Environment Canada Environnement Canada

Canada

# The assimilation of Radio Occultation Data at Environment Canada (EC)


Josep Aparicio, Godelieve Deblonde, Louis Garand, Stephane Laroche

Atmospheric Science & Technology Directorate

Environment Canada


ECMWF

Reading

16 June 2008

# EC analysis and forecast system I

- GEM Global Environmental Multi-scale Model
  - Global grid 800x600 or ~ 33 km resolution
  - 58 vertical levels
  - Model Lid at 10 hPa (~ 30 km)
  - Now under test, for operations in spring 2009:
    - 80 levels
    - "Hybrid" vertical coordinate
    - Lid at 0.1 hPa (~65 km)
- 4D-Var Incremental data assimilation system, with options to use
  - 3DVar
  - 3DVar with First Guess at Appropriate Time (FGAT)
- Drives a regional system
  - North America, ~ 15 km resolution
  - 58 vertical levels

# EC analysis and forecast system II

- Data assimilated (other than GPS-RO):
  - Radiosondes, aircrafts, surface, AMV,…
  - ATOVS (8 AMSU-A, 7 AMSU-B channels) –RTTOV8.7
  - GOES (1 channel -6.7 microns)
  - SSM/I (7 channels)
  - AIRS
  - QuikScat
- Dynamic (15 day sliding window) bias correction for radiances
- Assumed non-biased:
  - Radiosondes
  - Now also GPS-RO

# GPSRO Observables I : Options

- Level 1b
  - We have a 1-D operator
  - We are testing a 2-D operator
  - Vertical coordinate (impact parameter) is simulated from bg upwards from the surface.
  - Operator requires high model lid
- **Level 2 (Selected option)**
  - We have a 1-D operator
  - 2-D operator probably not feasible
  - Vertical coordinate (MSL altitude) is simulated from bg upwards from the surface.
  - **Operator accepts low model lid**
- Level 3
  - We have a quasi-local operator: $T_{Dry}(P_{Dry})$
  - Vertical coordinate ($P_{Dry}$) nearly identical to model coordinate. Simulated downwards from lid.
  - Operator accepts low model lid

# GPS-RO Observable I: The Refractivity

- Refractivity

Elasticity of electron clouds in atoms

$$N = 10^6 (n_{Air} - 1) = 77.6890 \frac{P_d}{T} + 71.2952 \frac{P_w}{T} + 375463 \frac{P_w}{T^2}$$

Induction of molecular rotation

- We assume the coefficients above (Rueger, 2002)

- We receive N(h) (i.e. Level 2 data)
  - N is a local quantity
  - However, h is not!
    - Model is built in p (or equivalent) coordinate, not h!
    - MSL Height is a nonlocal operator
    - Assimilation is nonlocal

# Observation operator (Level 2)

- We receive an array ($H_i$,$N_i$) (Height,Refractivity)
- We transform to ($GZ_i$,$N_i$) (Geop, Refractivity)
  - WGS84 Geopotential model
  - $H_i$ were already MSL (geoid not needed for H to GZ)

$$z = \frac{1}{g_0} \int_0^h g_{wgs84}(\lambda, h) dh$$

- The background provides (TT,LQ,PS)
  - Evaluate refractivity at each bg level: array **N**
  - Evaluate geopotential at each bg level: array **GZ**
  - $T_v$ virtual temperature
  - X=ln P

$$\Delta z = \frac{-R}{g_0} \left[ T_V \Delta \chi - \frac{1}{2} \frac{dT_V}{d\chi} (\Delta \chi)^2 \right]$$

- Find each $GZ_i$ within the bg **GZ** array
- Interpolate the bg **N** array to the observed $GZ_i$
  - Actually, better interpolate **log N**

# Observation operator (Level 1b, 1D)

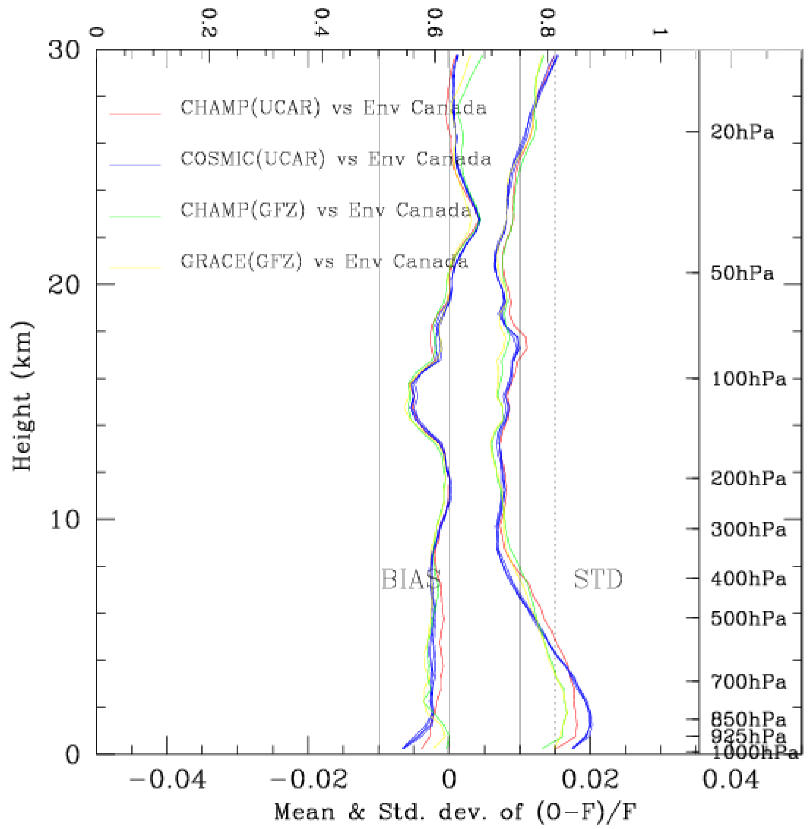
- We receive an array $(I_i,B_i)$ (Impact,Bending)
- The background provides (TT,LQ,PS)
  - Evaluate refractivity at each bg level: array **N**
  - Evaluate geopotential at each bg level: array **GZ**
  - Evaluate height at each bg level: array **H**
  - Evaluate impact at each bg level: array **I**
  - Evaluate bending at each bg level: array **B**
- Find each $I_i$ within the bg I array
- Interpolate the bg **B** array to the observed $I_i$
  - Actually, better interpolate **log B**

# Observation operator (Other)

- Derivatives evaluated through operator overloading
  - Automatic differentiation (analytical, not numeric)
  - Easier coding (very fast development, only fwd operator required)
  - Tangent, adjoint do not require new coding
  - Any code modification applies to all three operators
  - Gradients guaranteed
- For control vars CC=[TT, LQ, PS]
  - Quantity X: [X, dX/dCC]
  - Quantity Y: [Y, dY/dCC]
  - Quantity Z=X+Y: [X+Y, dX/dCC+dY/dCC]
  - Quantity Z=X*Y: [X*Y, X*dY/dCC+dX/dCC*Y]
  - Overloading allows writing only Z=X+Y or Z=X*Y
  - Chain rule evaluated without extra code

# The data I

- The different sources (satellites) of data are extremely similar (even if sampling is different)
- Differences between postprocessors are noticeable but small
- Shown all data in 2006.221-243 (last 3 weeks of Aug 2006) versus EC-Operational



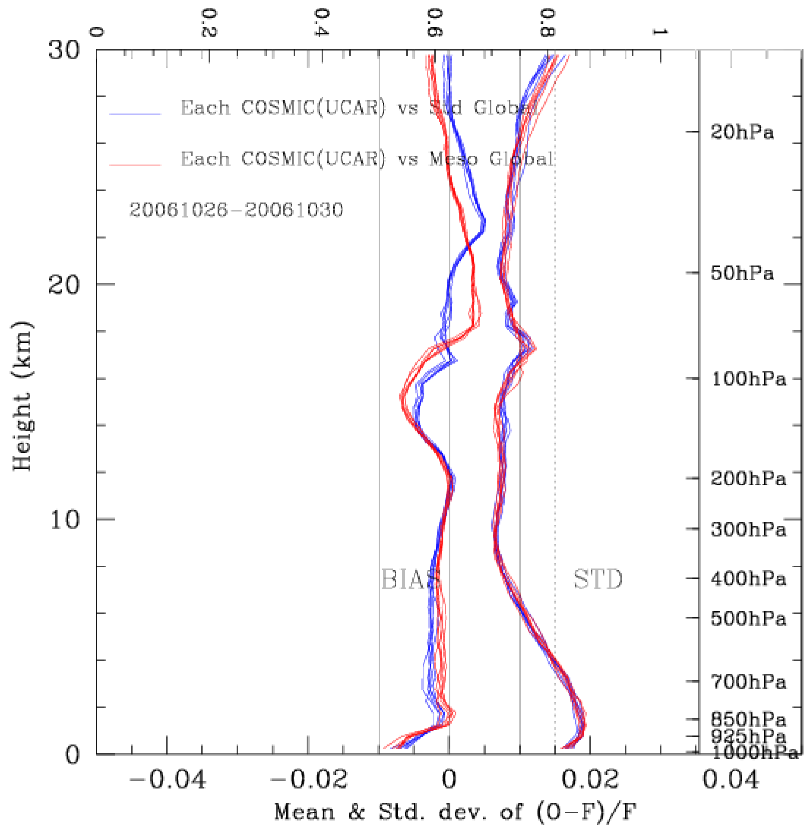



Environment Canada Environnement Canada

Canada

# The data II

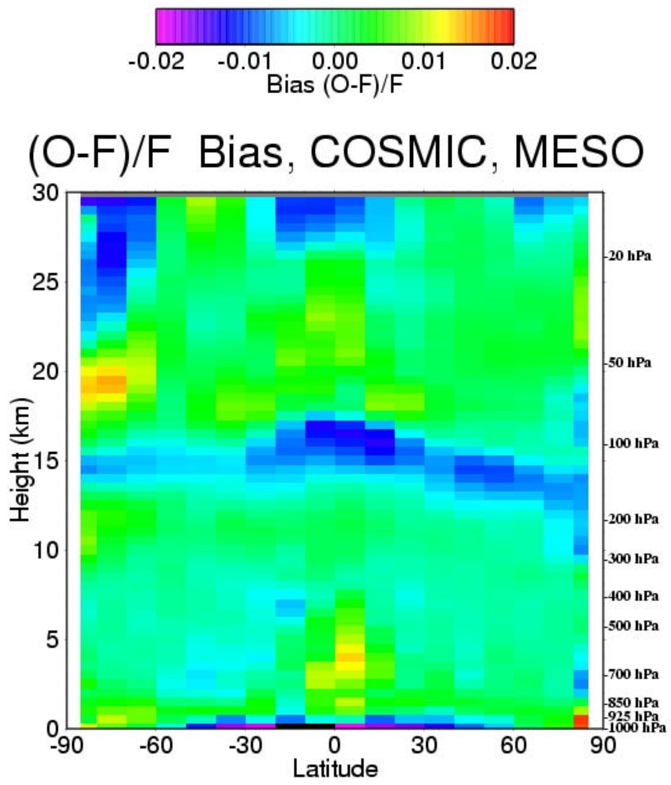
- Differences between models dominate the features
  - former operational 400x200L28 and current 800x600L58
- Each curve is shows data from one of the 6 LEO in COSMIC
- So
  - All LEO behave identically
  - All postprocessors are nearly identical
  - **Model issues are dominant over data issues**
  - **Only low tropo bias suspected to be real data bias**
  - **All RO data can be treated as statistically equivalent**



# The data III

- The observed bias (suspected to be largely model bias) is dominated by a tropopause feature.
- Qualitatively, all models show this tropopause bias
- A minor feature here will reappear later: areas of small negative bias in midlat tropo

- Shown: All Jan 2007 vs EC-Operational



# GPS RO data processing at EC (I)

- Refractivity profiles versus MSL height (Level 2)
- Data thinning:
  - No more than 1 profile within 45 minutes and 300 km.
  - No more than 1 observation/vertical km (approximately 1 datum out of every 5, as profiles are usually received with 200m resolution)
- Vertical Clipping: Use data
  - above 4 km
  - at least 1 km above background model surface
  - below background model lid (10 hPa, ~30 km)
- Background Error Check

  $-0.05 < (O-F6h)/F6h < 0.05$

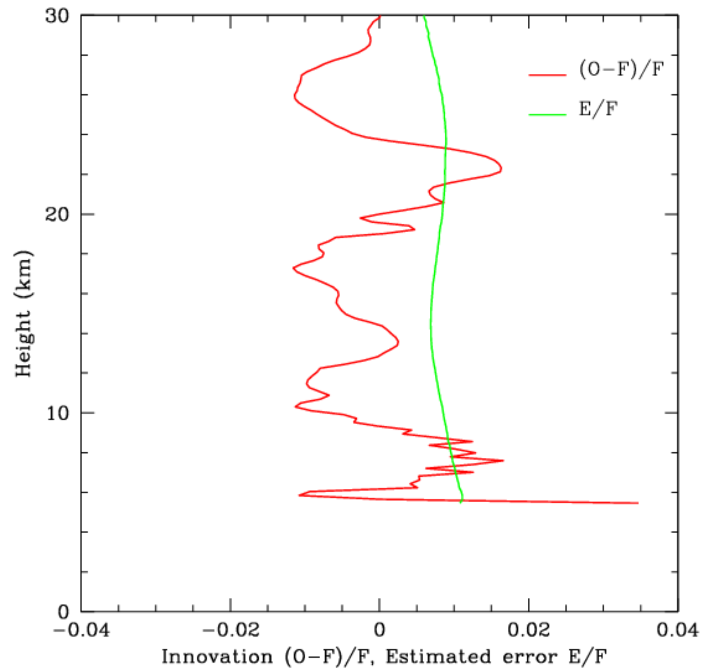
# GPS RO data processing at EC (II)

- Observation bias expected to be very small
  - no bias correction
  - data used as anchor to radiances
- Observation error: ~0.5-1.5%, known to vary with horizontal gradients (mostly with water content)
- Definition of observation error $E_i$: **adaptive dynamical**
  - O-F6h used online to estimate observation error
    - Within each profile:

    **Weighted RMS of O-F6h**

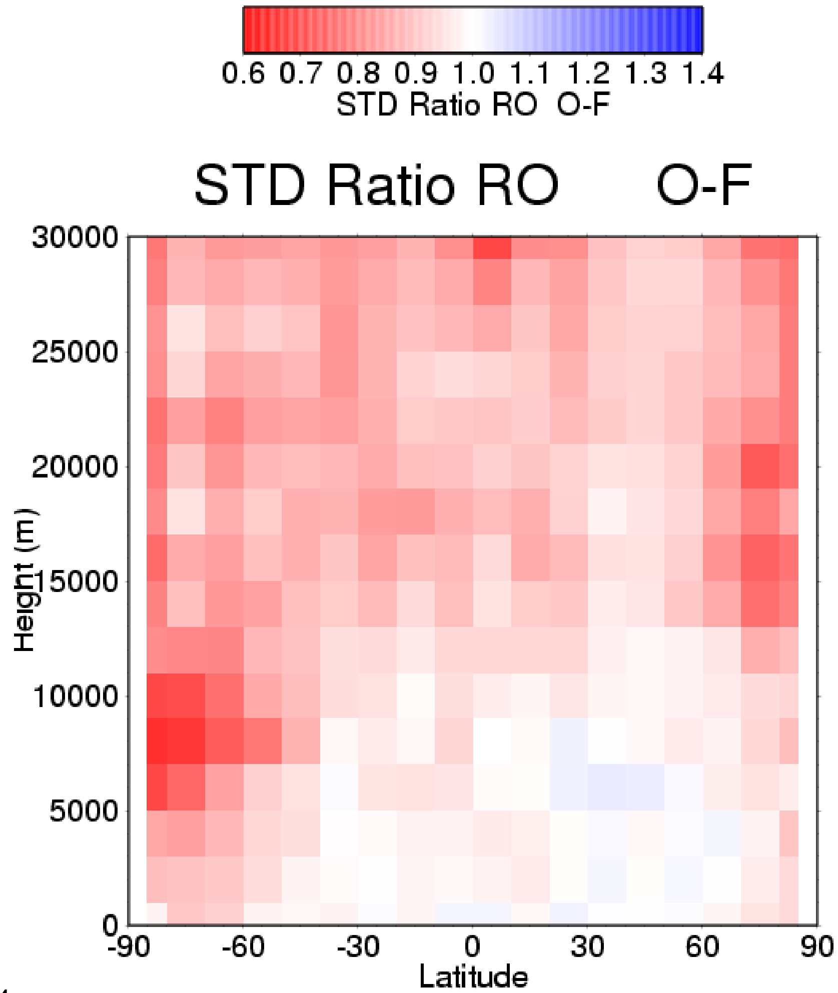
    Weighting: Gaussian, D=5km
  - This allows dynamical optimization
    - **Automatic adjustment of data weight**
    - Error tightens as cycle progresses (larger RO data weight ***if accuracy improves***)
    - Data weight automatically reduced if forecasts are less accurate

$$\left(\frac{E_i}{F_i}\right)^2 = \frac{\sum_j e^{-\left(\frac{h_i - h_j}{D}\right)^2} \cdot \left(\frac{O_j - F_j}{F_j}\right)^2}{\sum_j e^{-\left(\frac{h_i - h_j}{D}\right)^2}}$$



# GPS RO Impact on 6h forecasts

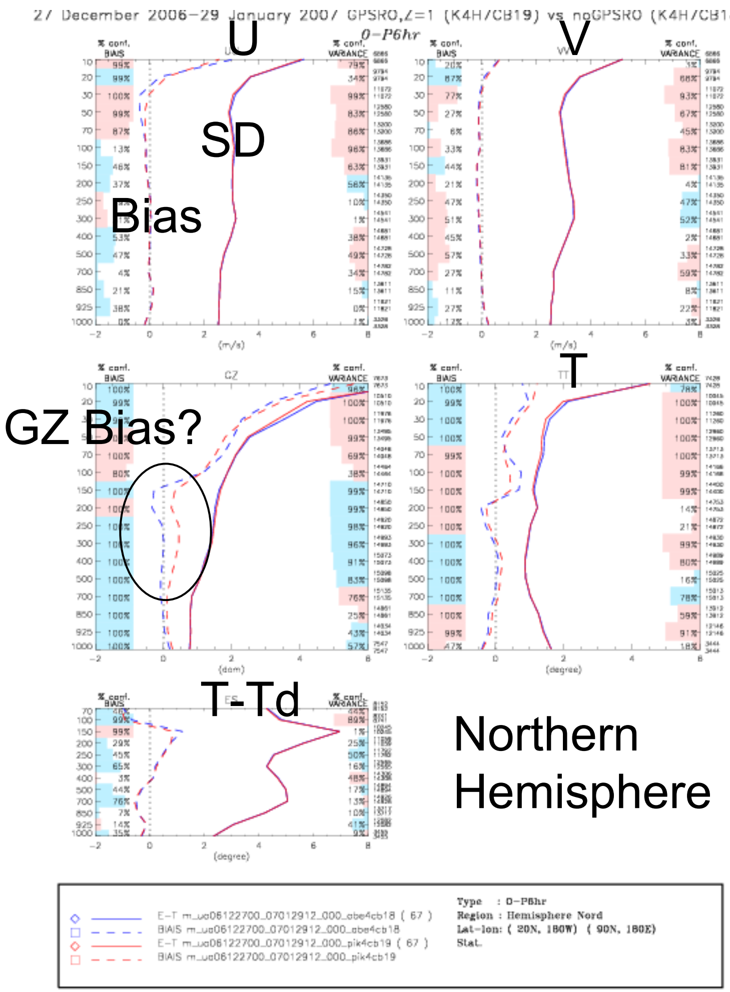
- ***TRUTH=GPS RO***
- Shown: ratio of refractivity forecast STD with/without assimilation of GPSRO
- Error in forecasting GPSRO observations reduced by 5-40% after assimilating RO.
- Red areas: improves with RO
- Blue areas: degrades with RO
- Similar in summer & winter

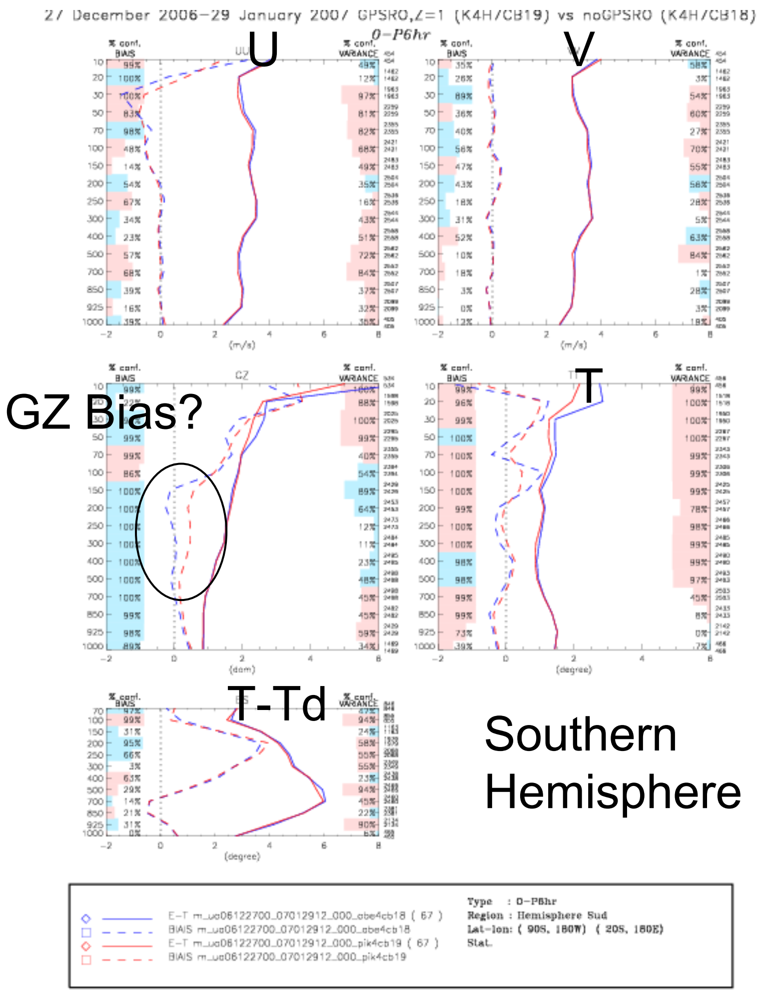


# GPS RO Impact on 6h forecasts Radiosondes (winter)

WITH GPS

WITHOUT GPS

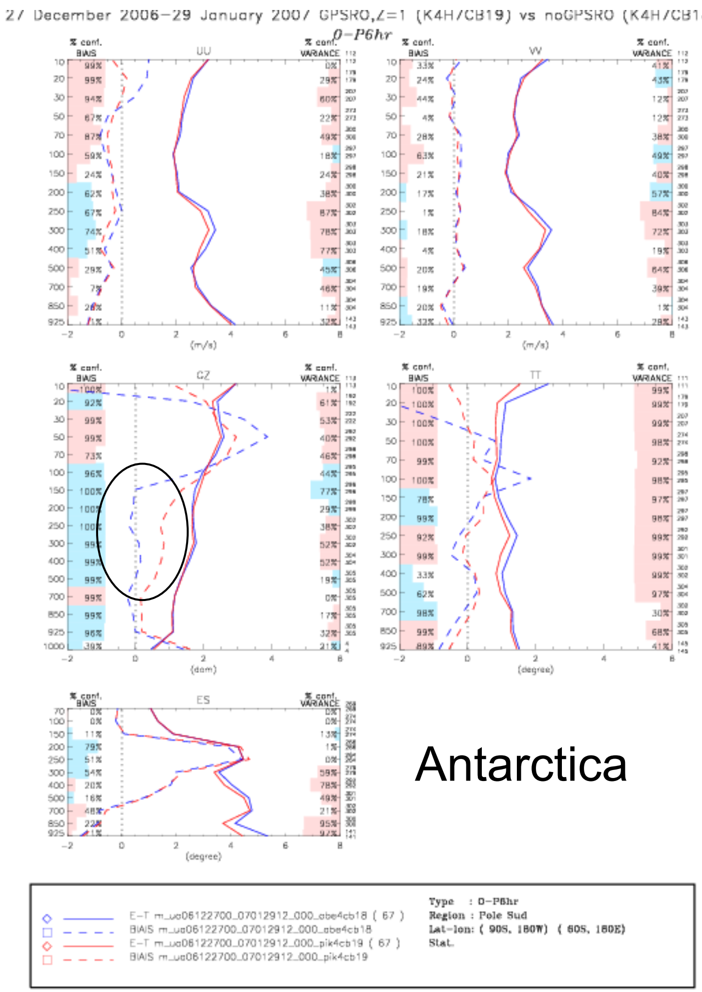


# GPS RO Impact on 6h forecasts Radiosondes (winter)

WITH GPS
WITHOUT GPS

- Very important impact in the Antarctica
- GZ Bias even bigger
- However, TT Bias and TT STD are improved
- Something is wrong with altitudes
- Even so, the effect is small (3-10m at 15 km) i.e. in the range 0.02%-0.1%



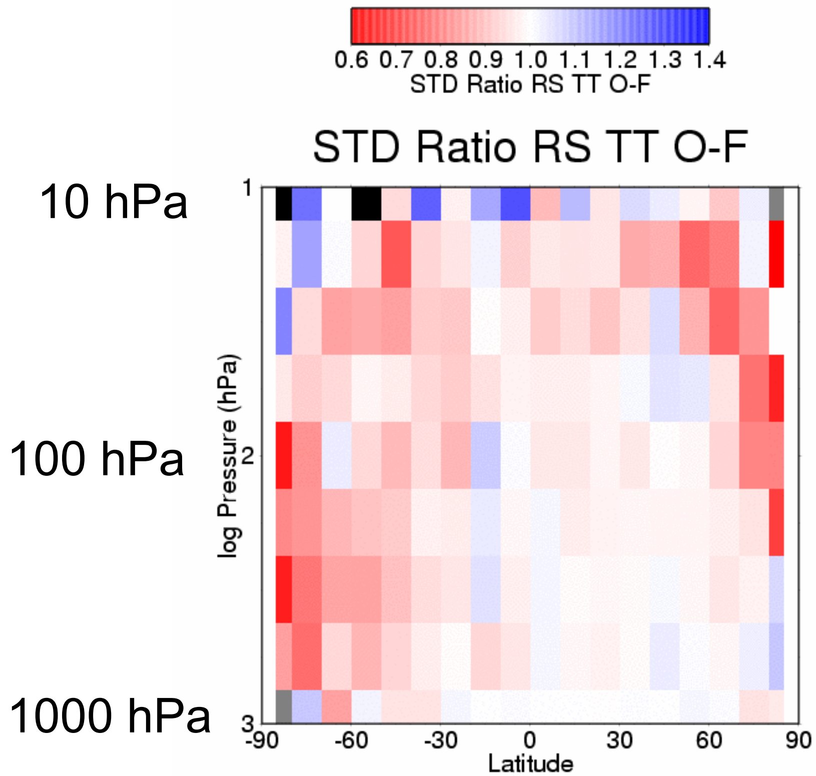



Environment Canada Environnement Canada

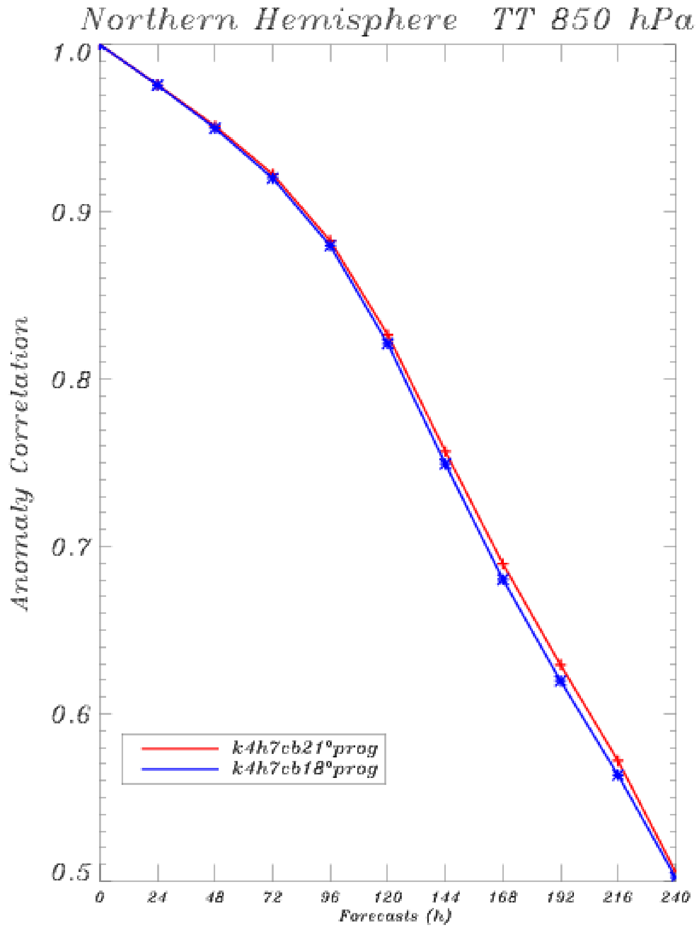
Canada

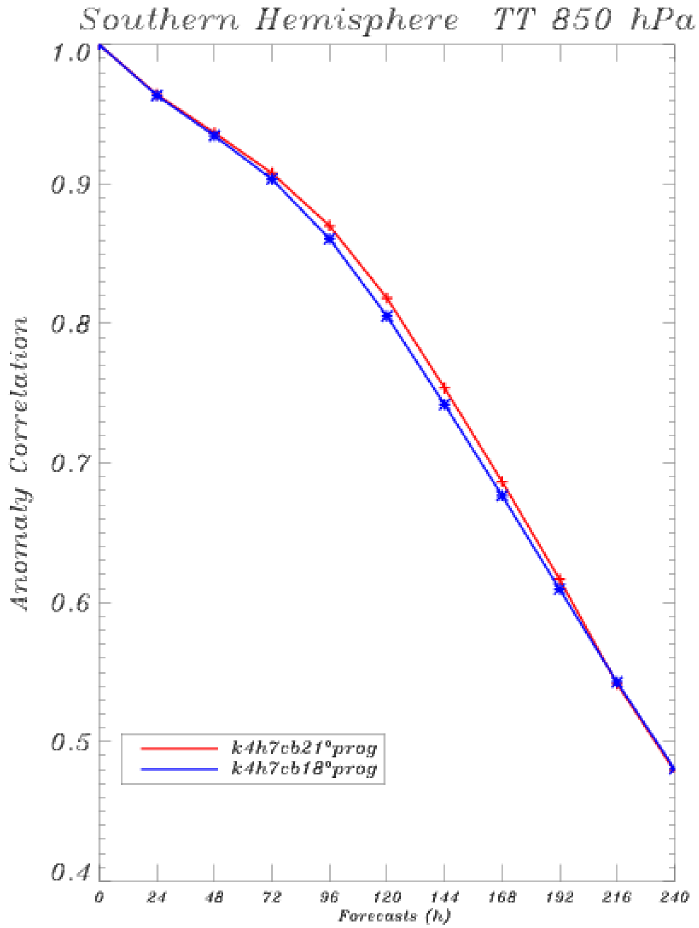
# GPS RO Impact on 6h forecasts Radiosondes

- ***TRUTH=Radiosonde Temperature***
- Error forecasting radiosonde temperature observations reduced on average by 5-10% after assimilating RO.
- Red areas: improves with RO
- Blue areas: degrades with RO

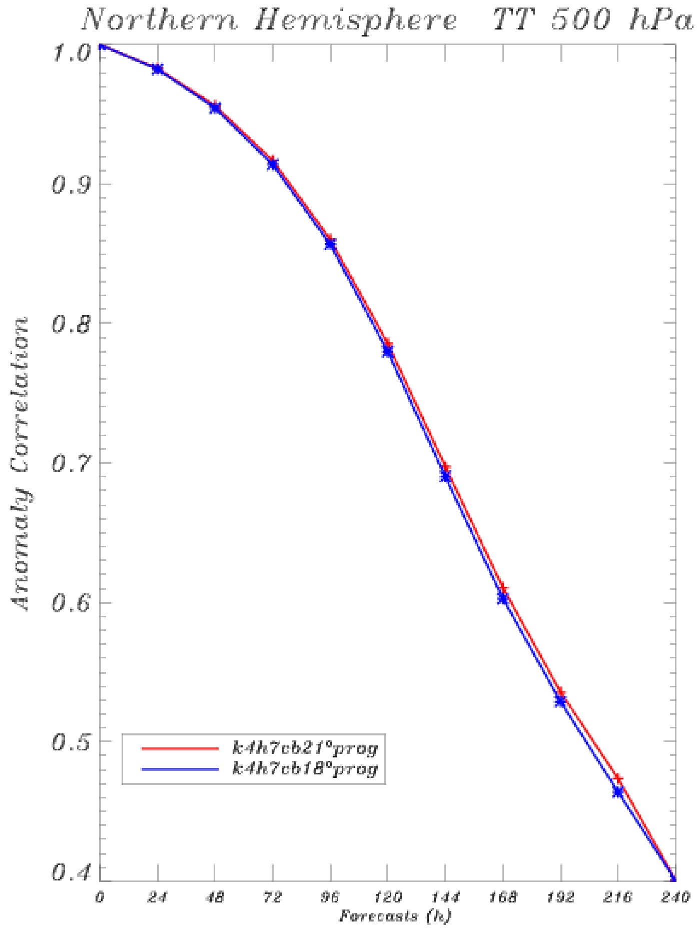


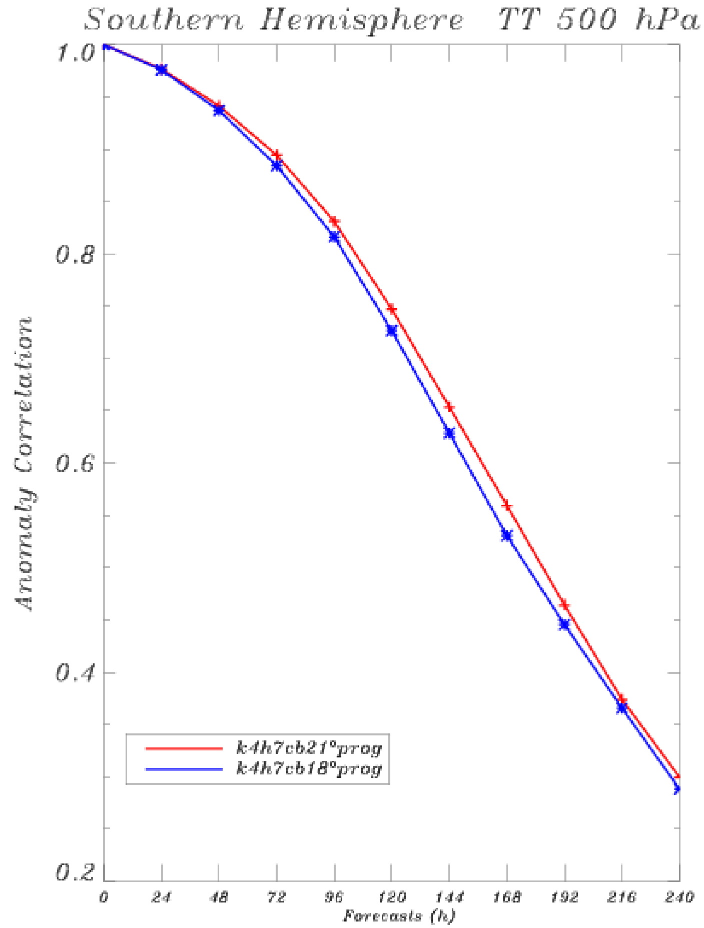
# GPSRO Impact on 0-10 day forecasts Anomaly Correlation: T at 850 hPa Winter



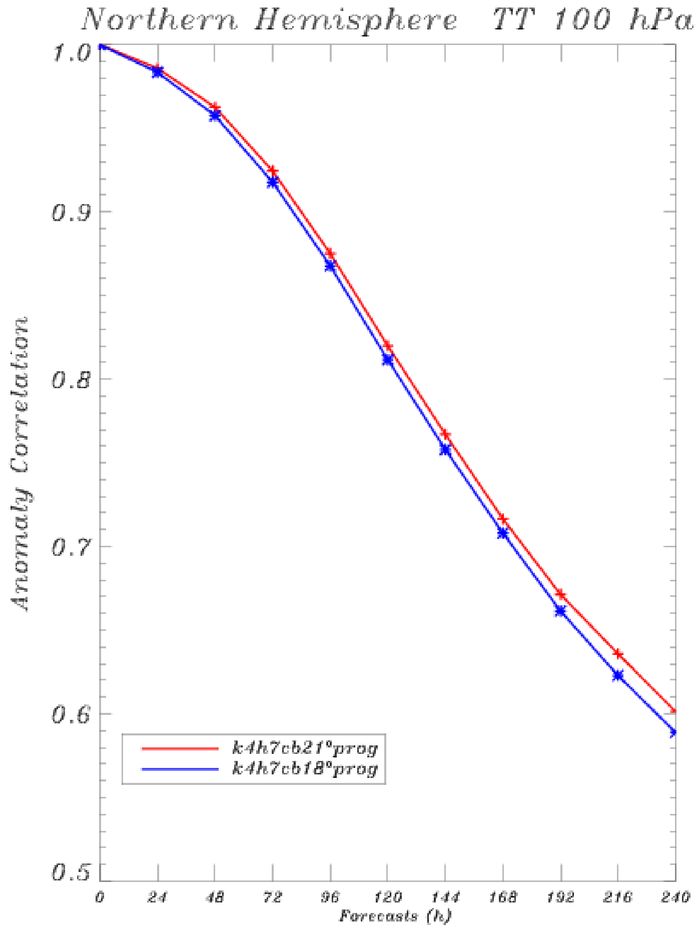



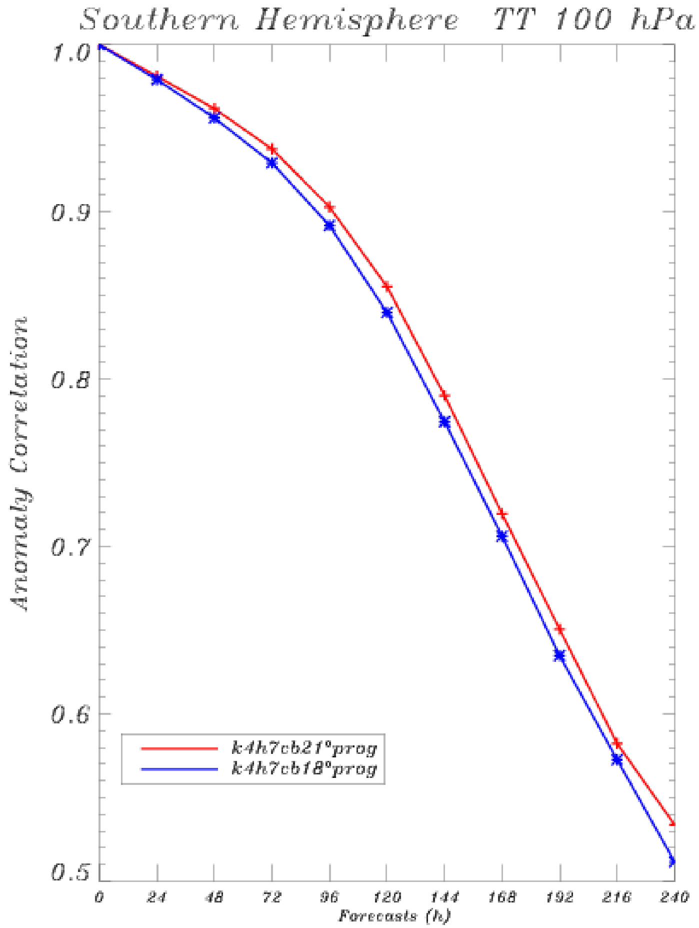
# GPSRO Impact on 0-10 day forecasts
# Anomaly Correlation: T at 500 hPa Winter



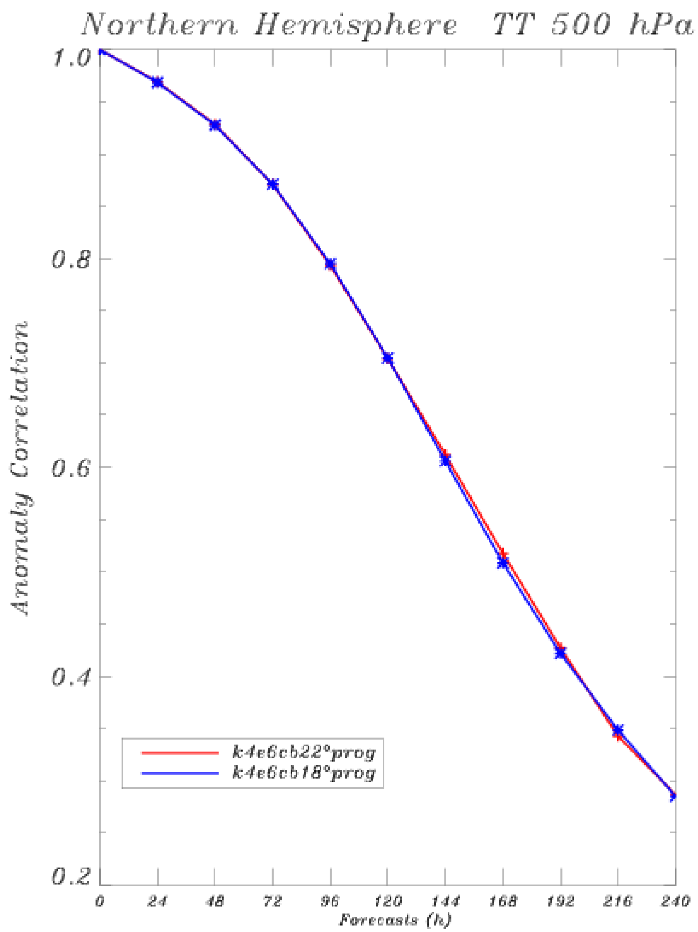



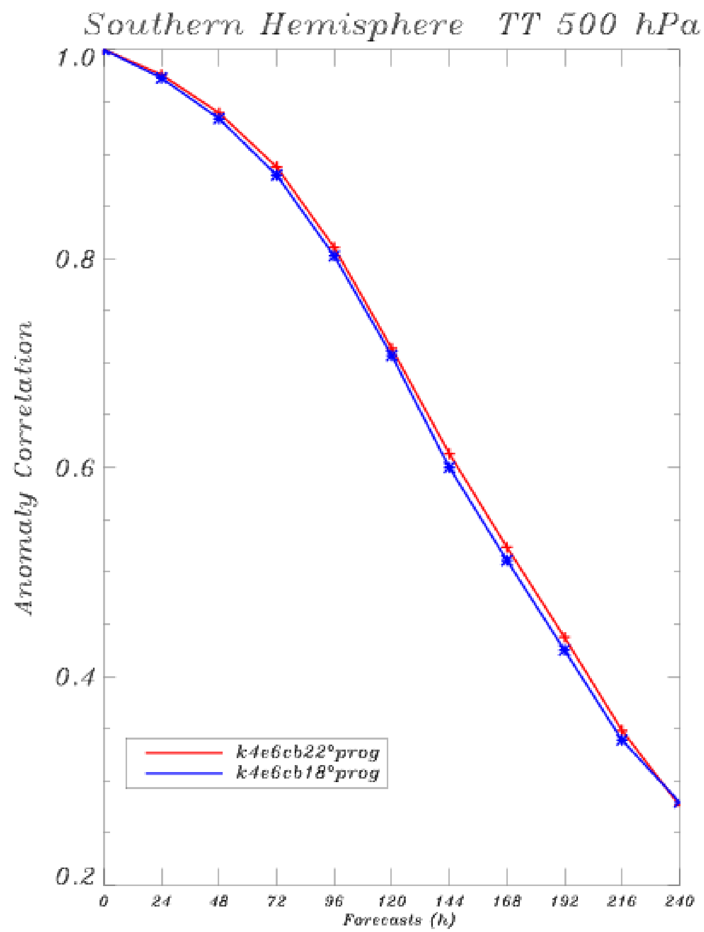
# GPSRO Impact on 0-10 day forecasts Anomaly Correlation: T at 100 hPa Winter



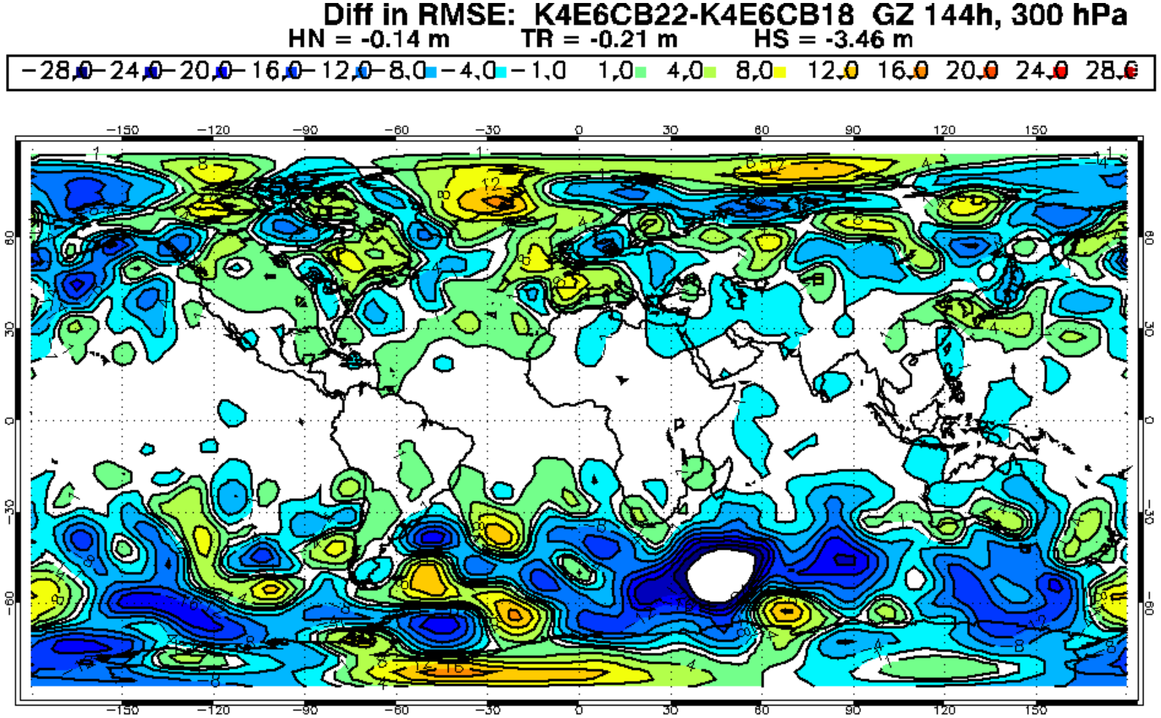
# GPSRO Impact on 0-10 day forecasts
# Anomaly Correlation: T at 500 hPa Summer



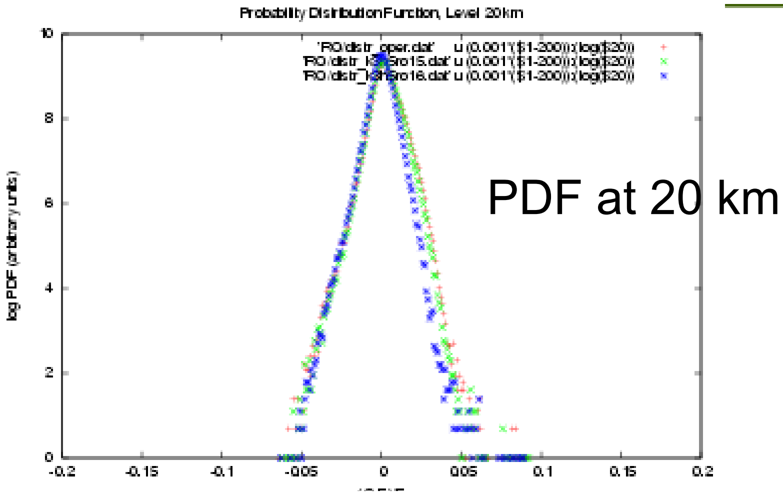
# GPS RO Impact on day 6 forecast
# Difference in RMS
# GZ at 300 hPa

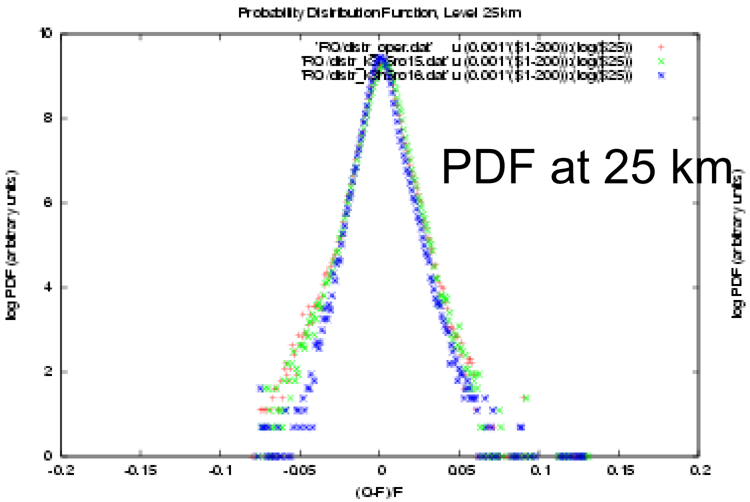
- ***TRUTH=Analysis***
- Summertime (2 mo)
- Blue areas: improves with RO
- Important in midlat
- Improved RMS in the 3 regions (HNorth, HSouth, TRopics)
- Small in tropics
- Major in S Hem

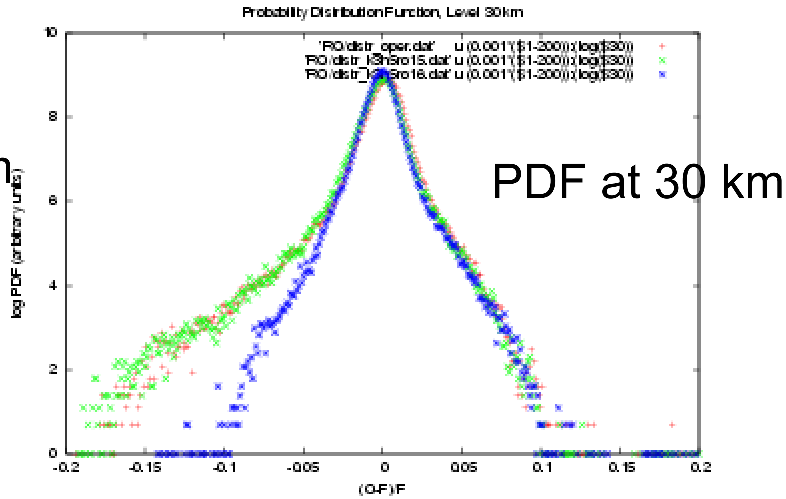


# GPS RO impact on extreme deviations

- GPSRO reduces the probability of extreme deviations
- Shown are Probability Distribution Functions (log) in stratosphere
- Red: EC Operational
- Green: Control (no GPSRO)
- Blue: GPSRO



# Conclusions up to this point

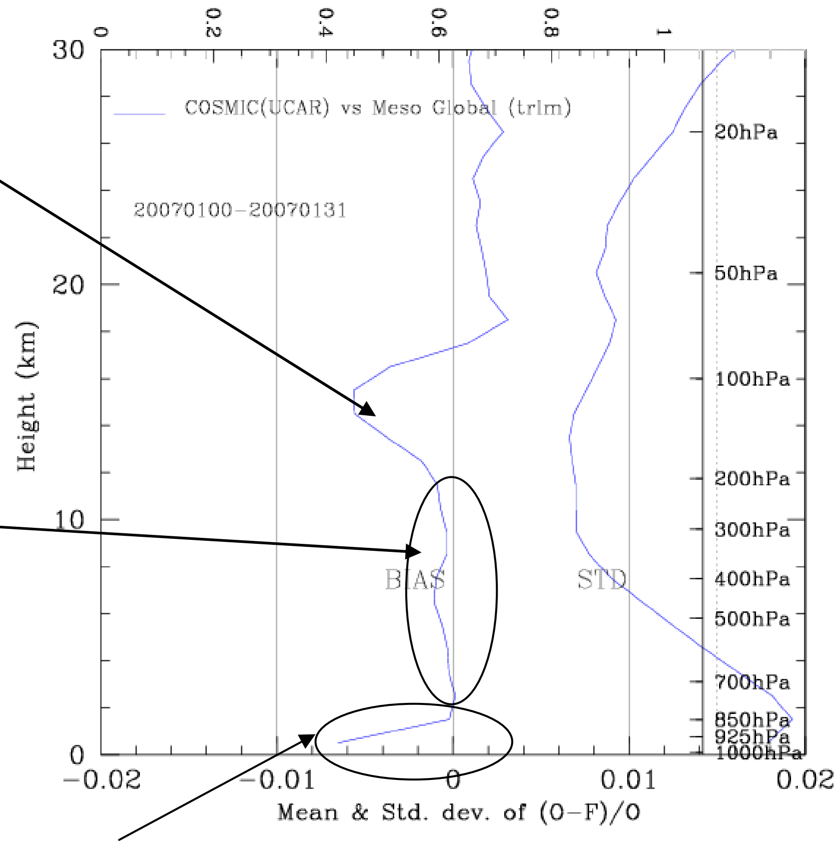
- Data arrives promptly in NRT. Quality is good.
- Assimilation of these data is in general positive
- Better temperatures, winds
- Particular coverage improvements in S Hem
- We expected more impact in tropics and moisture fields, but impact is small there (although not negative)
- Substantial (!!) impact in Antarctica (major data source)
- If everything is so nice… why does GZ have problems?

# The altitude mismatch

- RS indicates that there is an altitude mismatch of 0.02%-0.1% after GPSRO assimilation
- It is not big, but it is statistically significant
- GPSRO indeed uses a different vertical scale:
  - Nearly all measurements are vertically located in pressure (P)
  - Model, assimilation etc also work all in P-like coordinate
  - GPSRO is located in geometric altitude (H)
  - A mismatch P vs H of (fractional) ~5e-4 can lead to this GZ bias
  - *Many things in the chain were not designed with that accuracy*

# The bias in observation space I

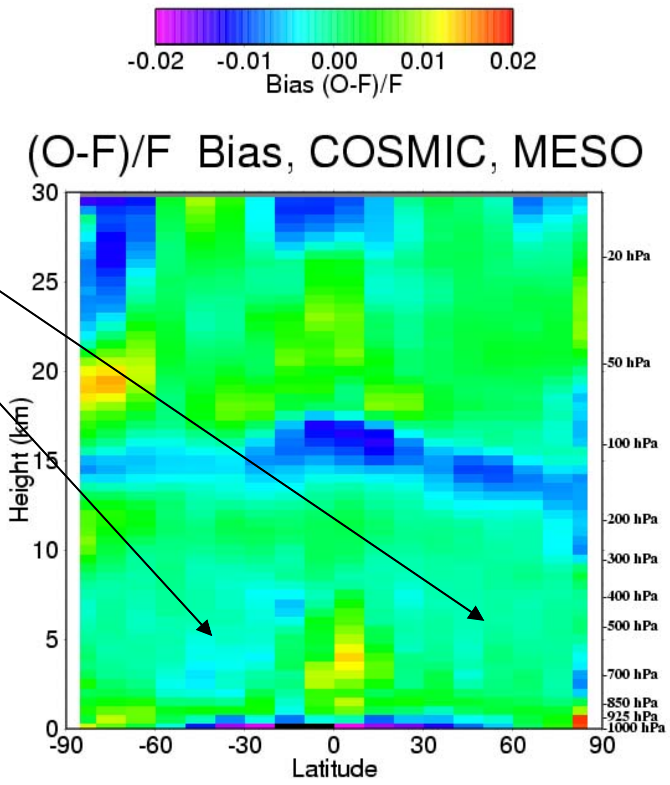
- The tropopause bias is NOT responsible (GZ bias appears even assimilating only below 12 km)
- The source of the GZ bias is a small tropospheric bias in refractivity. The forward model, when applied to the background field systematically forecasts a refractivity too big by ~5e-4.
- Effect is not noticeable in stratosphere, as there are other errors of bigger magnitude




Data below 4 km is not assimilated

# The bias in observation space II

- The latitude distribution actually showed this.
- Areas of small negative bias, sp in midlat troposphere
- Shown: All Jan 2007 vs EC-Operational



# The origin of the O-F bias

- The P vs H relationship is based on the hydrostatic equation

$$\frac{dp}{dH} = -\rho(H)\cdot\gamma(\lambda, H)$$

- And the ideal gas equation of state

$$p = \rho \cdot R_d \cdot T_v$$

- So (WMO standard)

$$dH = \frac{-1}{\gamma(\lambda, H)} \cdot R_d \cdot T_v \cdot d[\ln p]$$

$$dGZ = \frac{-1}{\gamma_0} \cdot R_d \cdot T_v \cdot d[\ln p]$$

- But a real gas is not ideal (small attraction/repulsion between molecules)

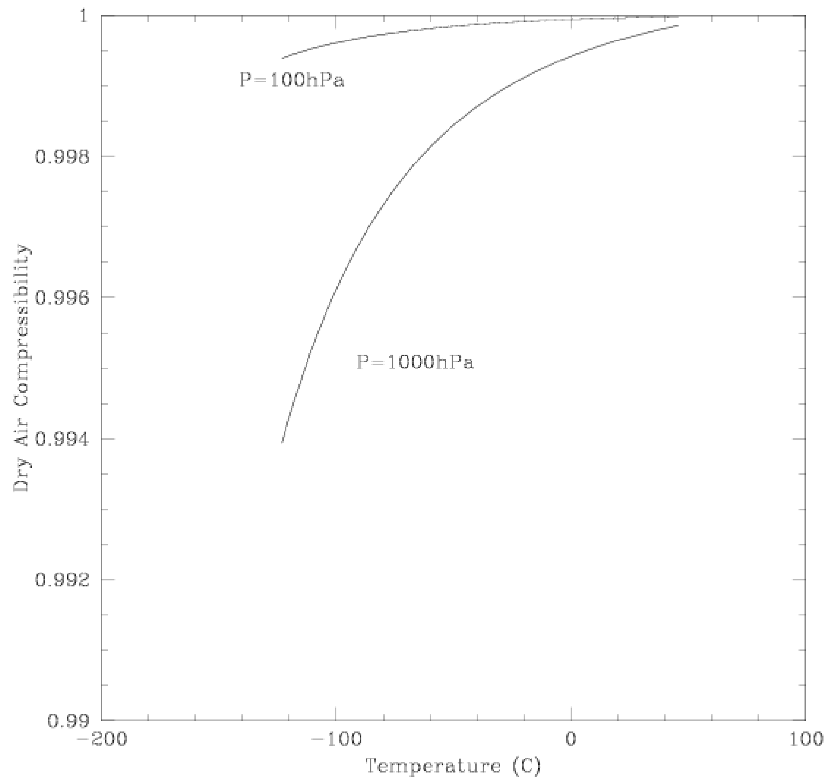
$$p = \rho \cdot R_d \cdot T_v \cdot Z$$

Should be added

Z~0.9990-0.9998 for air

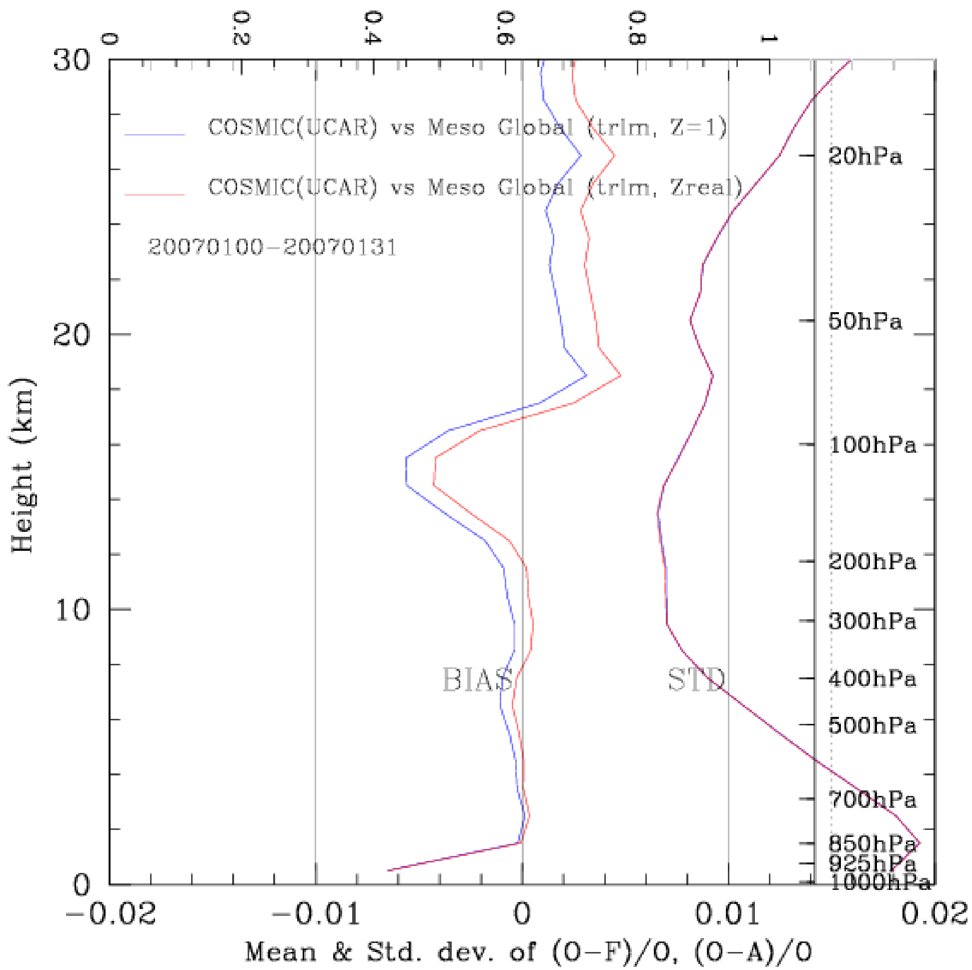
# Order of magnitude of the compressibility (Atmosphere)

- For dry air, Z-1=0 around 77C
- <1 at normal atmospheric conditions
- **Smaller with low temperature**
- ~ proportional to pressure
- Z(T) explains why GZ bias was larger in polar regions
- Effect never larger than ~0.1% for atmospheric P,T

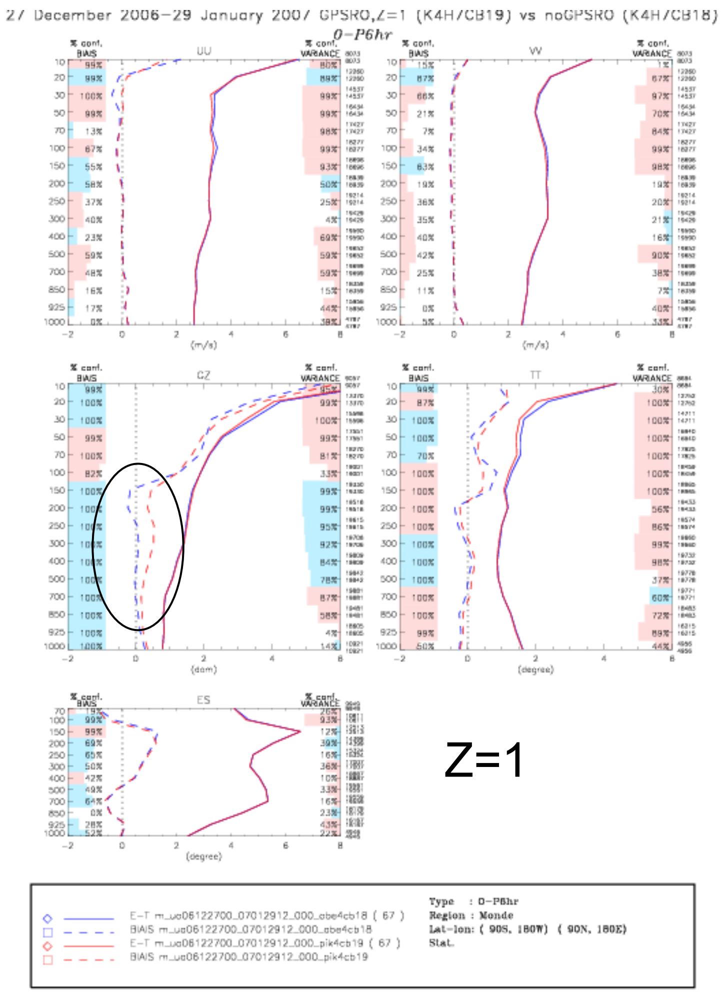


# Bias in observation space II

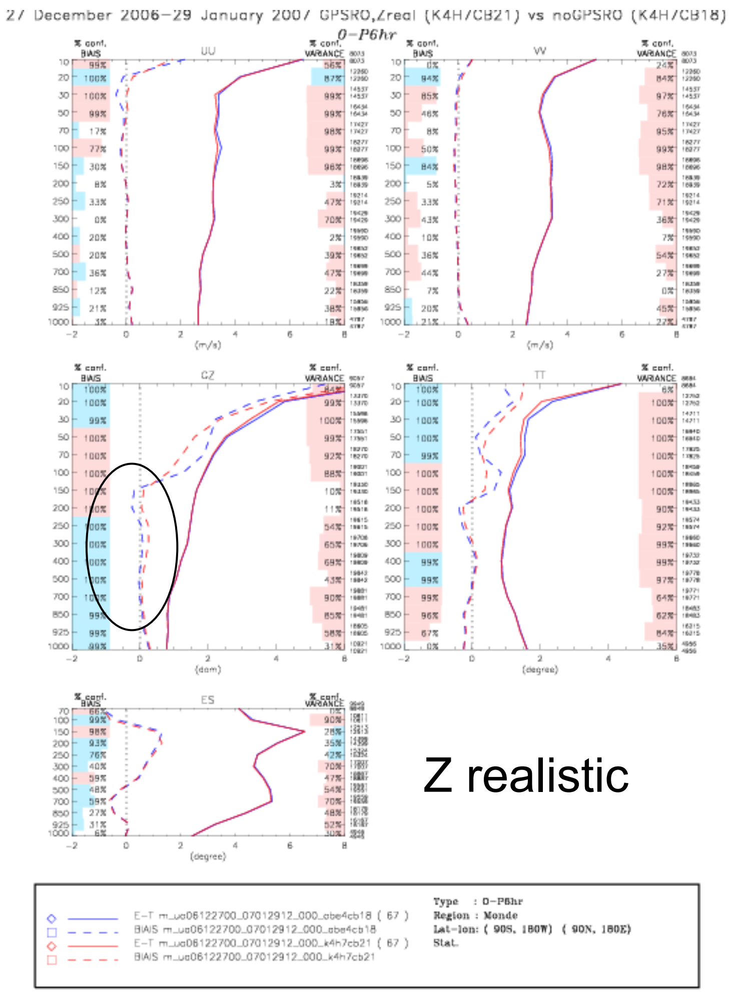
Once included the compressibility Z the bias is very small



# RS statistics after accounting for the compressibility (Winter, World)




Z=1




Z realistic

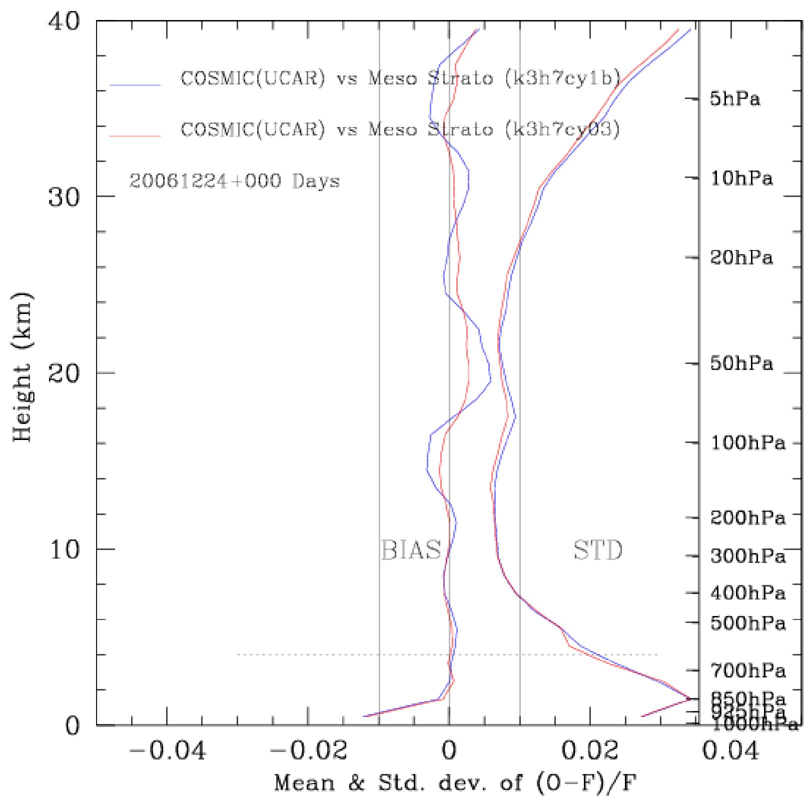
# Summary

- Impact of GPSRO data positive in general, especially in the stratosphere.
  - Verified against GPS RO:
    - Generally positive. Very positive in high troposphere & stratosphere.
    - Moderately in tropics.
  - Verified against Radiosondes:
    - Weak but positive impact in N Hemisphere
    - Very positive in S Hemisphere, especially high latitudes
  - Verified against analysis (e.g. anomaly correlations):
    - Weak impact in N Hemisphere
      - Neutral in troposphere
      - Moderately positive in stratosphere
    - Very positive in S Hemisphere, especially high latitudes
- No regions with degradation
- Tropical troposphere has been the most challenging
- Implemented dynamic weighting of GPSRO observations.
- Compressibility factor of air found not to be negligible. A GZ bias appears if it is not accounted for.

# Postscriptum: High-lid new model

- Winter 2007 (2 months)
- Bias & STD improved when measured against **GPSRO truth**
- Order of magnitude rule of thumb:
  - 0.01 corresponds to ~2.5K
- Blue no GPS assim
- Red with GPS assim

O-F6h for GPSRO data






Environment Canada Environnement Canada

Canada

# Some future lines

- High lid model allows L1b data
  - Desirable for low-altitude data
  - Not for high altitude (very sensitive to gravity waves)
- Can data be split?
  - Low altitude at L1b?
  - High altitude at L2?
- L1b at 1-D is only a marginal improvement
- L1b in 2-D?
  - Bending angle: integral with **Gaussian kernel**

$$\alpha = \int_{Path} \frac{\nabla n}{n} ds = \int_{Path} A(s) e^{-(s-s_{Tangent})^2 / R \cdot H} ds$$

  - A(s) is a slowly-variying function (but would contain h-gradients)
  - 1-D equivalent to Hermite quadrature of order 1
  - 2-D operator: Hermite quadrature of low order
  - Cost overhead can be as low as a factor of 2-3.

# *Thank you!*