

Indoors WIFI Location

Agenda

1 Goal

2 Analysis

Analysis pipeline

Data exploration and data cleaning

Insights from the data

Modeling and error location

3 Conclusions and suggestions

Goal

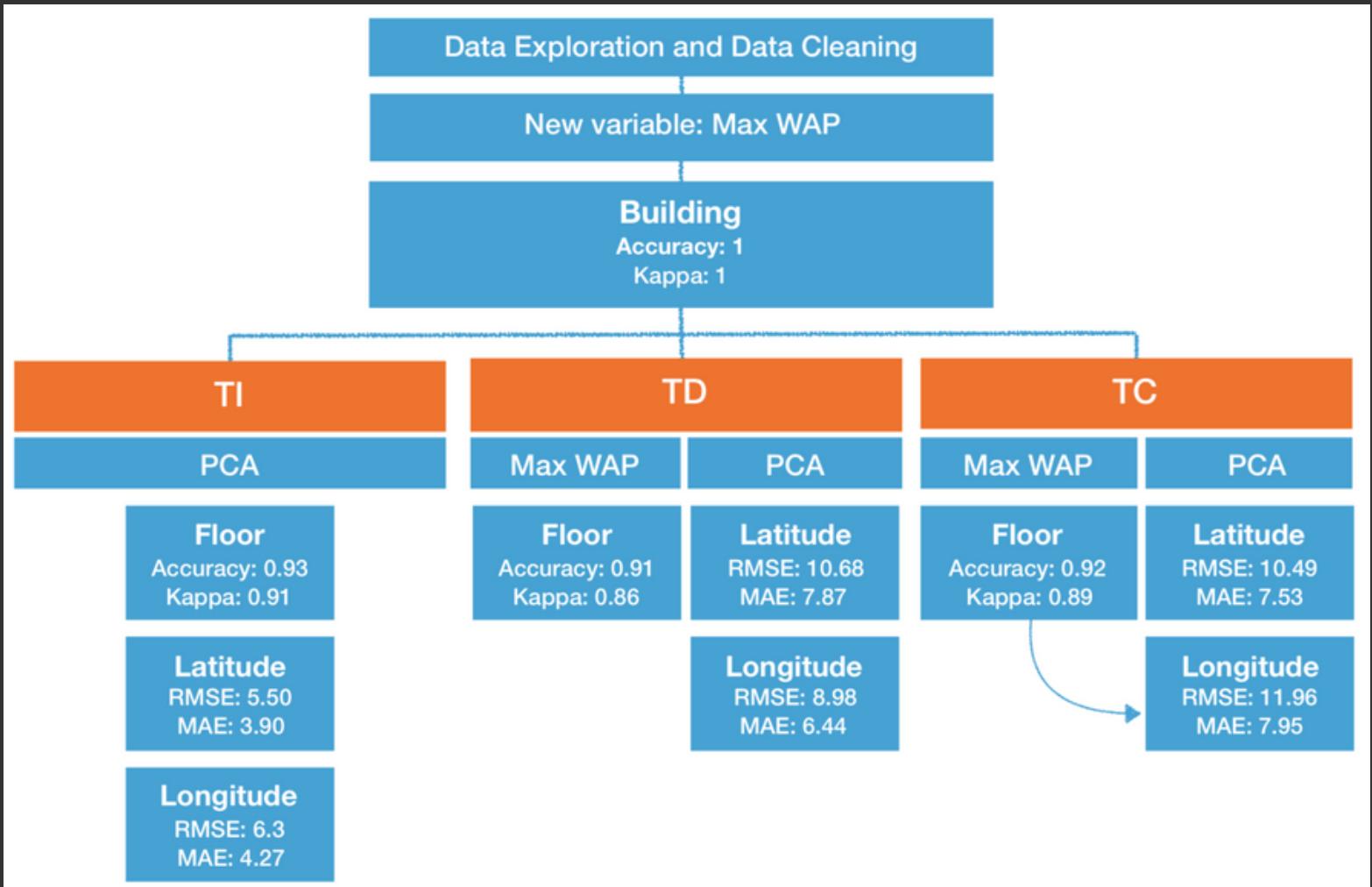
To develop a navigation system for interior spaces.

Use WIFI fingerprinting to determine a person's location in indoor spaces.



Analysis

Analysis Pipeline



Dataset



UJIIndoorLoc Dataset

- 1. Training: 19937 reference records
- 2. Validation: 1111 test records



RSSI levels of WAPs detected

520 WAPs



Position coordinates

Longitude and Latitude



Building

- 1. TI-ESTCE (TI)
- 2. TD-ESTCE (TD)
- 3. TC-ESTCE (TC)



Floor

- 4 floors TI and TD
- 5 floors TC



User ID

18 different users



Phone ID

- 14 Only in training
- 9 Only in validation
- 2 in both



Time stamp

- Training: May-June 2013
- Validation: September-October 2013

Data Cleaning and Feature engineering

1 Remove unique WAPs

- 153 in the Training not in the Validation
- 55 in the Validation not in the Training
- 39 in the Training low signal (< -80)

2 Remove duplicates

637 from the training

3 Remove fingerprints not detected any WAP

76 from the training

4 Remove phone 19

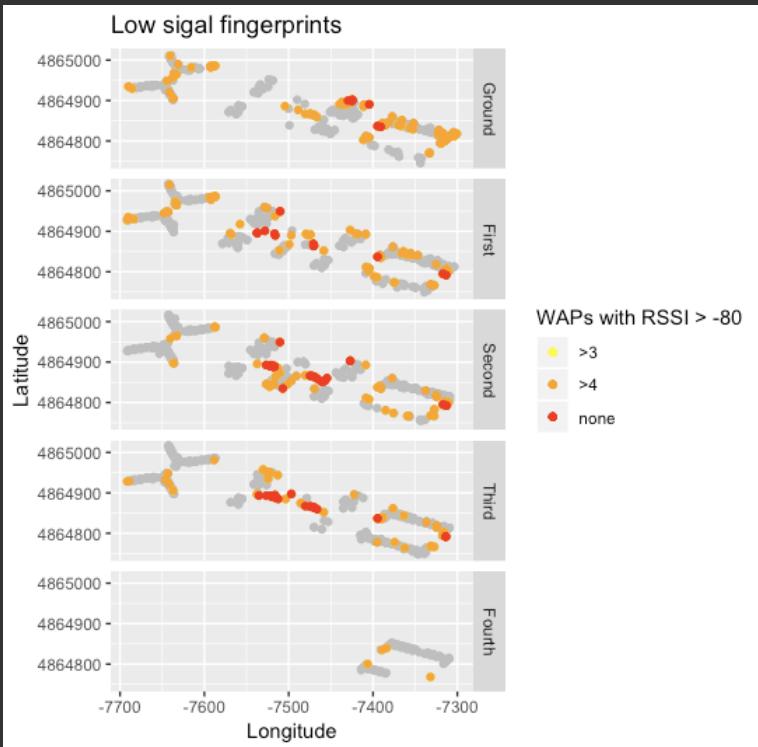
40% fingerprints values > -30
977 from the training

5 Group fingerprints from the same user/phone at the same location

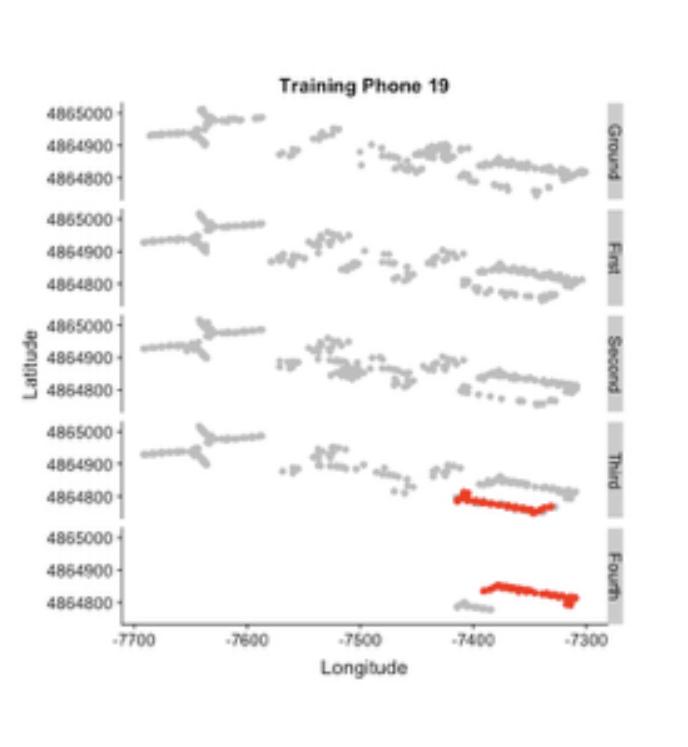
Median RSSI value of each WAP

6 New attribute: WAP with maximum RSSI value each fingerprint

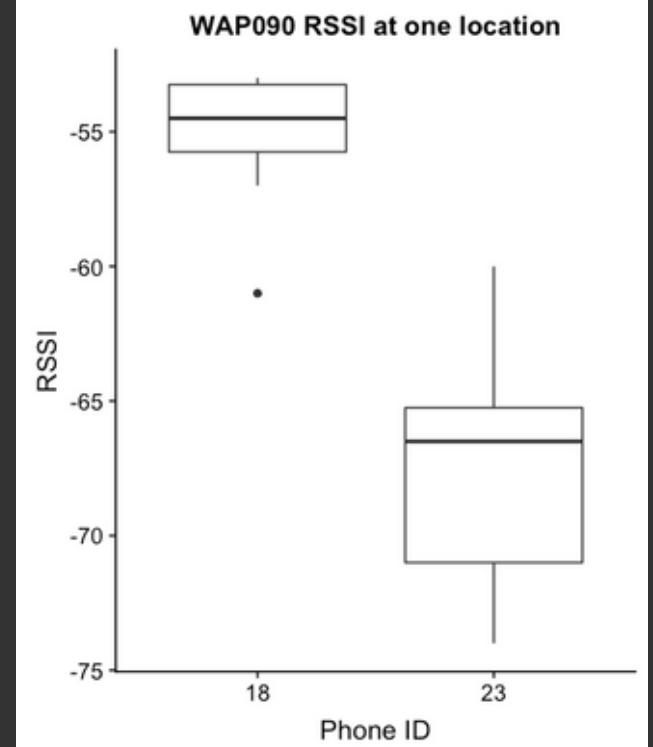
Insights from the data



Low RSSI value fingerprints

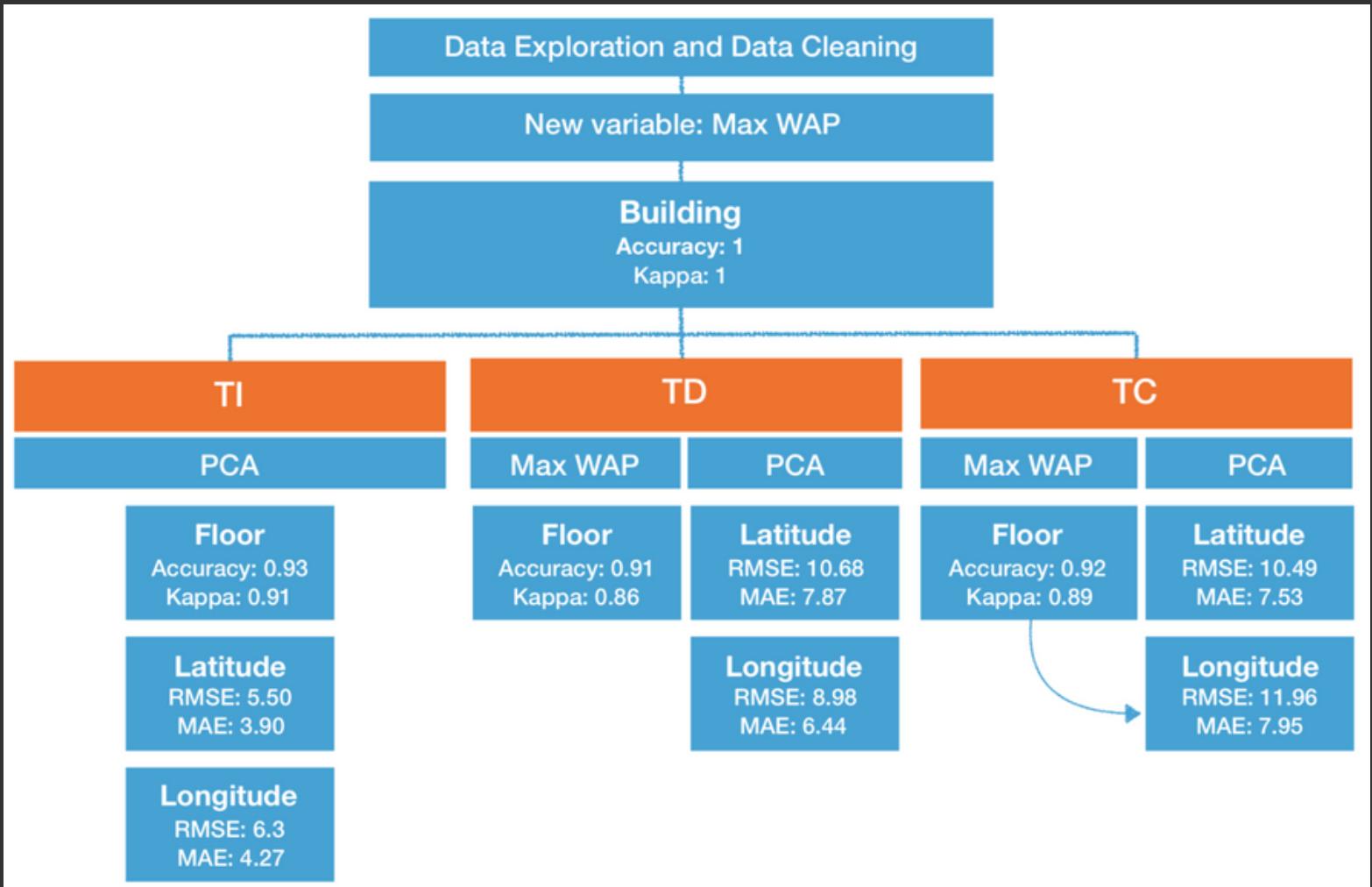


Regions not covered in the training



Specific regions each phone in training (also dates)

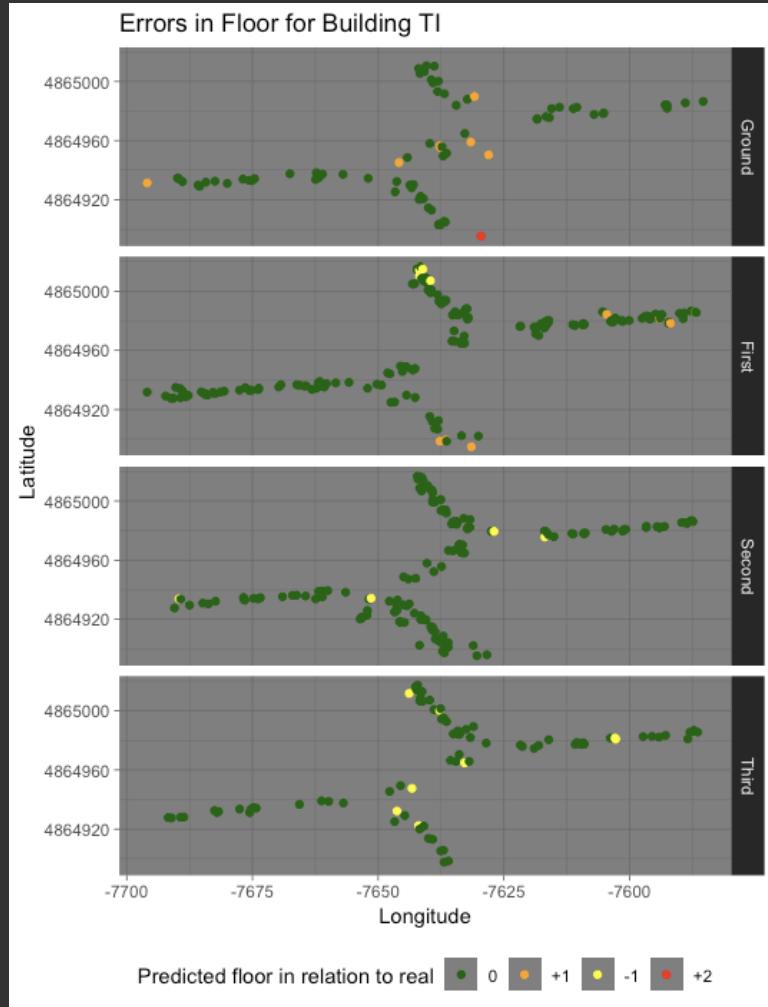
Analysis Pipeline



TI - ESTCE

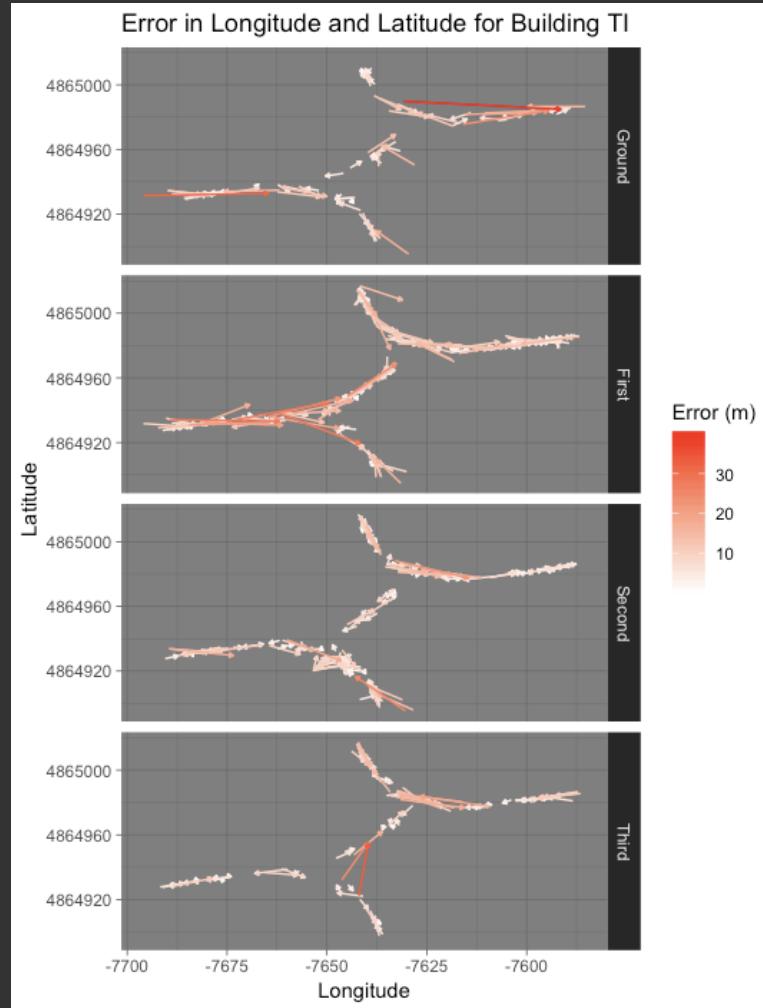


Predicting Floor at TI



- Random Forest
 - 18 Principal components
 - Accuracy: 93.7%
 - Kappa: 91.1%
- Errors due to phone ID
 - Phone 20 (10/33)
 - Phone 0 (9/33)

Predicting Longitude and Latitude at TI



- Longitude

KNN, PCA

RMSE: 6.33m

R-squared: 0.94

MAE: 4.27m

- Latitude

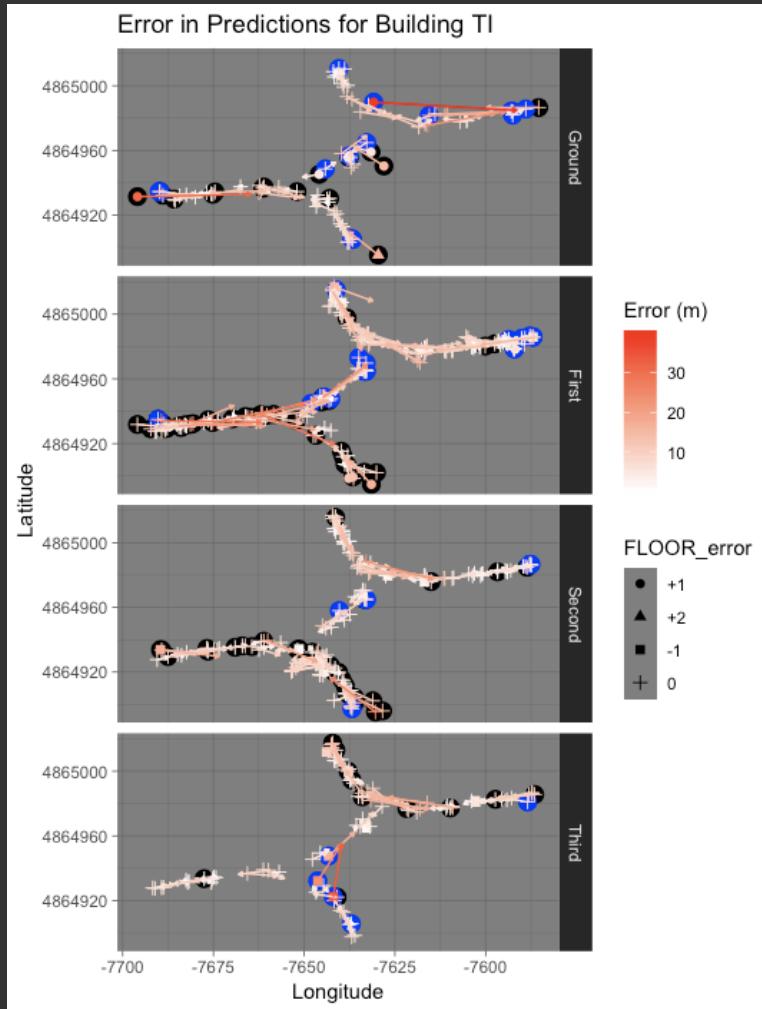
KNN, PCA

RMSE: 5.50m

R-squared: 0.97

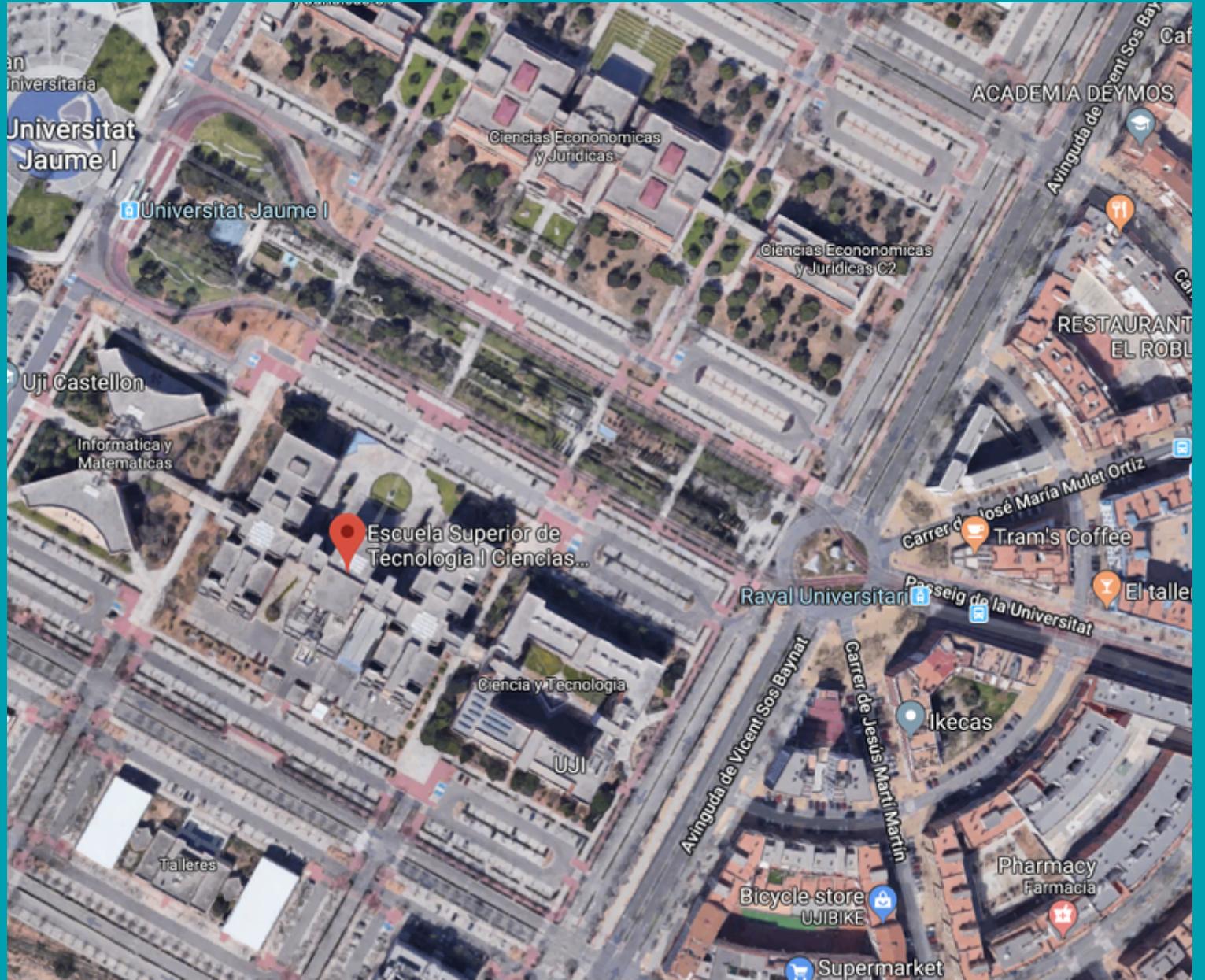
MAE: 3.90m

Predicting the position in TI

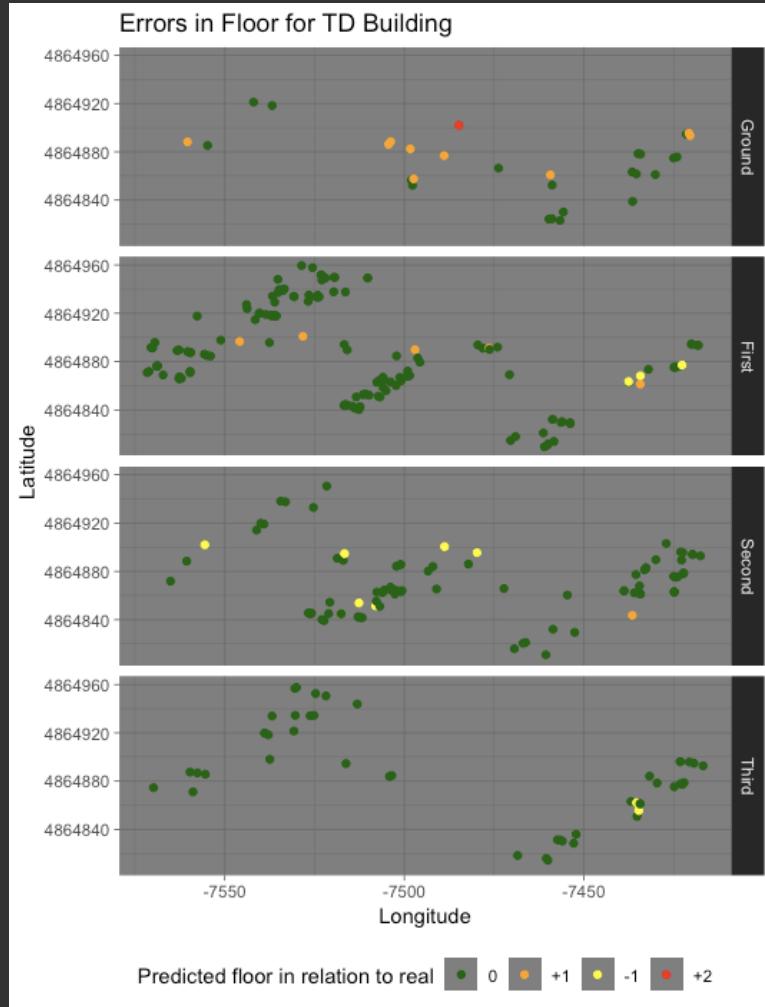


- Errors in floor also for longitude and latitude
 1. Phone 20 (Black dots)
 2. Less than 4 WAPs with RSSI higher than -80 (Blue dots)

TD - ESTCE

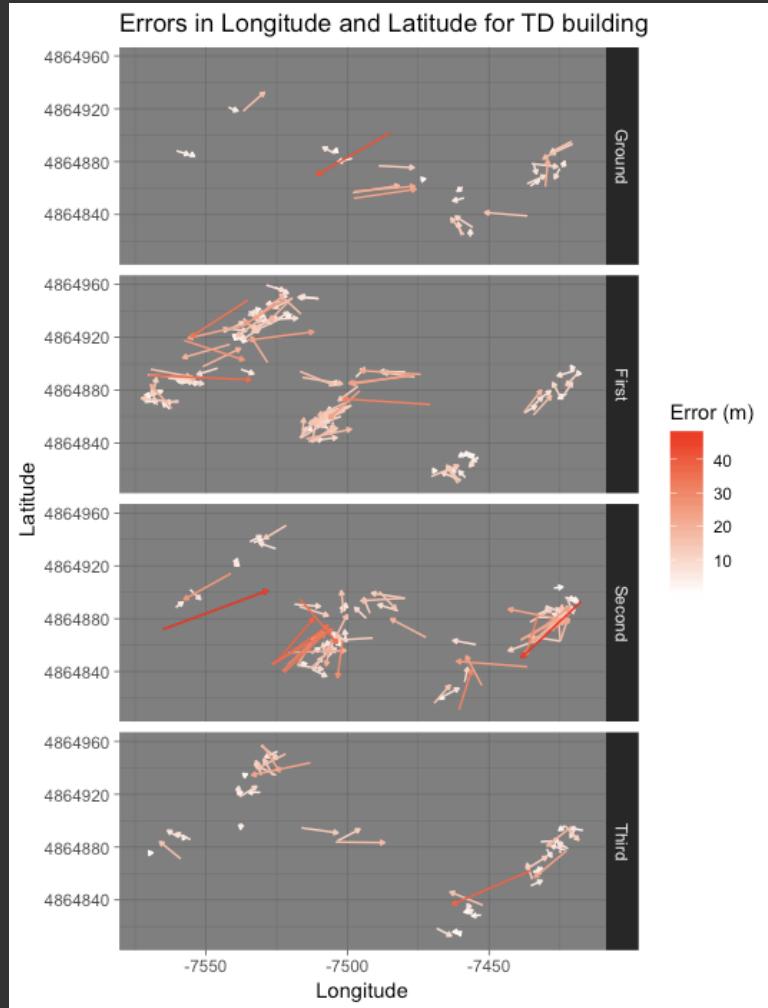


Predicting Floor at TD



- Random Forest
 - WAP with maximum RSSI
 - Accuracy: 90.88%
 - Kappa: 86.19%
- Phone ID
 - Phone 20 (6/28)
 - Phone 5 (6/28)
 - Phone 13 (6/28)
- High RSSI for First Floor

Predicting Longitude and Latitude at TD



- Longitude

RF, PCA

RMSE: 8.96m

R-squared: 0.96

MAE: 6.44m

- Latitude

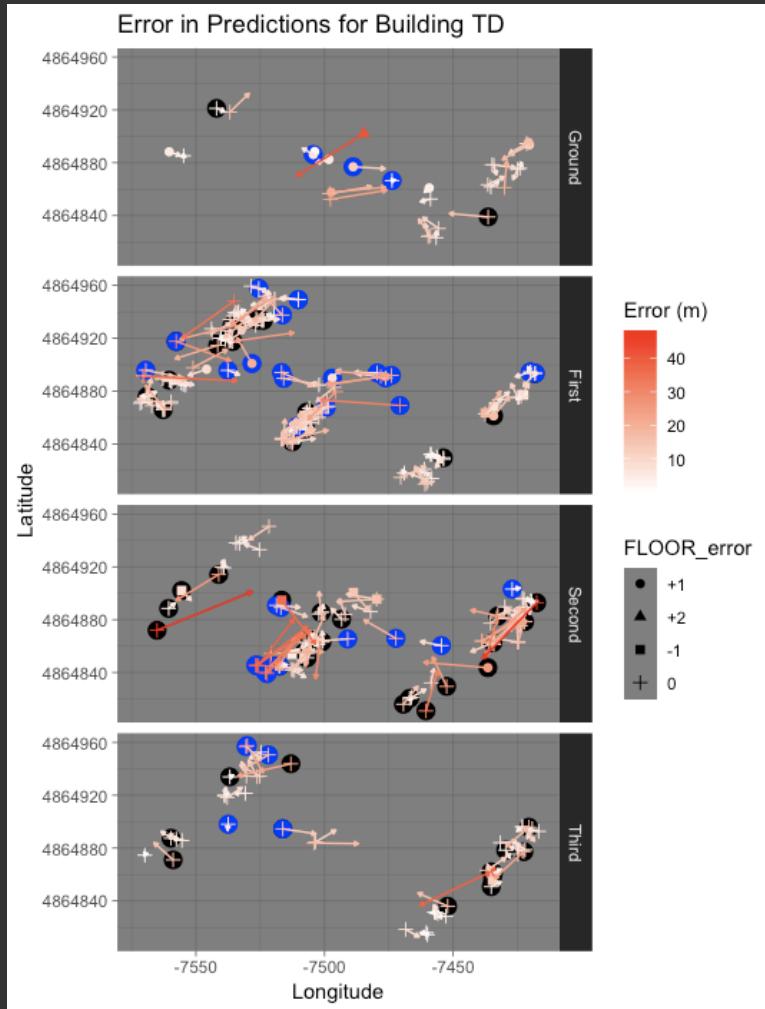
GBM, PCA

RMSE: 10.68m

R-squared: 0.91

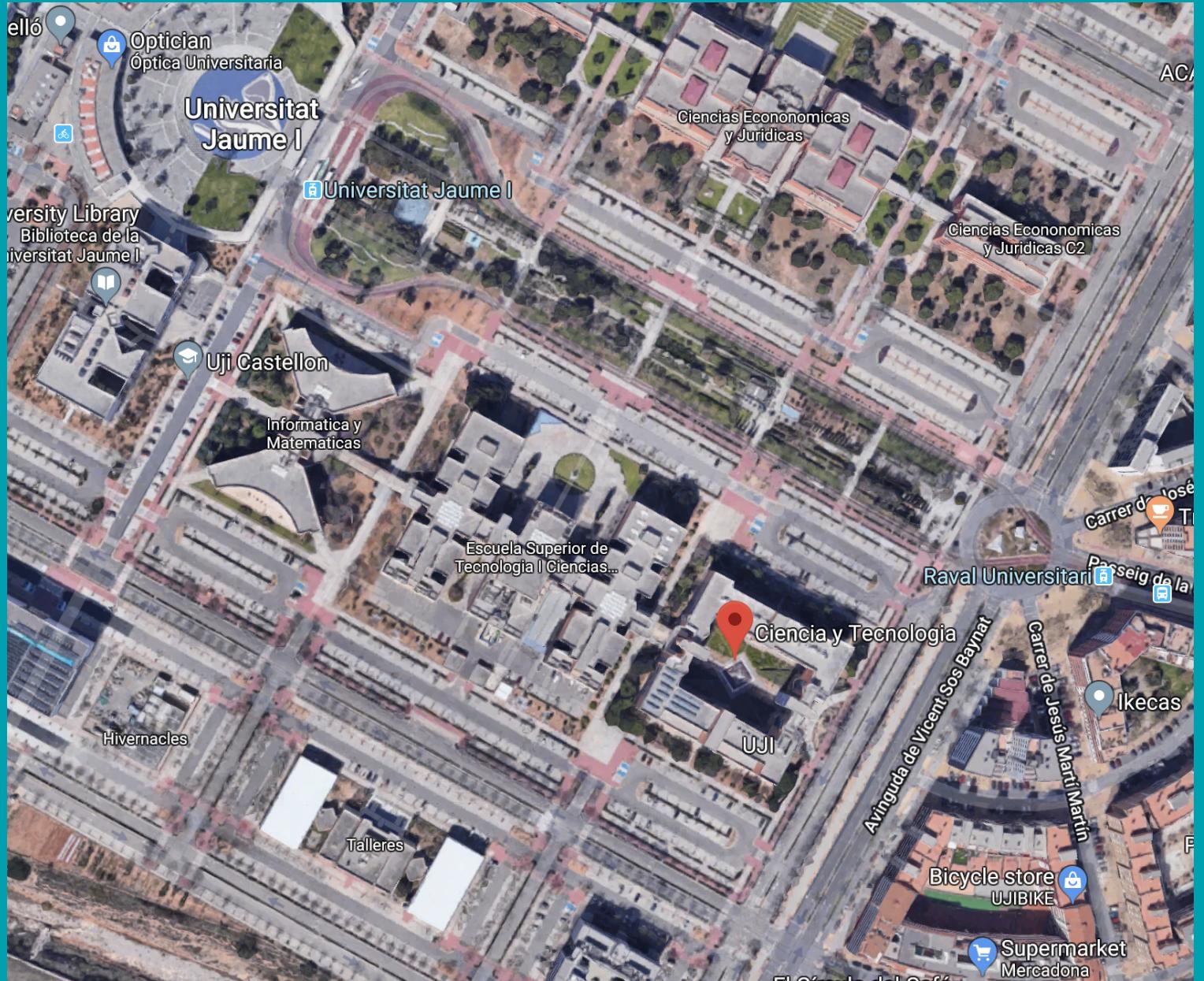
MAE: 7.87m

Predicting the position in TD

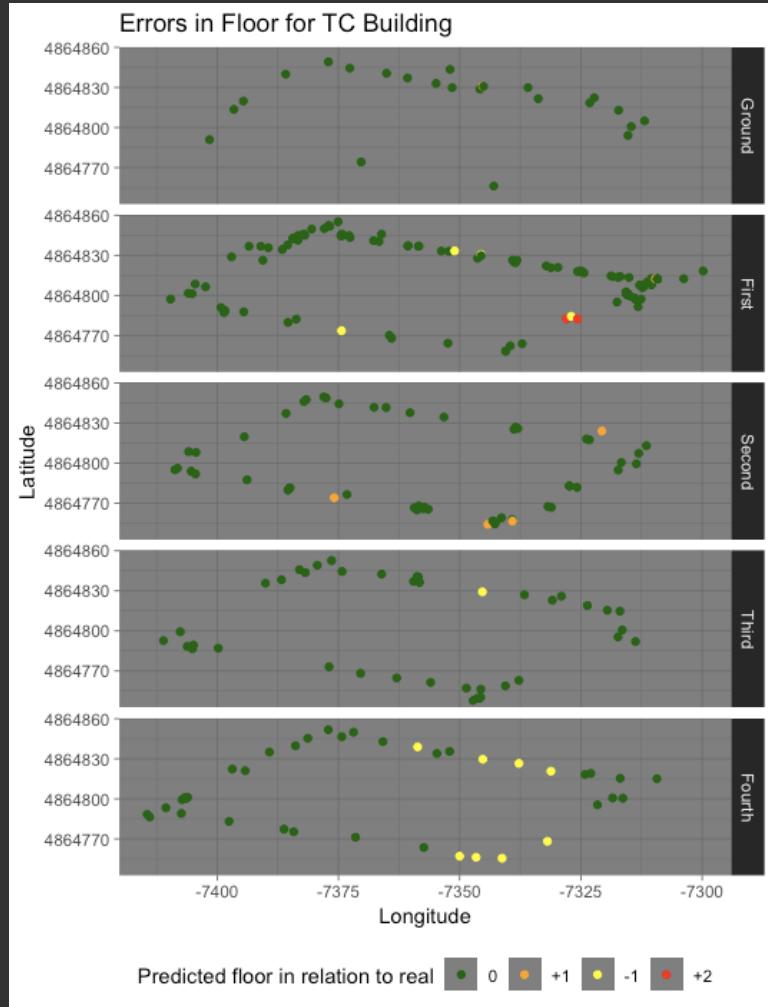


- Errors at middle wing second floor
Less than 4 WAPs with RSSI higher than -80
(Blue dots)
- Large errors
Phone 20 (Black dots)

TC - ESTCE

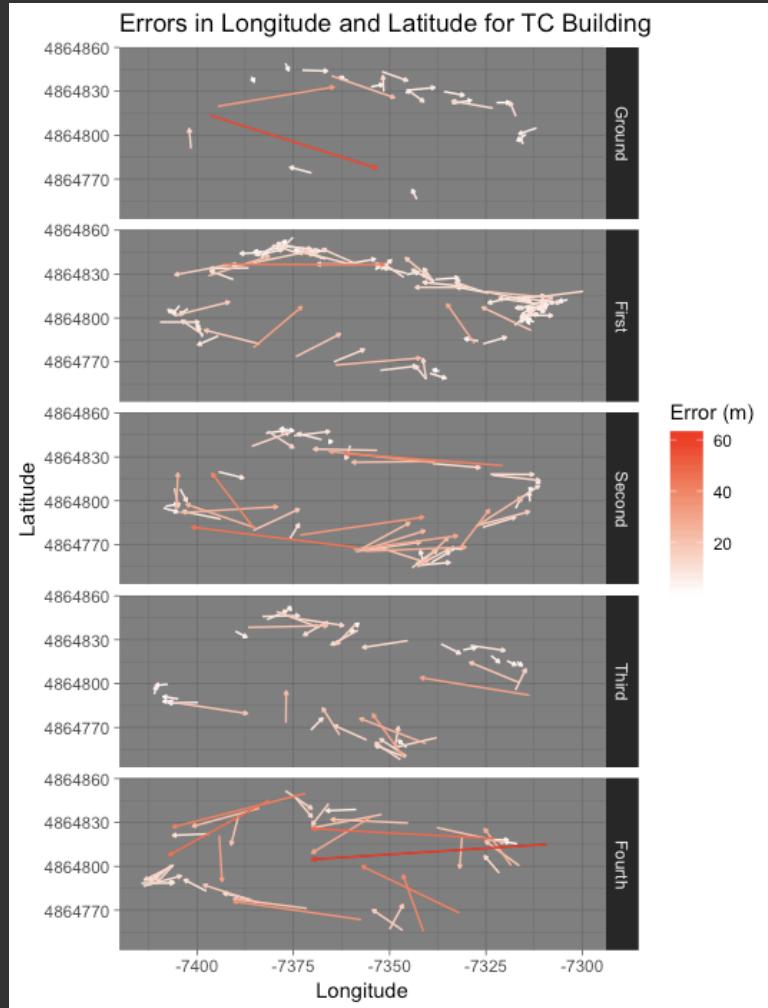


Predicting Floor at TC



- Random Forest
 - WAP with maximum RSSI
 - Accuracy: 91.98%
 - Kappa: 89.29%
- Phone ID
 - Phone 20 (8/21)
 - Phone 13 (6/21)
 - Phone 14 (4/21)
- Not covered in the training

Predicting Longitude and Latitude at TC



- Longitude

KNN, PCA+Floor

RMSE: 11.96m

R-squared: 0.86

MAE: 7.95m

- Latitude

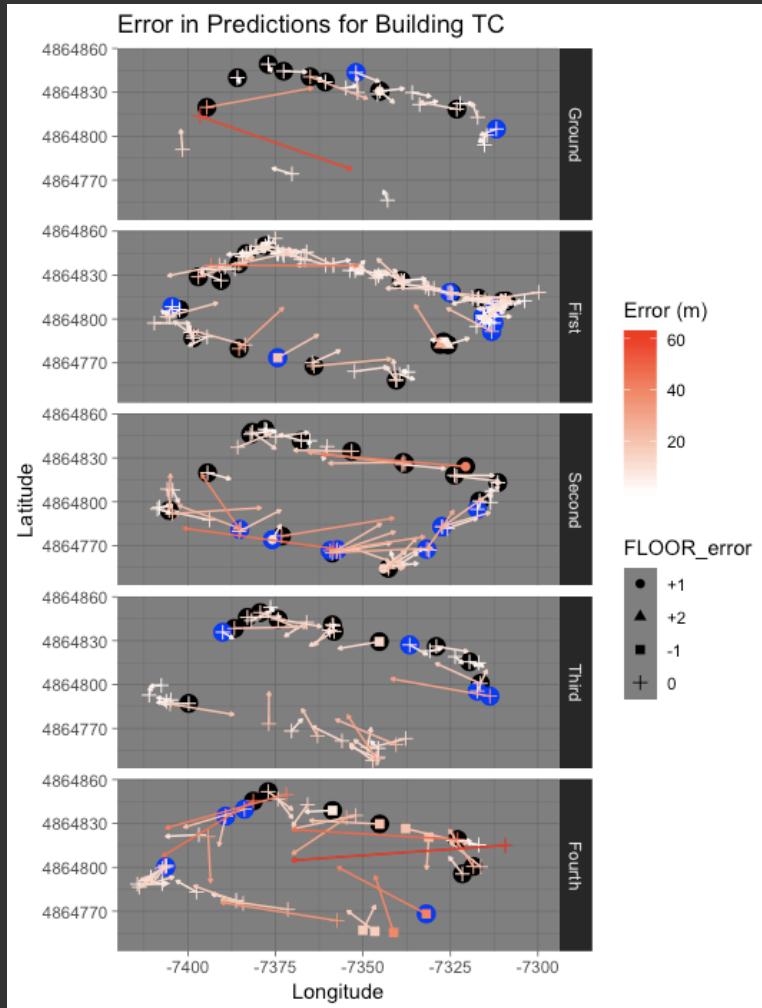
RF, PCA

RMSE: 10.49m

R-squared: 0.88

MAE: 7.53m

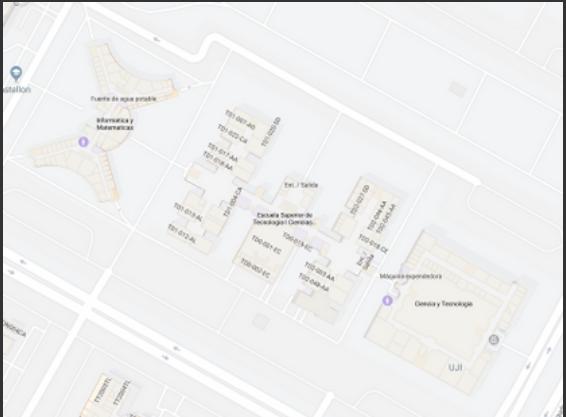
Predicting the position in TC



- Errors at second floor
Less than 4 WAPs with RSSI higher than -80 (Blue dots)
- Errors at fourth floor
 - Phone 20 (Black dots)
 - Not covered in the training

Conclusions and suggestions

Indoors location



Building Accuracy: 1 Kappa: 1

TI-ESTCE

Accuracy: 1

Kappa: 1

Floor: Accuracy=93%,

kappa=86%

Longitude: MAE 4.3m

Latitude: MAE 3.9m



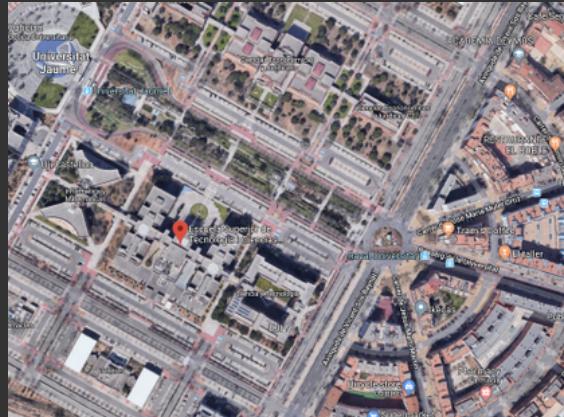
TD-ESTCE

Floor: Accuracy=91%

kappa=86%

Longitude: MAE 6.4m

Latitude: MAE 7.9m



TC-ESTCE

Floor: Accuracy=92%

kappa=89%

Longitude: MAE 8.0m

Latitude: MAE 7.5m

To some problems



Phone ID



Specific regions for phone and day in the training



Not covered in the training



Low RSSI values fingerprints



Normalize RSSI values to phone ID



New training set with phones (and dates) more distributed



Mix training and validation set



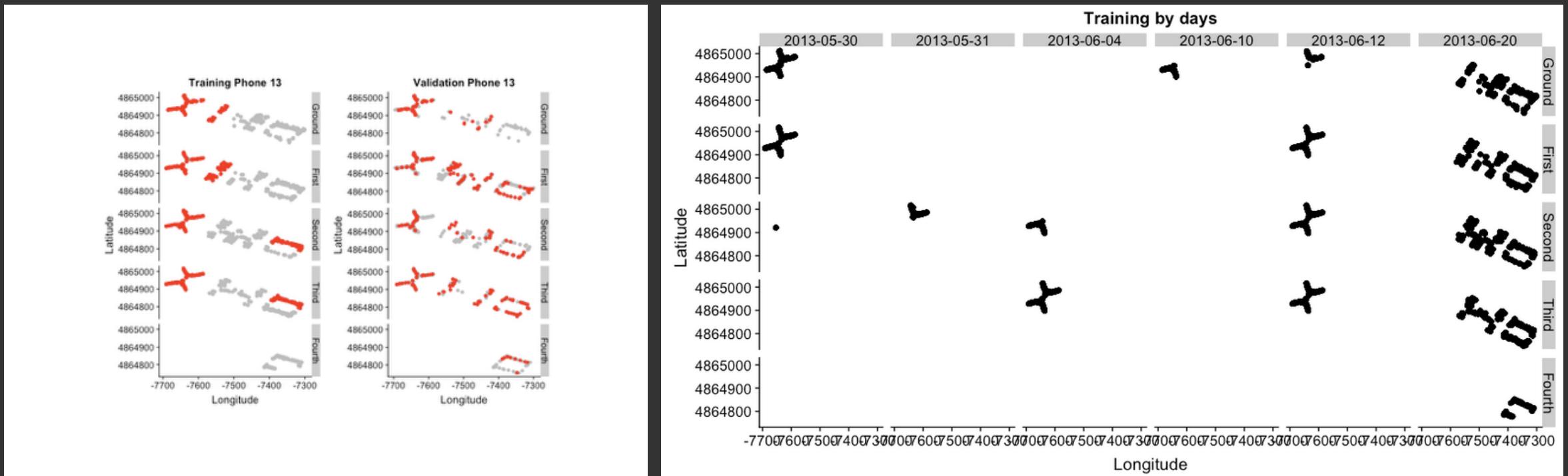
Add new WAPs

... some solutions



Indoors WIFI Location

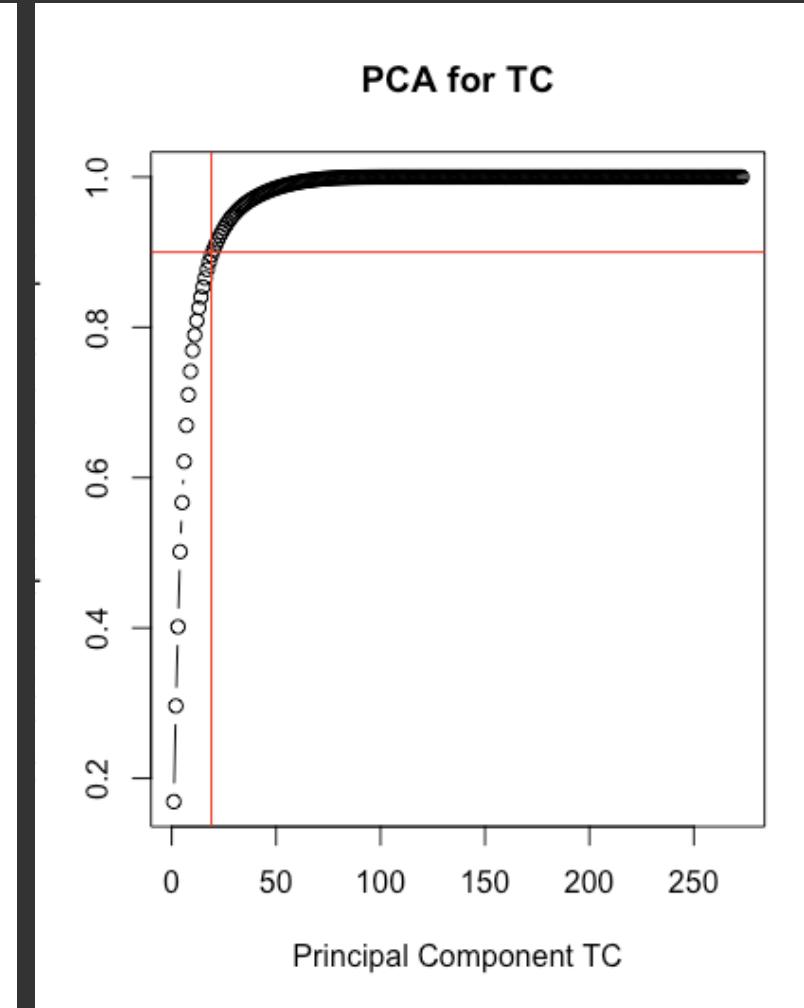
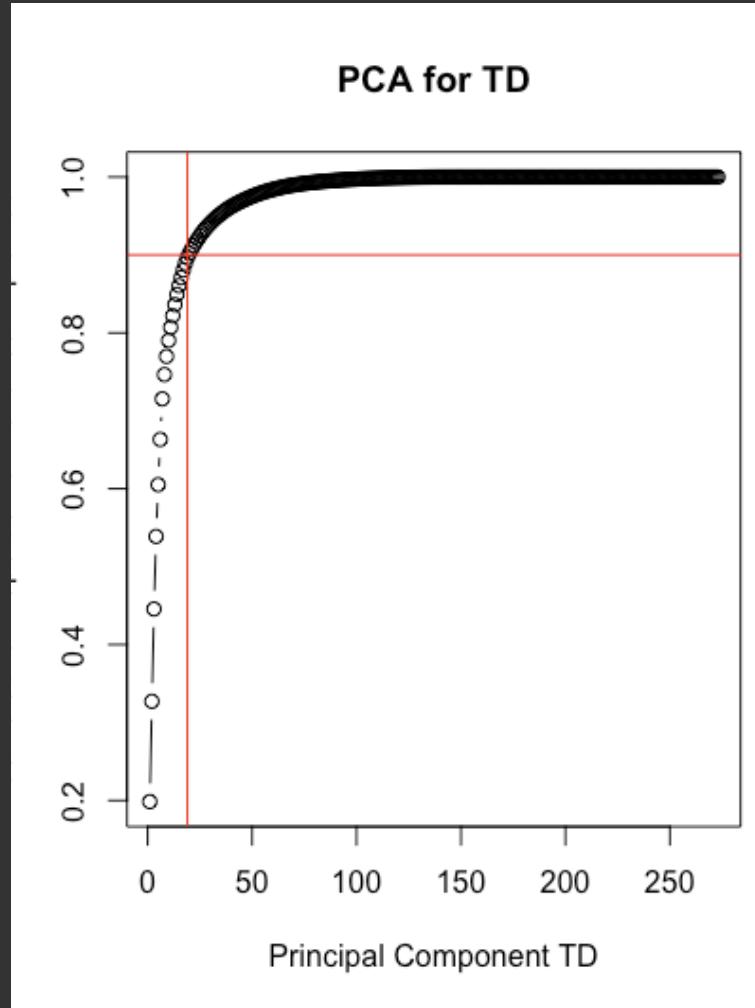
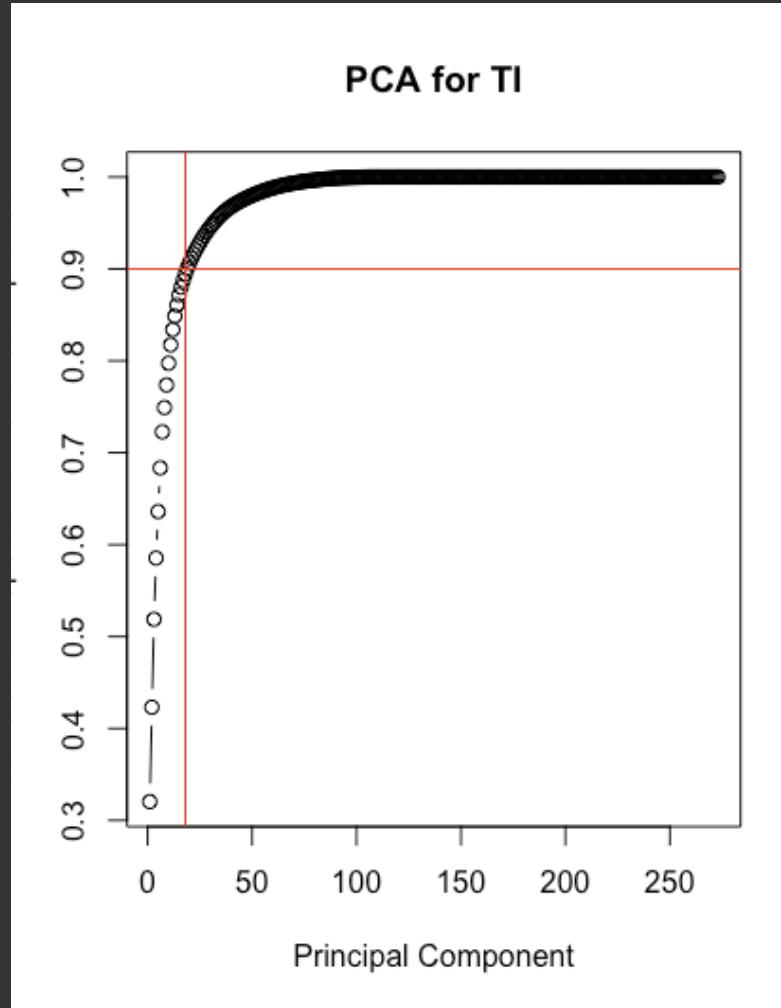
In training each phone has specific regions whereas in the validation are sparse



of WAPs / floor in each building

	Ground	First	Second	Third	Fourth
TI	15	28	32	22	-
TD	44	23	21	18	-
TC	12	19	14	17	8

Principal Component Analysis



Models tested to predict the floor

Building	Model	Accuracy	Kappa	Notes	
TI	floor_rf_ti_max	0.8645	0.8092	4 at 2nd predicted at ground	
TI	floor_rf_ti_all	0.9542	0.9352	3 in the ground predicted at the 3rd, 1 second at the ground	★
TI	floor_rf_ti_pca	0.937	0.911	1 in the ground predicted at 2nd.	★
TD	floor_rf_td_max	0.9088	0.8619	1 in the ground predicted at 2nd.	★
TD	floor_rf_td_all	0.772	0.6759	Lots errors from the first	
TD	floor_rf_td_pca	0.7655	0.6644	Lots errors from the first	
TC	floor_rf_tc_max	0.9198	0.8929	2 from first predicted third	★
TC	floor_gbm_tc_max	0.76	0.70	8-17 from 2nd, 3rd, 4th predicted at ground	
TC	floor_rf_tc_all	0.8359	0.7778	26+1 errors from the 4th.	
TC	floor_rf_tc_pca	0.8779	0.8335	All errors from the 4th floor (13+1+1)	

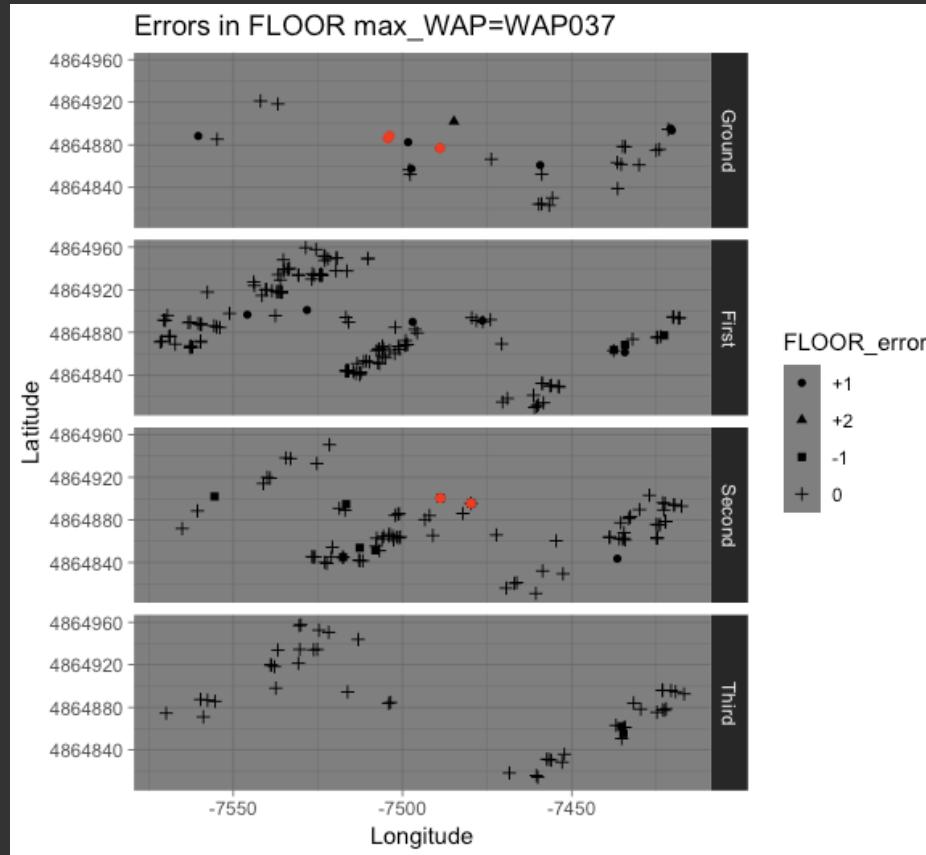
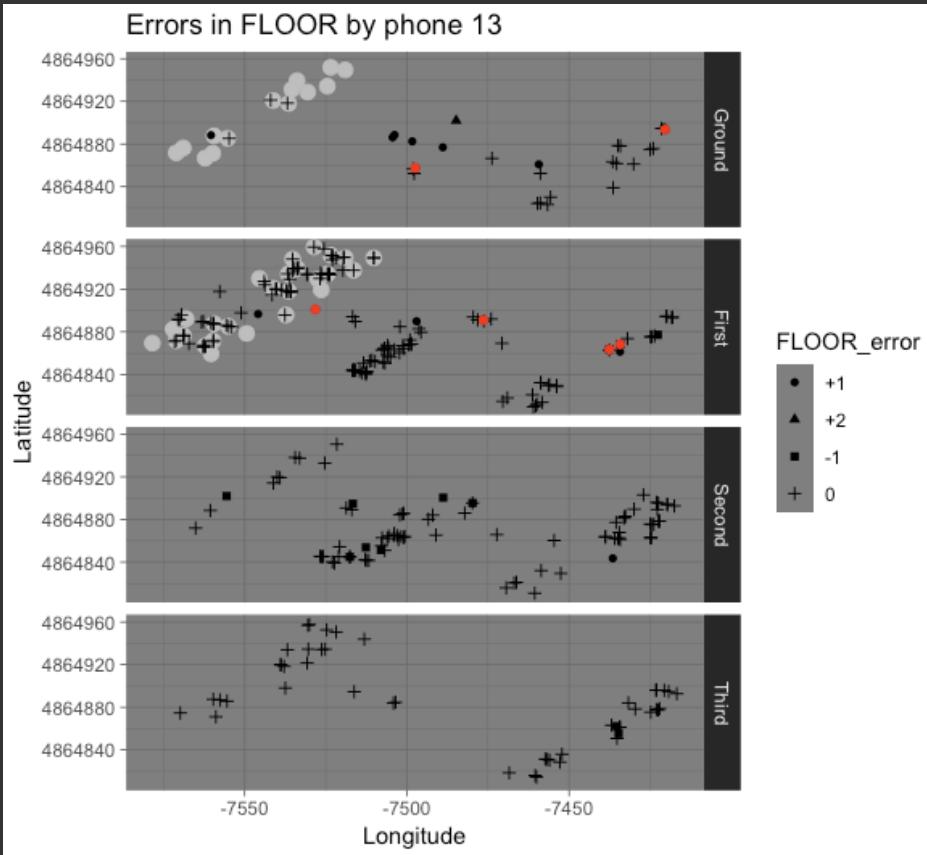
Confusion matrix floor in TI

	Ground	First	Second	Third
Ground	69	6	0	0
First	8	191	4	0
Second	1	6	154	8
Third	0	0	0	77

Confusion matrix for floor in TD

	Ground	First	Second	Third
Ground	20	3	0	0
First	9	135	7	0
Second	1	5	79	2
Third	0	0	1	45

Errors floor TD

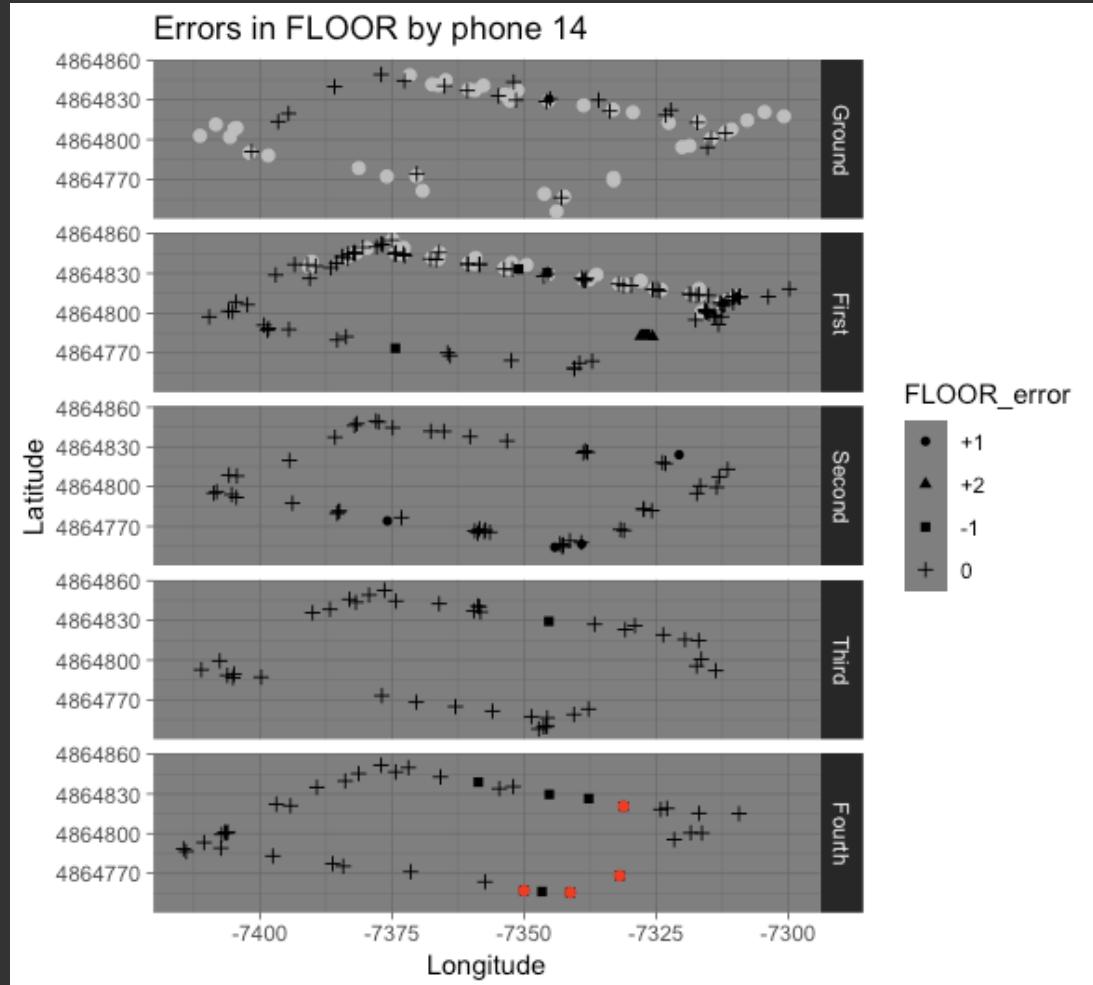
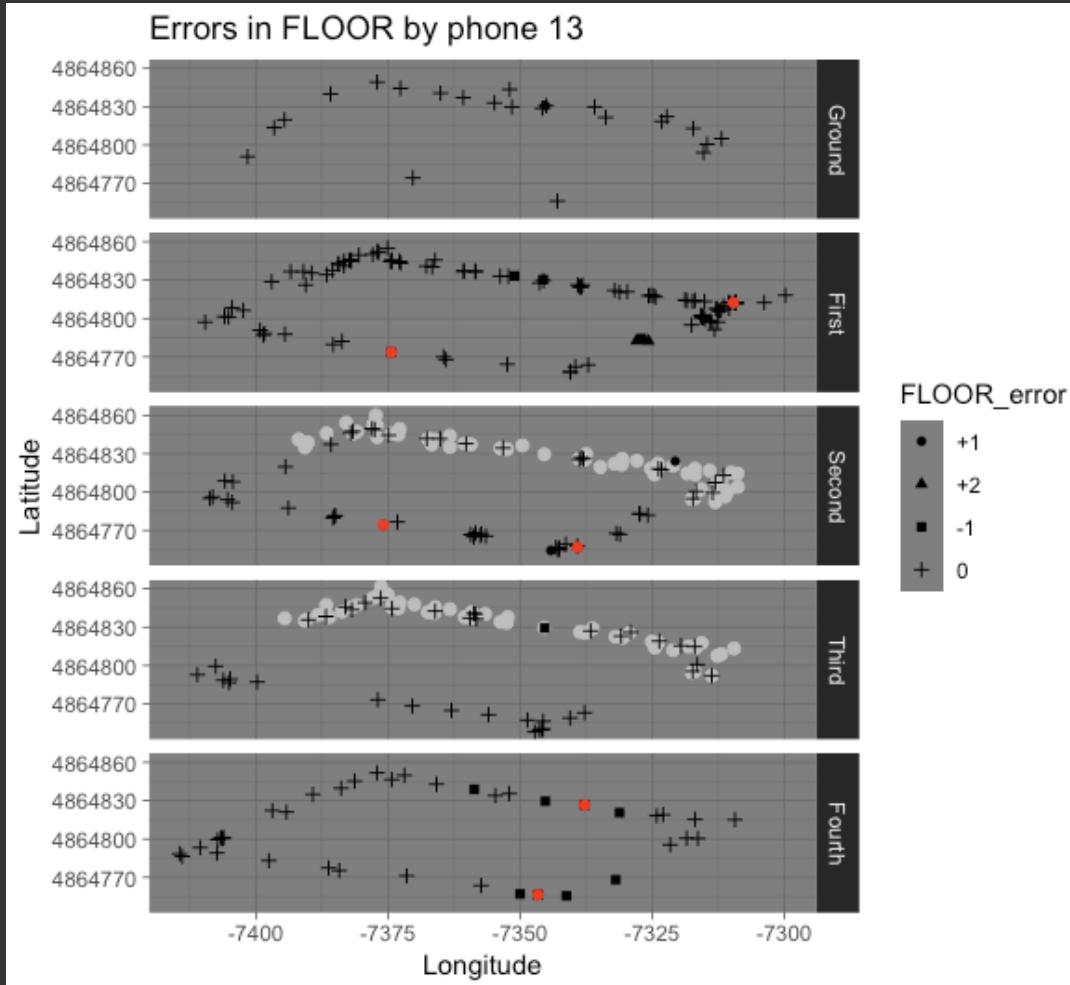


WAP037 located at the First floor

Confusion matrix floor for TC

	Ground	First	Second	Third	Fourth
Ground	23	4	0	0	0
First	1	99	0	0	0
Second	0	1	49	1	0
Third	0	2	4	39	8
Fourth	0	0	0	0	31

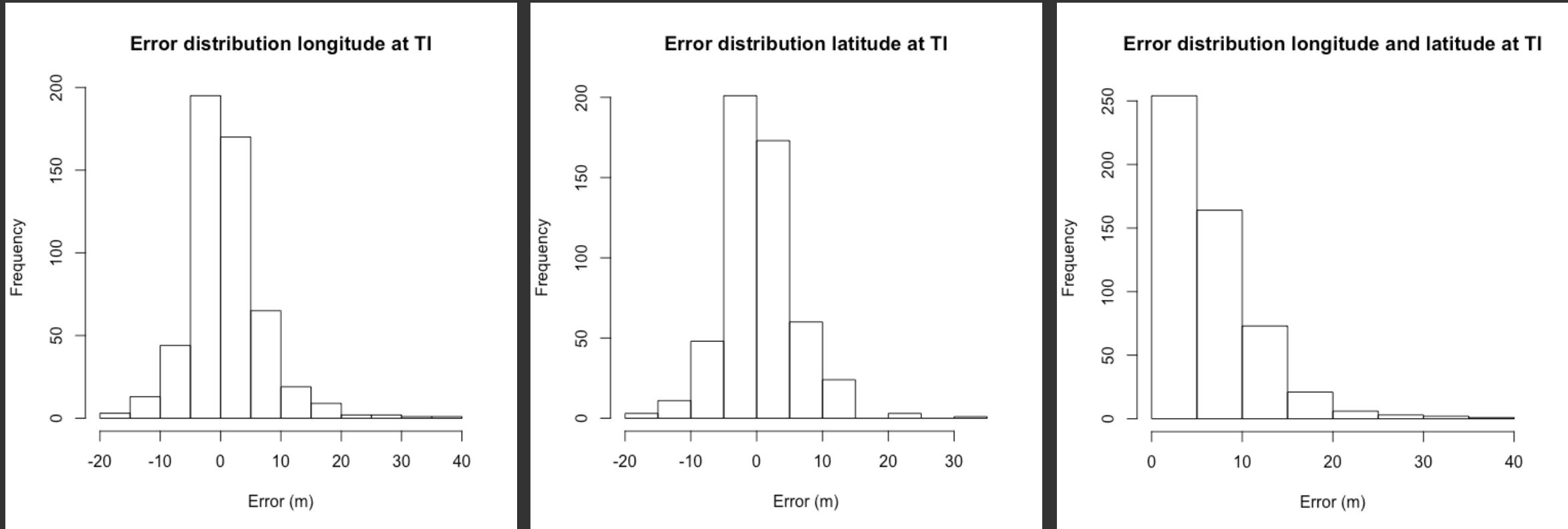
Errors floor TC



Models tested to predict longitude and latitude for TI

	Model	RMSE	squared-R	MAE	Notes	
Longitude	lon_knn_ti_pca	6.33	0.94	4.27	Good overlay, maybe a couple out of the building, Good distribution of errors.	★
	lon_svmr2_ti_pca	8.63	0.90	6.10	Some out, no peaks at the ends, best regarding distribution of errors.	
	lon_rf_ti_pca	7.51	0.92	5.10	Good overlay but some places out	
	lon_gbm_ti_pca	7.62	0.92	5.11	Good overlay, some places out	
Latitude	lat_knn_ti_pca	5.50	0.97	3.90	Some out, but others overlay well	★
	lat_svmr2_ti_pca	6.99	0.95	5.12	Lower values out of range	
	lat_rf_ti_pca	6.54	0.96	4.32	good overlay in general, but some places out of the building	
	lat_gbm_ti_pca	6.25	0.96	4.30	Good overlay, but few out of place	

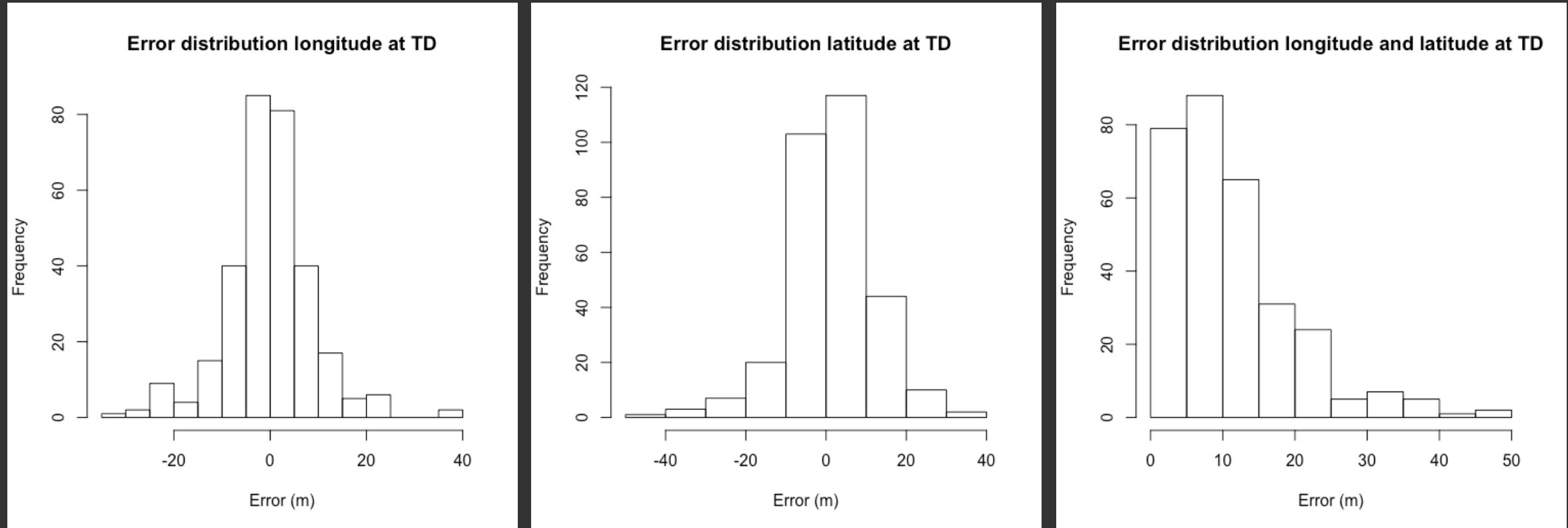
Error distribution longitude and latitude for TI



Models tested to predict longitude and latitude for TD

	Model	RMSE	squared-R	MAE	Notes	
Longitude	lon_knn_td_pca	10.20	0.95	6.93	Good overlay, maybe a couple out of the building	
	lon_svmr2_td_pca	12.03	0.93	9.18	Some out	
	lon_rf_td_pca	8.98	0.96	6.44	Good overlay, some places out	
	lon_gbm_td_pca	9.20	0.96	6.57	Good overlay, some places out	
Latitude	lat_knn2_td_pca	11.12	0.90	7.29	A lot out of the building	
	lat_svmr2_td_pca	10.95	0.91	8.18	Some out of the building	
	lat_rf2_td_pca	10.97	0.91	8.01	Good overlay, some places out of the building	
	lat_gbm2_td_pca	10.68	0.91	7.87	Good overlay, some places out of the building	

Error distribution longitude and latitude for TD



Models tested to predict longitude and latitude for TC

	Model	RMSE	squared-R	MAE	Notes	
Longitude	lon_knn2_tc_pca	13.68	0.81	8.84	Lot's error in the 4th floor	
	lon_knn2_tc_pca For 0-3 floors	13.10	0.82	8.44	Still some in the middle	
	lon_knn3_tc_pca	11.96	0.86	7.95	PCA + Floor. Best 4th floor	★
	lon_svmr2_tc_pca	13.75	0.81	10.39	Errors at the 2nd and 4th floor, wide error distribution	
	lon_rf_tc_pca	13.19	0.83	9.33	Errors in the middle. Not bad 4th floor	
	lon_rf_tc_pca for 4th floor	14.21	0.85	10.93	Right side all in the middle	
	lon_gbm_tc_pca	13.23	0.82	10.01	Errors 2nd and 1st floor, wide error distribution	
Latitude	lat_knn2_tc_pca	13.10	0.81	7.77	Errors at 4th and 2nd	
	lat_knn3_tc_pca	12.71	0.82	7.75	+floor, still bad 4th and 2nd	
	lat_svm2_tc_pca	12.67	0.81	9.07	Errors 2nd and 4th	
	lat_svm3_tc_pca	12.59	0.81	9.17	+floor, still bad	
	lat_rf_tc_pca	10.49	0.88	7.53	Still bad 4th but much better than the others	★
	lat_gbm_tc_pca	10.66	0.87	7.76	Similar to RF	

Error distribution longitude and latitude for TC

