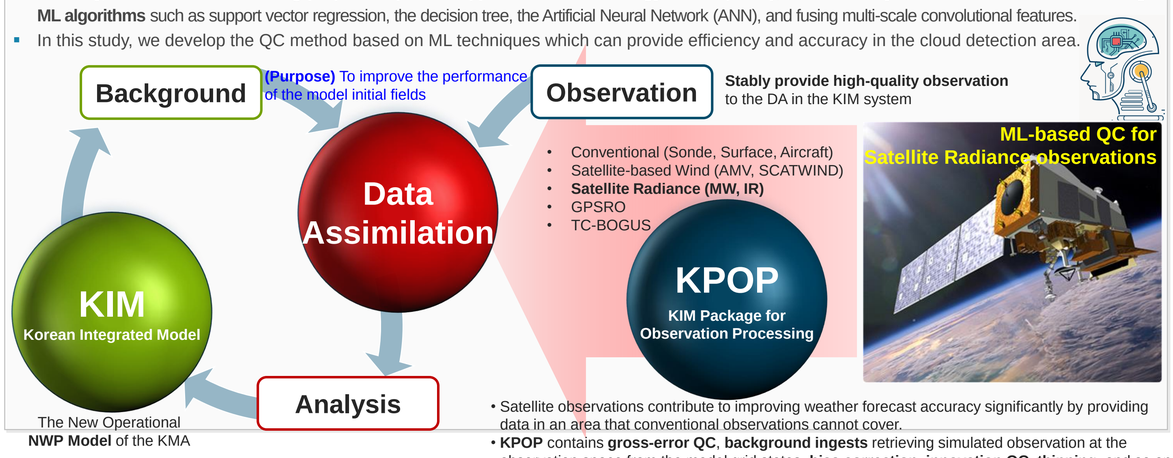
A Study on Machine Learning-Based Quality Control Techniques for the Satellite Radiance Data Assimilation

*Eun-jin Kim (ejkim@kiaps.org), Hyeon-Ju Jeon, Jeon-Ho Kang, In-Hyuk Kwon

Data Assimilation Group / Korea Institute of Atmospheric Prediction systems (KIAPS)

1 Introduction

- To improve the accuracy of weather forecasting, research on data assimilation (DA) with various qualified observations is being actively conducted.
- **Recently, various machine learning (ML) techniques are applied to the meteorology field.** Especially, there are **many studies for cloud detection using ML algorithms** such as support vector regression, the decision tree, the Artificial Neural Network (ANN), and fusing multi-scale convolutional features.
- In this study, we develop the QC method based on ML techniques which can provide efficiency and accuracy in the cloud detection area.



The New Operational **NWP Model** of the KMA

- Satellite observations contribute to improving weather forecast accuracy significantly by providing data in an area that conventional observations cannot cover.
- **KPOP** contains **gross-error QC**, **background ingests** retrieving simulated observation at the observation space from the model grid states**, bias correction, innovation QC, thinning**, and so on.

2 Previous Study (for MHS* humidity sounding)

*Microwave Humidity Sounder

- We developed the ML-based QC method using the estimating model of the MHS TB for the rigorous cloud detection.
- As a result of RMSD against IFS analysis, it shows a neutral impact compared to CTL.
  - ✓ Too strict to provide the sufficient water vapor information for the humidity sounding channels and analysis field.
  - ✓ Verified the applicability of the ML-based QC method, however, the limitations existed in the model design using the KPOP results only.

Performance of the DA experiments

2021. 05. 16. 18UTC ~ 06. 16. 00UTC

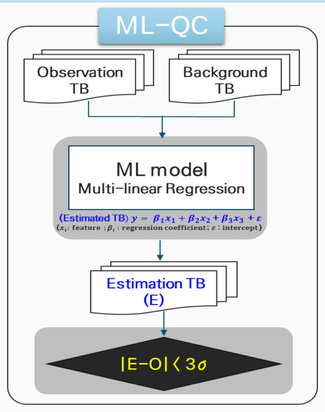


Fig 1. Flowchart of ML-QC method for MHS radiance

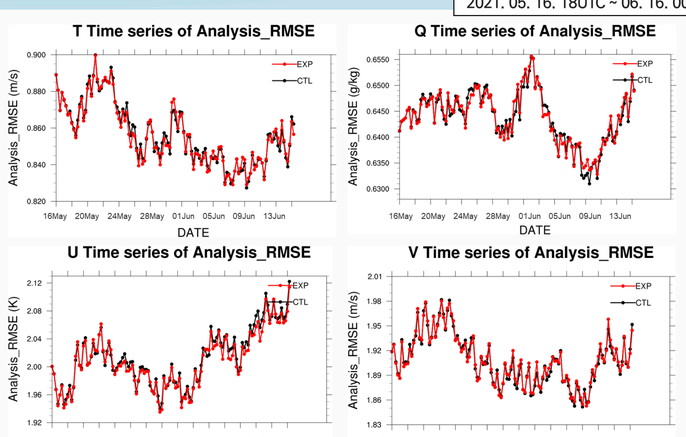


Fig 2. Time Series of Global RMSD against IFS analysis

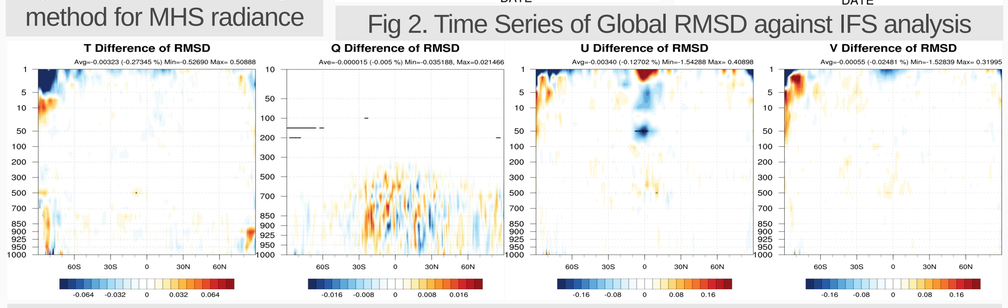


Fig 3. Zonal Mean DIFF (EXP-CTL)

3 Machine Learning Model (for ATMS*)

*Advanced Technology Microwave Sounder

- We develop the QC method based on ML techniques which can provide efficiency and accuracy in the microwave temperature sounding channels of ATMS using the information of observation (O), background (B), and analysis (A).

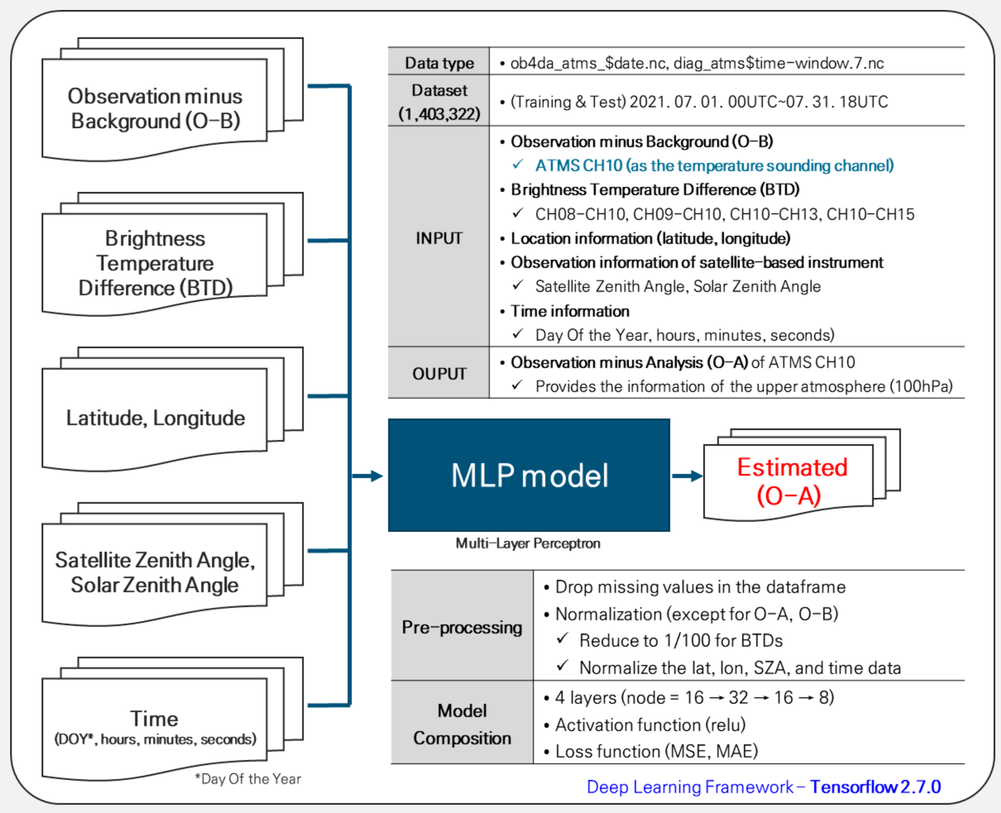


Fig 4. Description of the Machine Learning Model for the QC using the ATMS radiance

4 Results

- The MLP model is trained by the relationship between the O−A and the information of the upper atmosphere (over 100 hPa) in terms of the space-time with the satellite-based instrument.

Estimated O−A vs. O−A

- We use the temperature sounding channels as the features to predict the O−A values.
- Show the high performance to estimate the actual O−A (CH10) with the MSE of 0.06 K in the test dataset.

Table 1. ATMS channel characteristics

Channel	Frequency (GHz)	Peak WF (hPa)
1	23.8	Surface
2	31.4	Surface
3	**50.3**	**Surface**
4	**51.76**	**950**
5	**52.8**	**850**
6	**53.596 ± 0.115**	**700**
7	**54.4**	**400**
8	**54.94**	**250**
9	**55.5**	**200**
10	**57.29**	**100**
11	**57.29 ± 0.217**	**50**
12	**57.29 ± 0.322 ± 0.048**	**25**
13	**57.29 ± 0.322 ± 0.022**	**10**
14	**57.29 ± 0.322 ± 0.010**	**5**
15	**57.29 ± 0.322 ± 0.0045**	**2**
16	88.2	Surface
17	165.5	Surface
18	183.31 ± 7.0	800
19	183.31 ± 4.5	700
20	183.31 ± 3.0	500
21	183.31 ± 1.8	400
22	183.31 ± 1.0	300

Tian & Zou, 2018

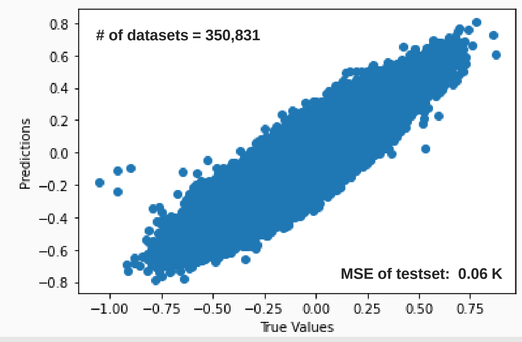


Fig 5. the MLP model's performance using the test datasets

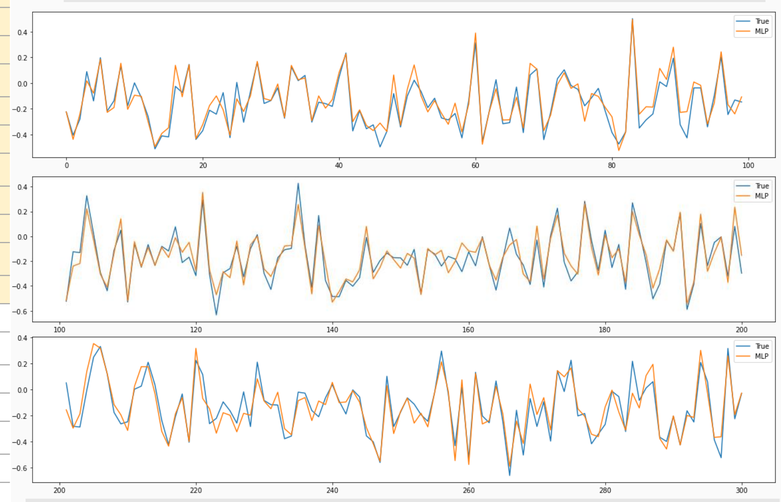
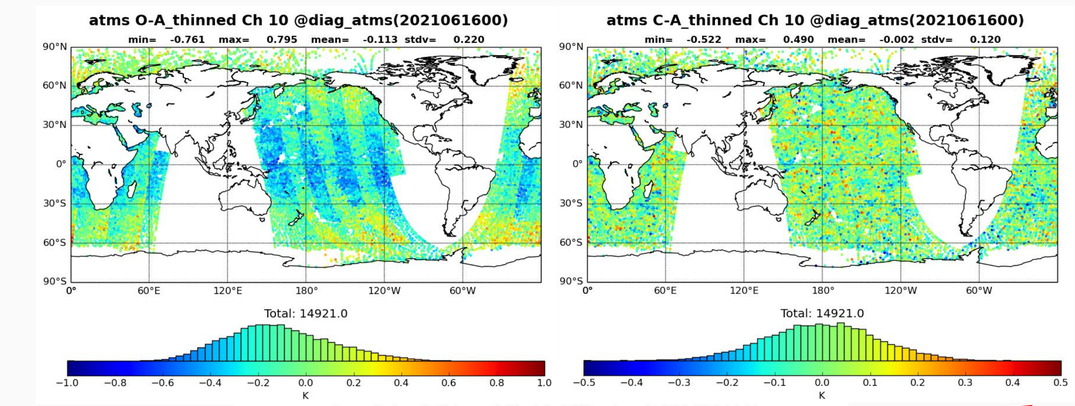


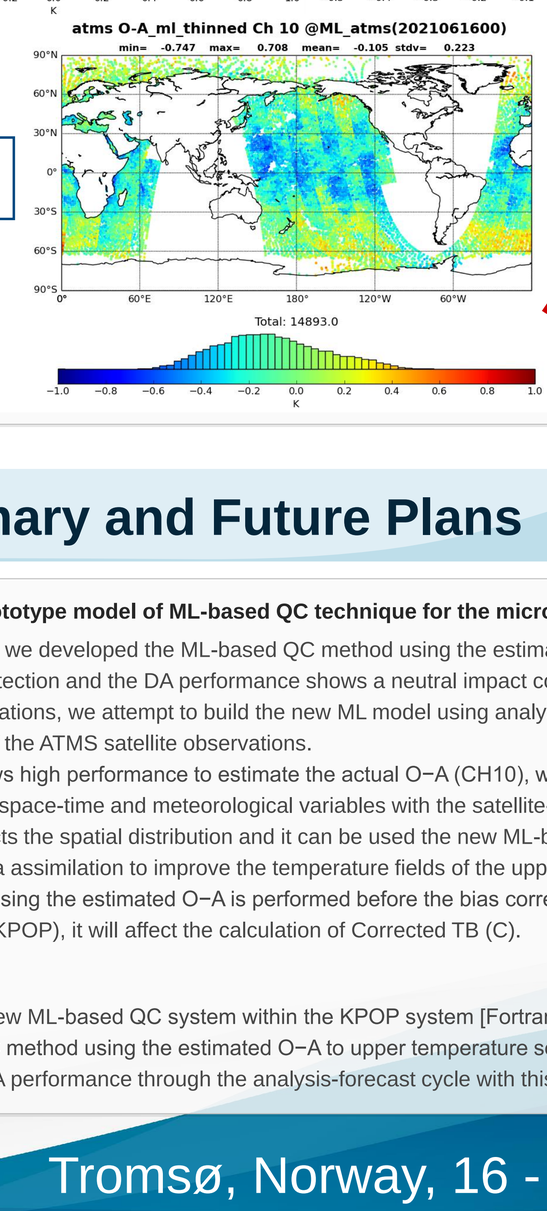
Fig 6. Time series of the O−A estimated (MLP) against O−A (True)

Availability of estimated O−A

- The ML model not only predicted well in terms of statistics, but also similarly predicted the spatial distribution.
- We can calculate the mean and standard deviation values of the estimated O−A in every scene and it can be used as the threshold for the new ML-based QC method in the satellite radiance (ATMS) data assimilation.



Estimated O−A (ATMS CH10)



Future work !!

5 Summary and Future Plans

Development of the prototype model of ML-based QC technique for the microwave satellite radiance DA

- In our previous study, we developed the ML-based QC method using the estimating model of the MHS TB for the rigorous cloud detection and the DA performance shows a neutral impact compared to CTL.
- To overcome the limitations, we attempt to build the new ML model using analysis (A) for the temperature sounding channels of the ATMS satellite observations.
- The MLP model shows high performance to estimate the actual O−A (CH10), which can consider the relationship between space-time and meteorological variables with the satellite-based instrument.
- It also similarly predicts the spatial distribution and it can be used the new ML-based QC method in the satellite radiance data assimilation to improve the temperature fields of the upper atmosphere.
- If the ML-based QC using the estimated O−A is performed before the bias correction in the Observation Processing System (KPOP), it will affect the calculation of Corrected TB (C).

Futrue Plans

- Establishment the new ML-based QC system within the KPOP system [Fortran ↔ Python]
  - ✓ Applying the QC method using the estimated O−A to upper temperature sounding channels.
- Verification of the DA performance through the analysis-forecast cycle with this new ML-based QC.