Precipitation Estimation using Himawari-8 in Vietnam

Nguyen Hung An

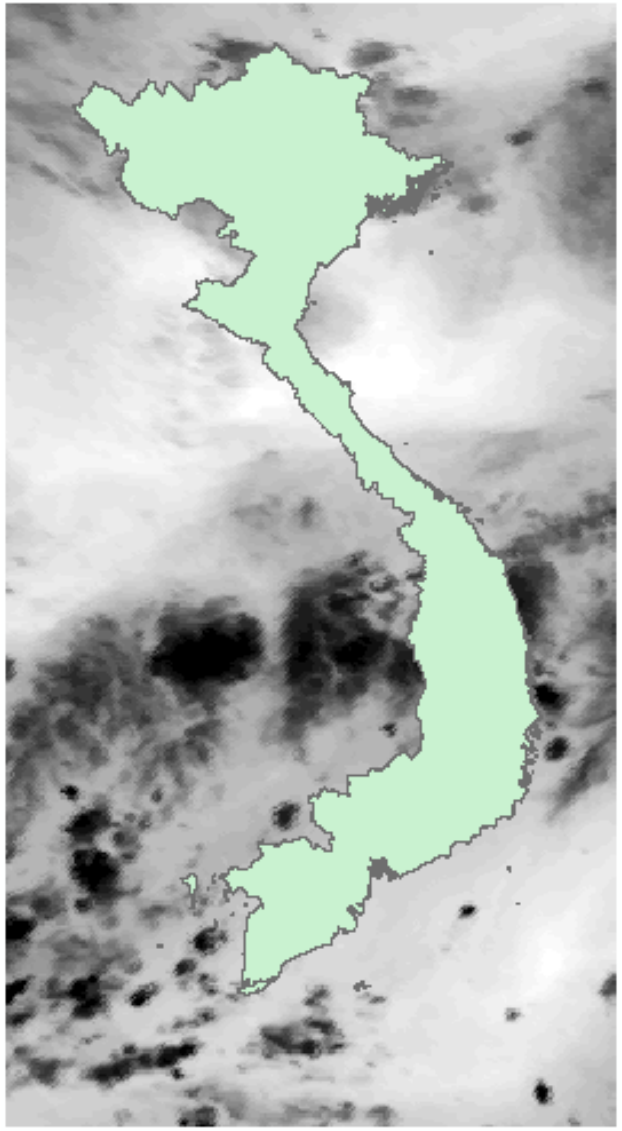
Le Quy Don Technical University

Outline

- Introduction
- Case studies
- Data
- Methodology
- Result and Discussion
- Conclusion

Introduction

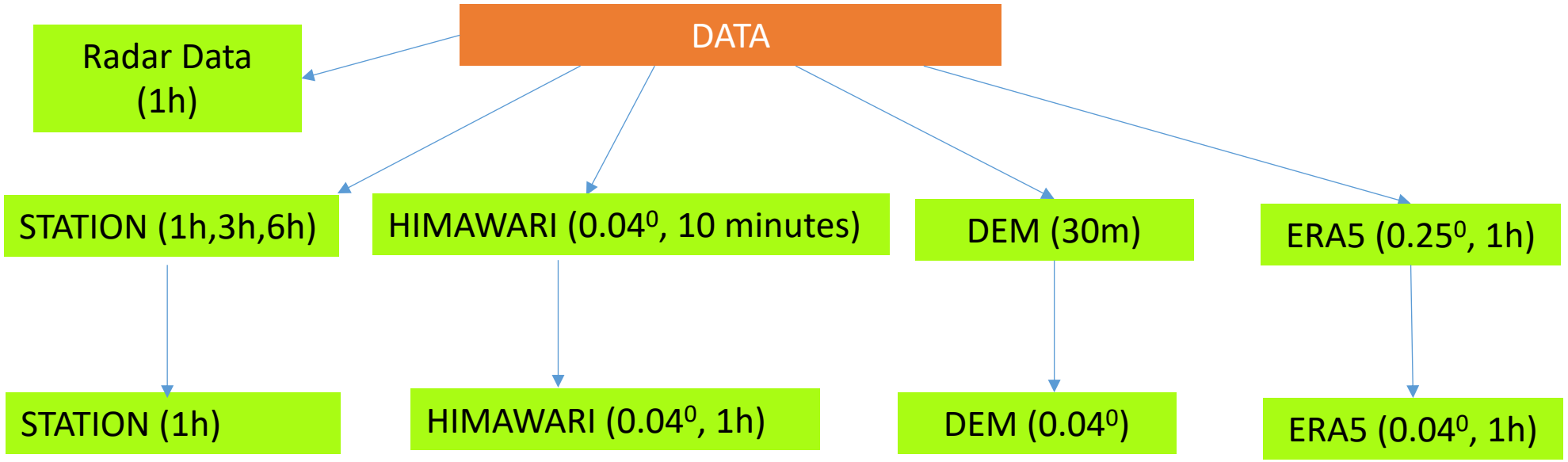
- Vietnam is an area of complex weather characteristics, which is clearly affected by climate changes.
- The increased occurrence frequency of strong storms and heavy rains for many days in the large extent caused floods, flash floods, landslides continuously, and severe consequences of properties and people's lives.
- Precipitation estimation for the nationwide range in real-time has an important significance



Case study

No	Period	Total Precipitation
1	02-10/8/2019	Đong Nai: 408 mm; Buon Ma Thuot (Đak Lak) 379 mm; Ea mat (Đak Lak): 349 mm; Đak Nong: 317mm... Especially, at Phu Quoc Island (Kien Giang) heavy rain, total precipitation measured from 2 to10/8 is 1160mm.
2	02 - 05/9/2019	Nghe An - Quang Ngai; Sa Pa (Lao Cai): 215 mm; Ngòi Thia (Yên Bai): 207 mm; Cô Tô (Quang Ninh): 259 mm; Ha Tinh (Ha Tinh): 266 mm; Ky Anh (Ha Tinh): 226 mm; La Khê (Ha Tinh): 225 mm; Cẩm Nhượng (Ha Tinh): 262 mm; Mai Hóa (Quang Binh): 222 mm
3	14-16/09/2019	The station Nguyễn Hữu Cảnh – TPHCM: 123,7mm Areas of Tây Nguyên, Nam Bộ và Khánh Hòa, Ninh Thuận, Bình Thuận, popular precipitation: 50 - 100 mm/24 h,
4	13;15 - 17/10/2019	Hương Khê: 362mm; Sơn Diệm: 383mm; Chu Lễ: 401mm; Linh Cảm: 466mm
5	20-21/07/2020	350mm for 10 h (từ 0h-10h/21/7) at Hà Giang province
6	01 - 02/08/2020	Thanh Hóa - Quảng Trị : 200 - 500 mm
7	06/08/2020	Station Mạc Đĩnh Tri- TPHCM:
8	14-15/09/2020	TPHCM, Tây Ninh, precipitation measured for 24h at Cát Lái station: 16,4mm; Củ Chi station: 26,8,mm; Hóc Môn station: 72,8mm...
9	17 - 19/9/2020	Thạch Đồng: 138mm; Kỳ Anh: 102mm; Hoành Sơn: 83mm; Cẩm Nhượng: 57mm
10	06 - 10/10/2020	Hương Trạch, Chu Lễ, Hòa Duyệt, Hương Khê : 434 – 529 mm
11	15 - 20/10/2020	Hoành Sơn, Kỳ Anh, Thạch Đồng, Cẩm Nhượng, Tp Hà Tĩnh: 759 - 1377mm
12	28 - 31/10/2020	Thạch Đồng: 588mm; Tp Hà Tĩnh: 697mm
13	14 - 16/11/2020	Kỳ Anh: 167mm; Hoành sơn: 147mm

Datasets



Study Grid:

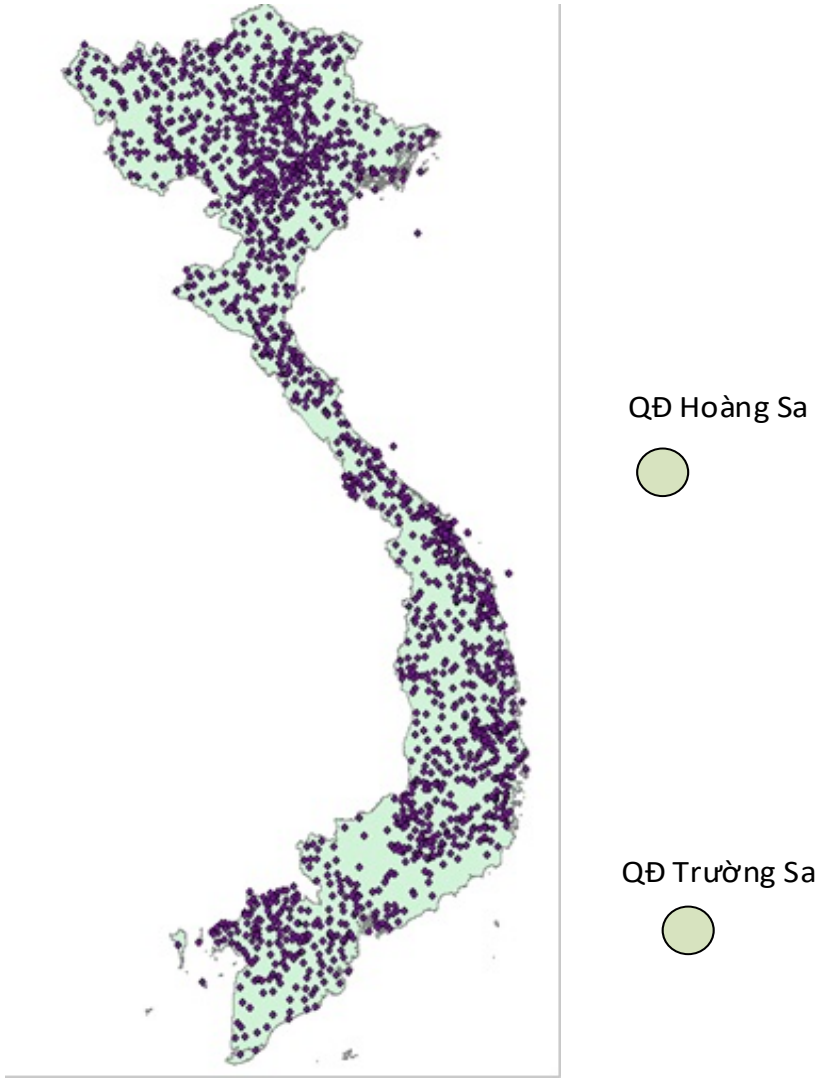
Spatial resolution: 4km x 4km (0.04^0 x 0.04^0)

Temporal resolution: 1 h

Applied grid: Left: 101.1^0; Top: 24.0^0; Right: 110.6^0; Bottom: 6.5^0

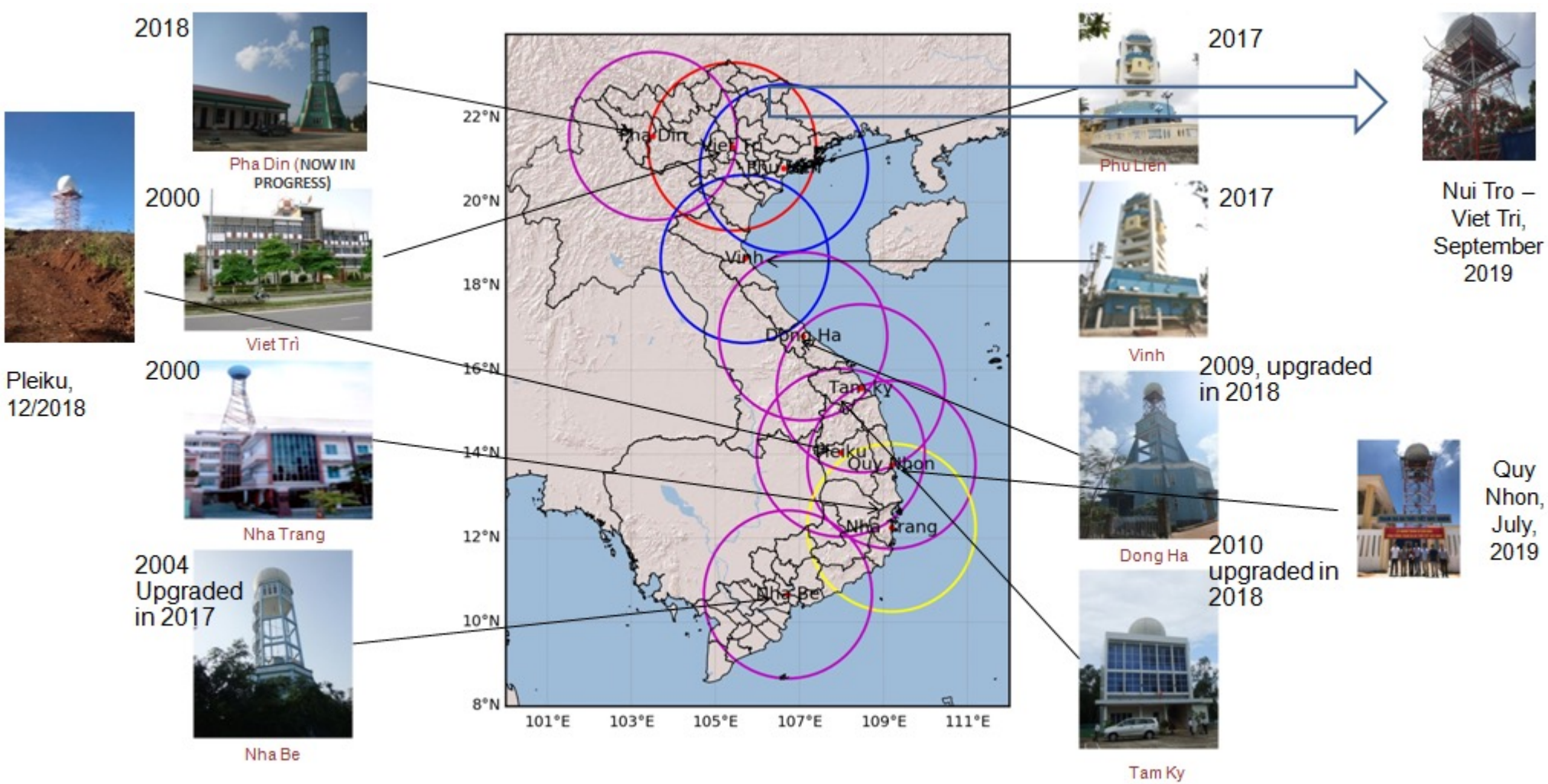
Rain gauges

- Number of stations: 1500 stations (~800 station/day)
- Resolution: 1h
- Precipitation unit: mm/h
- Data duration: 2019 - 2020



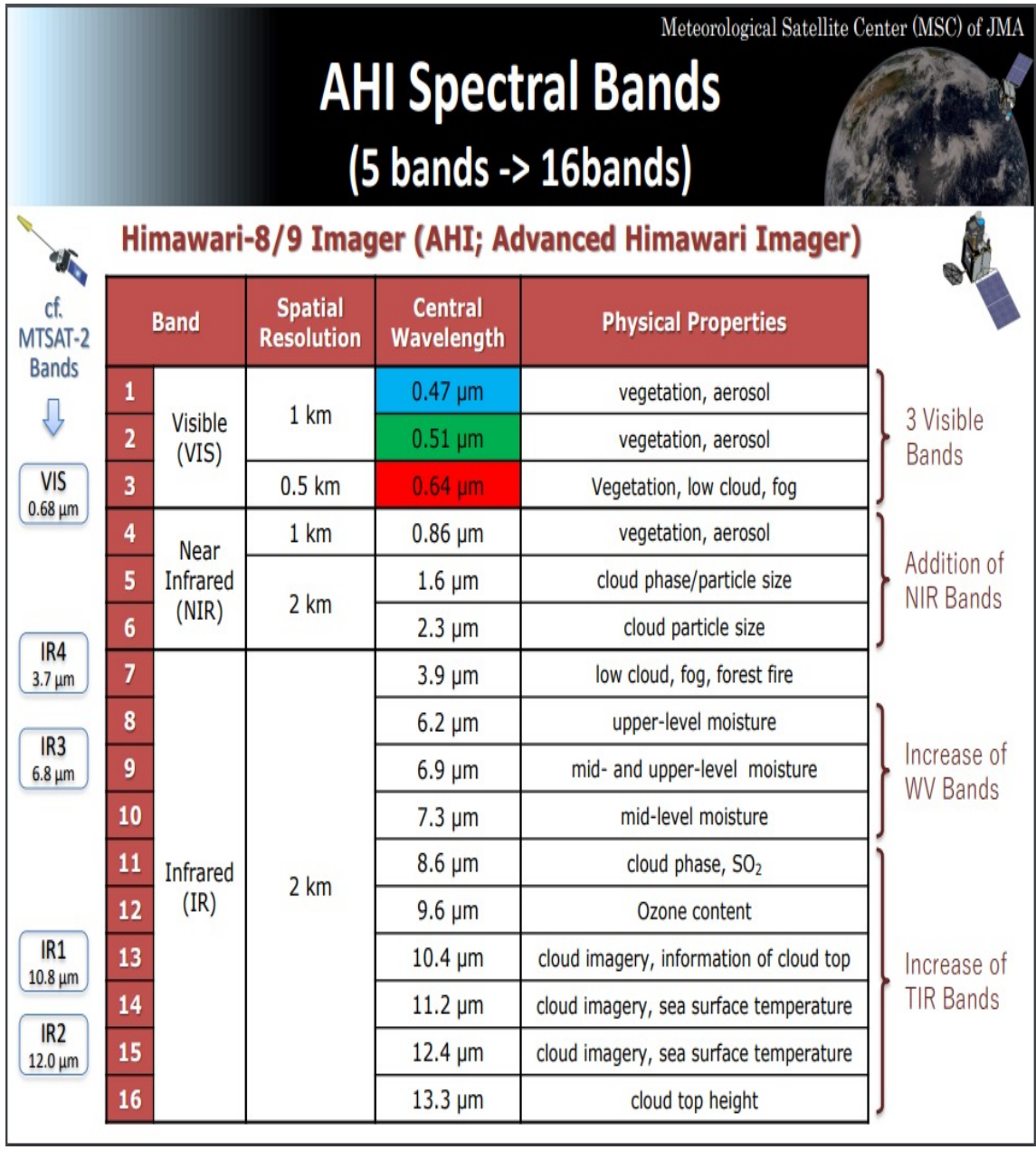
Distribution of rain gauges in Vietnam

Radar


2018
Pha Din (NOW IN PROGRESS)
2000
Viet Trì
2000
Nha Trang
2004
Upgraded
in 2017
Nha Be
Pleiku,
12/2018
2017
Phu Lien
2017
Vinh
2009, upgraded
in 2018
Dong Ha
2010
upgraded in
2018
Tam Ky
Nui Tro –
Viet Tri,
September
2019
Quy
Nhon,
July,
2019
Pha Din
Vinh
Dong Ha
Tam Ky
Pleiku
Quy Nhon
Nha Trang
Nha Be
22°N
20°N
18°N
16°N
14°N
12°N
10°N
8°N
101°E
103°E
105°E
107°E
109°E
111°E

Himawari-8

- Space resolution: 4 km
- Temporal resolution: 1h
- Data fields: 9x9x9
  - Brightness temperatures (BT) of 9 individual bands (6.2, 6.9, 7.3, 8.6, 9.6, 10.4, 11.2, 12.4, 13.3 μm)
  - Temporal features: 3 hour averaged BT of 9 bands
  - Spatial features: 3x3 pixel averaged BT of 9 bands

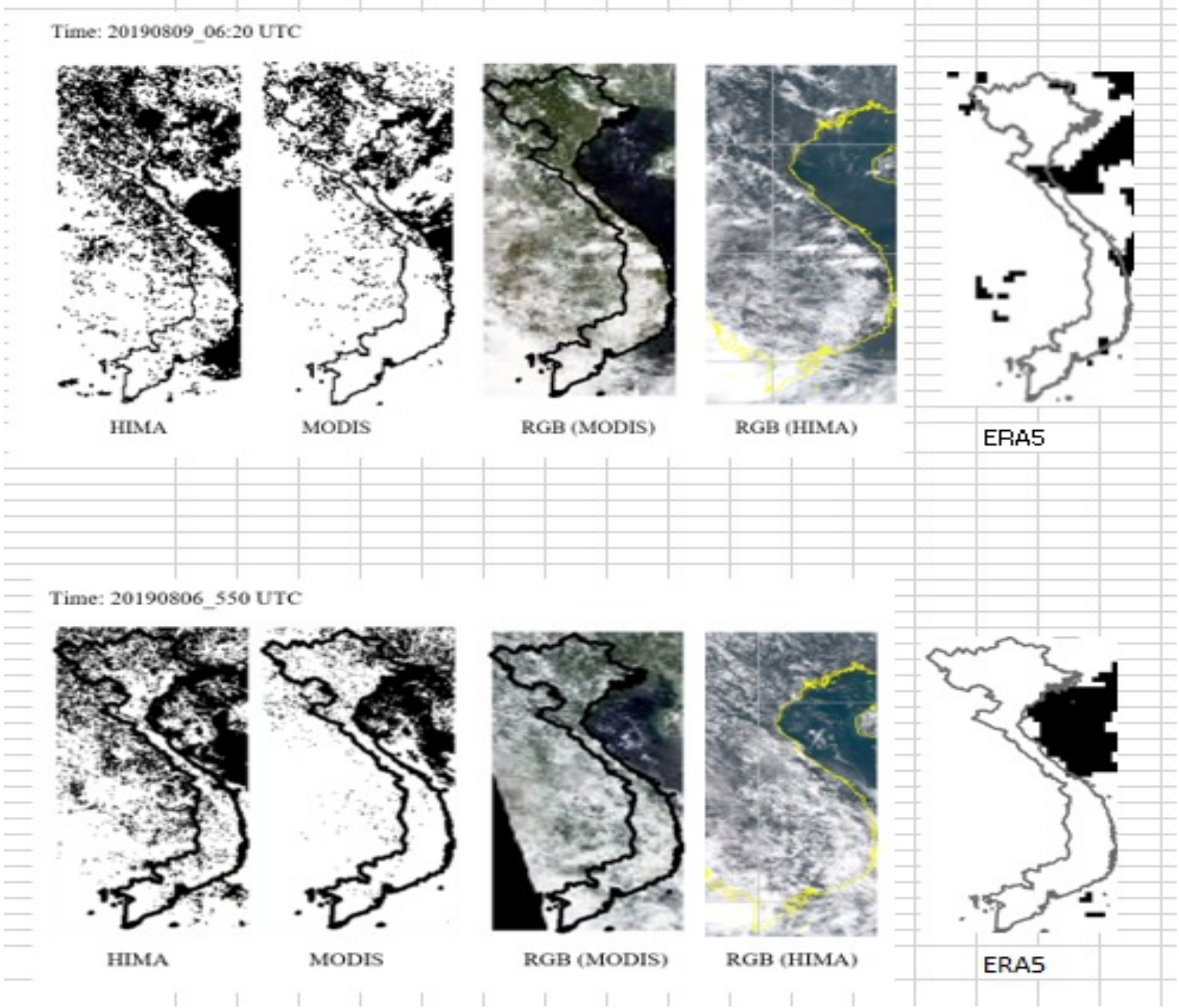


Band		Spatial Resolution	Central Wavelength	Physical Properties	
1	Visible (VIS)	1 km	0.47 μm	vegetation, aerosol	3 Visible Bands
2			0.51 μm	vegetation, aerosol	
3		0.5 km	0.64 μm	Vegetation, low cloud, fog	
4	Near Infrared (NIR)	1 km	0.86 μm	vegetation, aerosol	Addition of NIR Bands
5		2 km	1.6 μm	cloud phase/particle size	
6			2.3 μm	cloud particle size	
7	Infrared (IR)	2 km	3.9 μm	low cloud, fog, forest fire	
8			6.2 μm	upper-level moisture	Increase of WV Bands
9			6.9 μm	mid- and upper-level moisture	
10			7.3 μm	mid-level moisture	
11			8.6 μm	cloud phase, SO_2	Increase of TIR Bands
12			9.6 μm	Ozone content	
13			10.4 μm	cloud imagery, information of cloud top	
14			11.2 μm	cloud imagery, sea surface temperature	
15			12.4 μm	cloud imagery, sea surface temperature	
16			13.3 μm	cloud top height	

Cloud masks

- Himawari cloud products only for daytime
- Modis cloud products have low time resolution (1 image /day) in Vietnam
- ERA5 cloud product of "Total cloud cover" for the whole day and the whole area of Vietnam.

=> ERA5 cloud product was chosen to produce cloud mask



DEM & ERA-5 weather datasets

- DEM: ASTER: Global Digital Elevation Map: 30m spatial resolution
- ERA-5 weather datasets
  - ISOR : a measure of distortion of the shape of the terrain in the horizontal plane (from a bird's-eye view) from a circle
  - TCW (kg m^{-2}): the sum of water vapour, liquid water, cloud ice, rain and snow in a column extending from the surface of the Earth to the top of the atmosphere, but no precipitation included.
  - TCWV (kg m^{-2}): total amount of water vapour in a column extending from the surface of the Earth to the top of the atmosphere.
  - CAPE (J kg^{-1}): an indication of the instability (or stability) of the atmosphere.
  - Total cloud cover: the proportion of a grid box covered by cloud

Methodology

Approach

- Hirose, H., Shige, S., Yamamoto, M.K., & Higuchi, A. (2019). High Temporal Rainfall Estimations from Himawari-8 Multiband Observations Using the Random-Forest Machine-Learning Method. *Journal of the Meteorological Society of Japan. Ser. II.*

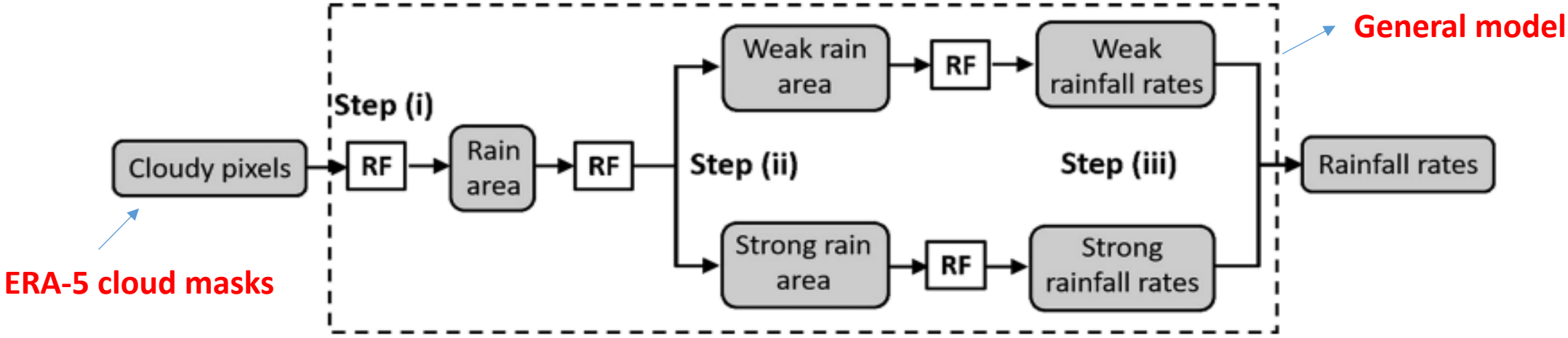


Fig. 1. Overview of the stepwise scheme for rainfall rate assignment.

General model design

STT	Variables	Sources	Model 1	Model 2
1	9 BT+ 9 temporal features + 9 spatial features	Himawari-8	X	X
2	DEM	nasa.gov	X	X
3	ISOR, TCW, TCWV, CAPE	ERA5		X

DATA						
Total samples	No of cloud samples	No of clear samples	No of rain samples	No of no-rain samples	No of weak rain samples	No of strong rain samples
722,708	624,585	98,123	144,264	480,318	95,067	49,197

Random Forest

- Using grid search on different values of number of trees, number of variables, and max depths to select best parameters for each RF model.

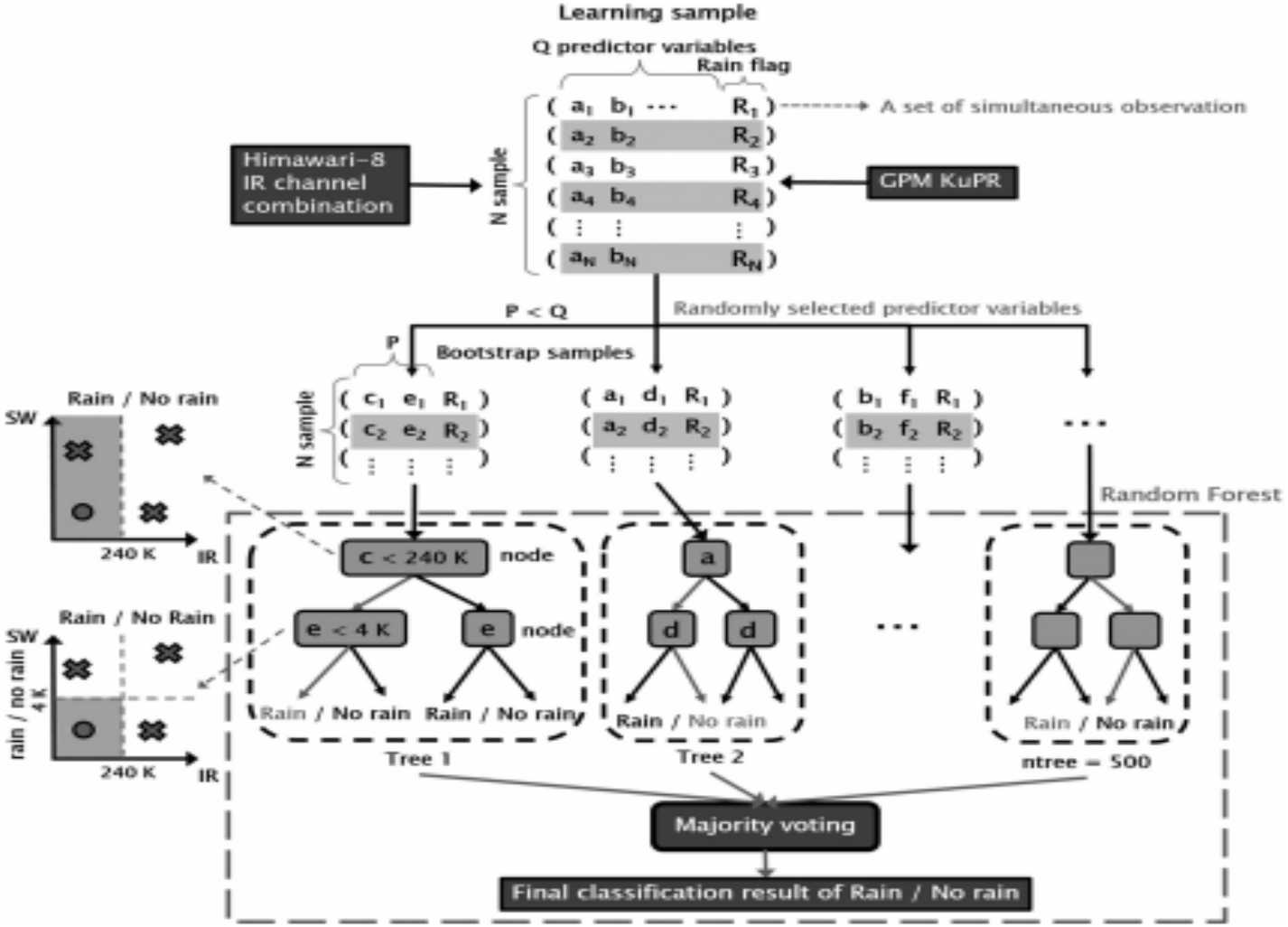
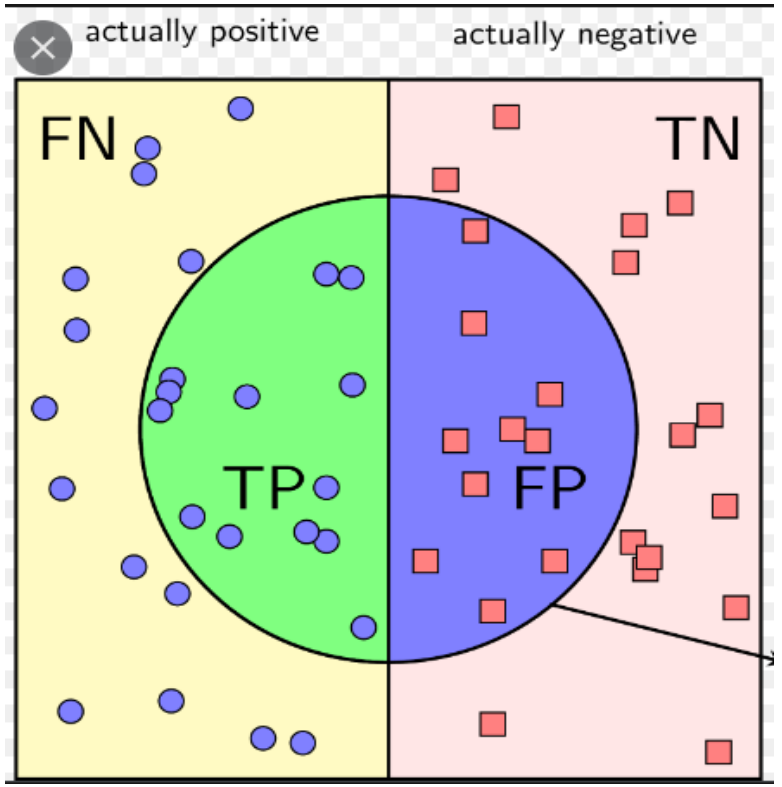


Fig. 2. Schematic diagram for making an RF model for the rain/no rain classification.

Evaluation

- Cross validation on 13 case studies
  - 12 cases for training, 1 case for testing and repeat 13 times to obtain the final average results
- Evaluation indexes



$$\text{Recall} = \frac{TP}{\text{total actual positive}} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{\text{total predicted positive}} = \frac{TP}{TP + FP}$$

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

$$\text{Accuracy} = \frac{TP + TN}{\text{total sample}}$$

Result and Discussion

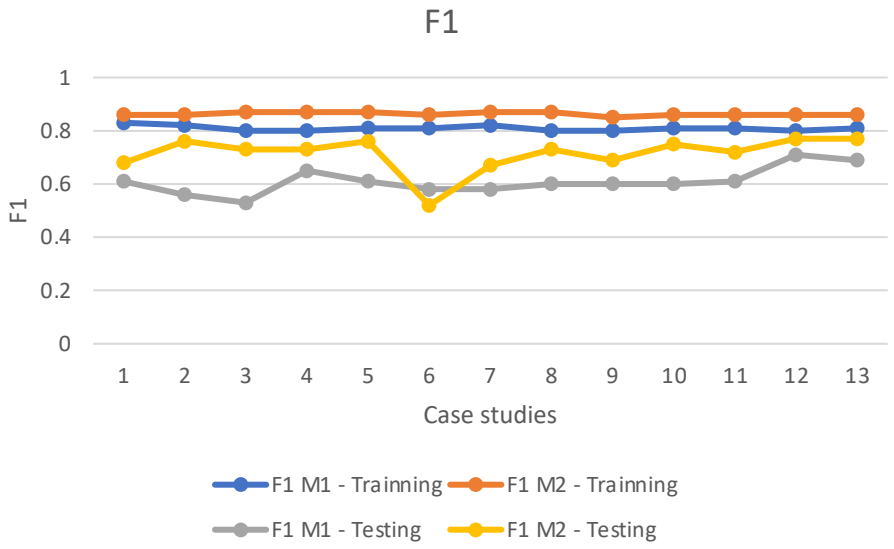
Model configuration

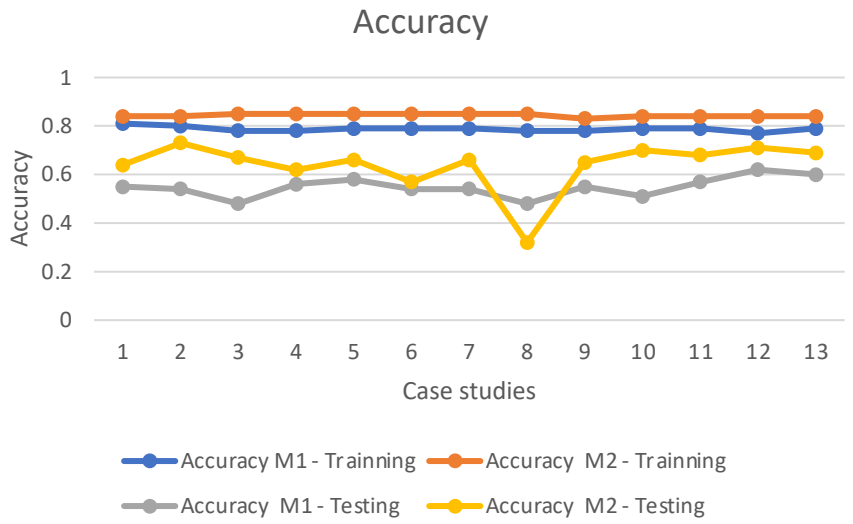
- RF1 is to classify no-rain/rain areas
- RF2 is to classify weak rain/strong rain areas
- RF3 is to estimate weak rain rates
- RF4 is to estimate strong rain rates

	RF1		RF2		RF3		RF4	
	M1	M2	M1	M2	M1	M2	M1	M2
No of trees	120	160	240	160	160	200	160	200
No of variables	16	26	22	24	24	22	22	22
Max depths:	16	16	16	16	16	16	16	16

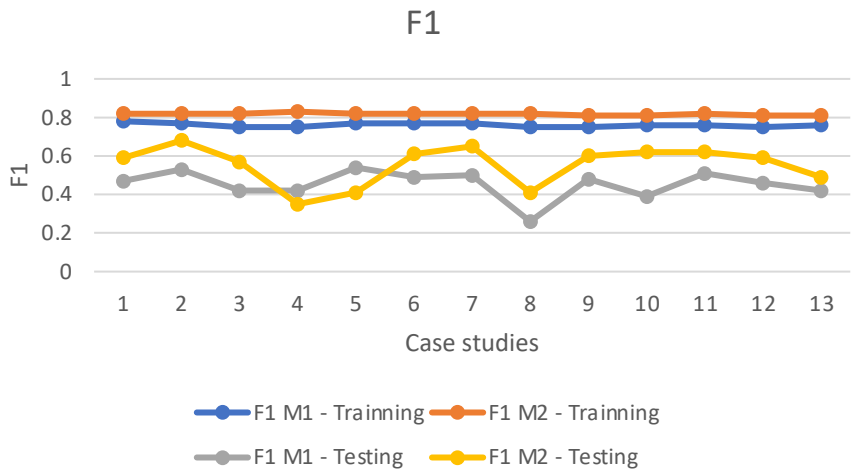
RF 1 – No rain /Rain classification

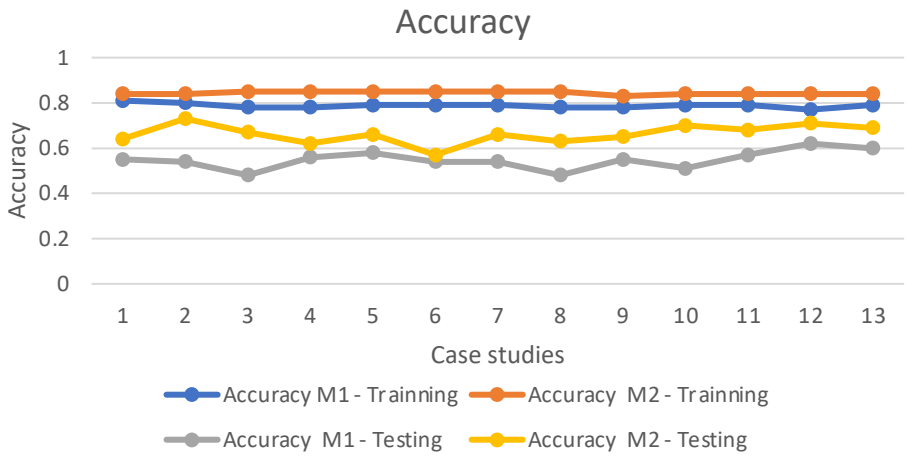
No rain





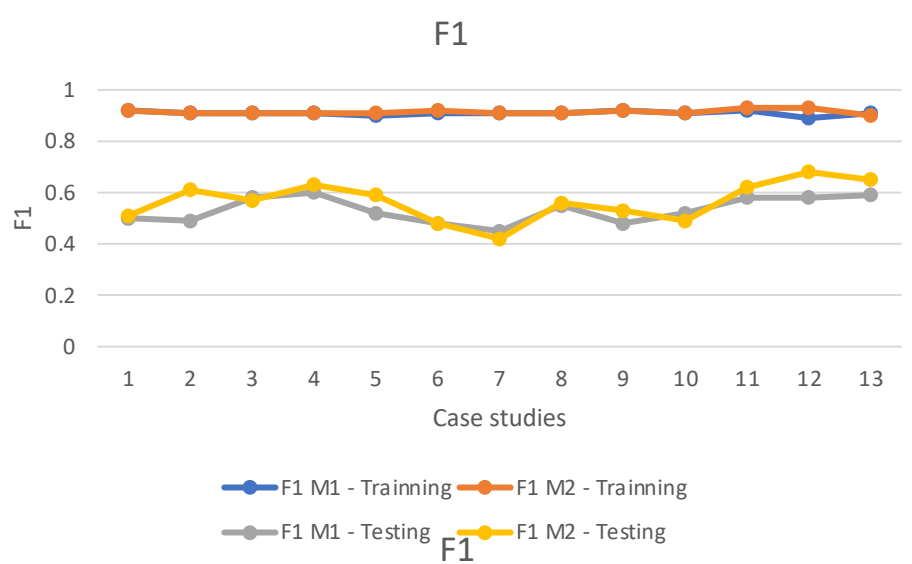
Rain

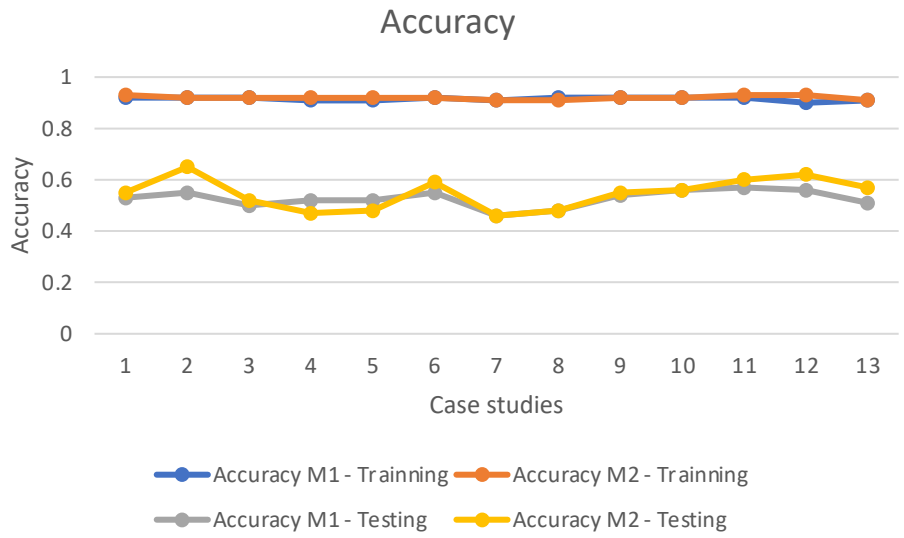




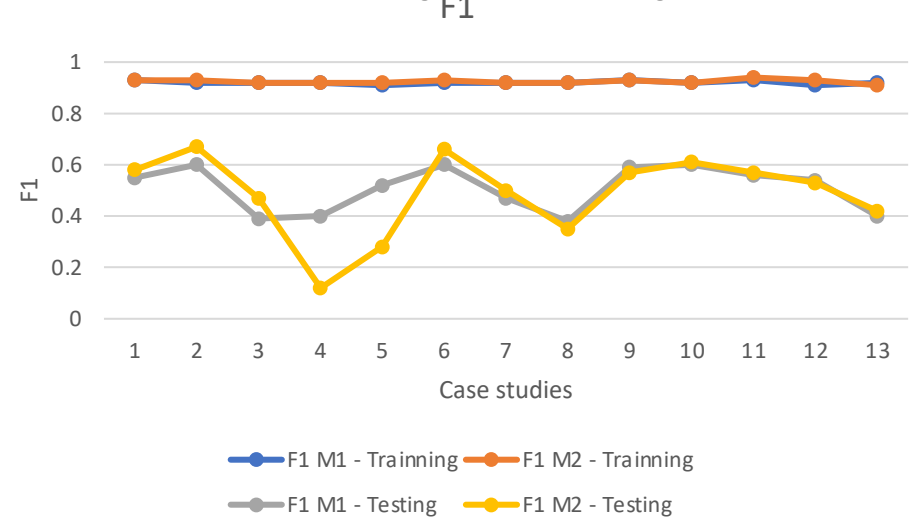
RF2- Weak rain/ Strong rain classification

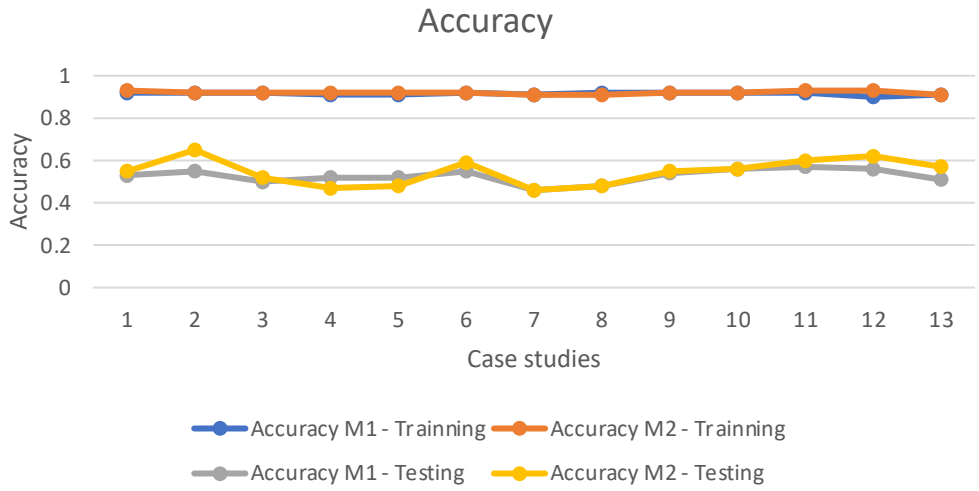
Weak rain



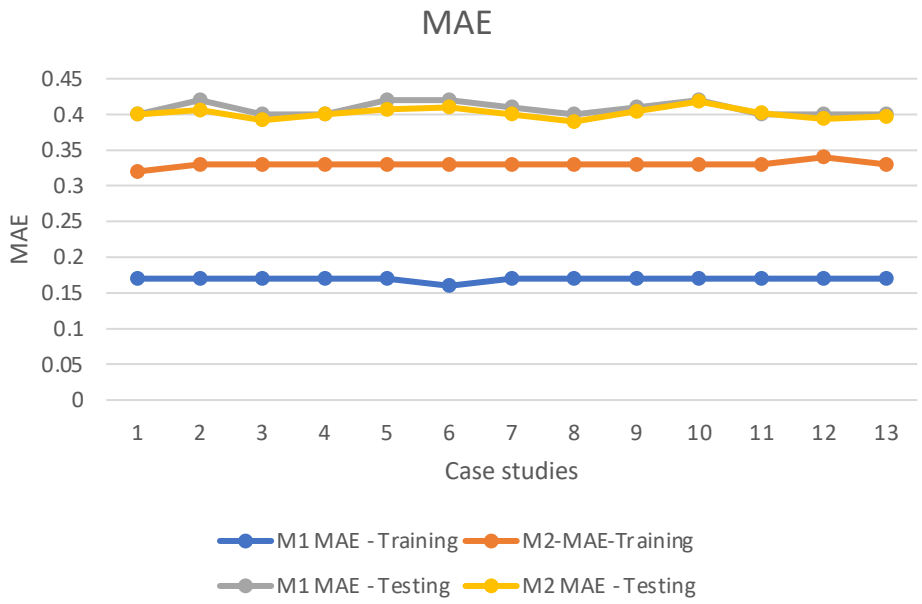


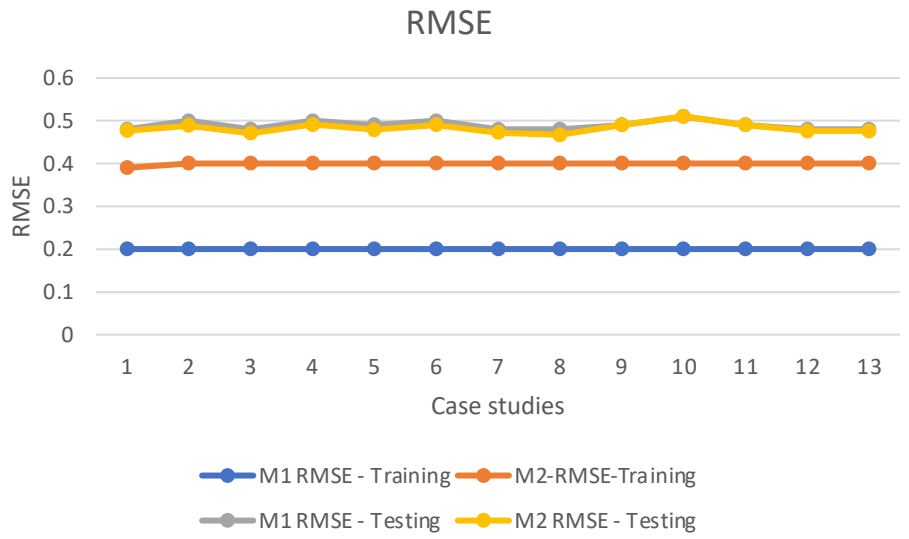
Strong Rain

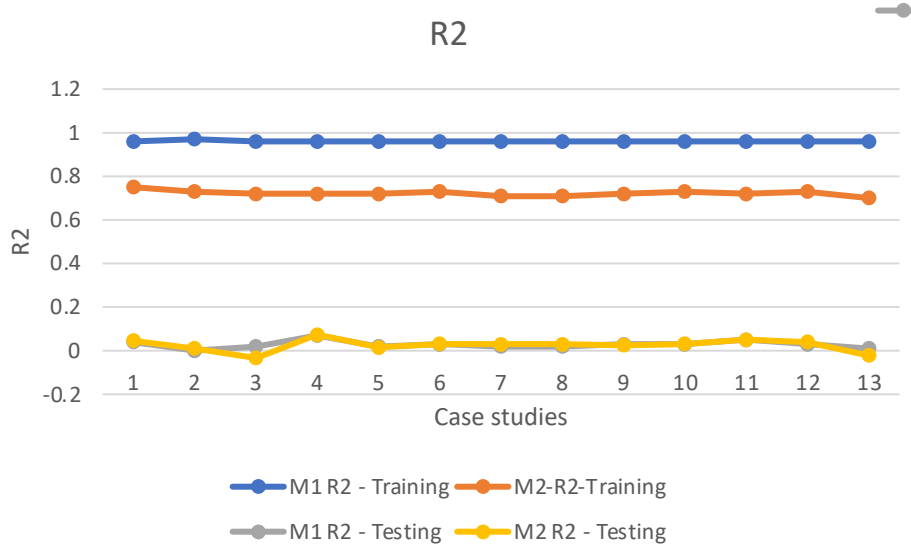




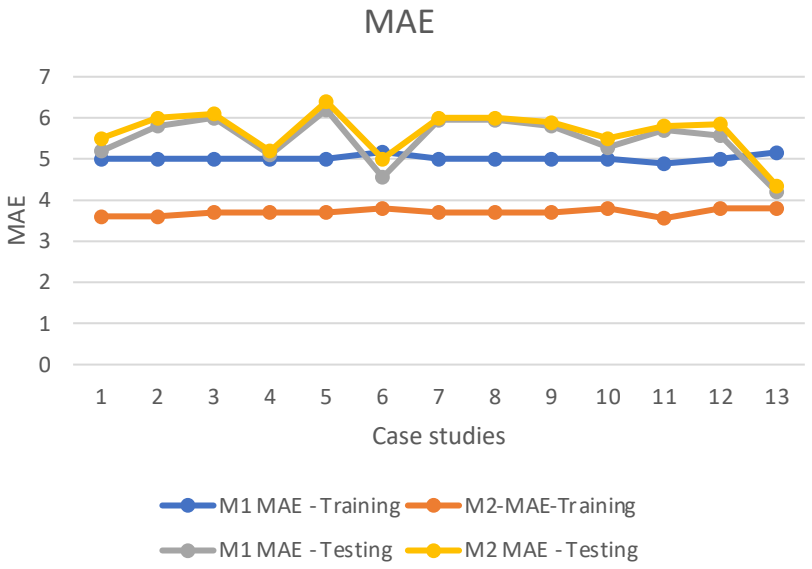
RF3- Weak rain rates

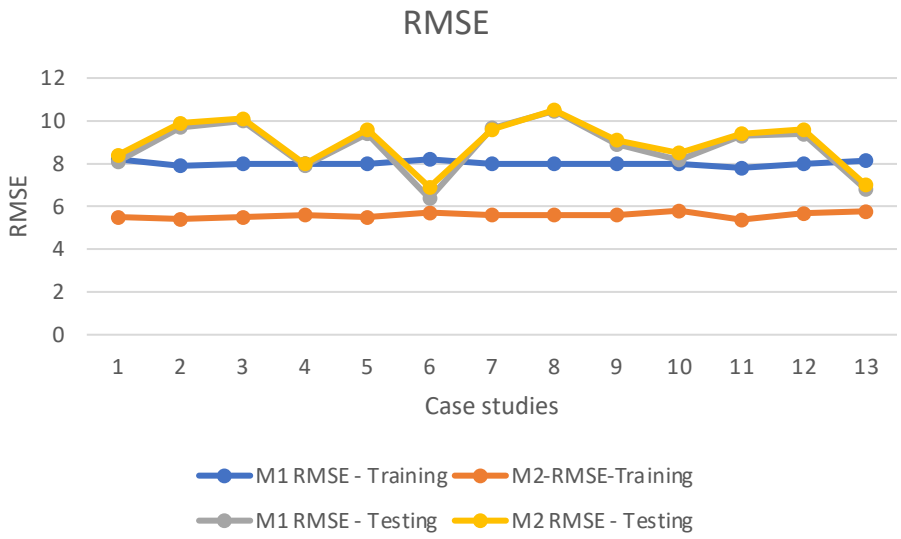


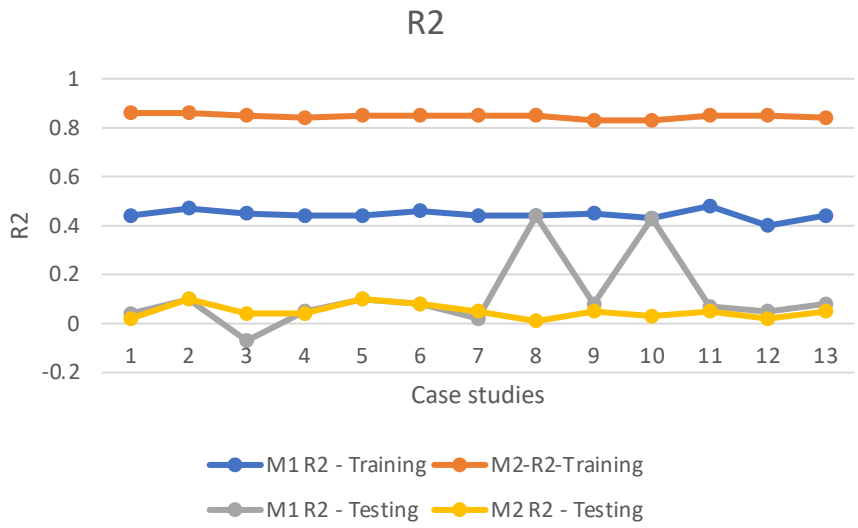




RF4- Strong rain rates







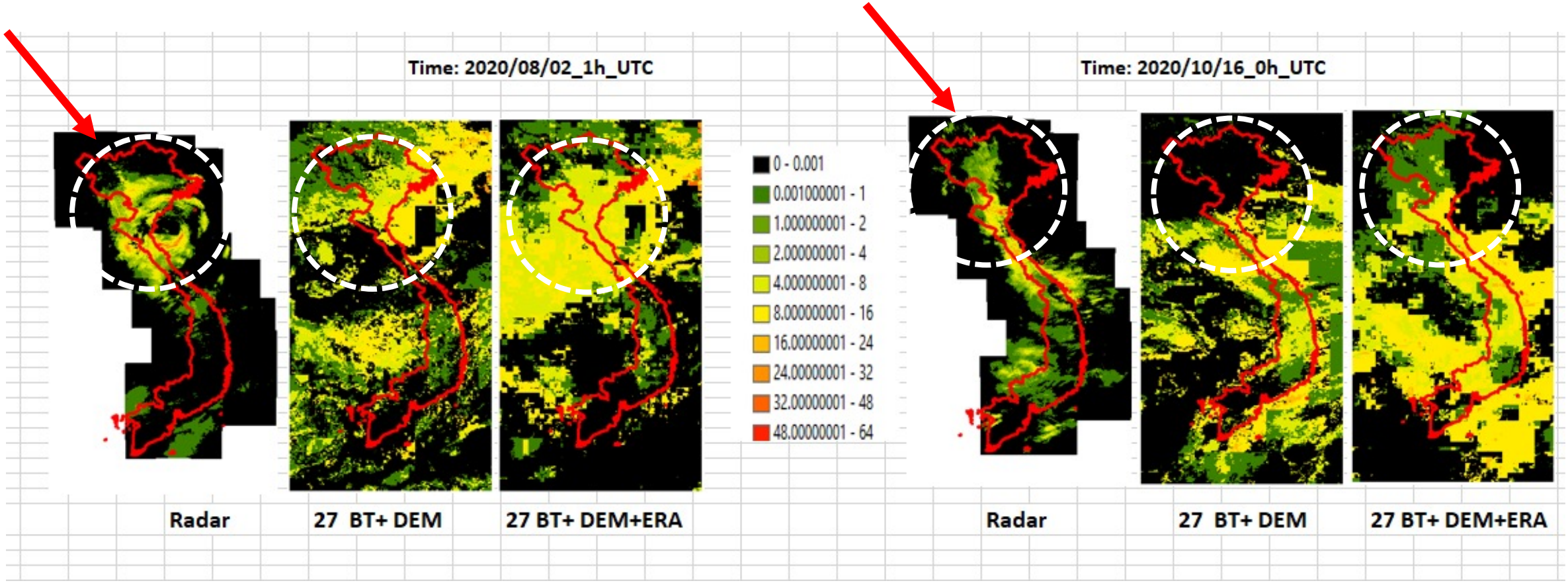
Synthetic results

- Classification: Model 2 performed slightly better than Model 1
- Estimation: Two model performance seems similar

Data test	
No of rain samples	No of no-rain samples
6176	9264

	MAE	RMSE	**R2**
M1	**3**	**5.8**	**0.01**
M2	**3.4**	**6.0**	**0.04**
Hitosi	**1.69**	**3.47**	**unknown**

Precipitation map



Conclusion

- The application of the model of Hirose (2019) to estimate precipitation in Vietnam using Himawari-8 data and auxiliary data obtained reasonable results.
- The use of ERA-5 weather datasets may improve the results on classification but rain rate estimation. The final rainfall map may be improved if using ERA-5 as input of the model.
- There are quite large differences between training and testing for RF1, RF2, RF3, RF4 model performances, which indicates the model over-fiting maybe resulted from large study area and small amount of datasets.
- In future, the model qualities are focused to be enhanced. The other machine learning and deep learning techniques will be investigated.

Thank you for your attention