# Updates on CrIS VIIRS Radiance Cluster Algorithm

**Likun Wang[1*], Haibin Sun[2], Chris Burrows[3]**

**1. Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, Maryland**

**2. GAMA-1 Technologies, Greenbelt, Maryland, USA**

**3. ECMWF, Reading, UK**

**Email address: wlikun@umd.edu;**

Acknowledgment: Lihang Zhou, Banghua Yan, Satya Kalluri, Mitch Goldberg, Qiang Zhao and Jim Jung

# Content

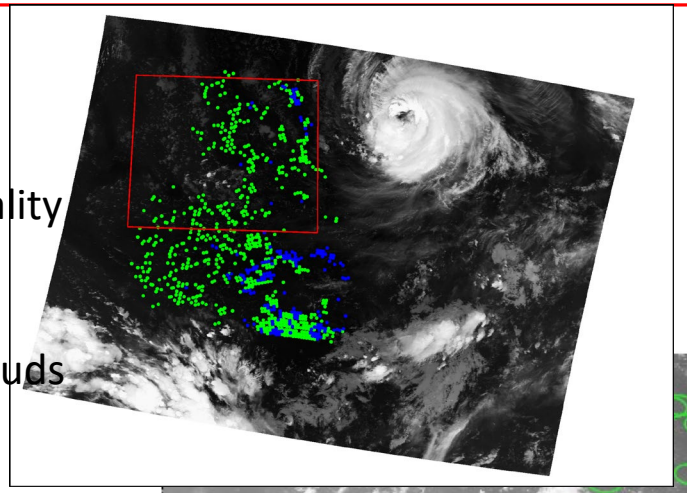
- Background
  - Why is imager subpixel information important for NWP data assimilations?
  - What have been done?
- CrIS VIIRS Cluster Analysis
  - Algorithm Description
  - Comparison with IASI cluster Analysis
  - Example of using CrIS VIIRS cluster analysis
  - Status on operational implementation
- Preliminary Results of Data Assimilation Experiment (collaborated with ECMWF)
- Conclusion

# Background

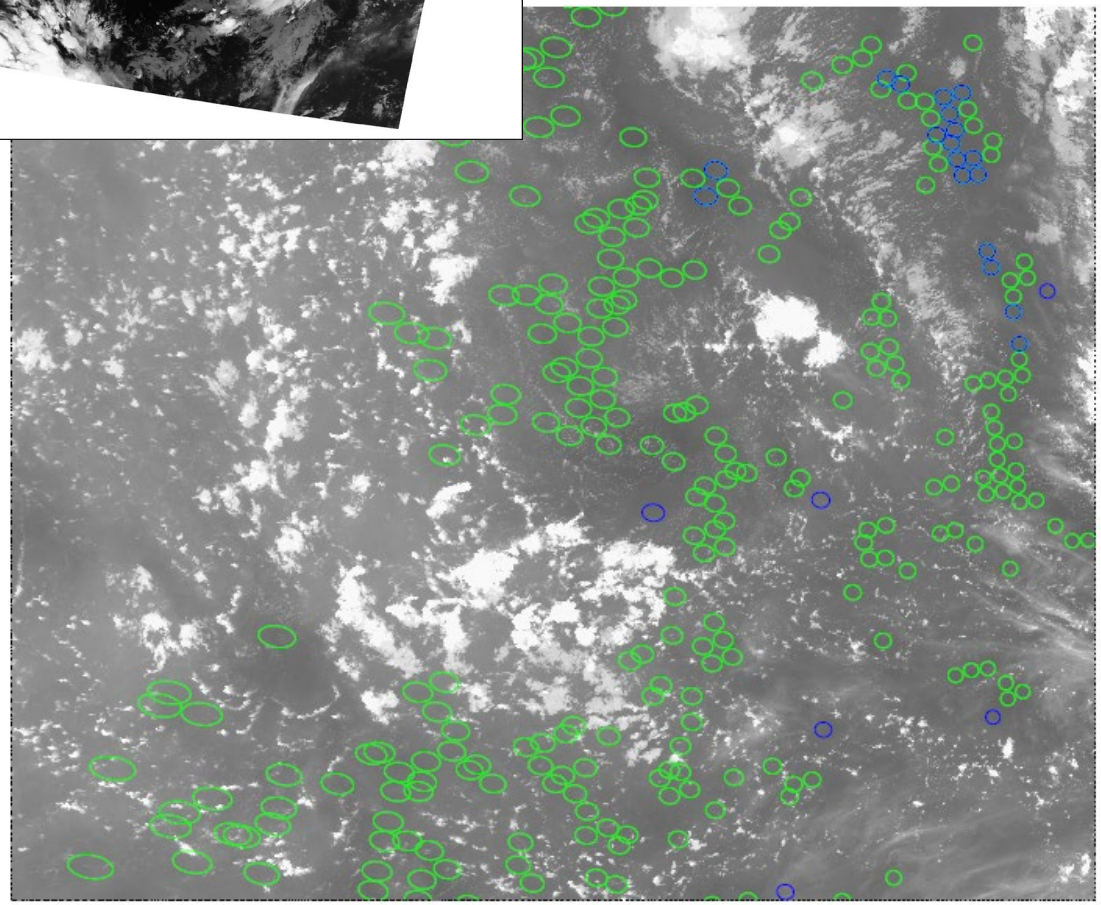
- Data Assimilation for IR Hyperspectral Data
  - ➢ Clear radiances above clouds

    Partial clouds – Subpixel imager information can help for Quality Control (Eresmaa 2014)
  - ➢ All Sky radiances (cloudy radiances)

    Subpixel imager information can distinguish homogenous clouds (Farouk et al. 2019)



The NWP O-B method is not good enough to guarantee “good data”



- IASI AVHRR cluster analysis have been implemented in level1c data (Cayla 2001)

- CrIS BUFR data
  - ➢ Collocated NDE VIIRS cloud products (already done)
    - VIIRS cloud mask, cloud top height, temperature, and pressure contained in BUFR data
    - Used NCEP GSI thinning routine (Jung et al. 2017).
  - ➢ VIIRS cluster analysis (this study)
    - Filled with blank fields

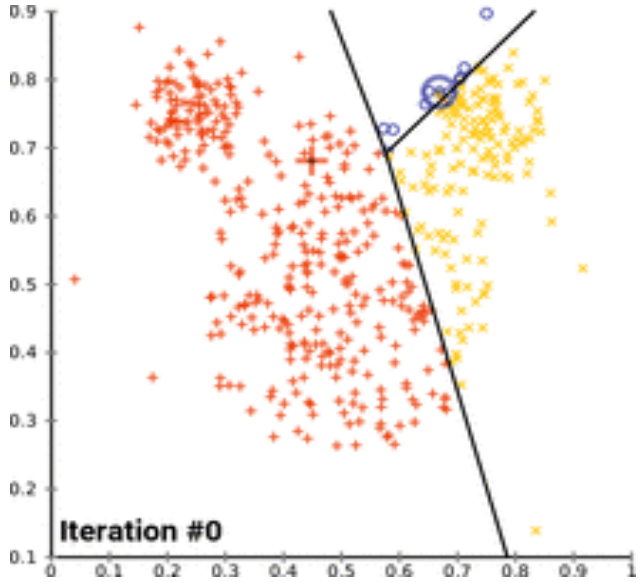
# CrIS Cluster Analysis Algorithms

- Using the standard K-means algorithm (Lloyd, 1957; MacQueen, 1967).

- VIIRS channels are grouped together as a vector
  - Can be extended into other VIIRS channels
  - Spatial information is not used

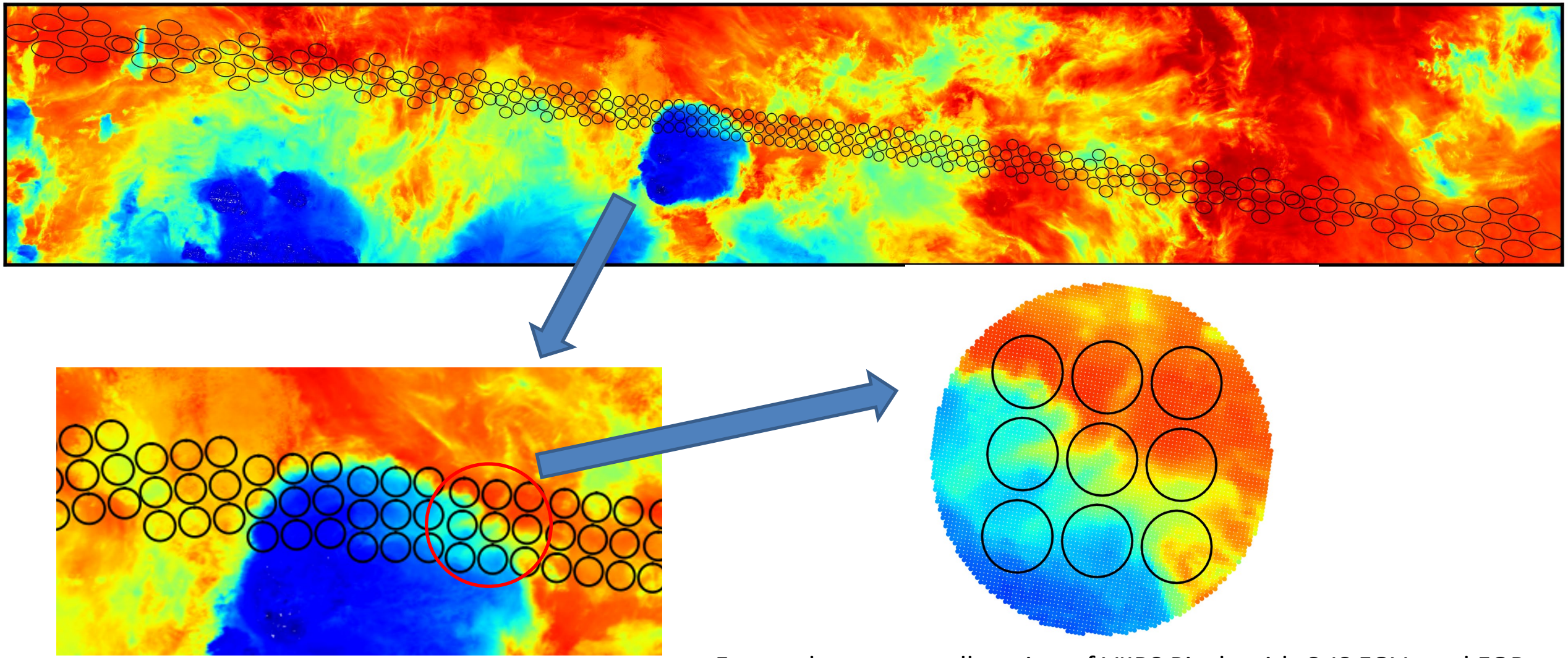
- It aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean.

- Output:
  - ✓ Mean radiances of each cluster
  - ✓ Standard Deviation of each cluster
  - ✓ Coverage of each cluster




https://en.wikipedia.org/wiki/K-means_clustering

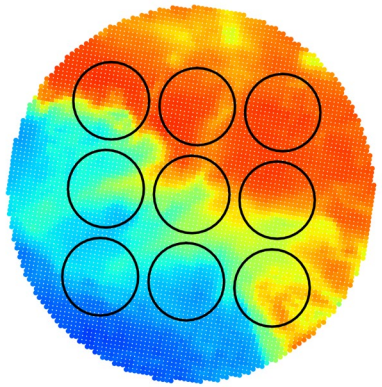
# CrIS FOV overlapped with VIIRS Image

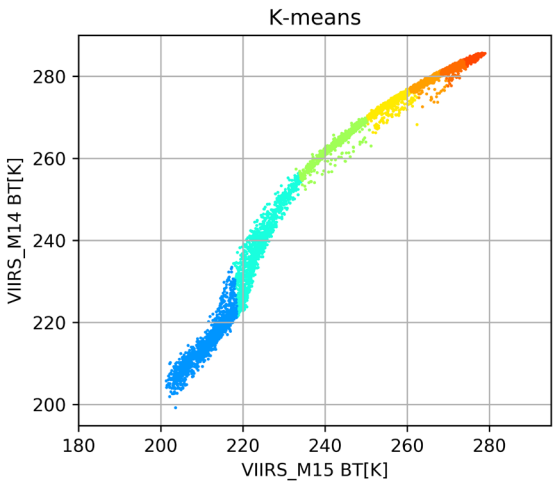


Fast and accurate collocation of VIIRS Pixels with CrIS FOVs and FORs (Wang et al. 2016)

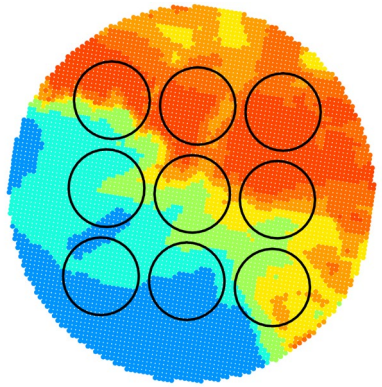
# CrIS Cluster Analysis Procedures

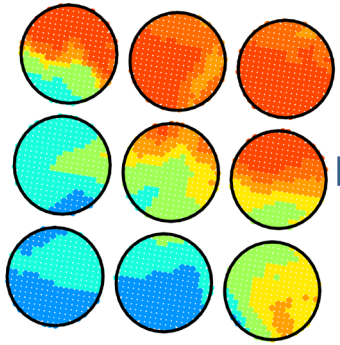
**Pixels are indicated by BTs in colors (continuously)**

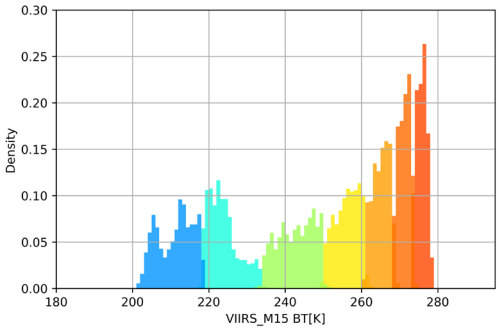





**7 clusters, the pixels in the same cluster are indicated by the same color**








Extracting VIIRS data with CrIS Field of Regard (FOR)

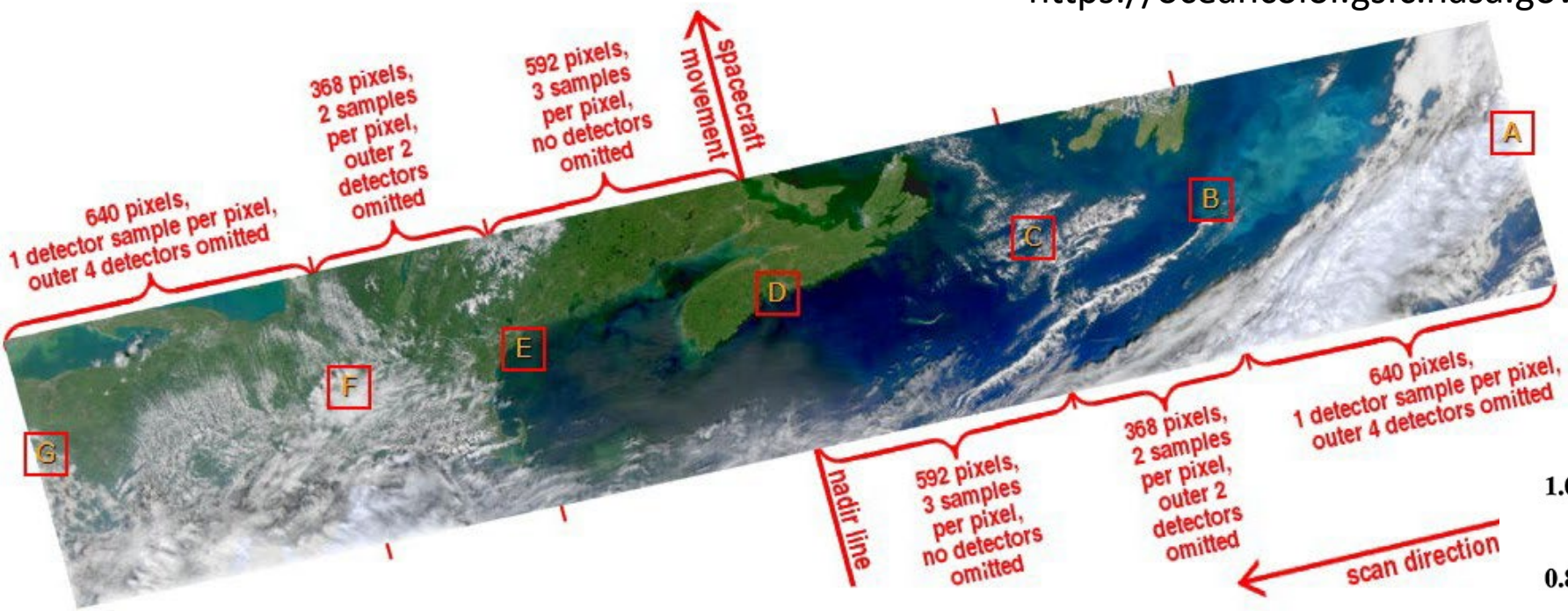
Applying K-means Cluster Analysis with 7 clusters in VIIRS radiance domain

Mapping cluster analysis results into each FOV

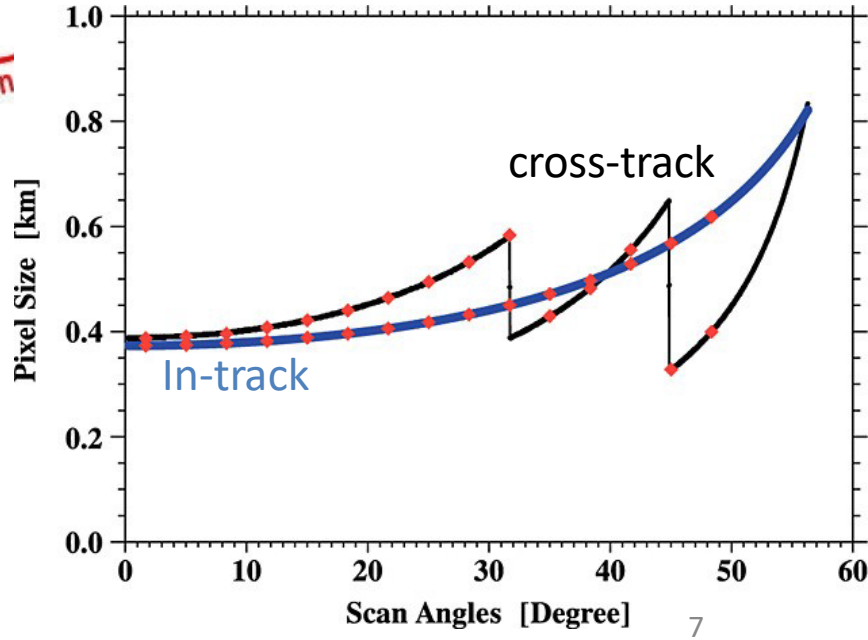
Output the Mean, Stdev, and Coverage for each cluster for each band within each FOV

# VIIRS Bow-Tie Deletion

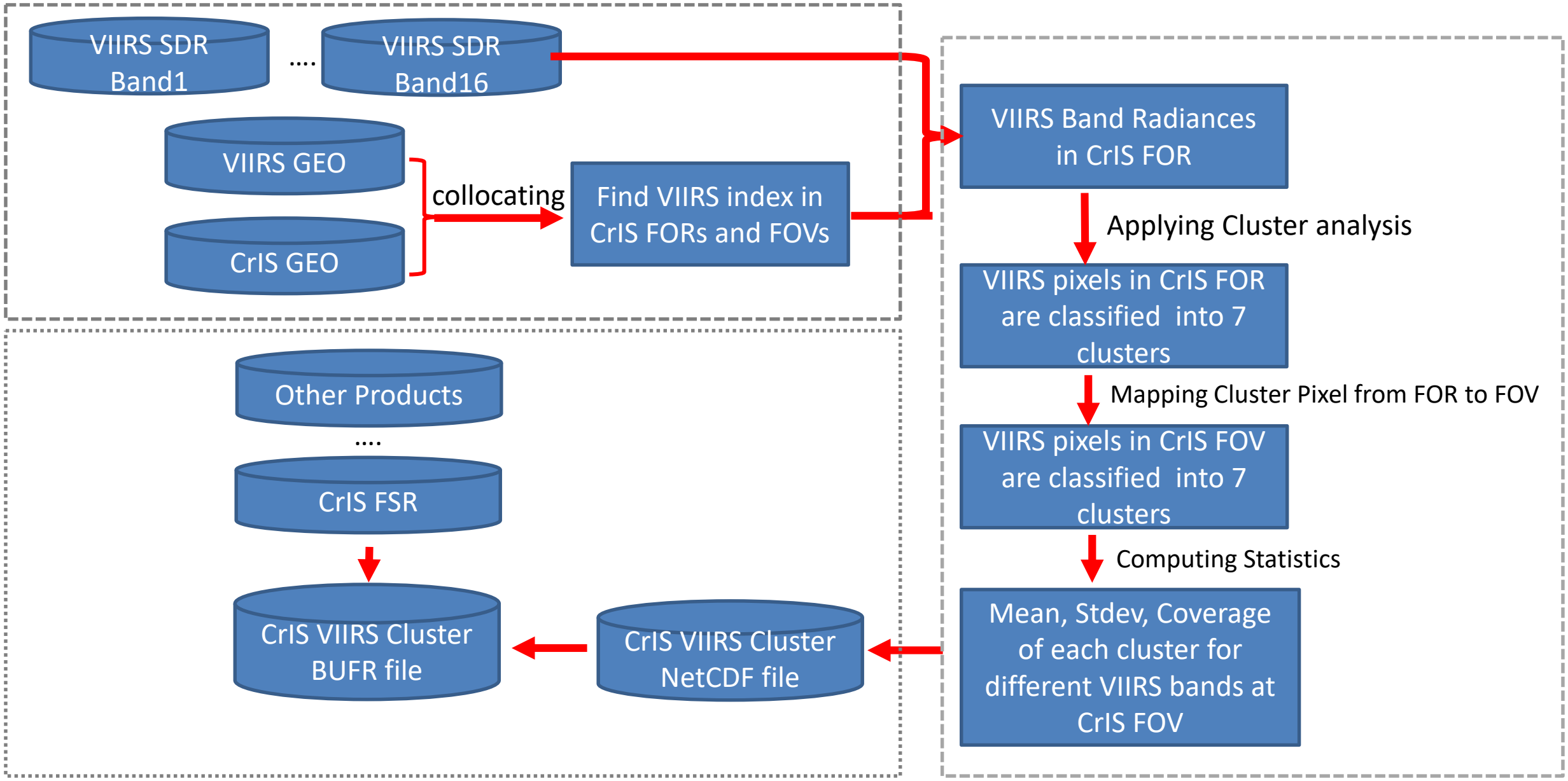
https://oceancolor.gsfc.nasa.gov/docs/bowtie/viirs/




The bow-tie deletion along VIIRS scan line should be carefully taken care by assigning the weights to each VIIRS pixel, when calculating the cluster mean and standard deviation.



# Flowchart for Data Generation

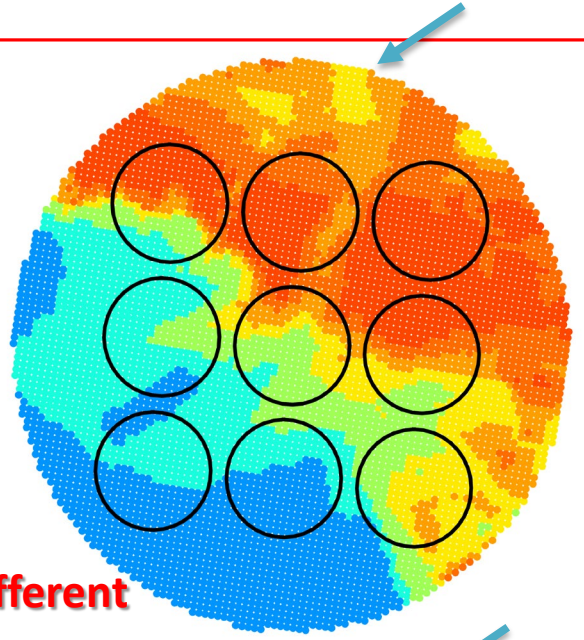


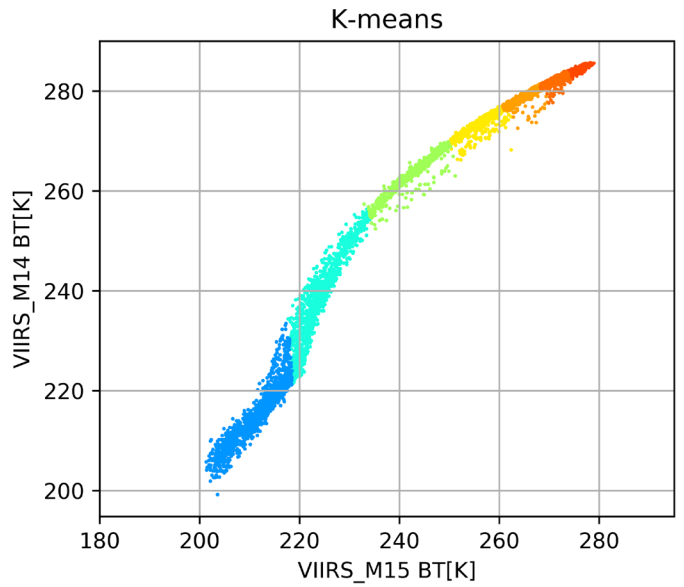
# CrIS-VIIRS vs. IASI-AVHRR cluster analysis

| | CrIS VIIRS | IASI-AVHRR |
|---|---|---|
| **Collocation** | Need to independent collocation software | Using IASI-imager to align AVHRR together. |
| **Spatial Resolution** | VIIRS Bow-Tie effects in cross-track directions | AVHRR pixel increase smoothly with scan angle |
| **Cluster Analysis Algorithm** | K-means clustering | Hierarchical clustering |
| **Cluster number** | 16 bands; 7 cluster per band<br>Mean, Stdev, Coverage per cluster | 5 bands; 7 cluster per band<br>Mean, Stdev, Coverage per cluster |
| **Data format** | Separated BUFR data with other cloud information | Packed in IASI level1-c data in EPS format |

# Difference between IASI and CrIS Cluster Analysis

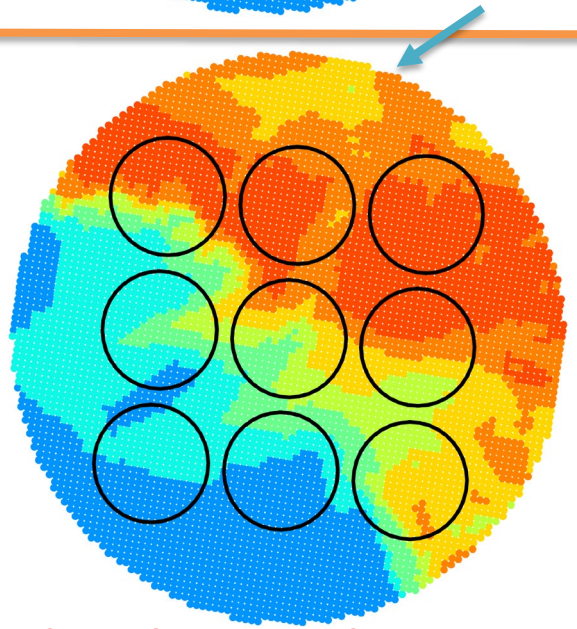
- CrIS uses **K-means clustering**
  - centroid-based clustering
  - Need a fixed number (k) of the cluster
  - Fast

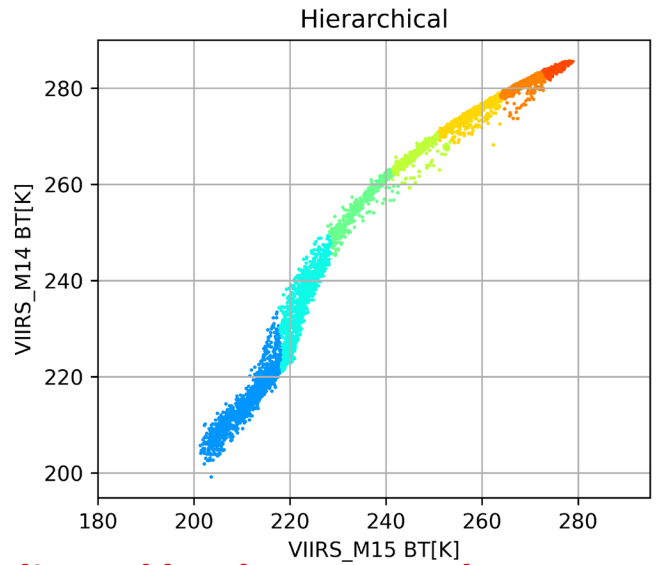





**The results are very similar, while details are a little different**

- IASI uses **Hierarchical clustering**
  - Connectivity-based clustering
  - Need to build hierarchical tree structures
  - Slow






**7 clusters, the pixels in the same cluster are indicated by the same color**

# Same IASI data using CrIS and IASI algorithm
# Inter-cluster Departure in Unit [K]

To evaluate the interclass homogeneity, the cluster separation are calculated using the following formula from the collocated VIIRS Cluster

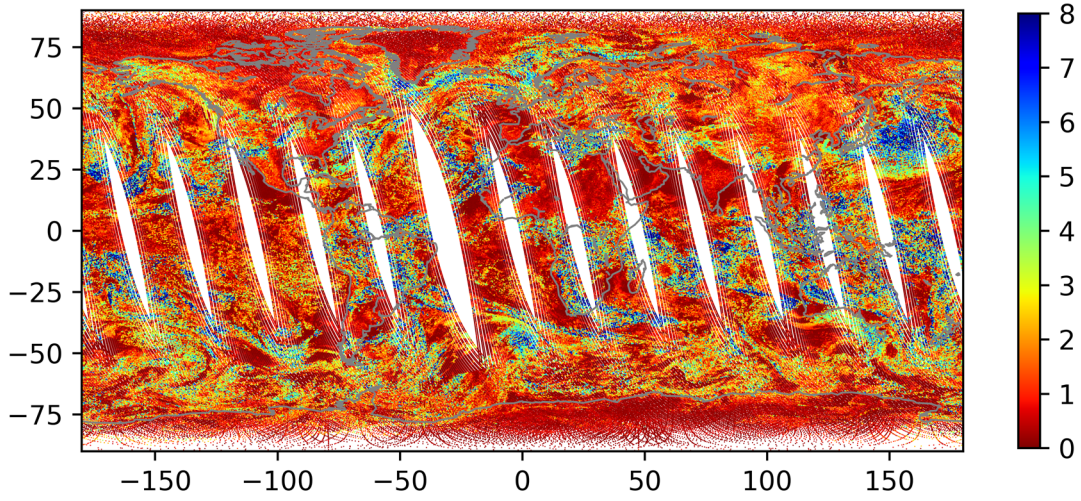
$$\sigma_{\text{inter}} = \sqrt{\frac{1}{\sum C_j} \sum_{j=1}^{N} C_j \left( R_i^j - R_{i,\text{mean}} \right)^2},$$

Mean VIIRS BT

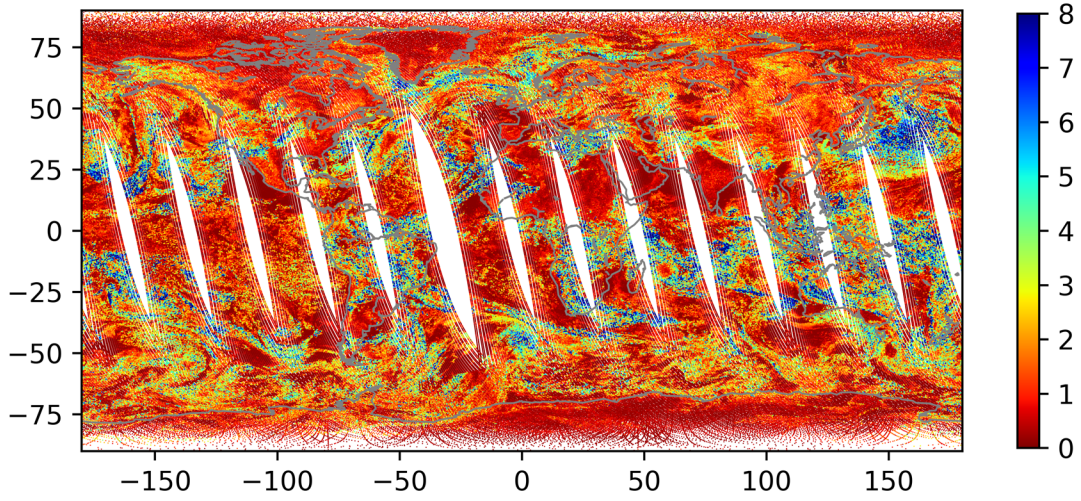
Each Cluster Coverage

Each Cluster Mean BT

IASI cluster analysis **from Level-1C IASI data**

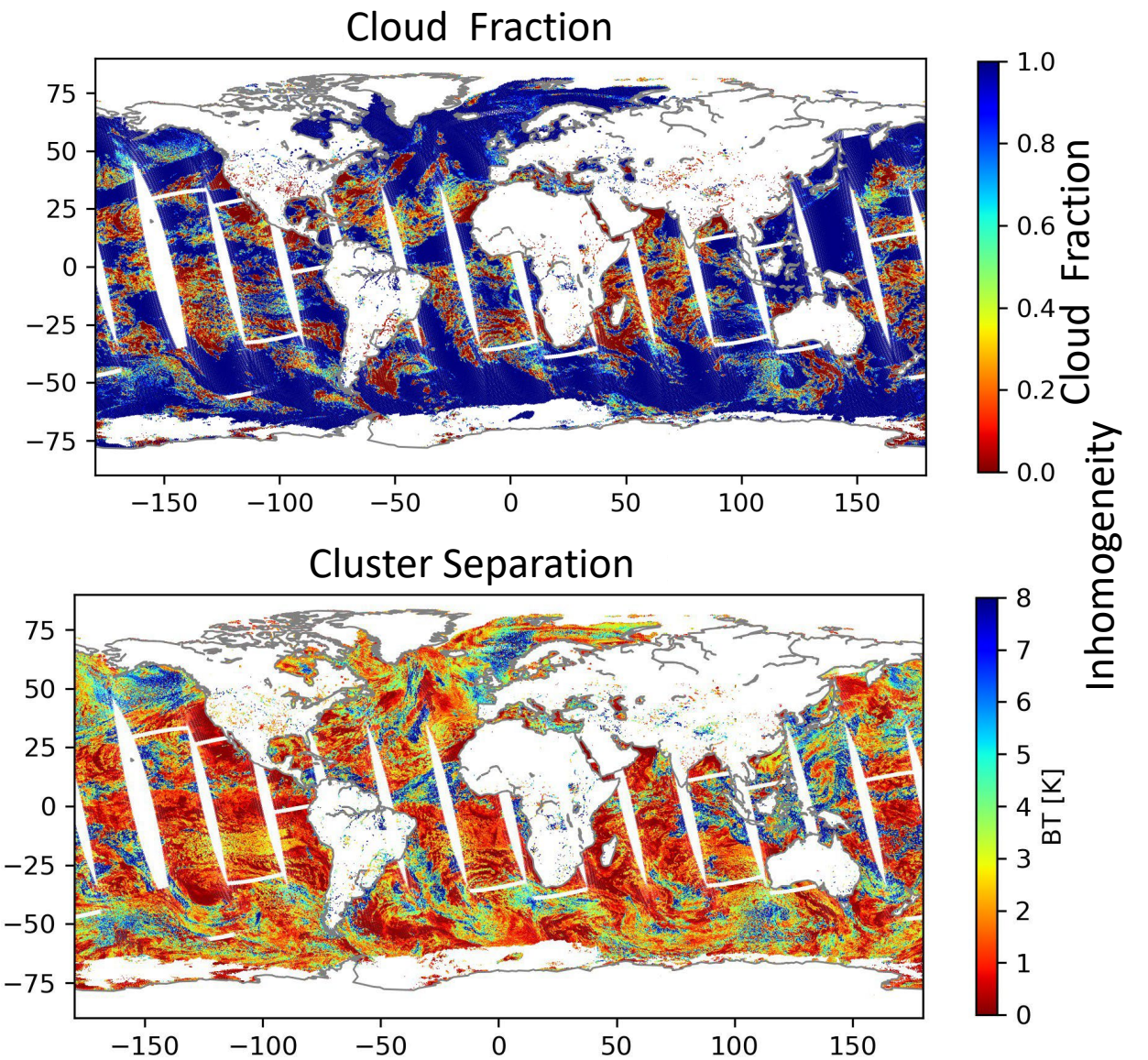



IASI cluster analysis **using the CrIS Algorithm**




**Both IASI and CrIS cluster analysis gives similar patterns of Inter-cluster Departure**

# Inhomogeneous Clouds
# Cloud Fraction vs Cluster Separation



# Current Status

- VIIRS cluster analysis code in Python has been delivered for R2O transition.
  - https://drive.google.com/drive/folders/1D56Zn1GsUiWmSw1YNYL_IJKTRn4LQ3-p?usp=share_link
  - Both python 2.7 and 3.9 version
  - Working closely the ASSISTT team for operational implementation

- One day's BUFR test data (03/25/2022) have been generated and distributed to NWP centers for the BUFR format testing.
  - ftp://ftp.star.nesdis.noaa.gov/pub/smcd/spb/lwang/cluster_bufr_test/

- Long-term CrIS VIIS Cluster data from 10/18/22 to 01/31/23 have processed to test the software for operational transition and data assimilation experiment.
  - ftp://ftp.star.nesdis.noaa.gov/pub/smcd/spb/lwang/cluster_op/
  - Some anomaly granules are found in some days, and we are investigating the root cause case by case to improve the software(missing files, bad pixels, QA…)
  - NCEP is now checking the data (Iliana Genkova and Jim Jung)
  - ECMWF is now running a long assimilation experiment with the VIIRS cloud detection (Chris Burrows)

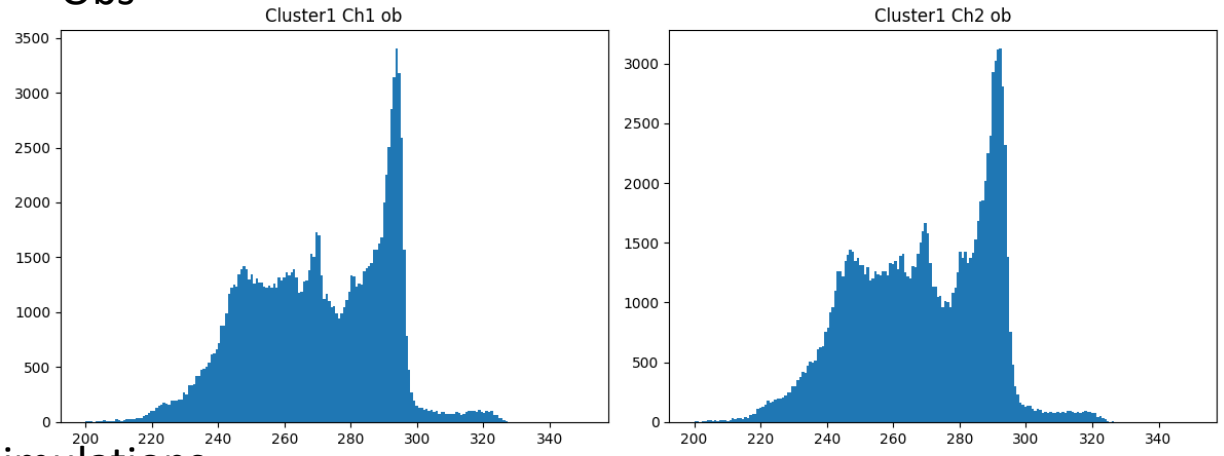
# BUFR Table Feedbacks

| | |
|---|---|
| The VIIRS cluster analysis can be directly filled into the CrIS BUFR data by filling into the data to the current blank fields. | Already contain blank fields for VIIRS cluster analysis |
| A small change on the one variable description is proposed. | “FCPH | 025085 | Fraction of clear pixels in HIRS FOV” => Fraction of VIIRS radiance cluster in CrIS FOV |
| CrIS VIIRS cluster units different from IASI AVHRR units | defined in the BUFR table as variable "SCHRAD”<br>AVHRR units: W M**-2 SR**-1 CM<br>VIIRS units: W M**-2 SR**-1 UM**-1 |
| 7 clusters for each VIIRS band: | No cluster values should be filled as 0.0 for turn-on channels |
| Turn-on and Turn-off channels (visible channels during nighttime) | Filled with missing Values for all Turn-off channels. |

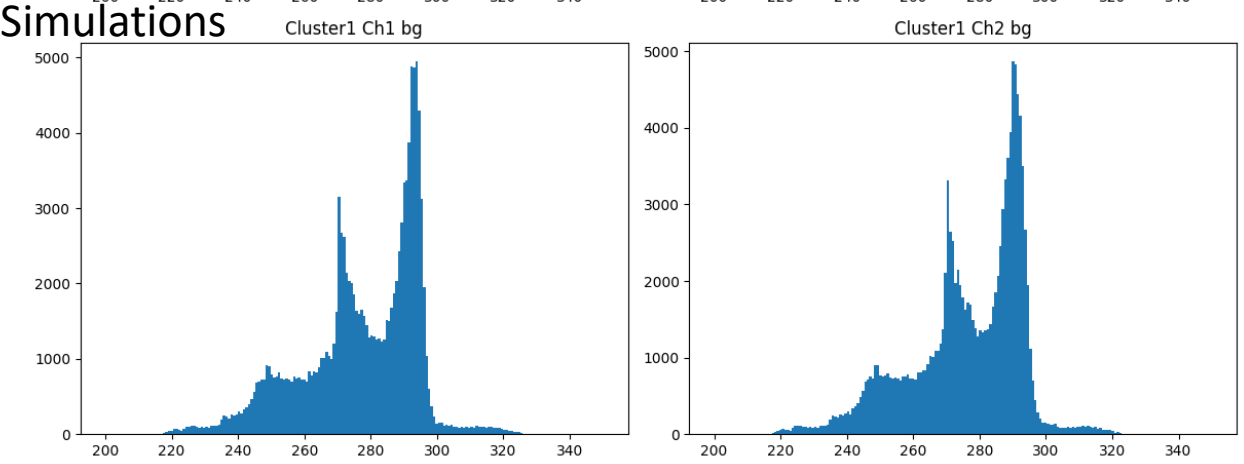
# Cluster Histogram Comparison in ECMWF DA System
# CrIS-VIIRS vs IASI-AVHRR

## CrIS VIIRS

Obs

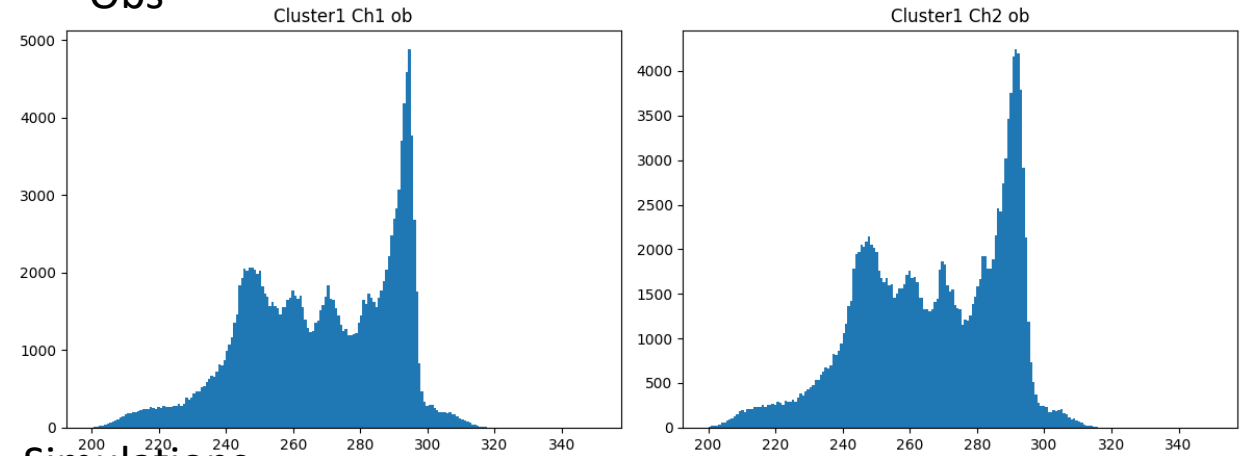



Simulations

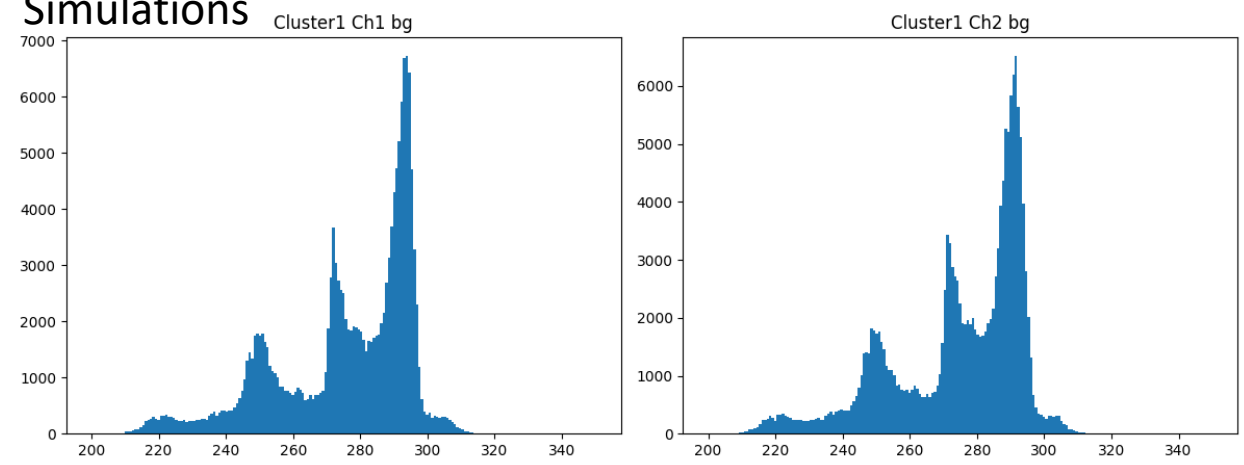



## IASI AVHRR

Obs



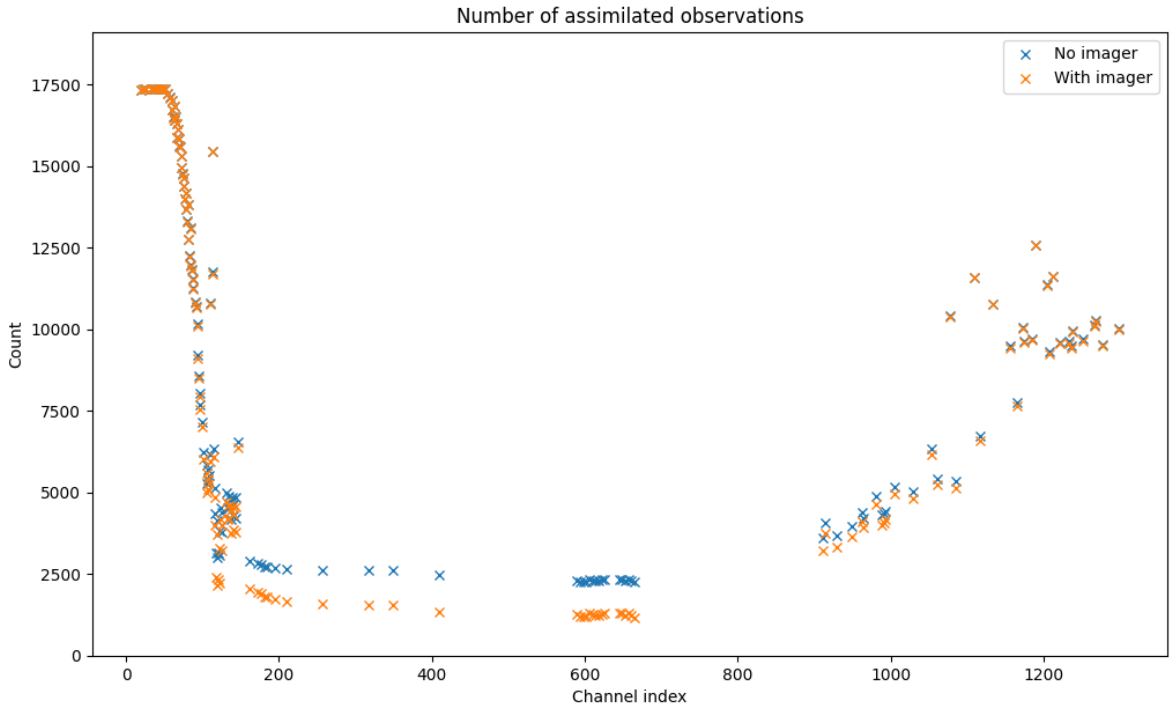

Simulations



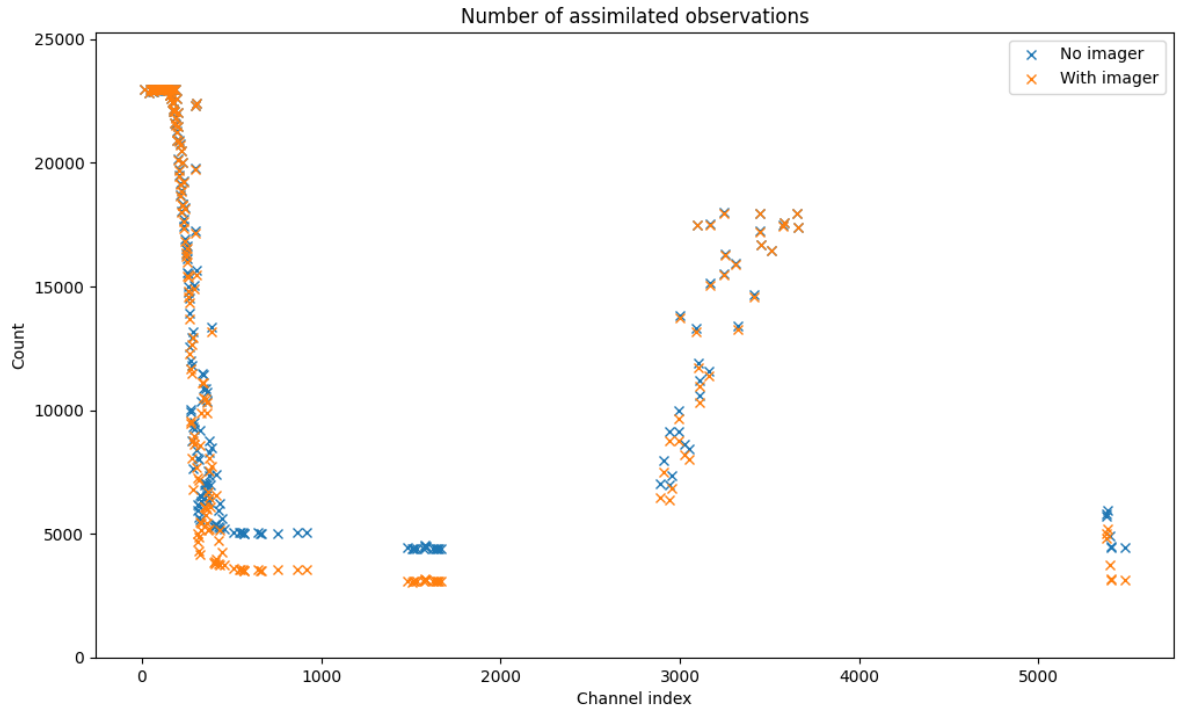

This is a small sample, but it shows that VIIRS and AVHRR have broadly similar PDFs.

# Number of observations assimilated with and without the imager cluster checks

## CrIS using VIIRS




## IASI using AVHRR

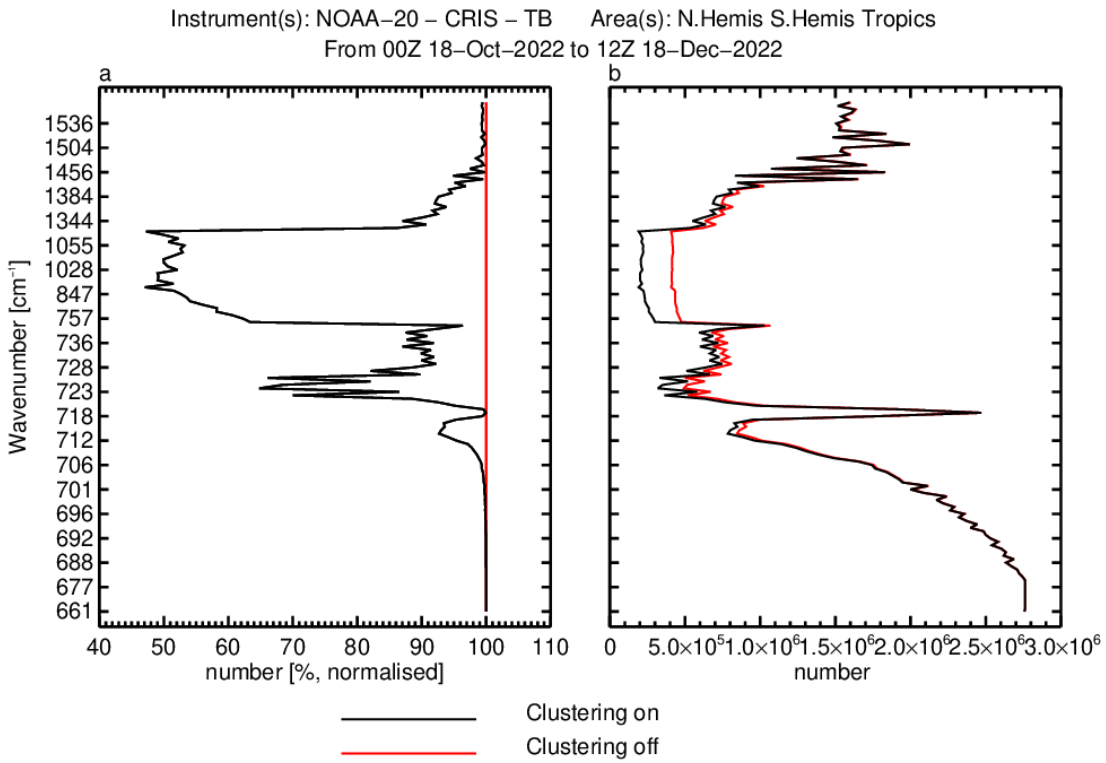



The reduction in counts when the imager check is applied is very similar between IASI and CrIS.

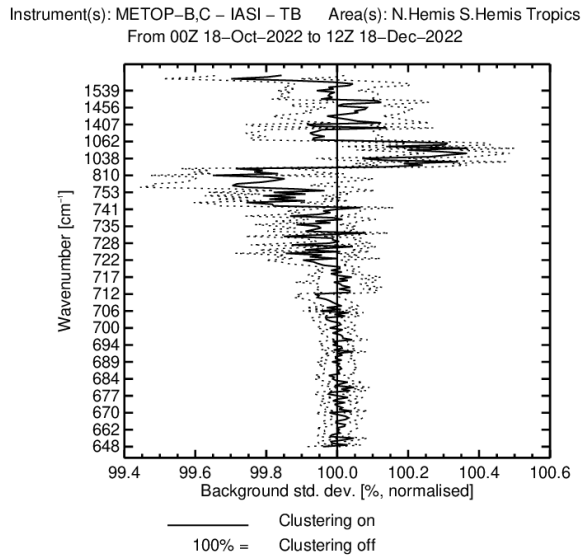
Lower-peaking channels show the largest reductions.

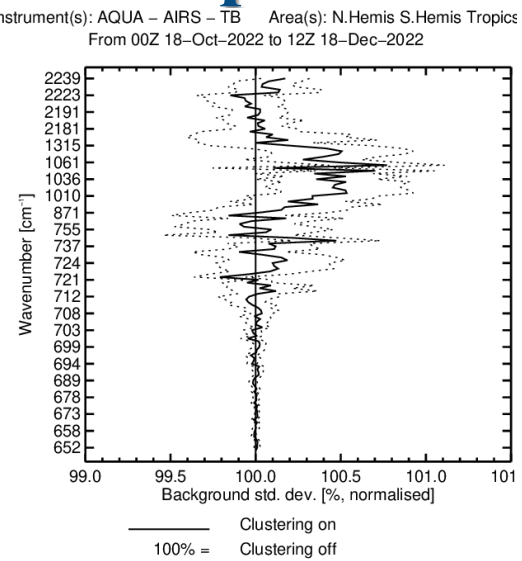
# Data Assimilation Experiment in ECMWF system

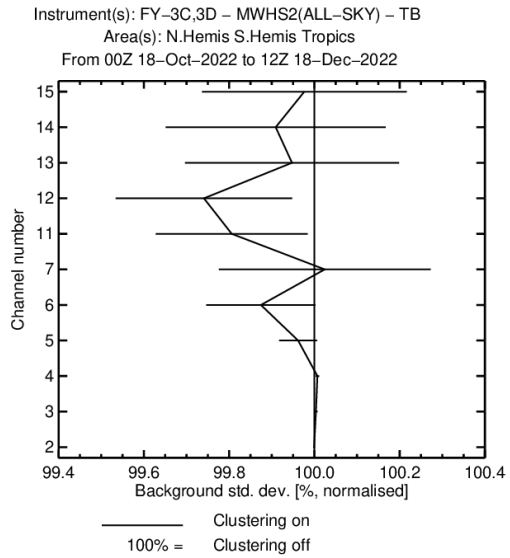
- So far, the experiment from 18th October to 18th December 2022.
- The thresholds applied are the same as for AVHRR based on the similarity of the statistics – no tuning has been done The experiments are based on a full observing system.
- Only NOAA-20 CrIS is used in either experiment.
- For a clean comparison, the “control” uses NOAA-20 CrIS data from the BUFR data stream but without the collocated imager information.

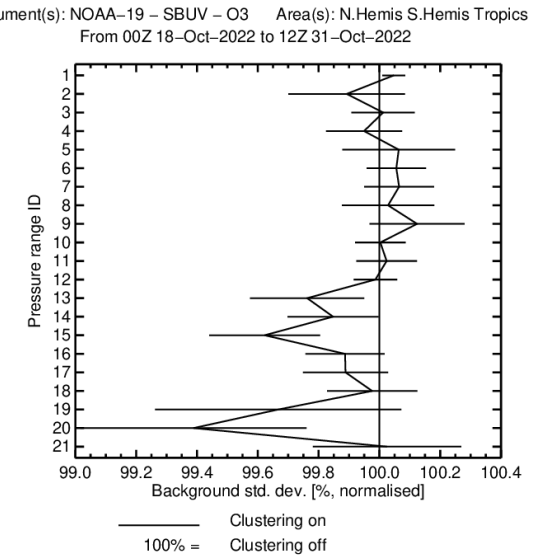


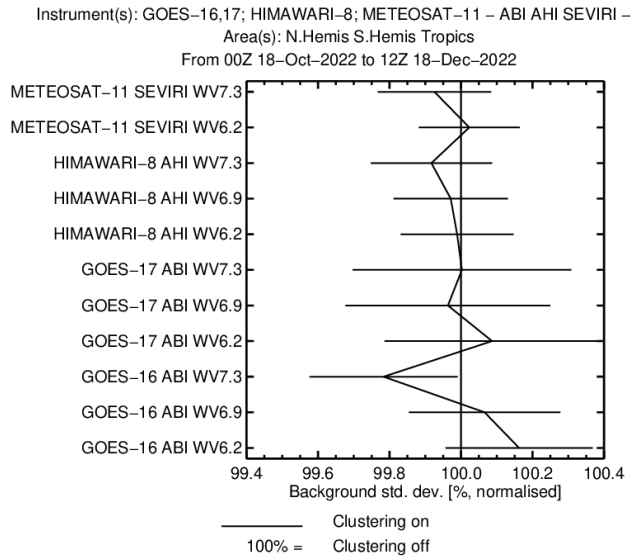
# O-B fit to independent obs

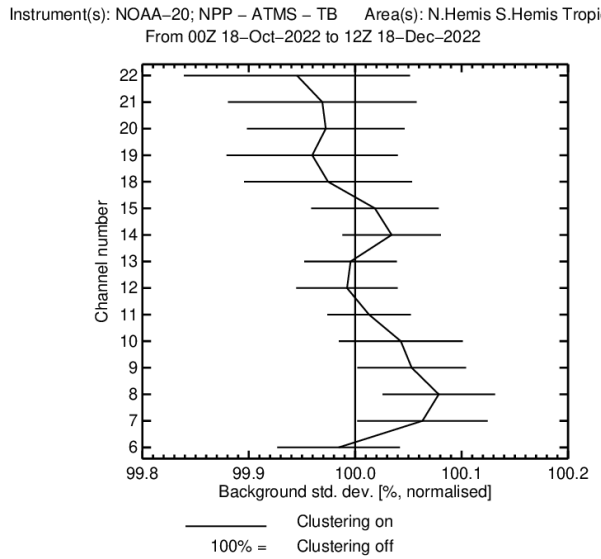


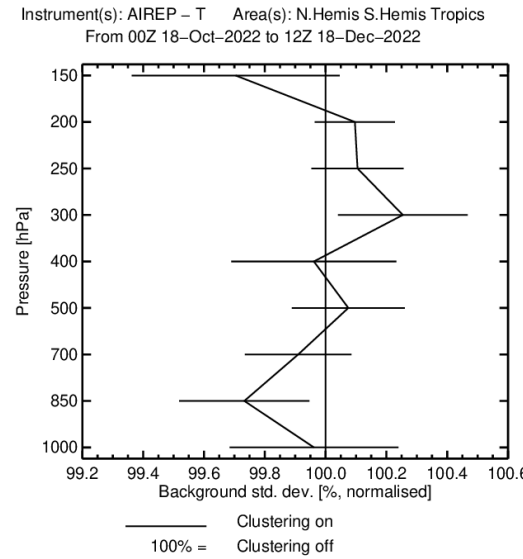



# Conclusion

- To support NWP model data assimilation, we implemented VIIRS radiance cluster analysis within the CrIS FOV to characterize its scene homogeneity.

- We are working closely with the NOAA ASSISSTT team for operational implementation.

- Preliminary data assimilation experiment using the test VIIRS cluster datasets indicates:
  - CrIS VIIRS cluster analysis is consistent with IASI-AVHRR in the statistical analysis
  - Applying the clustering check results in sensible reductions in used observations for lower-peaking CrIS channels as expected.
  - The impacts of CrIS-VIIRS cluster analysis is neutral with improvements at lower IASI channels.