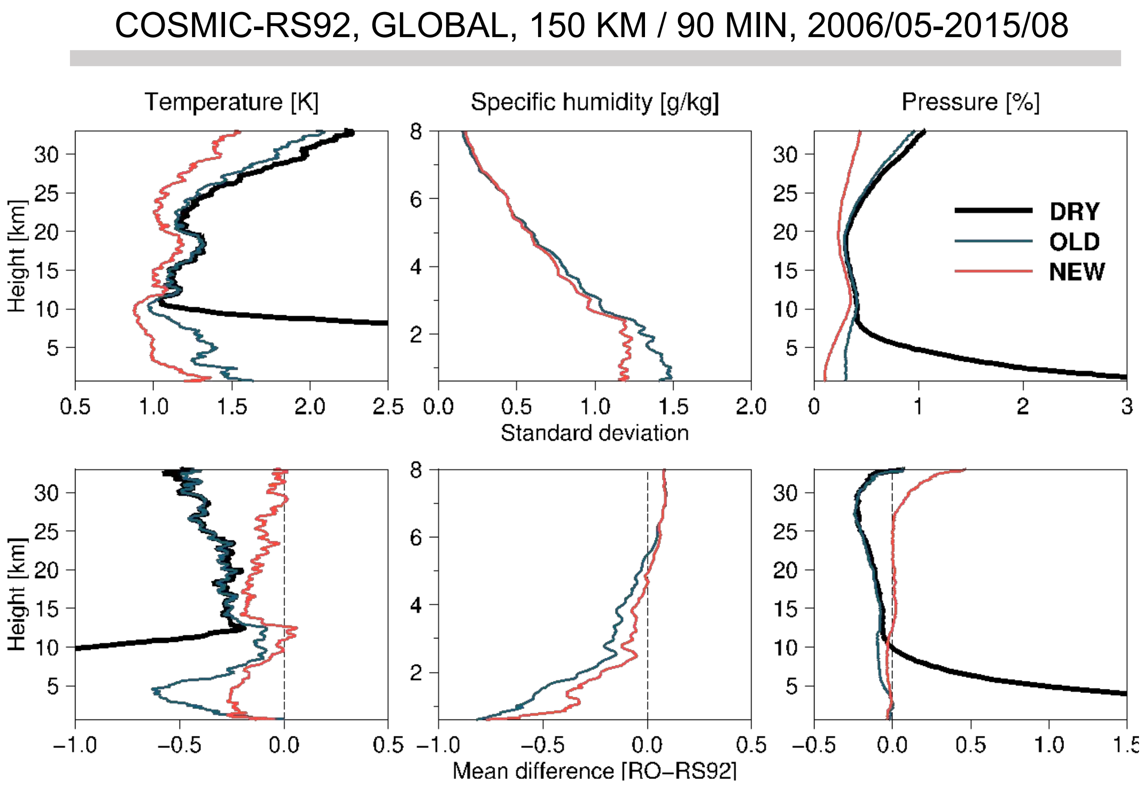


Synthetic: RMS error wrt the known truth



- 1D-Var is significantly smaller than FG in the error. The aspect, however, depends on atmospheric condition
- RO observation is very sensitive to moisture; moisture error reduction dominates in moisture-abundant atmosphere

COSMIC compared to (1-s) RS92



DRY – atmPrf
OLD – wetPrf
NEW – wetPf2

- OLD is close to DRY in the heights above 12 km (as intended)
- NEW agrees significantly better with RS, compared to OLD and DRY
- As a minimum-error-variance estimator, 1D-Var can be smaller than FG in the error (shown later)

Observation types (operators)

- Refractivity (N) and bending angle (BA) are currently admissible (others under testing)
- N versus BA boils down to inverse versus forward Abel transforms
- Use of BA suffers from the fact that it belongs to a space different from N
- **COSMIC 1D-Var**

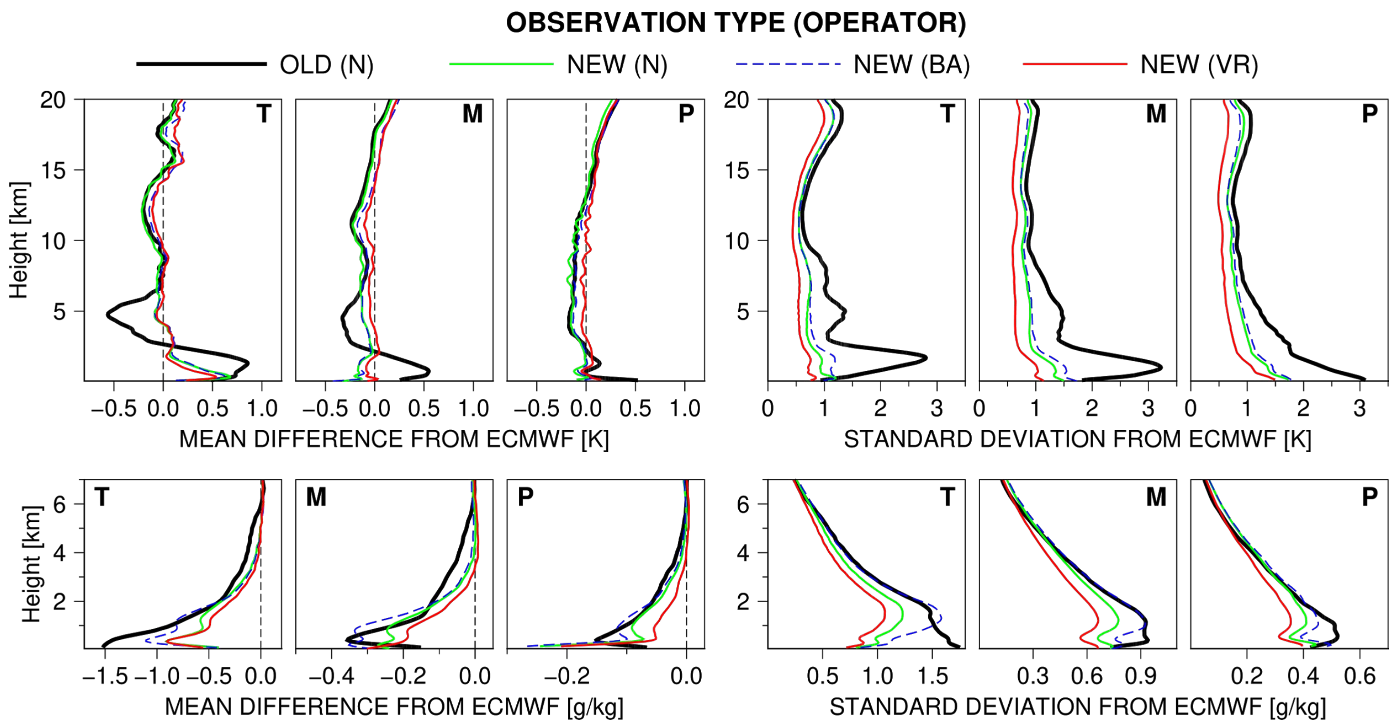
    Uses BA through a variational Abel transform

    Carries out two minimizations, one in the observation space (impact parameter) and the other in the model space (height coord.)

    Rationale & Support are given by Wee (AMT, 2018)

Comparison of observation operators

- Stats for 7-y COSMIC (FG: NCEP fcst, Verification: ECMWF anal)
- Uses consistent OEs between BA and N (same method applied to a common data set)



OLD (N) Prior 1D-Var version which takes refractivity (N) as the observation

NEW (N) Uses N as the ob
NEW (BA) Uses BA
NEW (VR) Uses BA through Var Abel transform (default)

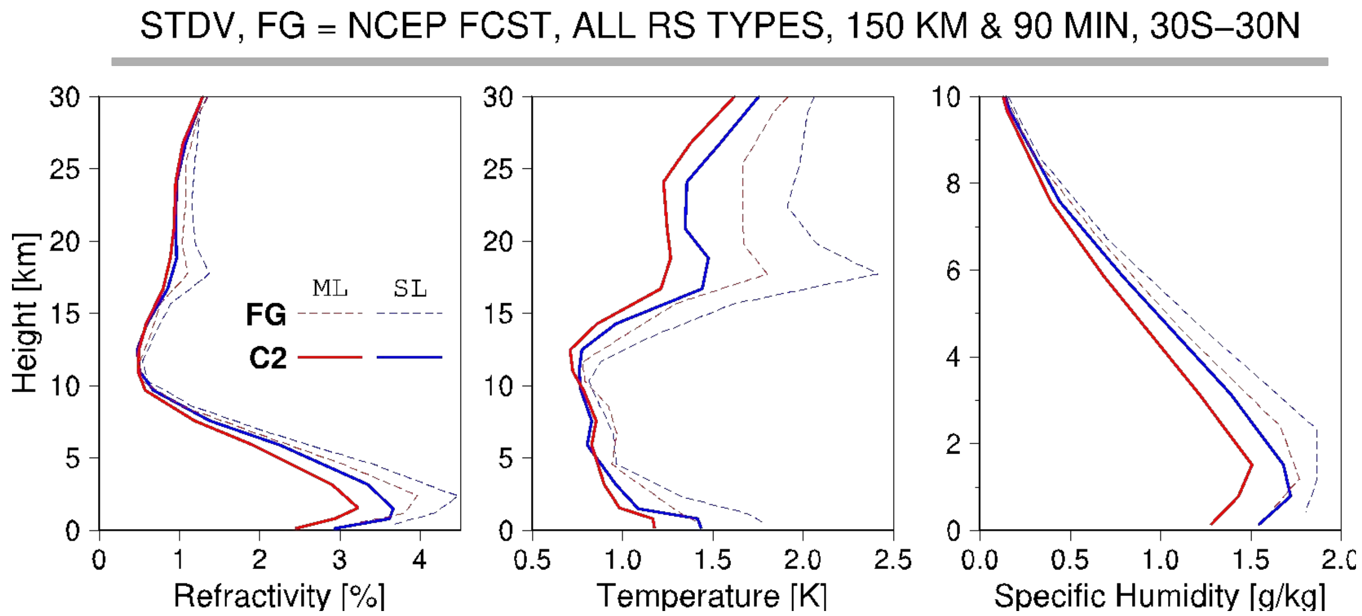
T Tropics
M Middle lats.
P Polar lats.

- NEW(N) v. OLD(N) shows the gross improvement brought by the recent upgrade (other than the ob type)
- NEW(N) v. NEW(BA) relates to the relative effectiveness btw inverse v. forward Abel transforms. Disputable, but N serves better in lower troposphere, because of large fwd modeling error and high volatility of BA there
- NEW(BA) is the most effective

Can 1D-Var add any value to 1st guess?

C2 comparison to collocated (operational; GTS) radiosondes

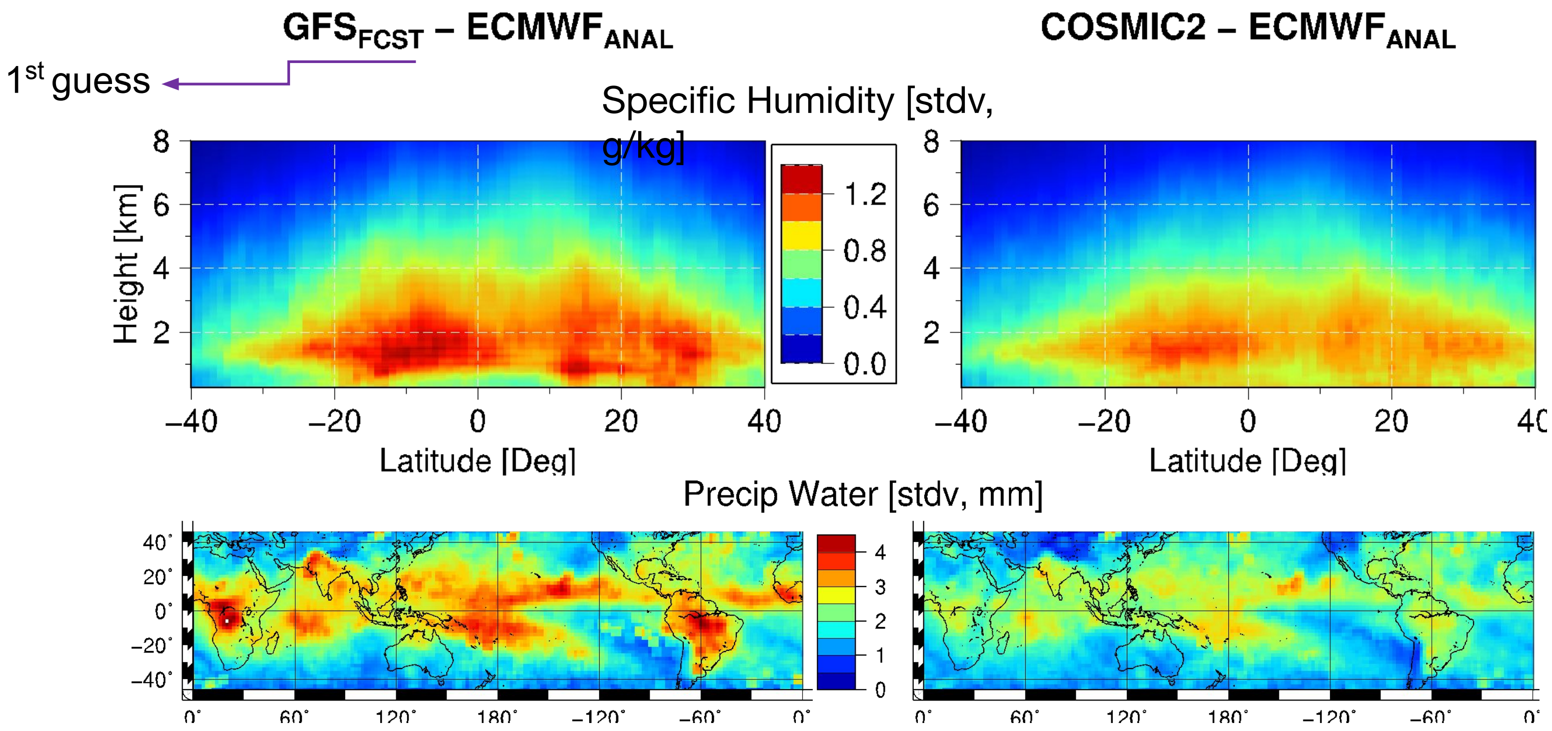
- 1D-Var 1st guess (FG) used here is NCEP GFS short-term forecasts
- FG and 1D-Var (C2) are compared to nearby radiosondes (RS) on ML (Mandatory p. levs) and ST (Significant T. levs)



1D-Var (**C2**, heavy solid lines) agrees better with RS than FG (dashed) does

Can 1D-Var add any value to 1st guess? (II)

FG (GFS_{FCST}) v. 1D-Var retrieval (C2) wrt ECMWF anal (2019.197 —2020.032)



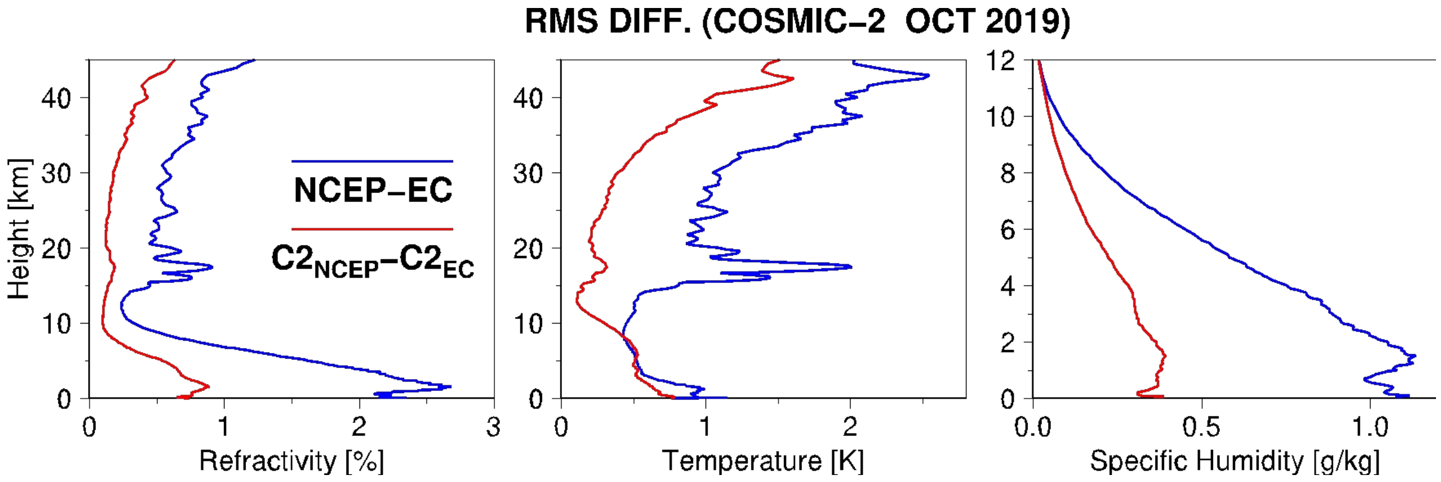
- 1D-Var shows a significantly better agreement than its FG
- This exemplifies the benefit that C2 can bring to weather forecasting/analysis

C2 dependence on FG

- While 1D-Var needs FG to deal with the underdeterminancy and to better use the observation, its dependence on FG is unavoidable and not necessarily bad. It is, though, important to quantify/understand the dependence
- The dependence is measured by the diff. btw two 1D-Var retrieval sets that make use of different FGs (NCEP and ECMWF), relative to the diff. btw the FGs:

$$dep. = \frac{\|C2_{NCEP} - C2_{EC}\|}{\|NCEP - EC\|}$$

NCEP	NCEP forecast
EC	ECMWF forecast
$C2_{NCEP}$	1D-Var using NCEP as FG
$C2_{EC}$	1D-Var using ECMWF as FG



- 1D-Var retrieval sets are closer to each other than the distance between FGs
- The weak FG dependence indicates that 1D-Var retrieval is constrained well by RO
- If RO observation is accurate, the uncertainty of 1D-Var retrieval is lower than FGs

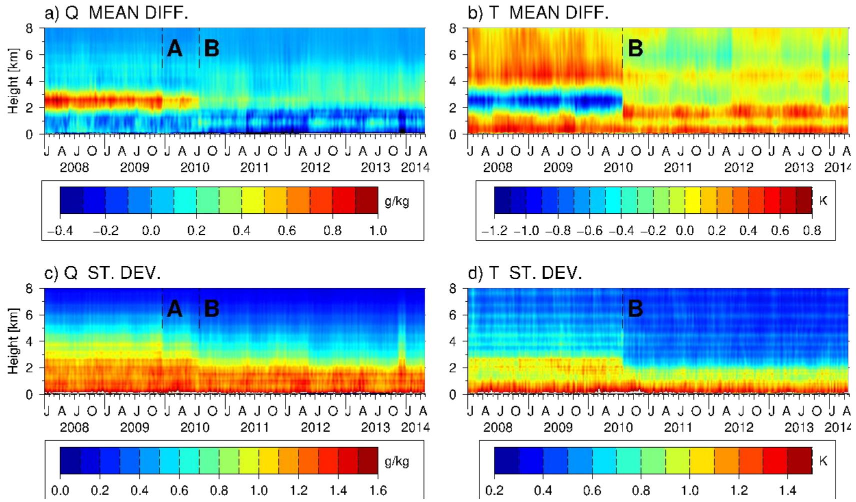
Weak FG dependence enables long-term stability

(COSMIC data record 2008-2014)

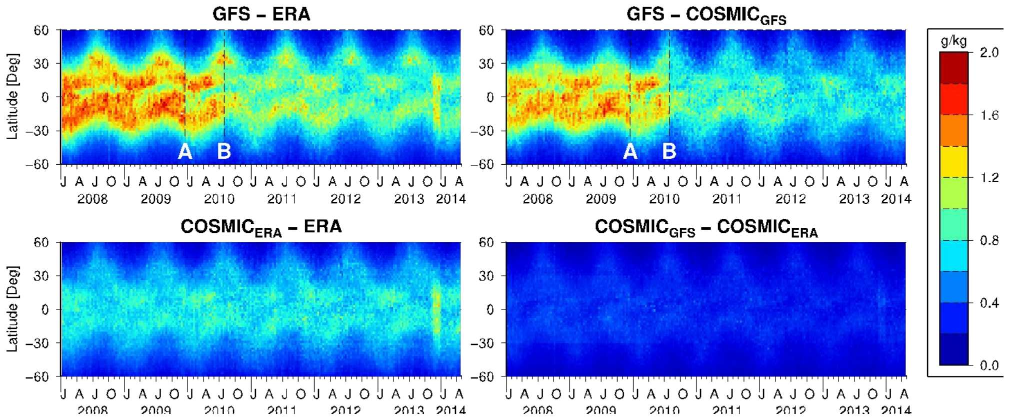
GFS	NCEP GFS analysis
ERA	ERA-Interim reanalysis
$COSMIC_{GFS}$	1D-Var with GFS as FG
$COSMIC_{ERA}$	1D-Var with ERA as FG

GFS departure from ERA-Interim

GFS – ERA (±25° LAT.)



Specific humidity (rmsd, g/kg) at 2.5 km



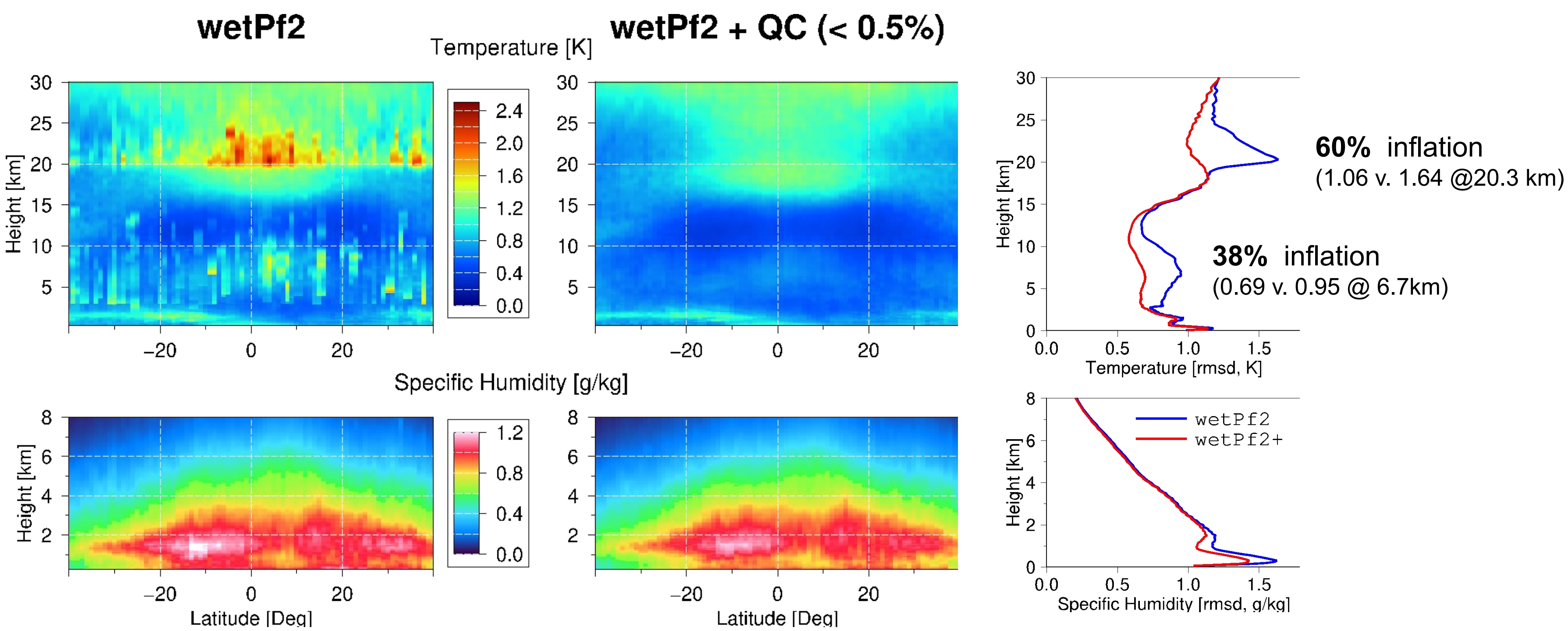
Past GFS major upgrades (marked **A** and **B**) result in discontinuities

- 1D-Var data record is stable enough to show the GFS breaks
- 1D-Var shows no obvious dependence on FG (lower-right p.)

Outlying C2 soundings (not flagged bad)

- Existence of few obvious C2 outliers were informed (S.-Y. Chen, NCU)
- Currently, no QC is applied within 1D-Var on purpose. With QC enabled, a small number of unphysical outliers (< 0.5% for Oct 2019) were detected (considered not bad by lower-level data processing)
- Despite the tiny fraction, these outliers inflate RMSD significantly by up to 60%
- “Robust” statistics, rather than straightforward evaluation, is advised

Standard deviation from ECMWF analysis (w/ NCEP as FG)



Conclusions

- COSMIC 1D-Var (v2) behaves as designed
- Recent changes made to the 1D-Var brought in significant overall improvement
- 1D-Var retrievals show good agreements with radiosonde and global forecasts/analyses
- Dependence on background is weak, enabling long-term stability of 1D-Var data records
- Some issues (new with C2 or unnoticed earlier) remain such as few outliers, dry bias (due to biased observation) in the lowest 2 km, and larger than expected C2 temperature error in the middle troposphere