SFP^2 : Evaluation of Allsky Single Footprint retrievals using Artificial Intelligence methods and Optimal Estimation methods

Sergio DeSouza-Machado[1], Eric Maddy[2], Ruben Delgado[3]

September 24-27, 2024

[1] UMBC JCET

[2] Riverside technology

[3] Hampton University

Outline and Motivation

- Compare boundary layer parameters from single footprint (SFP) retrievals versus L2 cloud cleared retrievals
- We are focusing primarily on boundary layer height
- SFP^{AI} Eric has developed a very fast Artificial Intelligence Single Footprint Retrieval (last 3 years or so)
- SFP^{OEM} Sergio has been upgrading his Optimal Estimation Single Footprint Retrieval
- Ruben estimates boundary layer height (PBLH) based on using wavelets to study transitions of profiles (*T(z),WV(z),lidar backscatter(z)*) near the surface

SFP^{AI} : Multi-Instrument Inversion and Data Assimilation Preprocessing System - AI, MIIDAPS-AI

- Fundamentally MIIDAPS-AI is a neural network - several layers deep, 100s-1000s of hidden units per layer, which runs on GPU or CPU
- Given a set of millions of satellite radiometric observations, ***MIIDAPS-AI is trained to produce an estimate of the atmospheric state***
  - Thermodynamics such as Temperature, moisture, SKT and clouds parameters (ice/water cloud height, optical depth).
  - Surface or near surface parameters such as emissivity, surface temperature, surface pressure, near surface winds, *PBL height*.
- Target atmospheric state provided by Global models such as ***FV3GFS/GDAS*** (this study) or ECMWF; Regional models such as HRRR, RRFS; etc.
- ***We have successfully applied MIIDAPS-AI to passive IR, MW, IR+MW, as well as active MW sensors in simulation and with real data.***
  - Maintains the accuracy of traditional approaches
  - Is computationally efficient - ***On a single Linux CPU server, and using all full spectral resolution CrIS channels, we can generate > 60,000 profiles per second in all-surface, all-weather conditions; even faster on GPU***

MIIDAPS-AI Is Sensor Agnostic and has been demonstrated on Operational and Research Satellite Data

Instrument	Modes of Operation	Status	Observation Types
ATMS	Remote Sensing, DA/DF	Validated	Simulated and Real Obs
ATMS/CrIS	Remote Sensing	Validated	Simulated and Real Obs
CrIS	Remote Sensing	Validated	Simulated and Real Obs
AMSU-A/MHS	Remote Sensing, DA/DF	Validated	Simulated and Real Obs
ATMS/AMSU-A/MHS	DA/DF	Validated	Simulated and Real Obs
AIRS	Remote Sensing	Provisional	Simulated and Real Obs
TROPICS PATHFINDER	Remote Sensing, DA/DF	Provisional	Simulated and Real Obs
TEMPEST-D	Remote Sensing, DA/DF	Validated	Simulated and Real Obs
ATMS/AMSU-A/MHS/TEMPEST-D	DA/DF	Validated	Real Obs
HyMS ($\nu \sim$ 1GHz – 330GHz at $\Delta\nu \sim$ 1MHz – 100MHz)	Remote Sensing	Validated	Simulated
GOES ABI	Remote Sensing	Experimental	Simulated
GPM DPR Ku/Ka, GPM DPR/ATMS, GPM DPR/TROPICS	Remote Sensing, DA/DF	Experimental	Simulated and Real Obs
EPS STERNA	Remote Sensing	Experimental	Simulated
TMS	Remote Sensing	Experimental	Simulated

SFP^{OEM} : Physical Retrieval

- Initialized using NWP thermodynamics and “best cloud swap”
- Can now read in Eric’s retrieved parameters; hence SFP^2
- Have developed SARTA-Cloudy Analytic Jacobian recently
- Typically use all 97 layers or so for retrievals
- Takes less than about 1.5seconds per retrieval (faster if initialized by Eric’s clouds)
- Retrieves and output thermodynamic info (SKT, T(z),WV(z)) and cloud parameters (ice/water cloud height, optical depth)
- Plan to do error analysis characterization (AI vs OEM)
- Same code is used for AIRS, CrIS, IASI (channel information, NeDT etc is updated)

Some results of *SFP*^{2}

Have run off a granule for 2019/04/15 and 2022/09/03; have the following datasets for JPSS CrIS

- ECM initialization
  - retrieval from ECM first guess (with cloud swap)
- AI CrIS only retrieval
  - retrieval from ECM cloud swap, AI thermodynamics
  - retrieval from AI cloud (no cloud swap), ECMWF thermodynamics
- CLIMCAPS

QC filter : profile discarded if ST,T(800 mb) or WV(800 mb) is more than 2σ away from mean/stddev ST,T(800 mb),WV(800 mb) of ECMWF

Validation data : NASA Langley High Altitude Lidar Observatory (HALO)

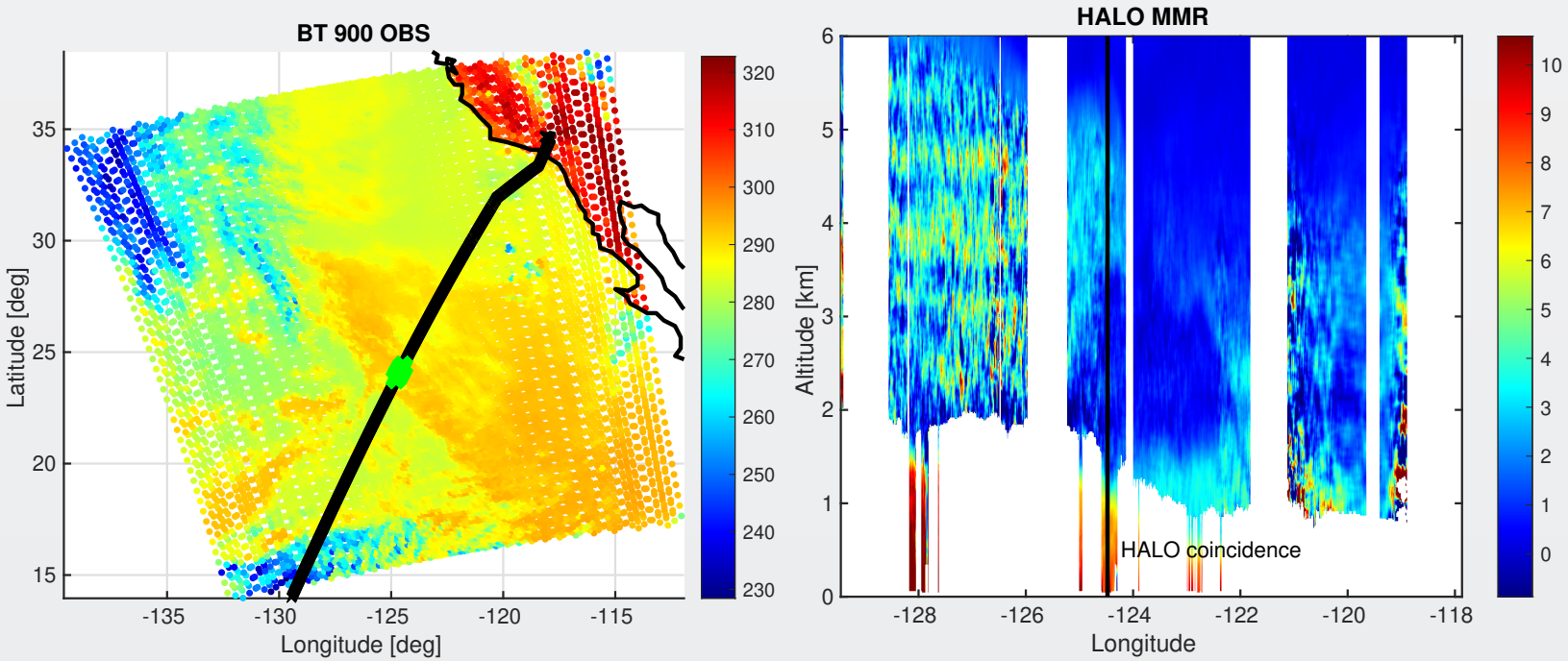
Computing PBLH

Preliminary work : we can use different methods to compute the heights

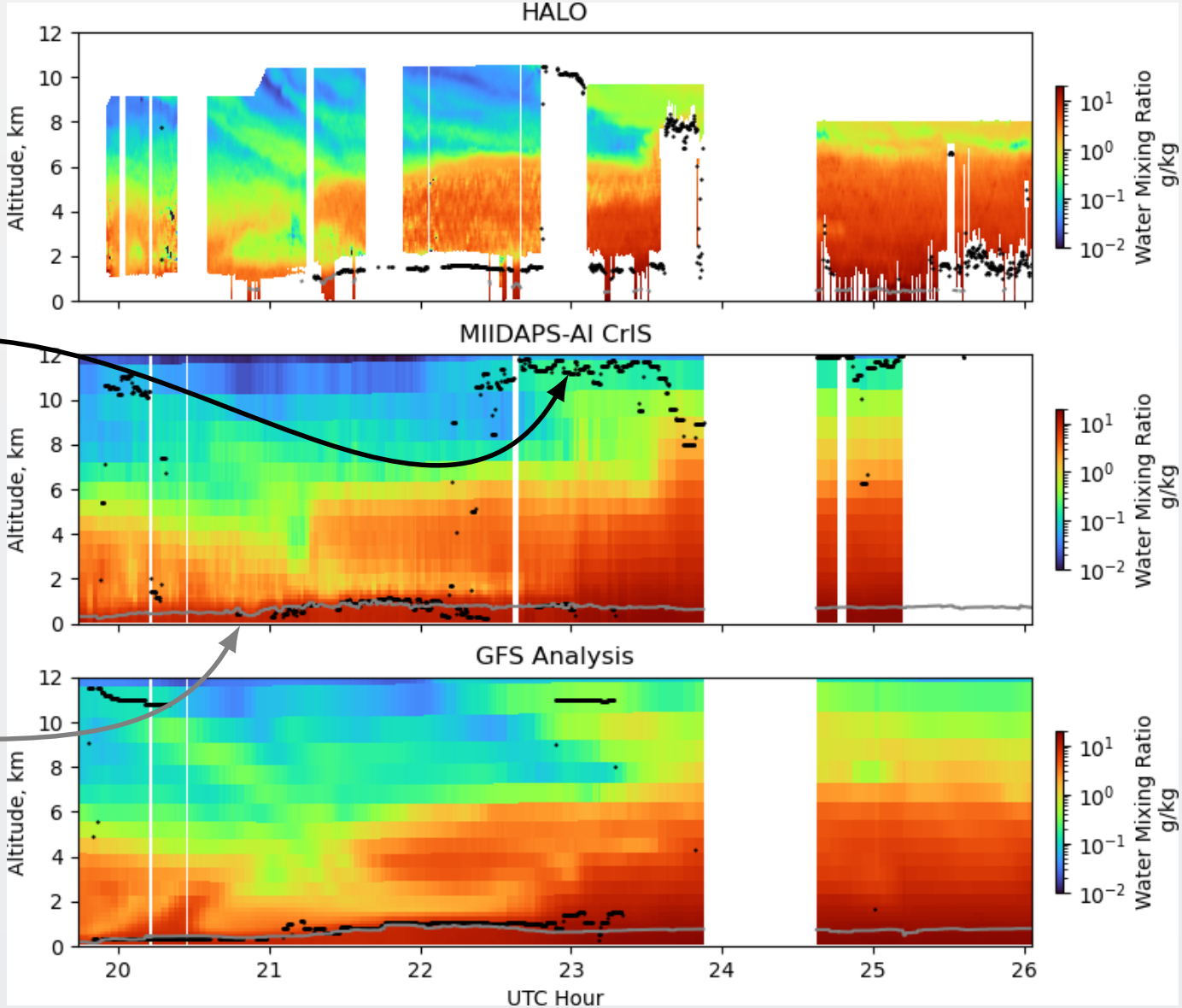
- OEM (Sergio)
  - Use AIRS 00 layers (250 m vertical spacing in lowest few layers)
  - look for lowest extrema (max or min as needed) of *RH(z),SH(z),Tvirtual(z),Tpotential(z)*
- AI (Eric)
  - GFS uses windspeed to compute PBLH whenever Richardson Number > 0.25
  - MIDAPS-AI can directly retrieve PBLH only (targeted model output) or from *d/dz(T/Q)* profiles as done by SFP^{OEM}

JPSS 2019/04/25 g214

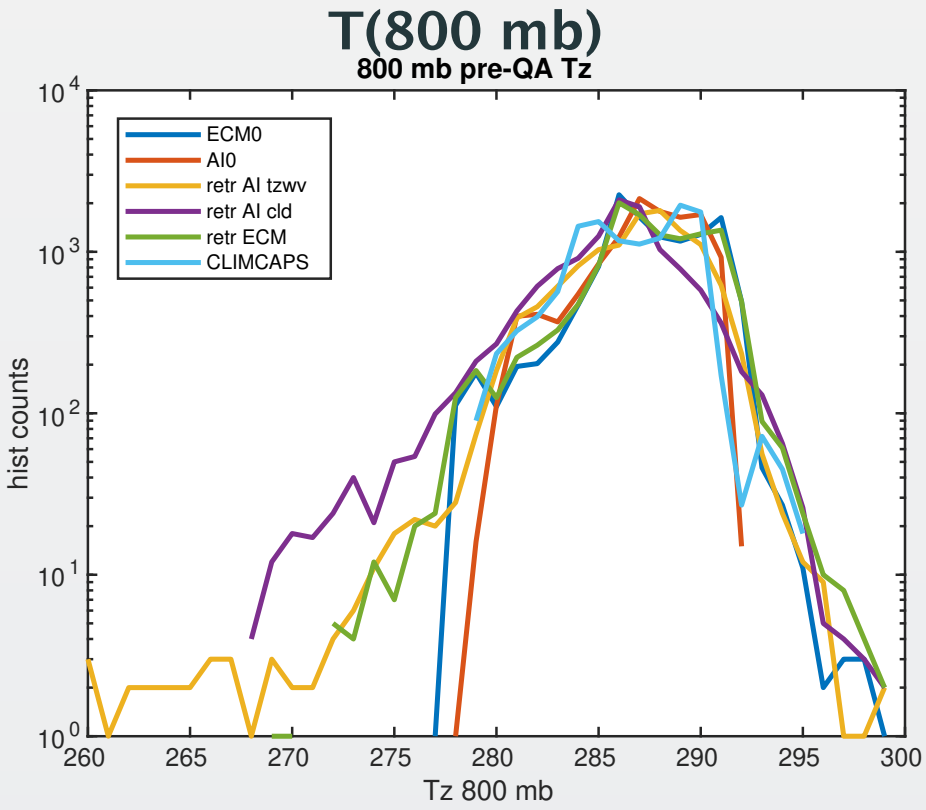
Worked with d/dz (and d/dp!!!) RH, WVgg, Tvirtual,T
Code finds max (or min) derivative closes to ground
Black is HALO track, green is JPSS/HALO coincidence

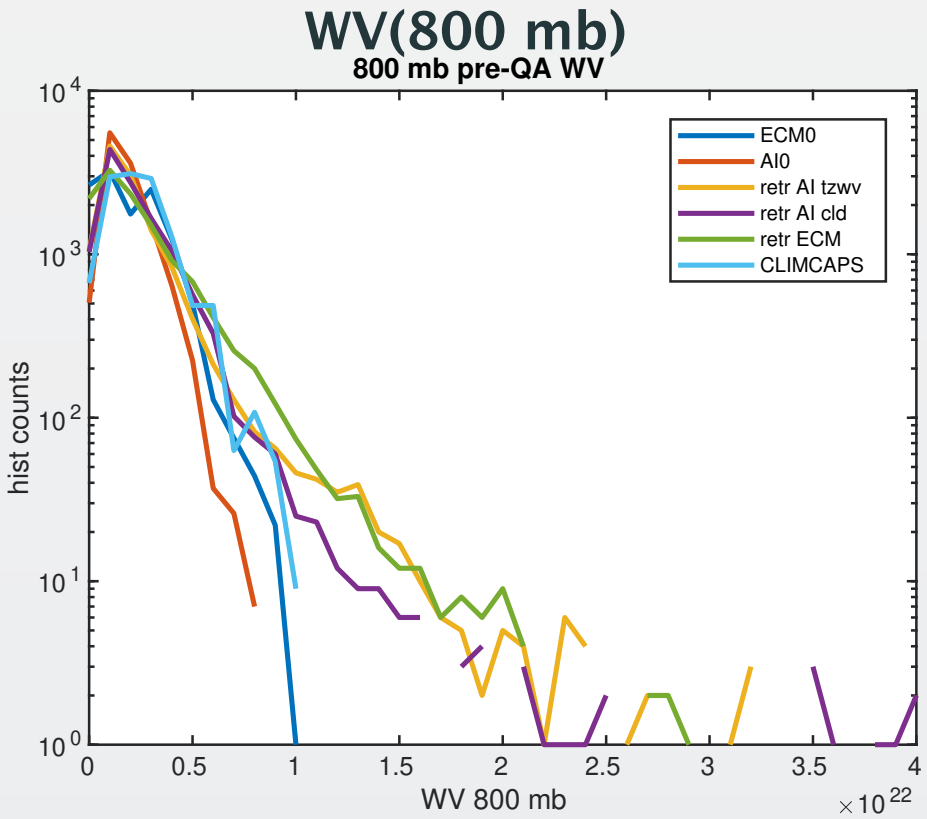


HALO 2019/04/25 vs AI

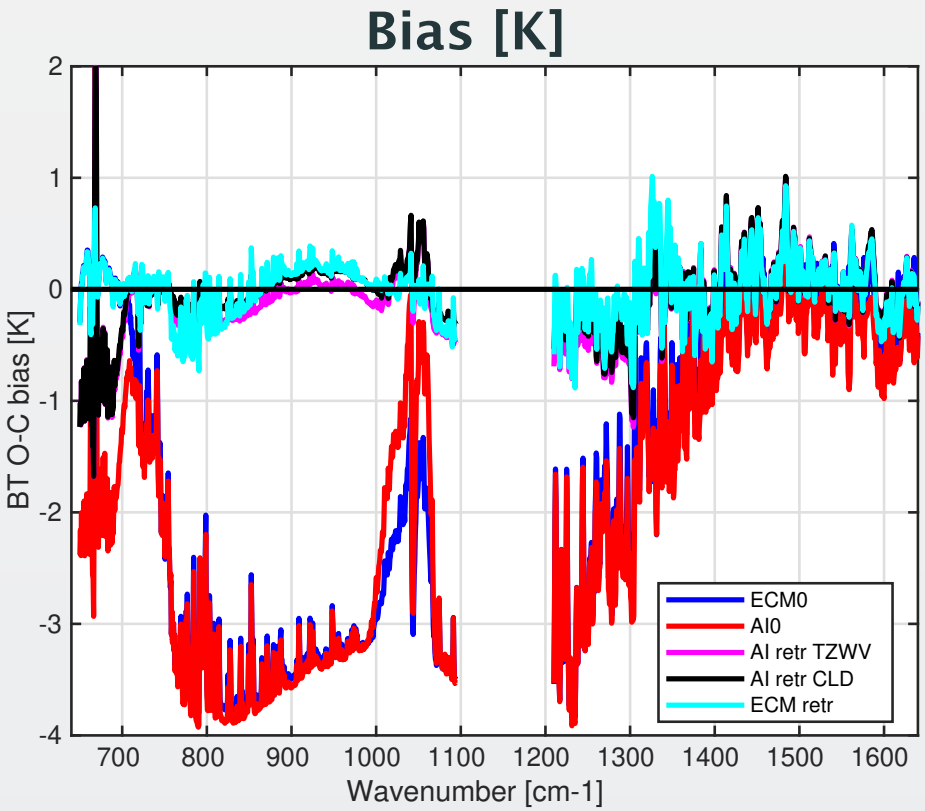


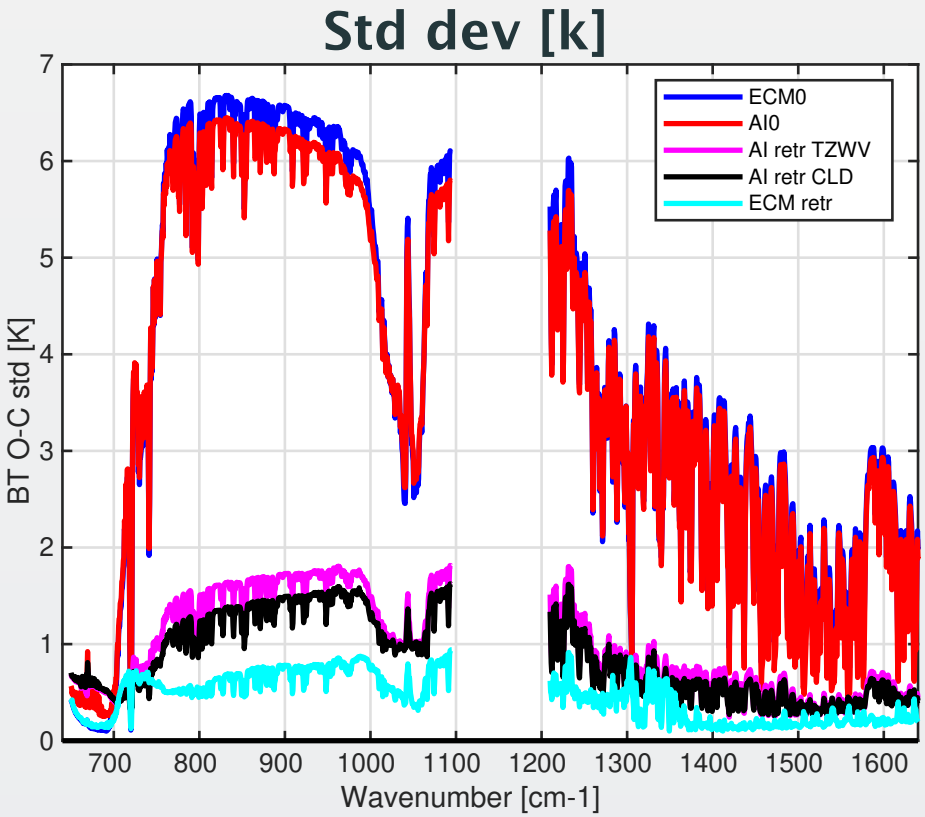
QC : 2019/04/25 g214



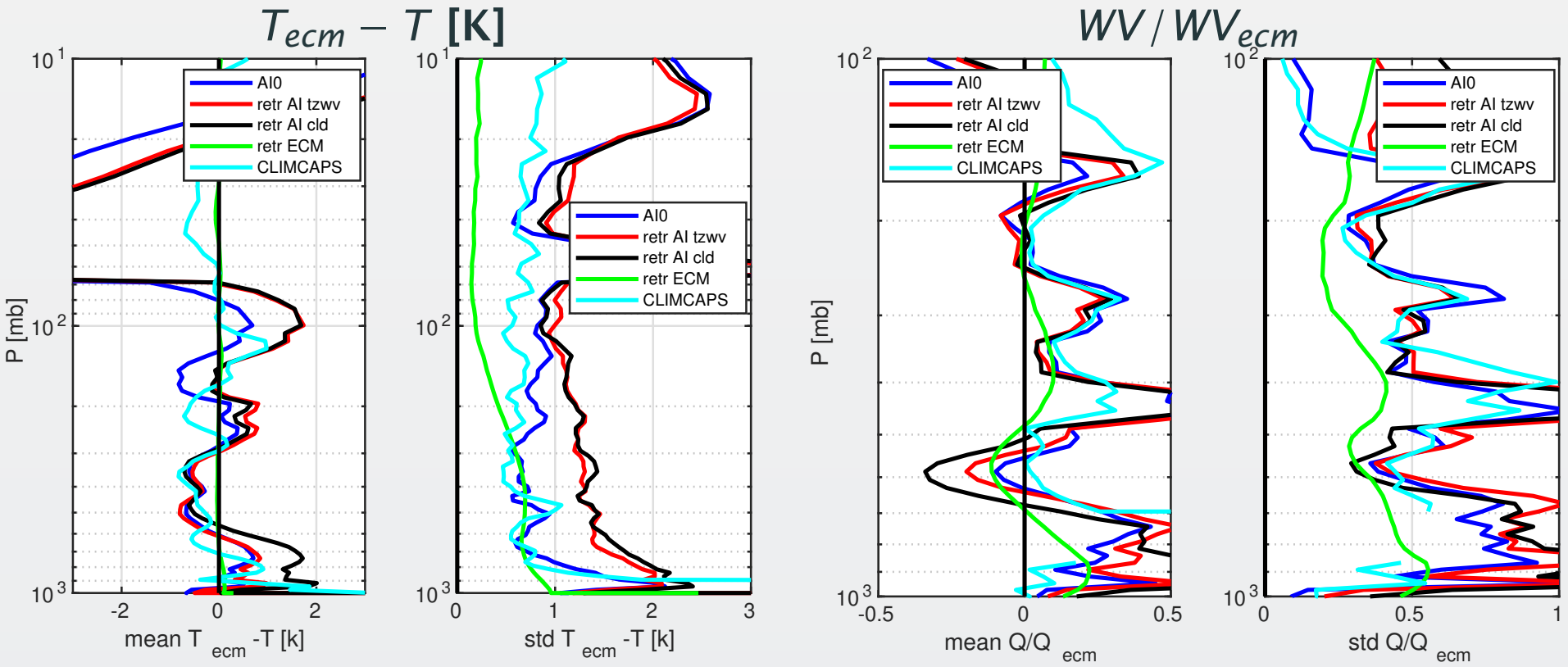


QC : 2019/04/25 BT bias and std dev



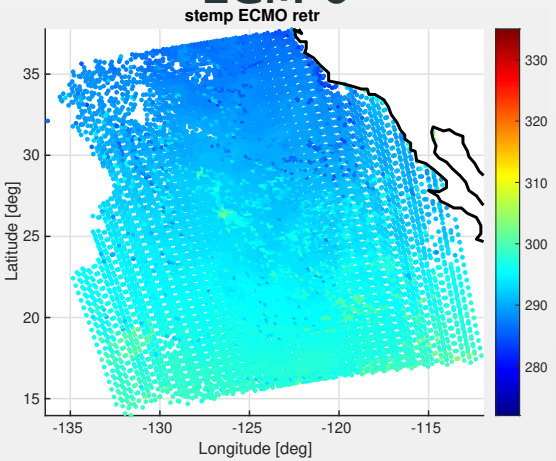


QC : 2019/04/25 T(z) and WV(z)

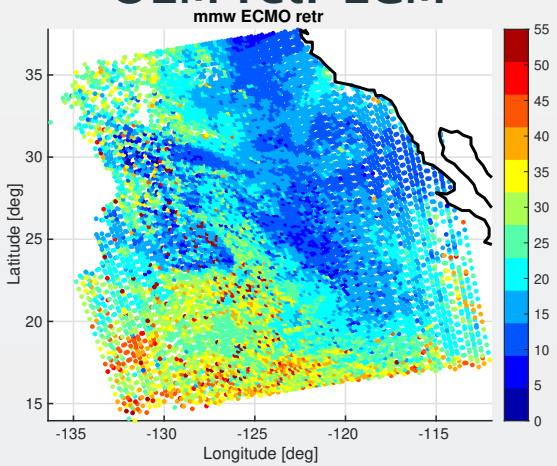


2019/04/25 g214 : Column Water(mm)

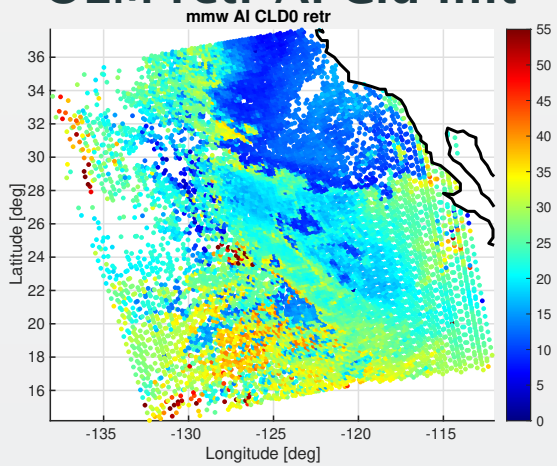
ECM 0



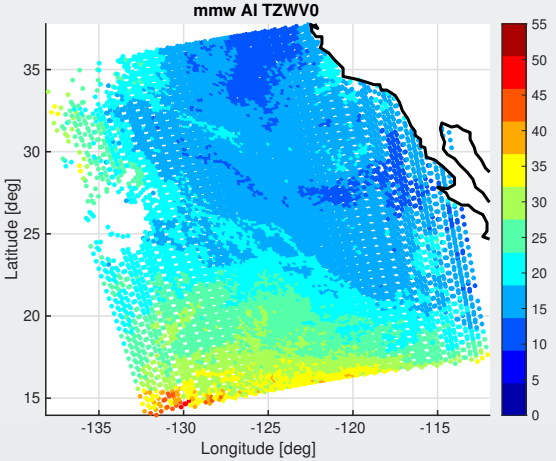
OEM retr ECM



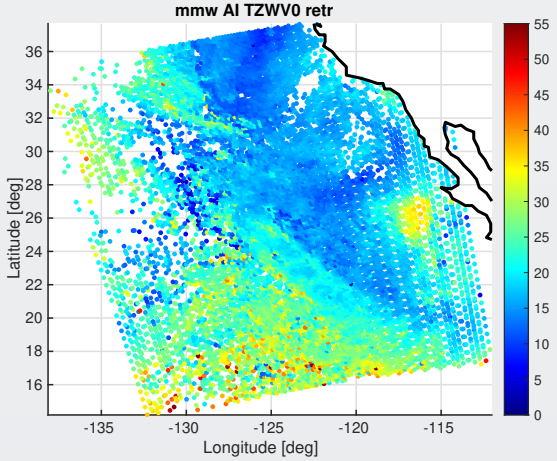
OEM retr AI Cld init



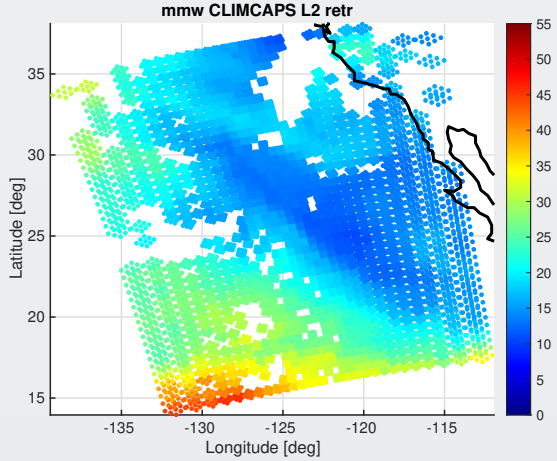
AI 0



OEM retr AI TzWV init

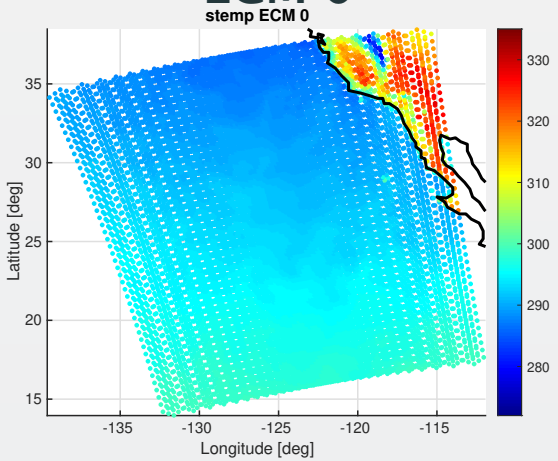


CLIMCAPS

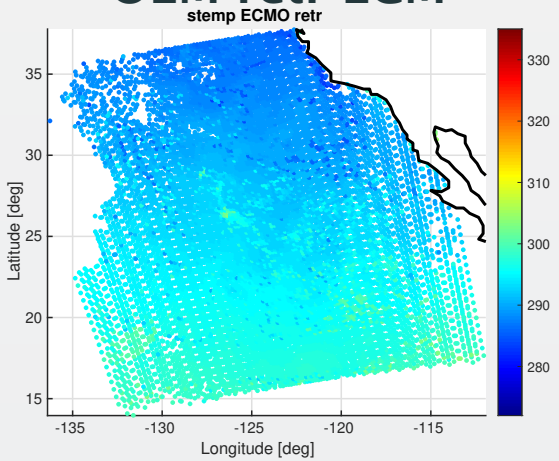


2019/04/25 g214 : Surf Temp [K]

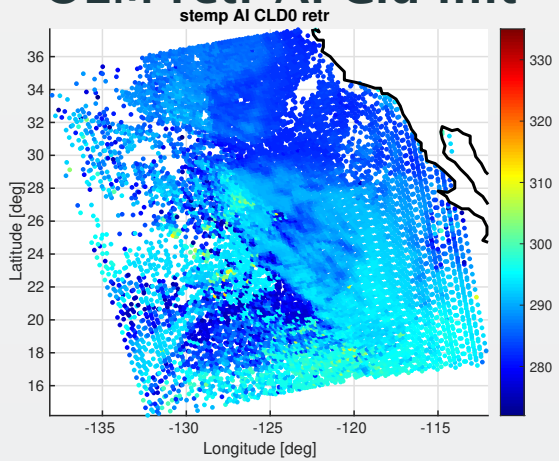
ECM 0



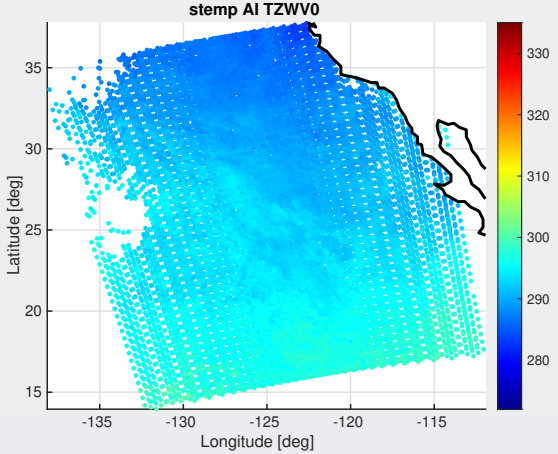
OEM retr ECM



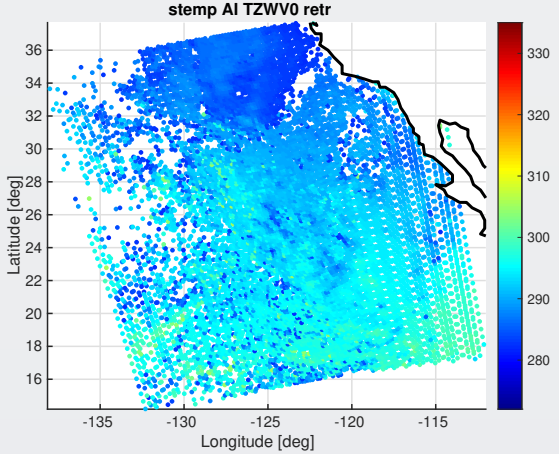
OEM retr AI Cld init



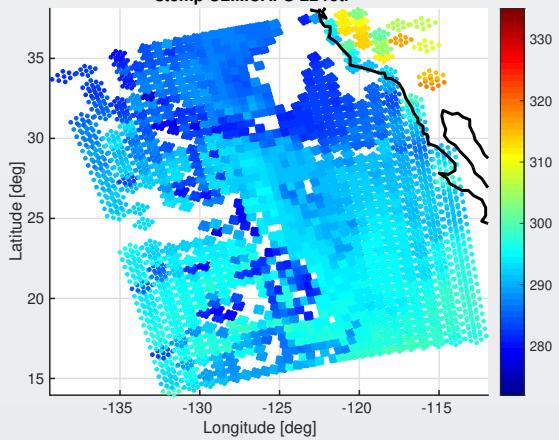
AI 0



OEM retr AI TzWV init

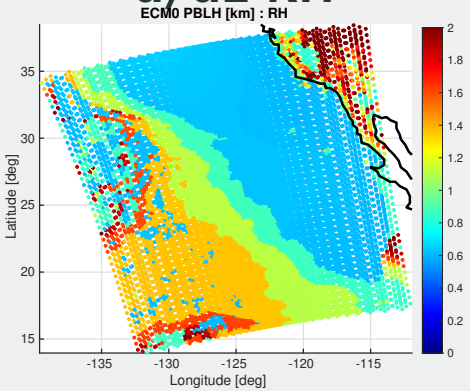


CLIMCAPS

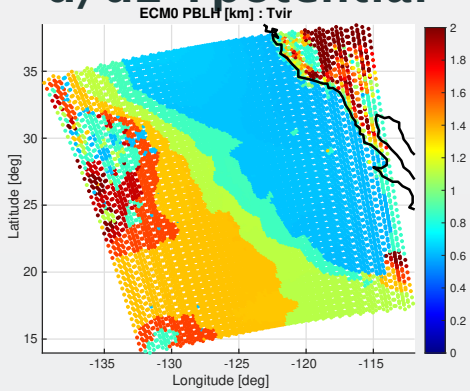


2019/04/25 g214 : Comparing PBLH [km] for ECM and AI

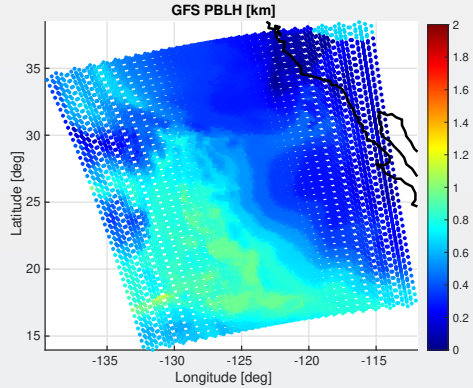
d/dz RH



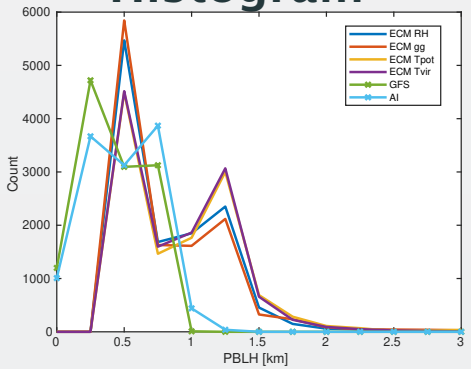
d/dz Tpotential



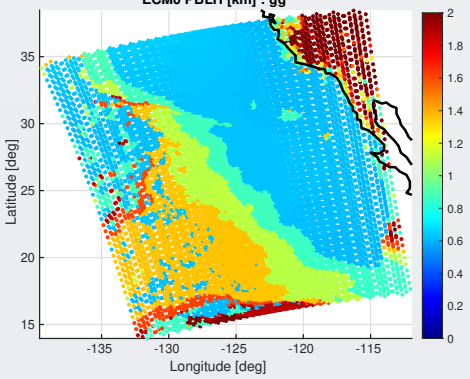
GFS



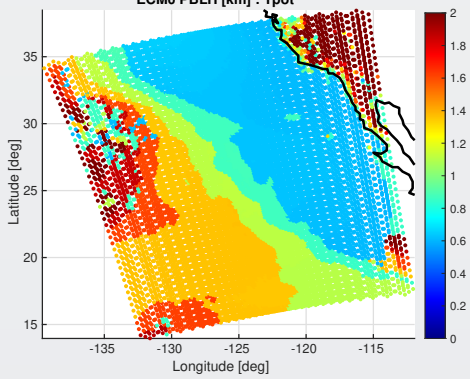
Histogram



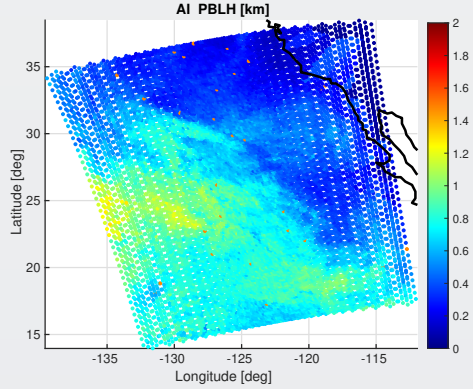
d/dz g/g



d/dz Tvirtual

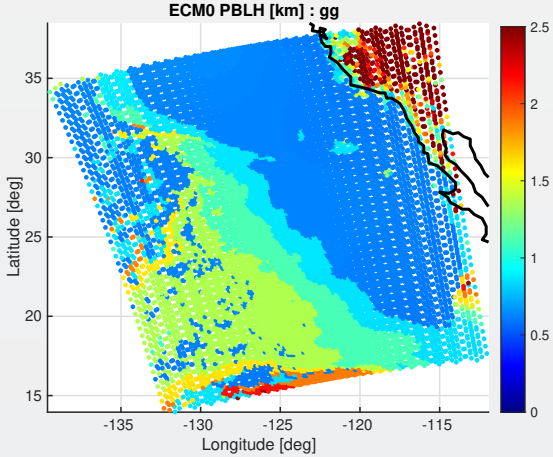


AI 0

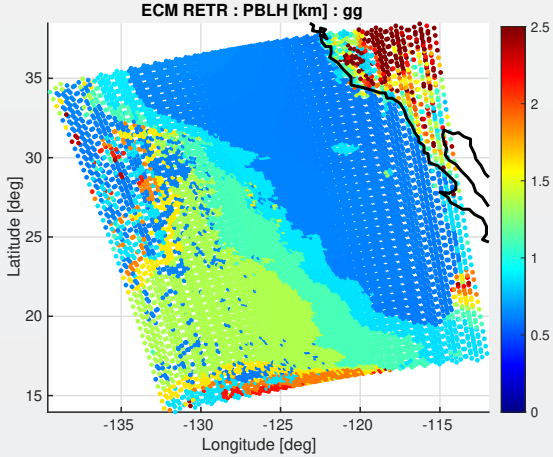


2019/04/25 g214 : PBLH [km]

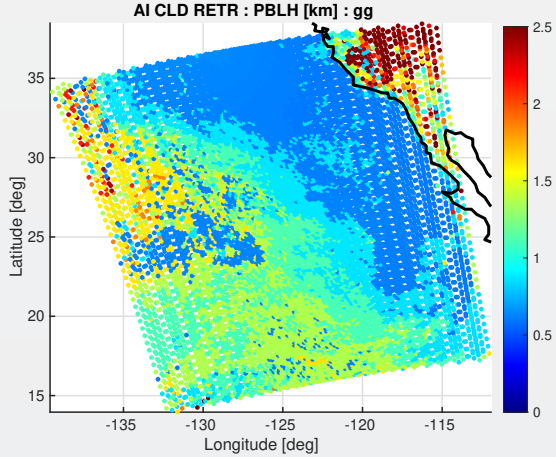
ECM 0



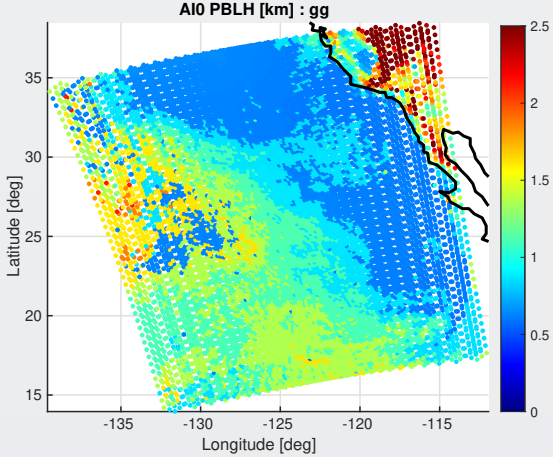
OEM retr ECM



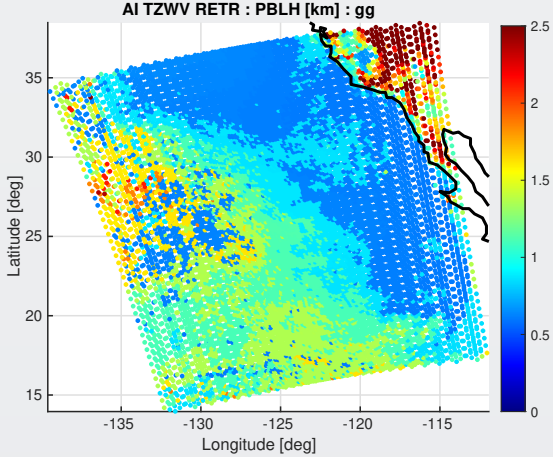
OEM retr AI Cld init



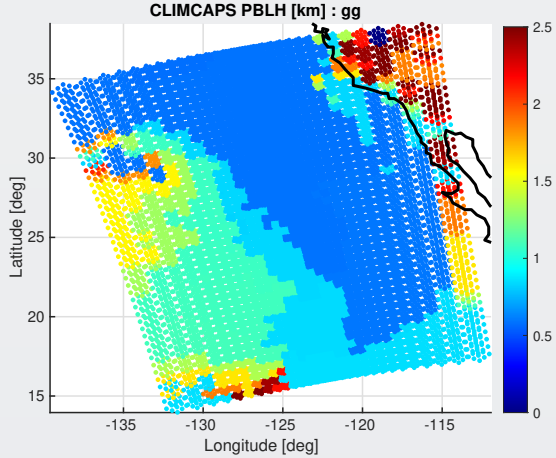
AI 0



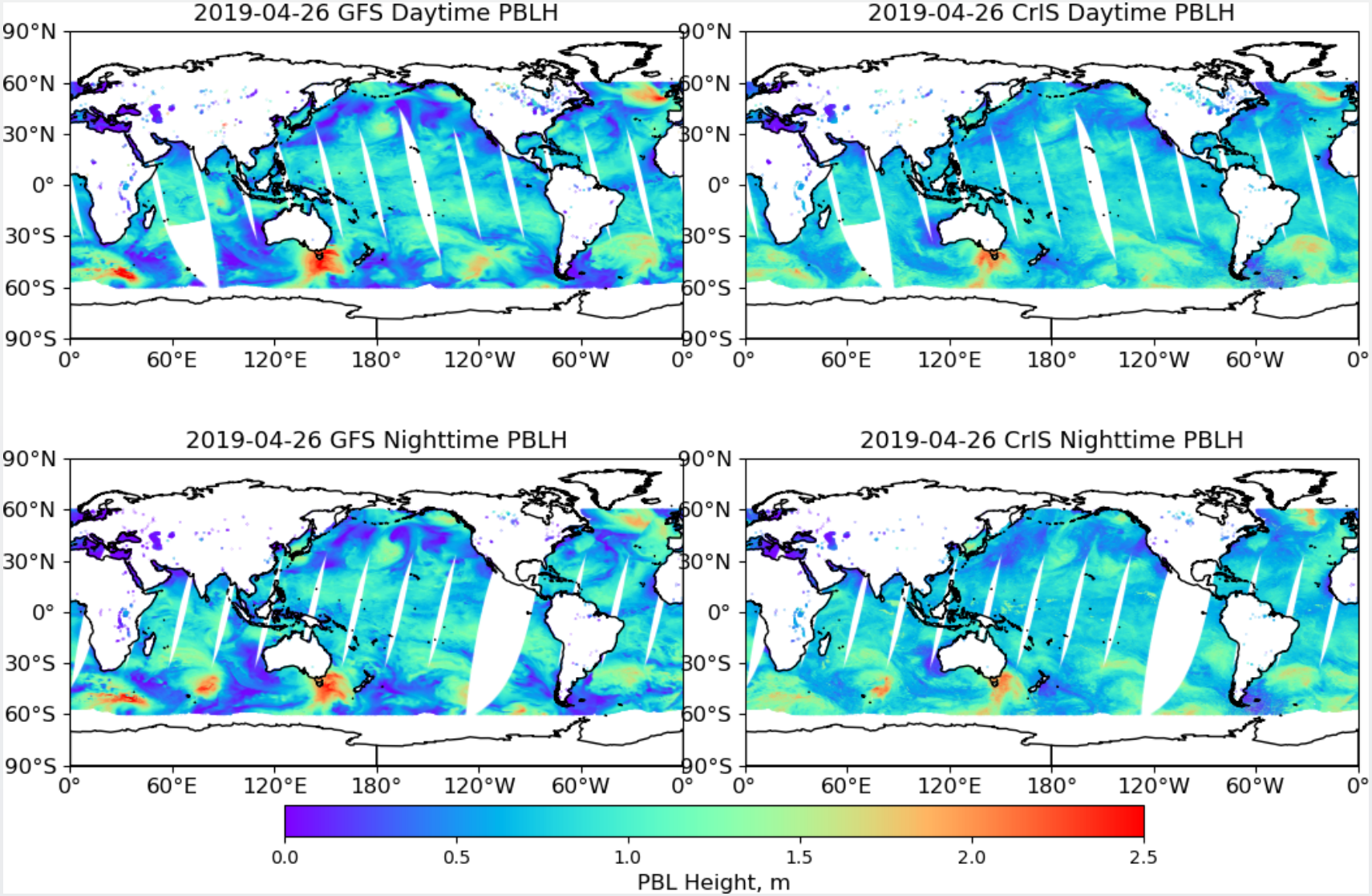
OEM retr AI TzWV init



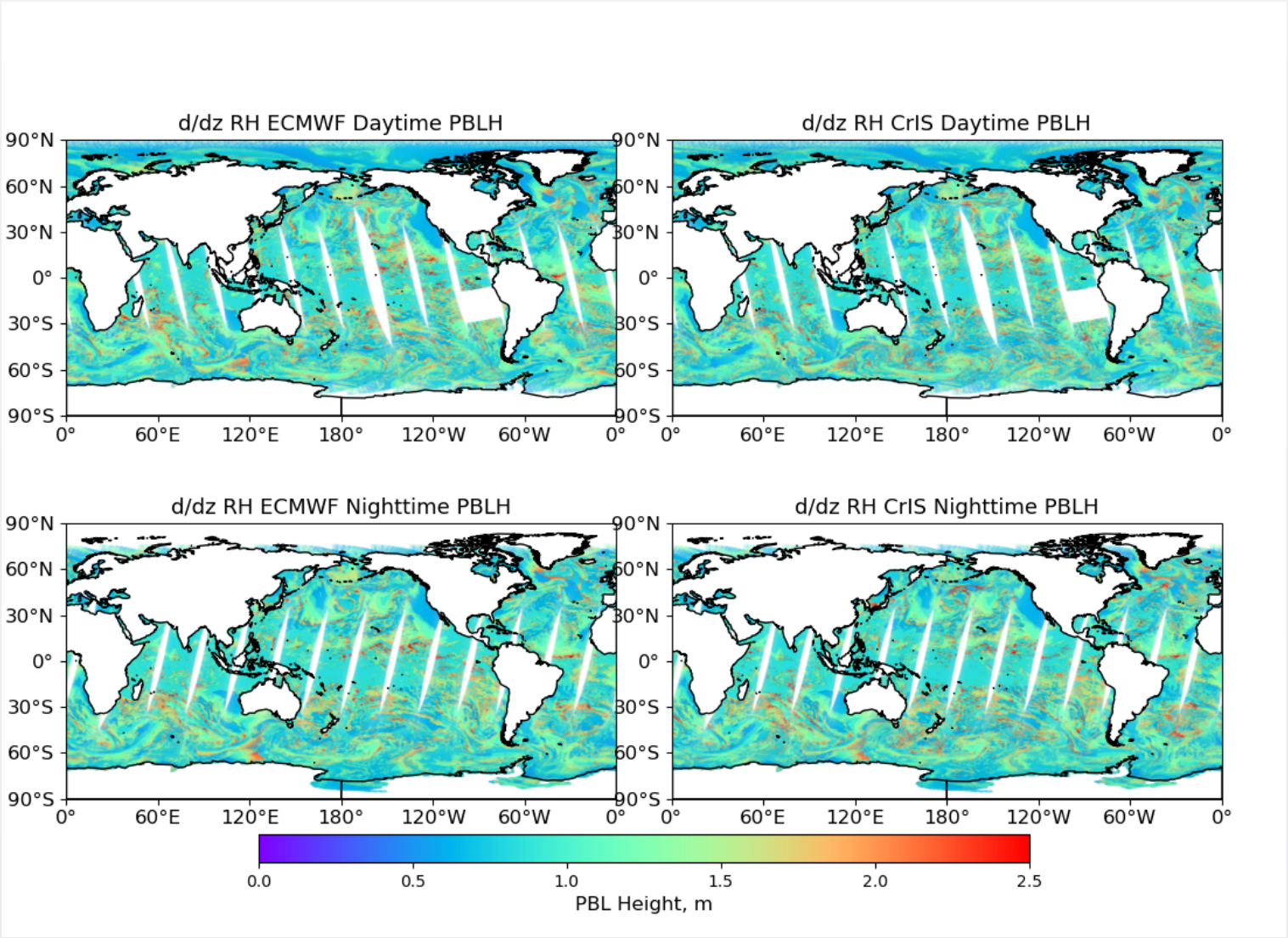
CLIMCAPS



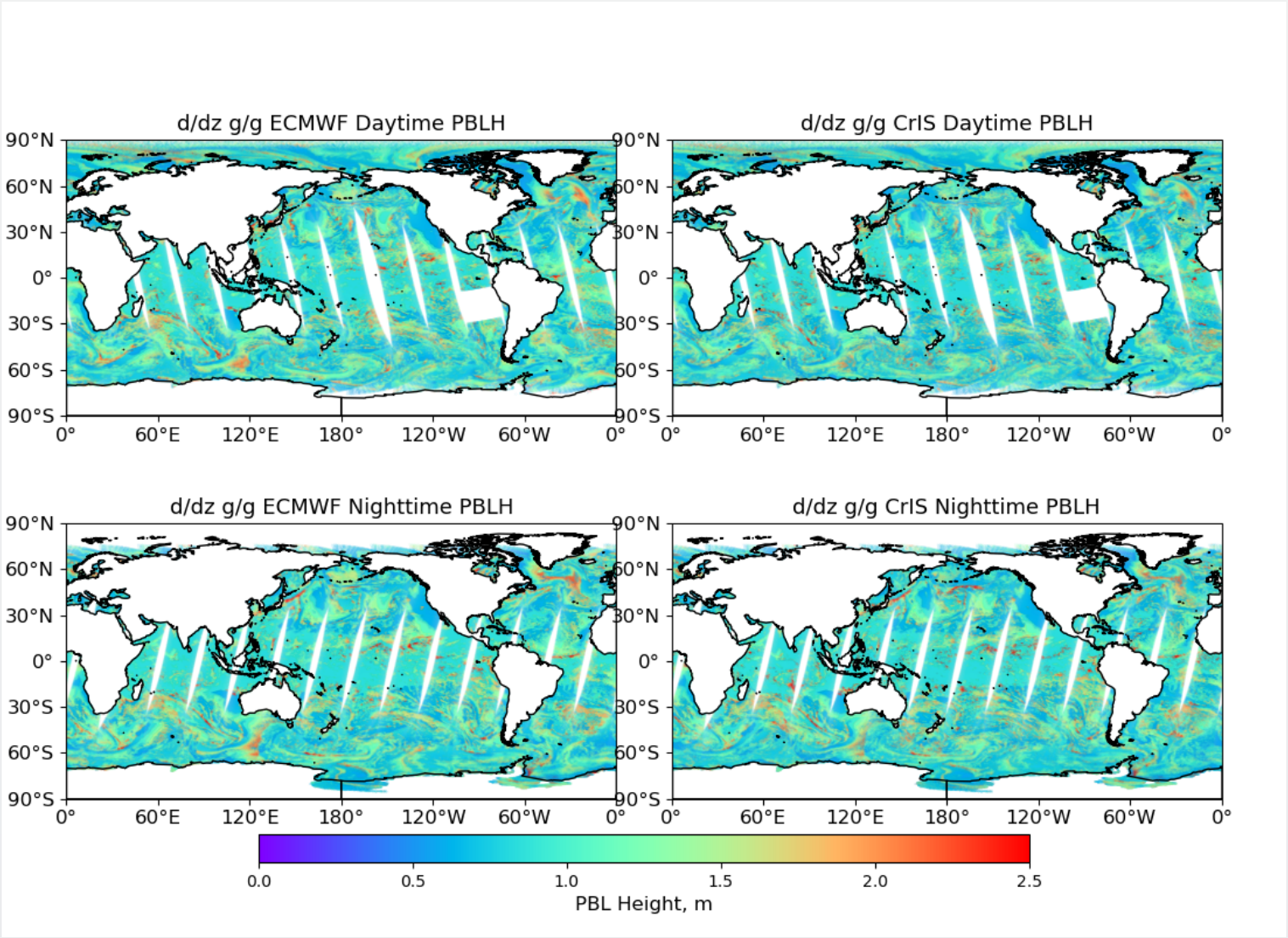
Global PBL height retrieval at 15km IR sounder res - 2019/04/26 AI



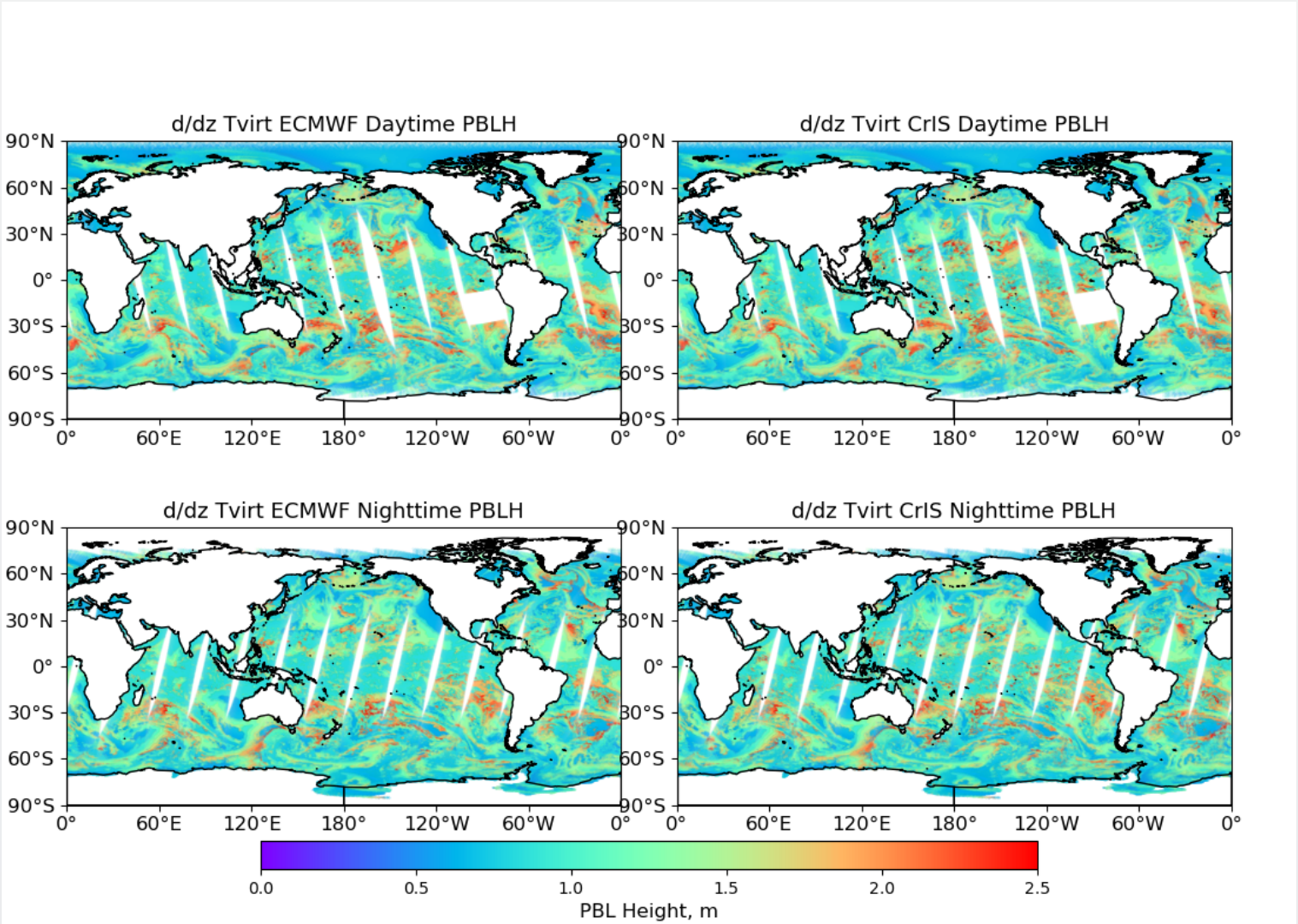
Global PBL height retrieval at 15km IR sounder res - 2019/04/26 OEM : dRH/dz



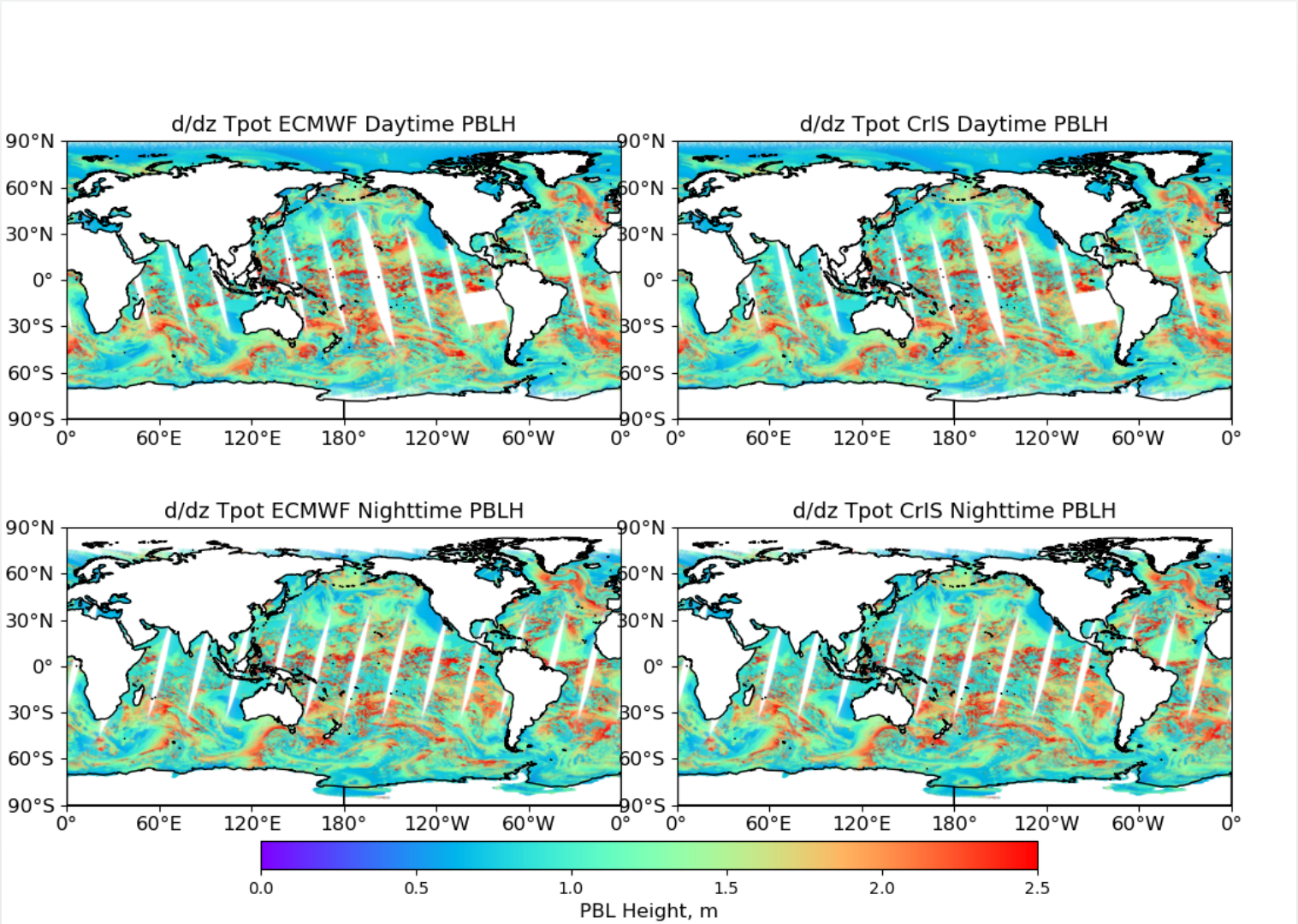
Global PBL height retrieval at 15km IR sounder res - 2019/04/26 OEM : dSH/dz



Global PBL height retrieval at 15km IR sounder res - 2019/04/26 OEM : dTvir/dz



Global PBL height retrieval at 15km IR sounder res - 2019/04/26 OEM : dTpot/dz



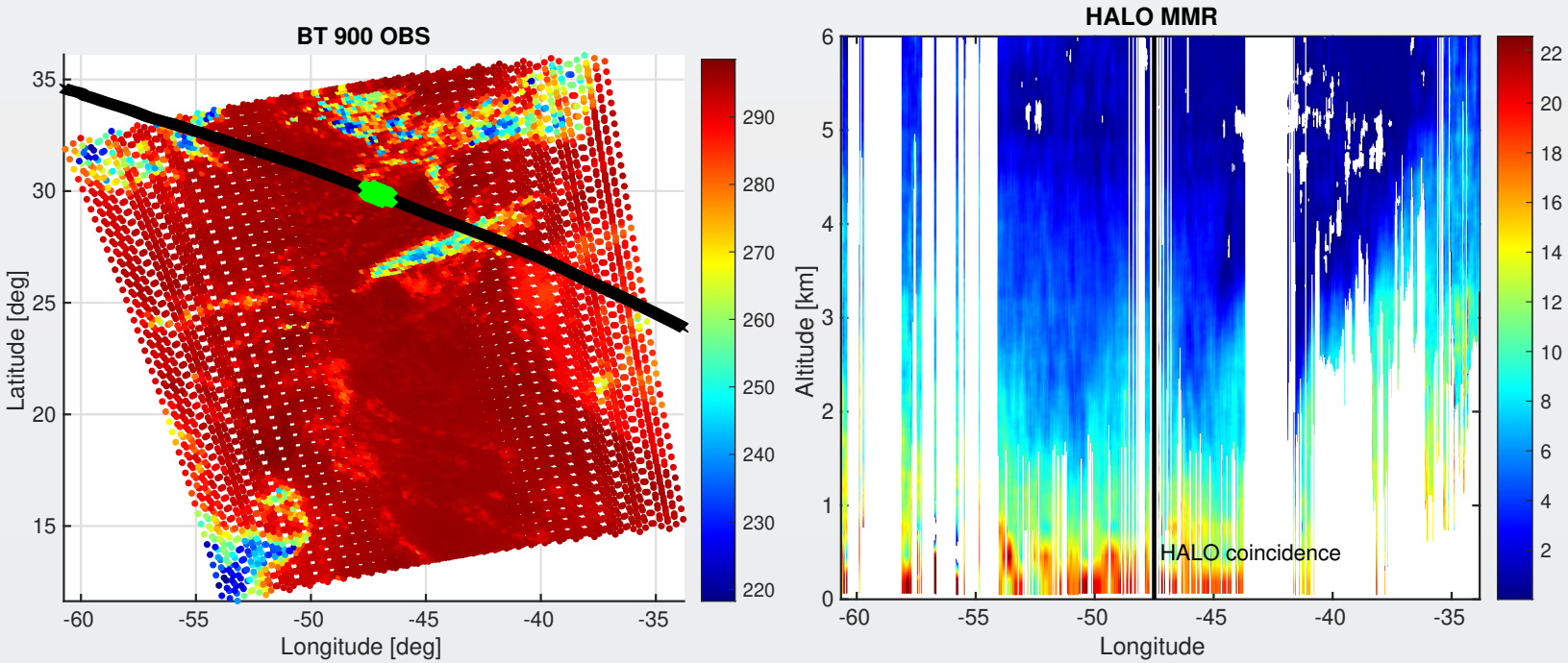
Conclusions

- MIIDAPS-AI based CriS or ATMS or CrIS/ATMS code is working
  - Very fast, about 10000 profiles/second
  - Looked at CrIS only result, over ocean
  - Eric will work on improving retrievals - PBLH, cloud retrievals, T(z), Q(z)
  - *Also provides PBLH* using RIchardson number
- Can ingest the profiles into UMBC-OEM
  - Can parse out what to read (clouds or thermodynamic profiles, or both)
  - Eliminating Cloud Swap means no time needed to pre-process
  - Use SARTA 100 layer allsky analytic jacobians : sequential or simultaneous retrieval
  - One granule (12150 FOVS) takes 3-5 hours (1-2 seconds per FOV)
  - *Also provides PBLH* using d/dz(RH,SH,Tvir,Tpot)
- More testing/tweaking of both codes (AI clouds, OEM covariance matrices etc)
- Will work on Validations of T(z),Q(z),stemp and PBLH

Thanks to NASA-Langley HALO scientists!

JPSS 2022/09/03 g162

Worked with d/dz (and d/dp!!!) RH, WVgg, Tvirtual,T
Code finds max (or min) derivative closes to ground
Black is HALO track, green is JPSS/HALO coincidence



Global PBL height retrieval at 15km IR sounder res - 2022/09/03