

Project 4

Group C

Chu-Chun Chen, Xinchang Li, Sarah Szymborski

Importing and handling the data

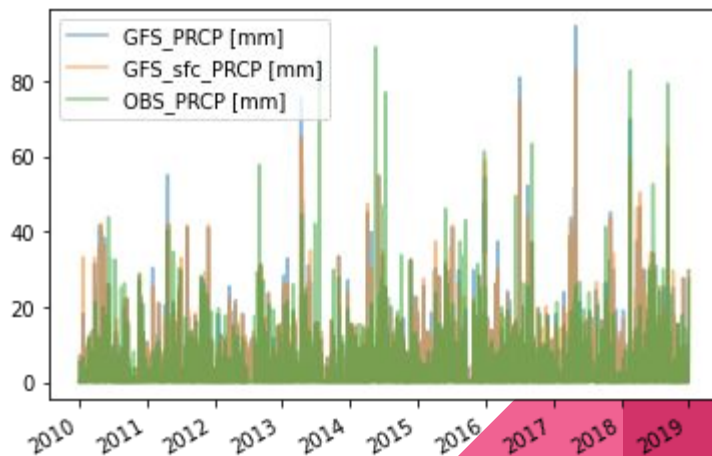
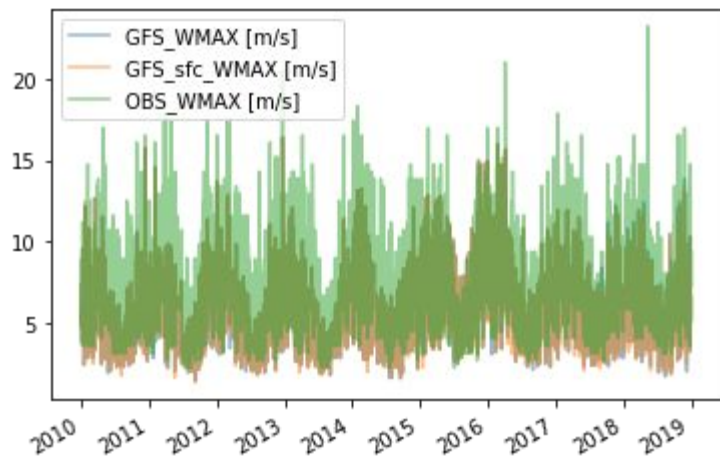
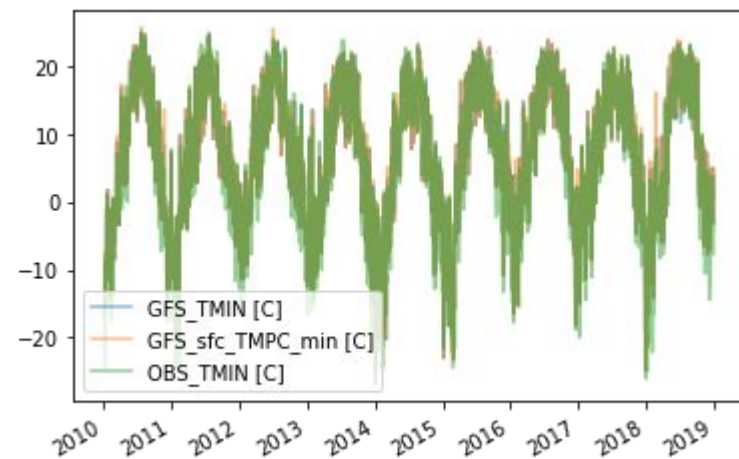
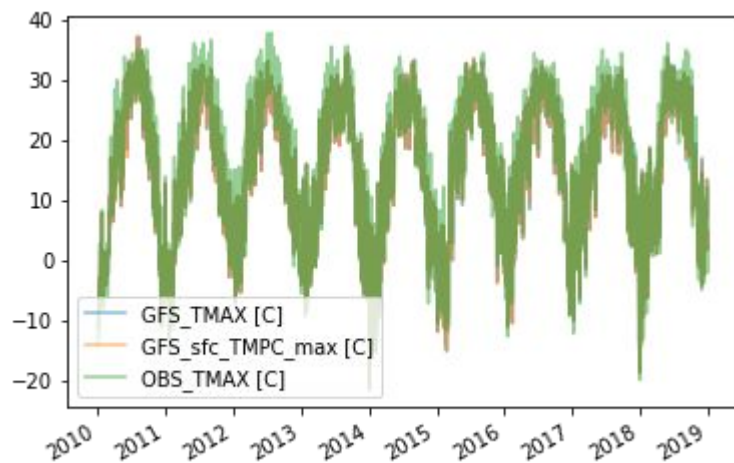
- What data we used
 - GFS daily [daily.tar.gz]
 - GFS surface [sfc.tar.gz]
 - KCMI_daily.csv
 - KCMI_hourly.csv (precipitation only)
- Missing data
- Units
 - Temperatures: °C
 - Wind speed: m/s
 - Precipitation: mm
- GFS surface data
 - Double 06UTC values
 - We calculated the extremes from 3-hourly data
 - may be different than the true values



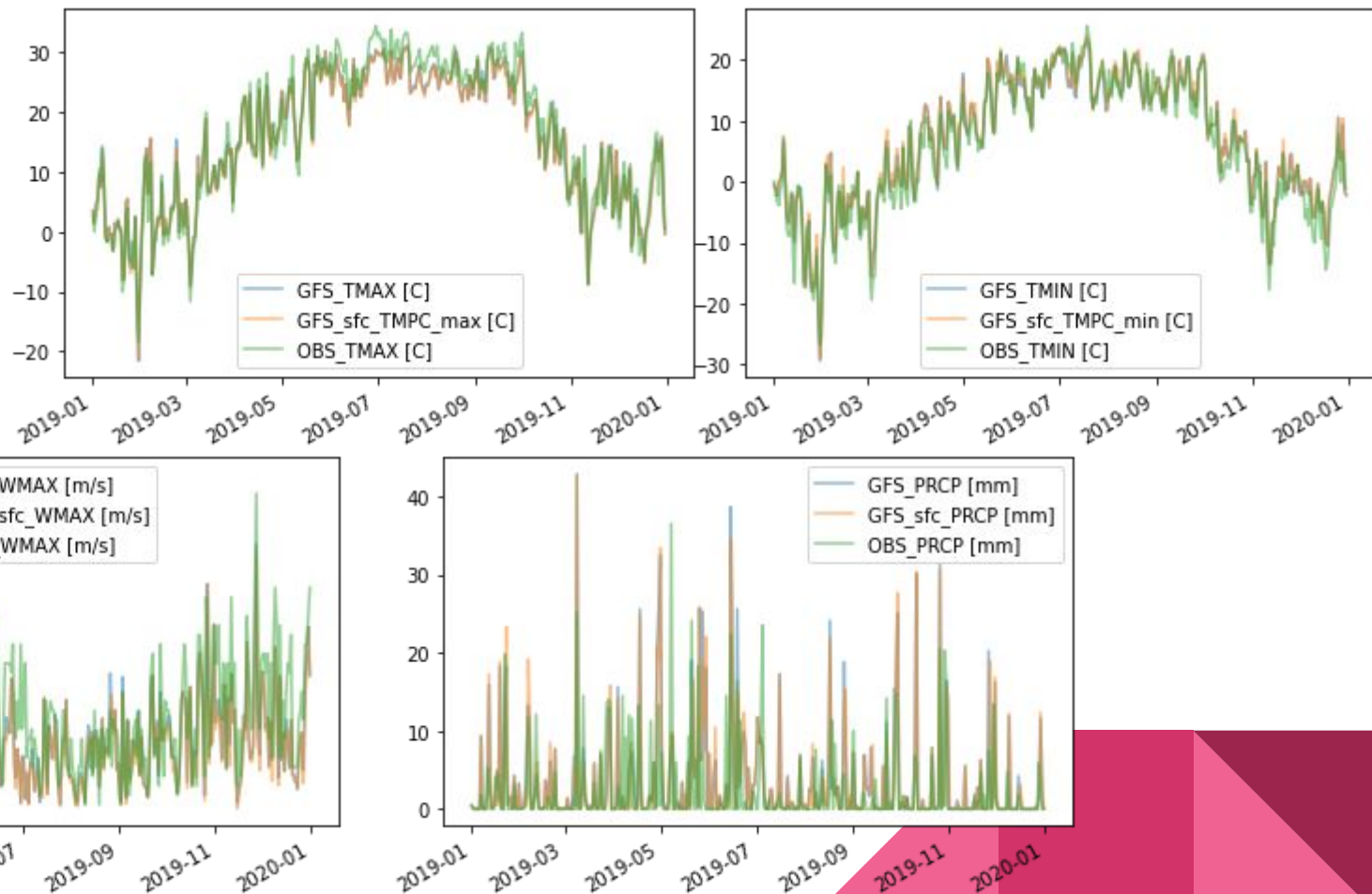


Preparing Data for Training and Validation

Training



Validation

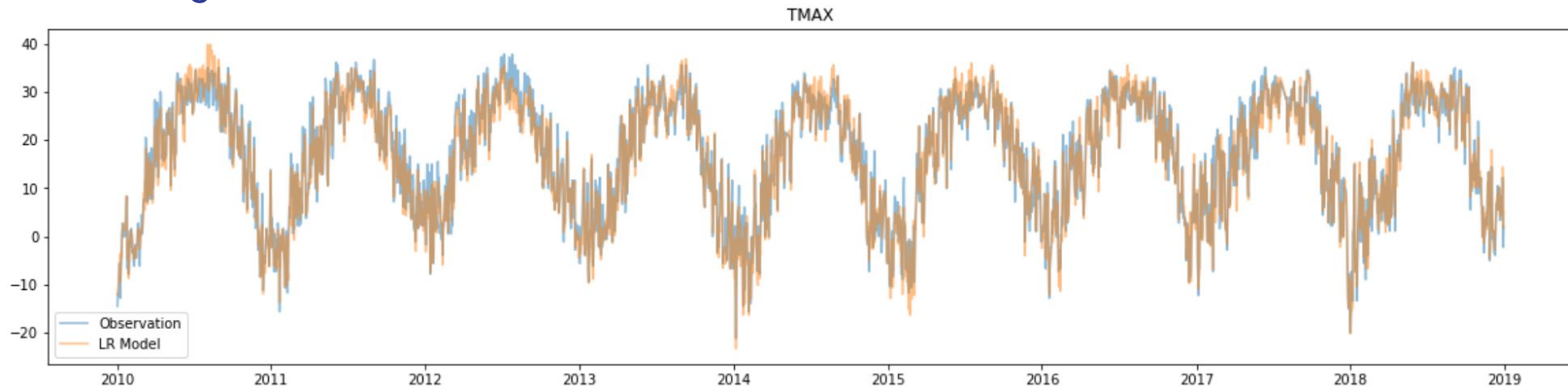
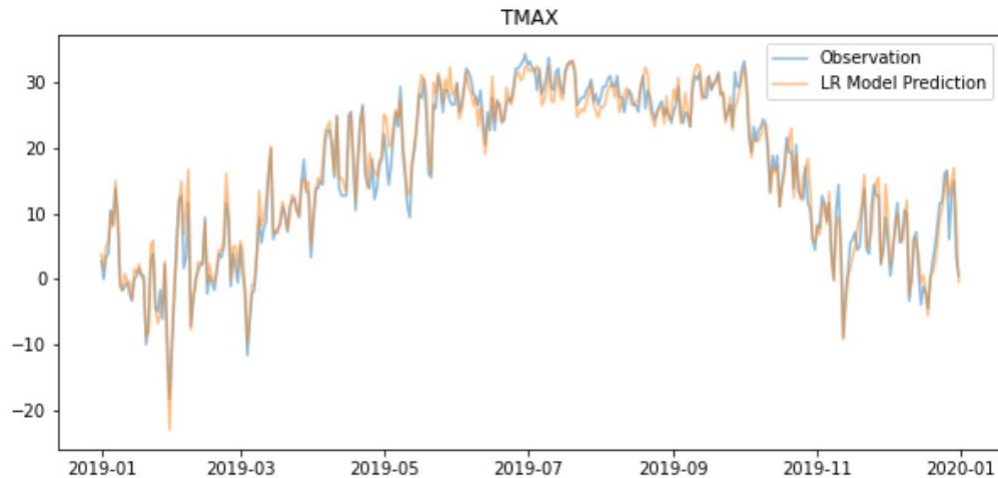


Linear Regression Models

TMAX (°C)

Validation
RMSE = 2.02

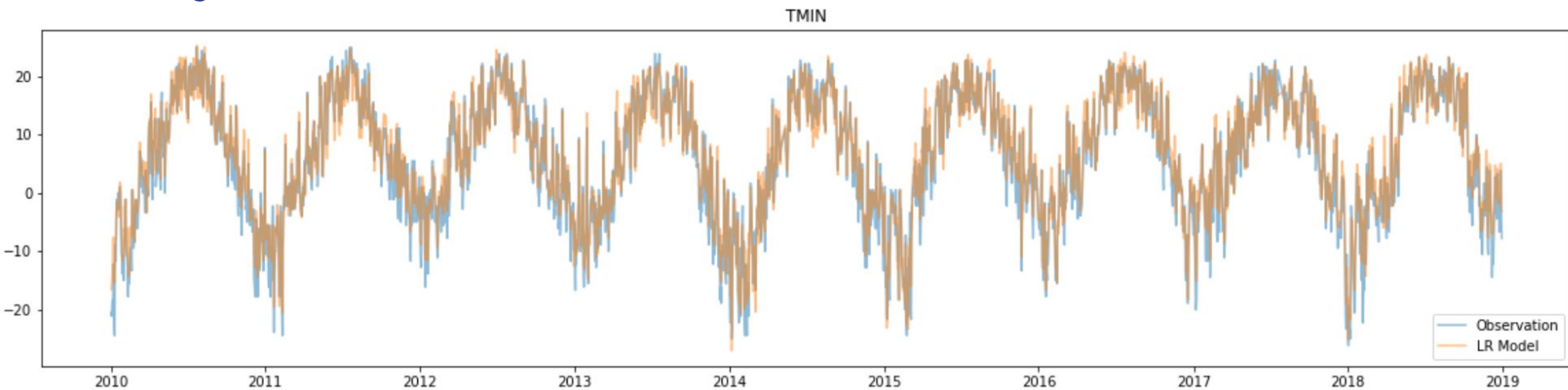
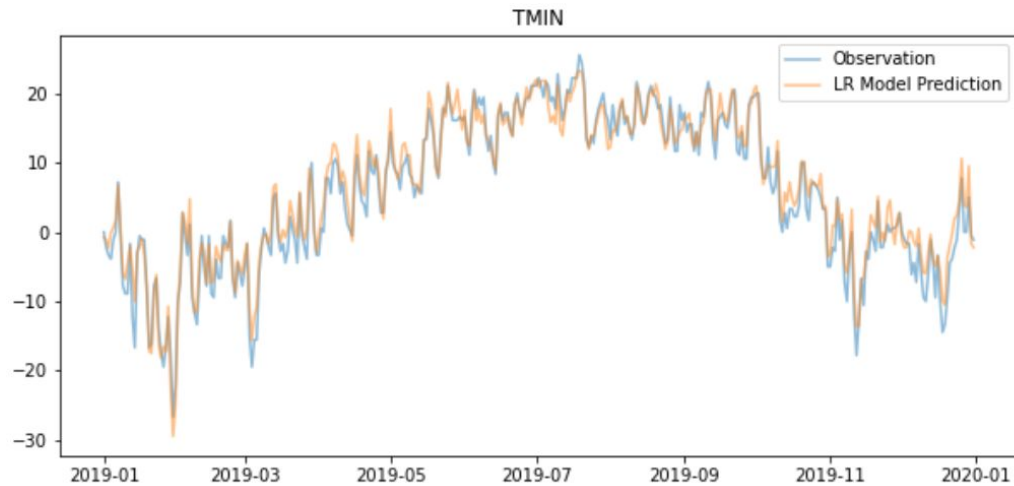
Training



TMIN (°C)

Validation
RMSE = 2.15

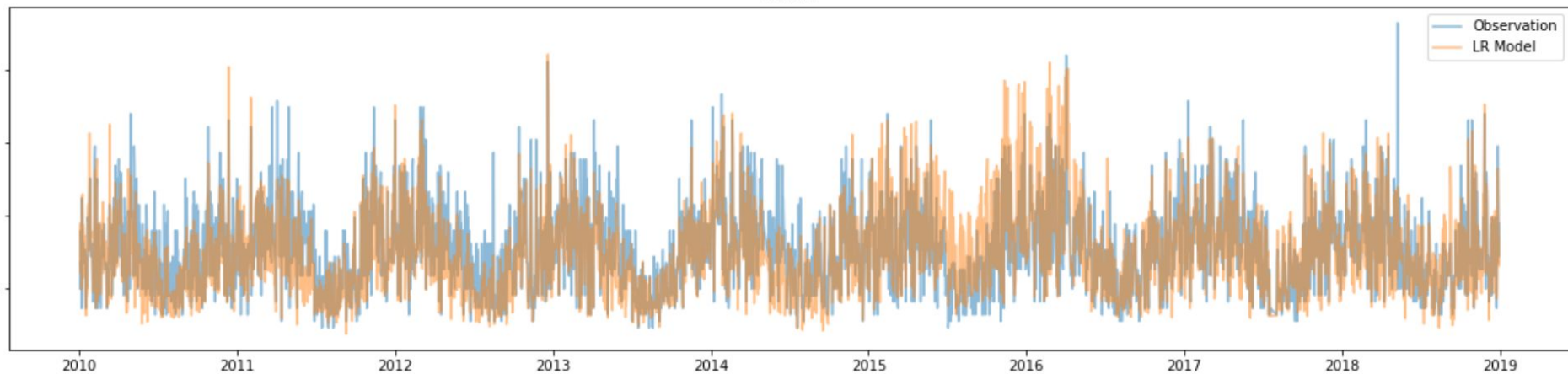
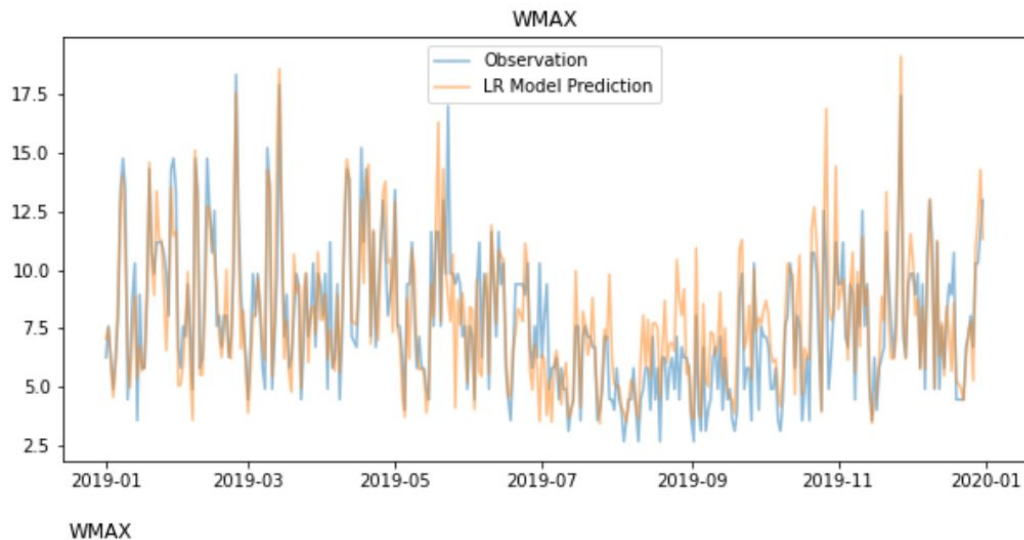
Training



WMAX (m/s)

Validation
RMSE = 1.64

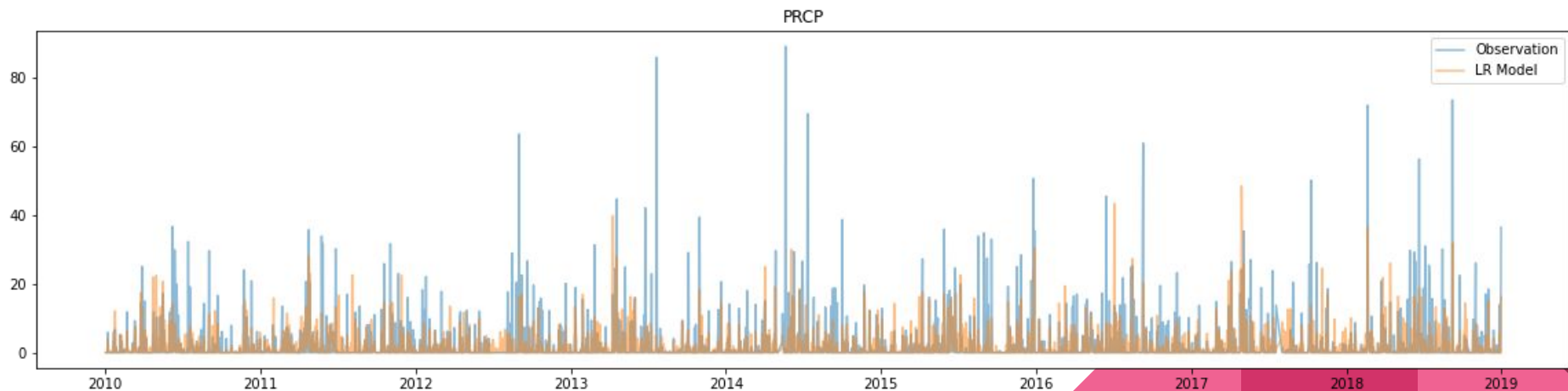
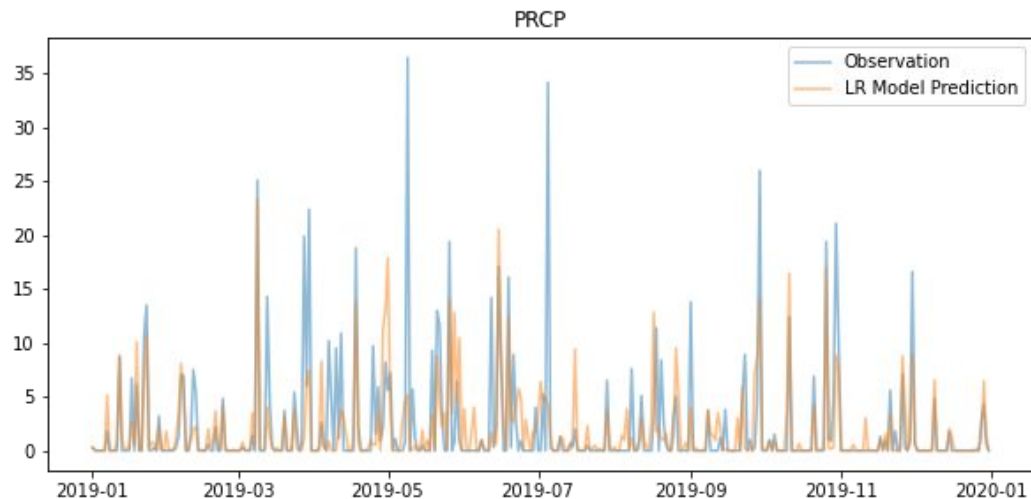
Training



PRECIP (mm)

Validation
RMSE = 3.88

Training

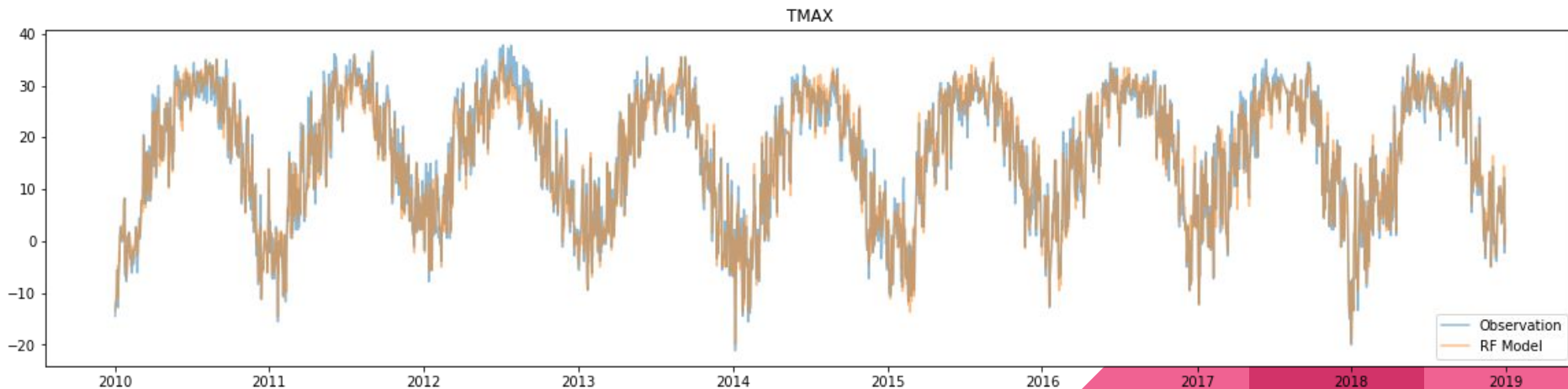
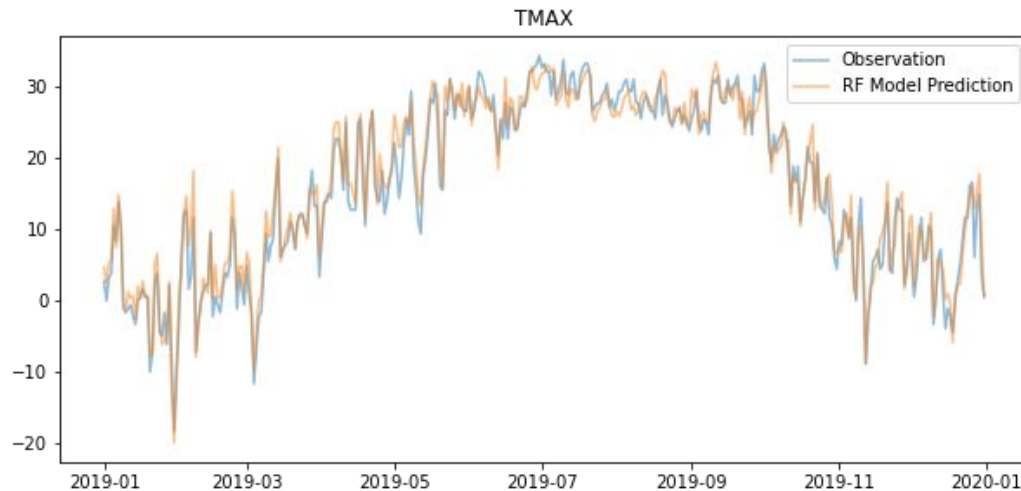


Random Forest Regression Models

TMAX (°C)

Validation
RMSE = 2.22

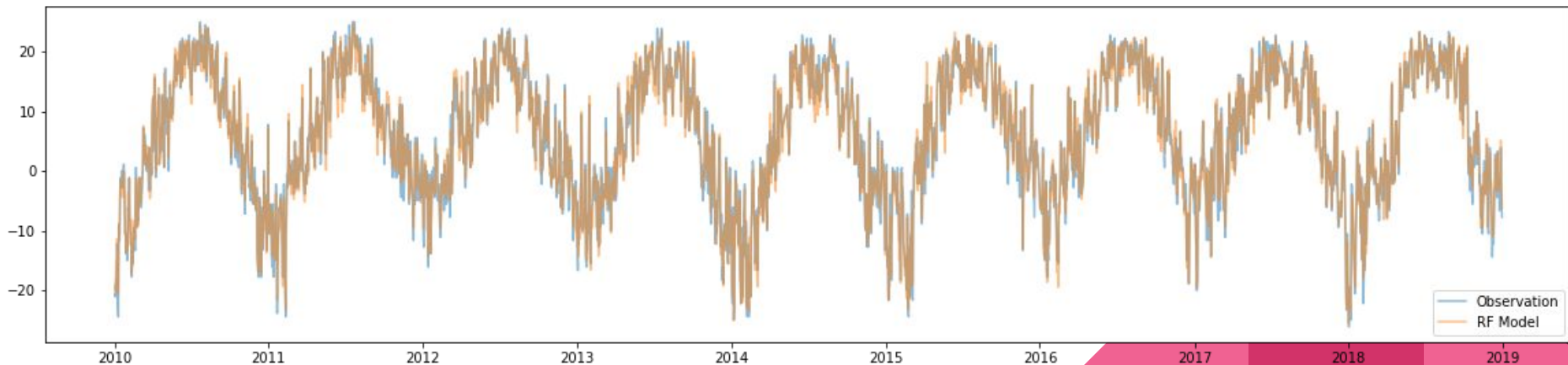
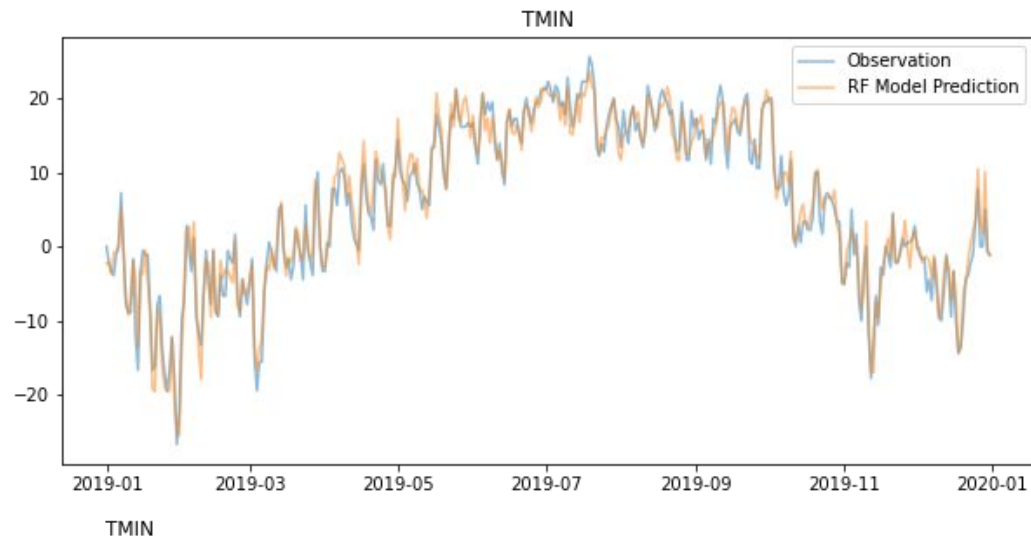
Training: 225 estimators



TMIN (°C)

Validation
RMSE = 1.96

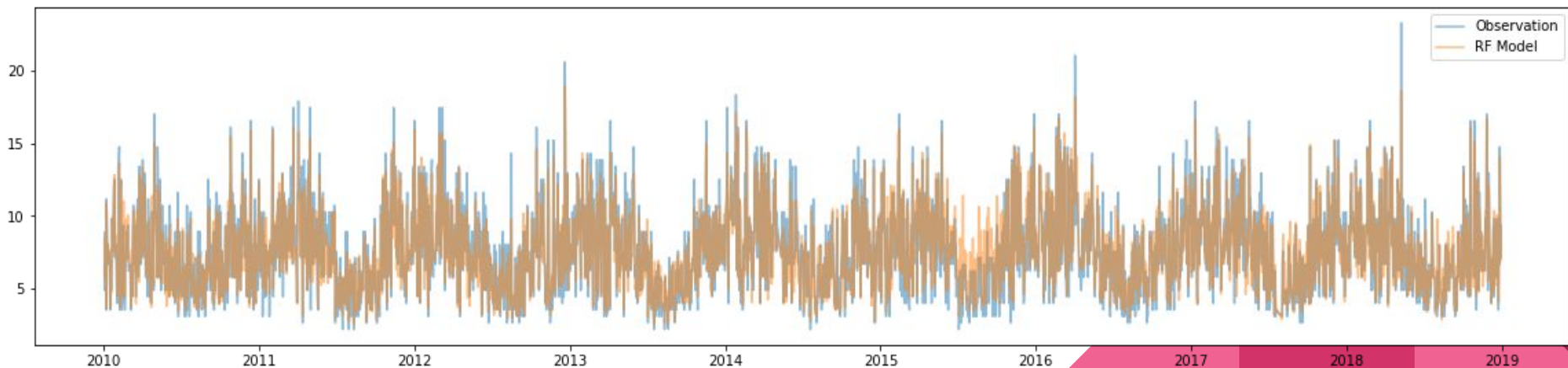
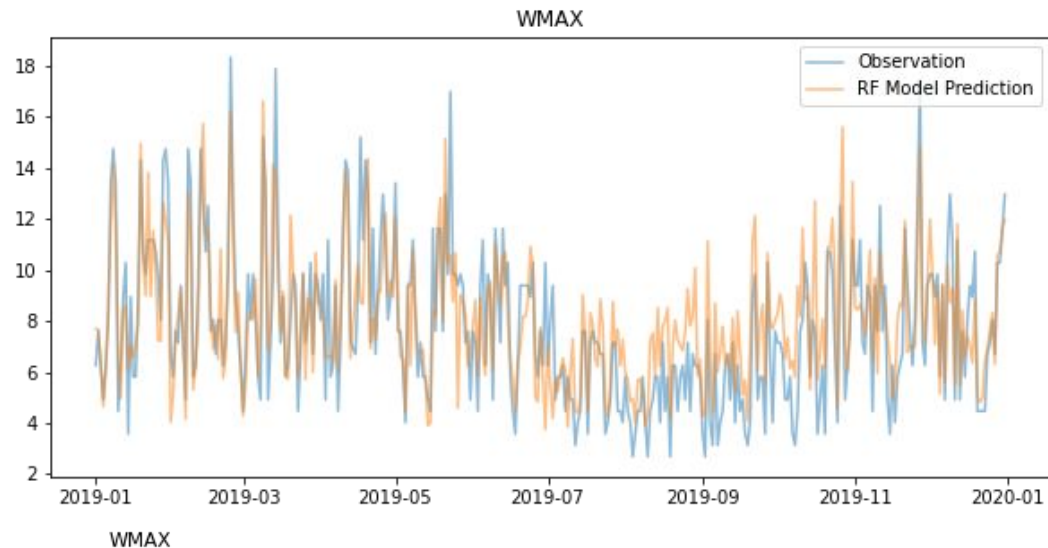
Training: 250 estimators



WMAX (m/s)

Validation
RMSE = 1.81

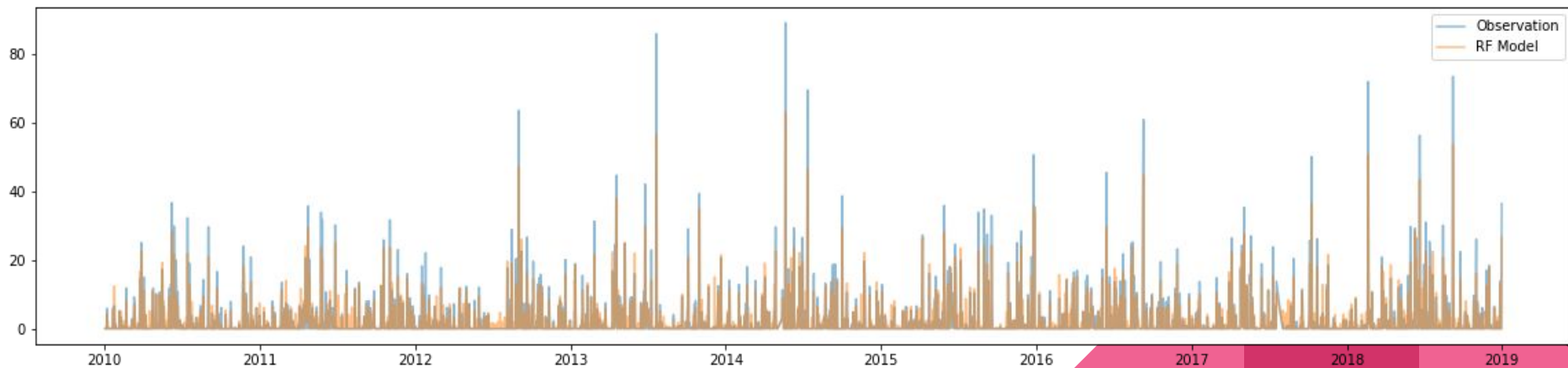
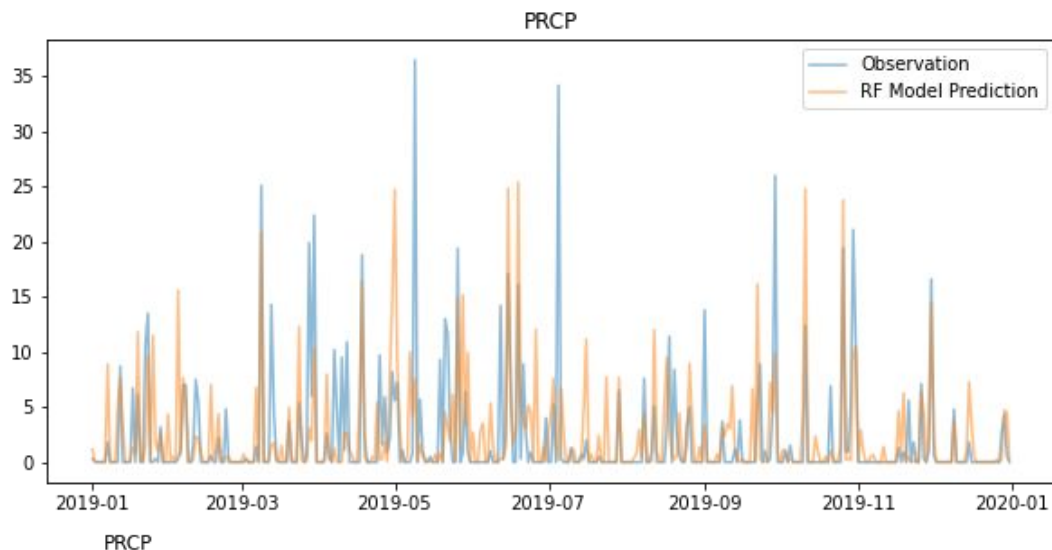
Training: 100 estimators



PRECIP (mm)

Validation
RMSE = 4.58

Training: 300 estimators



Comparison between LR and RF Models

	TMAX (°C)	TMIN (°C)	WMAX (m/s)	PRECIP (mm)
Linear Regression	2.02	2.15	1.64	3.88
Random Forest	2.22	1.96	1.81	4.58

