Supplements for RC1

Supplement 1 Revised Fig. 9 (now Fig.10 in the revised manuscript)

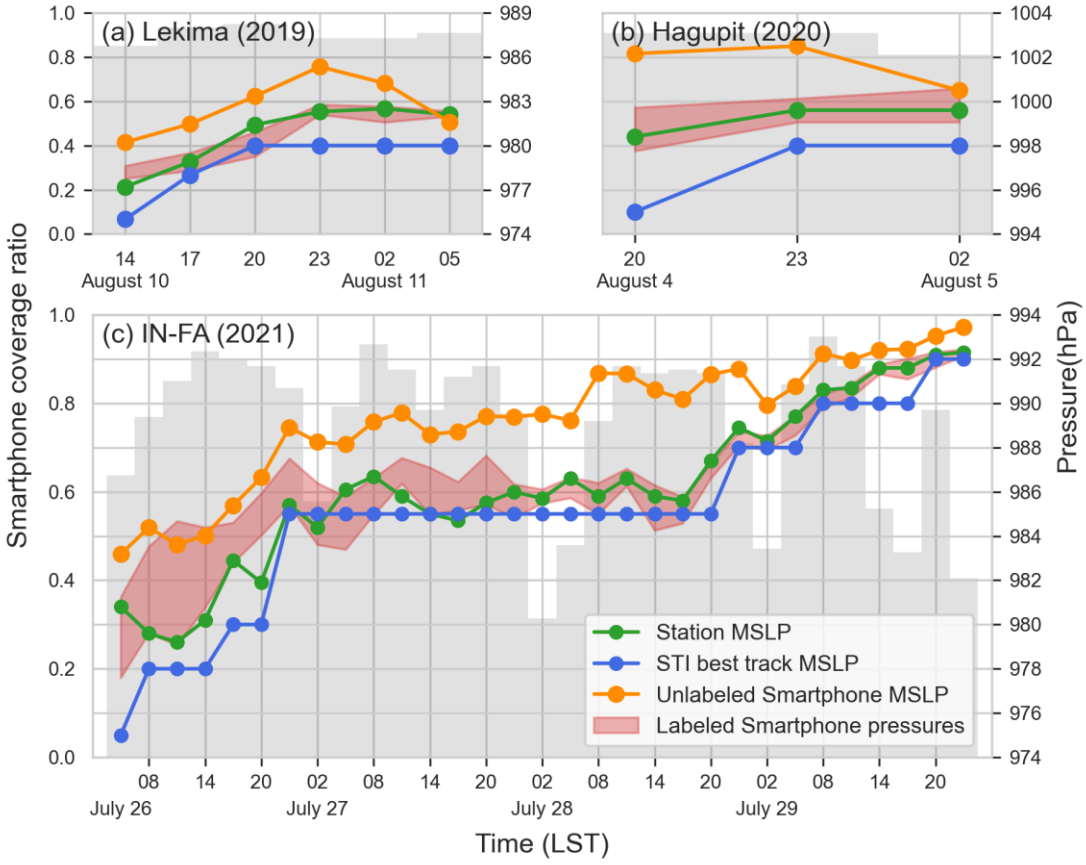

Figure 10 Variation of the MSLPand smartphone coverage ratio during (a) TC Lekima from 14:00 LST on August 10 to 05:00 LST on August 11, 2019, (b) TC Hagupit from 20:00 LST on August 4 to 02:00 on August 5, 2020, and (c) TC IN-FA from 05:00 LST on July 27 to 23:00 LST on July 29, 2021. Green, blue and orange dots represent the MSLP from weather stations, STI best track and unlabeled smartphones, with a temporal resolution of 3, 3 and 6 hours respectively. Red shaded areas represent the lowest 10% labeled smartphone pressure. Gray bars represent smartphone coverage ratio. All the statistics were done in the area of 1.2º x 1.2º surrounding the TC center.

Supplement 2 Revised Tab. 1

Table 1 Descriptive features of the two machine learning models

Unlabeled data	Labeled data
Longitude	Longitude
Latitude	Latitude
Month	Month
Date	Date
Moment	Time
Land-use type	Day of the Week
Gridded pressure	Smartphone pressure
Observations number	
Pressure standard deviation	

Supplement 3 Revised Tab. 2

Table 2 Hyperparameter settings of the two machine learning models

	Unlabeled data	Labeled data
max_depth	9999	9999
max_samples	0.7	0.7
min_samples_leaf	1	1
max_features	log(M+1)	M
n_estimators	100	30

All parameters are from the function “RandomForestRegressor” of the Scikit-learn machine learning library in Python (Pedregosa et al. 2011).

max_depth: The maximum depth of the tree (also known as “the base estimator”).

max_samples: The proportion of samples to draw from the training set to train each tree when bootstraping.

min_samples_leaf: The minimum number of samples required to be at a leaf node.

max_features: The number of features to consider when looking for the best split. M represents the number of features used by the model.

n_estimators: The number of trees in the forest.

Reference:

Pedregosa, F., et al. (2011). "Scikit-learn: Machine Learning in Python. " Journal of Machine Learning Research 12: 2825–2830.

Supplement 4 Revised Fig. 10 (now Fig.11 in the revised manuscript)

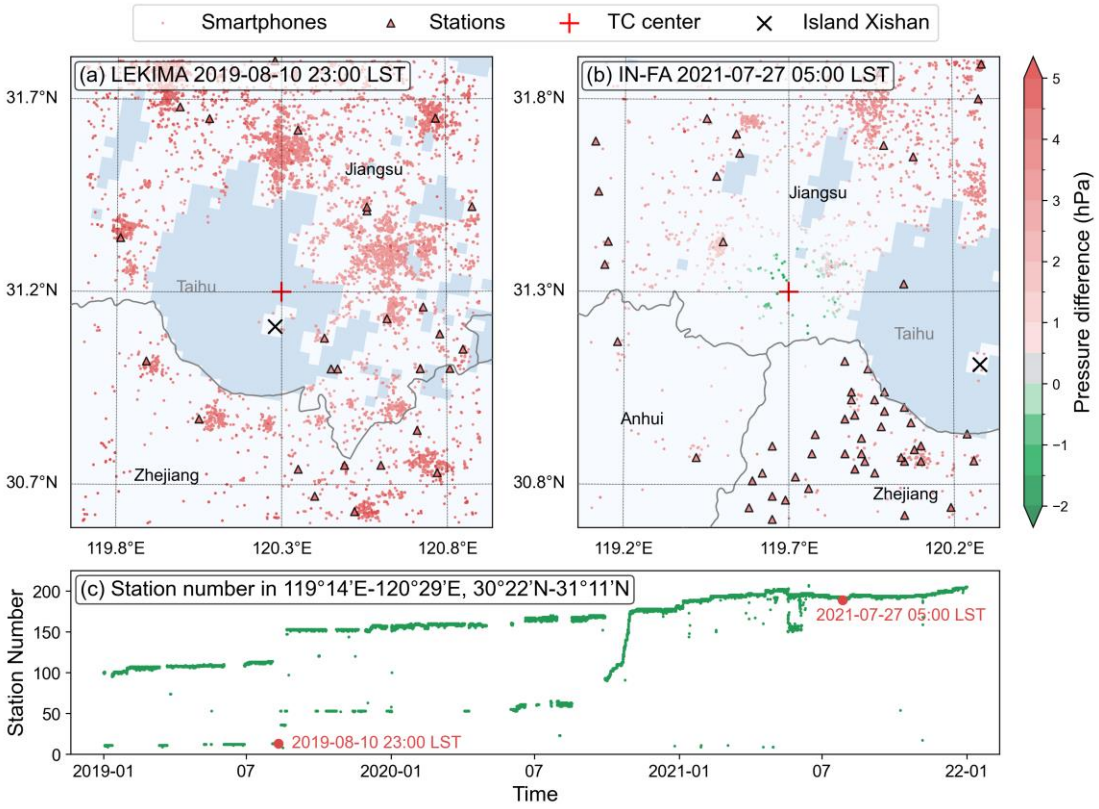

Figure 11 Distributions of weather station and smartphone observations from two examples during (a) TC IN-FA and (b) TC Lekima, in the area of 1.2º×1.2º surrounding the TC center. The coloring represents the difference between the pressure observations and the STI best track MSLP. (c) Changes in the number of weather stations providing pressure observations from 2019 to 2021, in 119°14'E-120°29'E, 30°22'N-31°11'N (the geographical scope of Huzhou, Zhejiang Province).