Assimilation of data from remote and in-situ platforms for improved numerical weather prediction

MD. NURUL HUDA
DOCTORAL STUDENT
CROFTON DEPT. OF AEROSPACE AND OCEAN ENGINEERING, VIRGINIA TECH

APRIL 08, 2021

VIRGINIA TECH™

WRFDA Overview

WRFDA is a Data Assimilation system built within the WRF software framework, used for application in both research and operational environments....

What WRFDA can do?

Provide Initial conditions for the WRF model forecast
- Verification and validation via difference observations and model
- Observing system design, monitoring and assessment
- Reanalysis
- Better understanding:
  – Data assimilation methods
  – Model errors
  – Data errors
  – Physical process interactions

Assimilation methods

Empirical methods

– Successive Correction Method (SCM)

– Nudging

– Physical Initialisation (PI), Latent Heat Nudging (LHN)

Statistical methods

– Optimal Interpolation (OI)

– 3-Dimensional VARiational data assimilation (3DVAR)

– 4-Dimensional VARiational data assimilation (4DVAR)

Advanced methods

– Extended Kalman Filter (EKF)

– Ensemble Kalman Filter (EnKF)

– Hybrid VAR/Ens DA

WRFDA Observations

In-Situ:
- SYNOP
- METAR
- SHIP
- BUOY
- TEMP
- PIBAL
- AIREP, AIREP humidity
- TAMDAR

Virginia Tech

WRFDA Observations

Remotely sensed retrievals :

- Atmospheric Motion Vectors (geo/polar)
- SATEM thickness
- Ground-based GPS TPW or ZTD
- SSM/I oceanic surface wind speed and TPW
- Scatterometer oceanic surface winds
- Wind Profiler
- Radar data (enhancements in V3.7)
- Satellite temperature/humidity/thickness profiles
- GPS refractivity (e.g. COSMIC)
- Stage IV precipitation/rain rate data (4D-Var)

WRFDA Observations

Radiances:

– HIRS NOAA-16, NOAA-17, NOAA-18, NOAA-19, METOP-A
– AMSU-A NOAA-15, NOAA-16, NOAA-18, NOAA-19, EOS-Aqua, METOP-A, METOP-B
– AMSU-B NOAA-15, NOAA-16, NOAA-17
– MHS NOAA-18, NOAA-19, METOP-A, METOP-B
– AIRS EOS-Aqua
– SSMIS DMSP-16, DMSP-17, DMSP-18
– IASI METOP-A, METOP-B
– ATMS Suomi-NPP
– MWTS FY-3
– MWHS FY-3
– SEVIRI METEOSAT

Challenges for NWP Models

- Improving the accuracy of numerical solution of the atmospheric hydro thermodynamics equations.
- Improving the evaluation of the initial state of the atmosphere.
- Account of random properties of atmospheric processes.

Goal

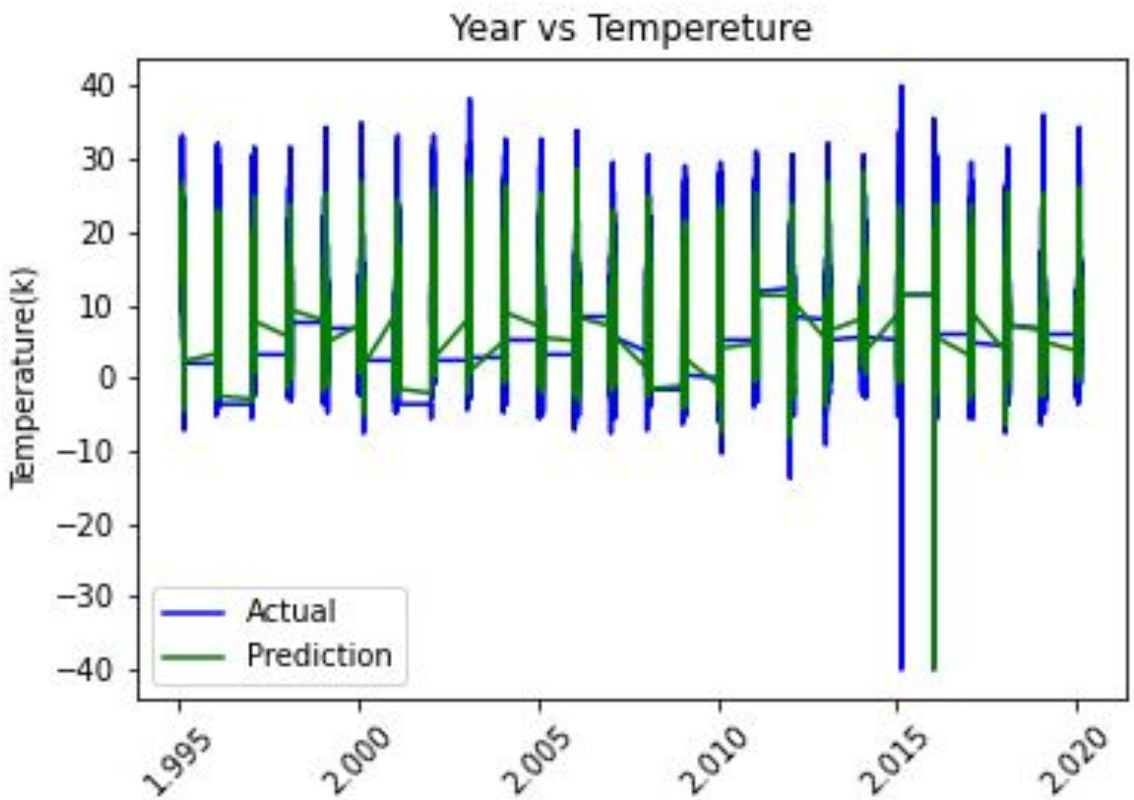
- To use machine learning and relevant data assimilation techniques to find and exploit relationships between atmospheric elements like water vapor, temperature, and clouds. These extracted relationships will then be input into data assimilation systems in order to improve NWP.

Data Description

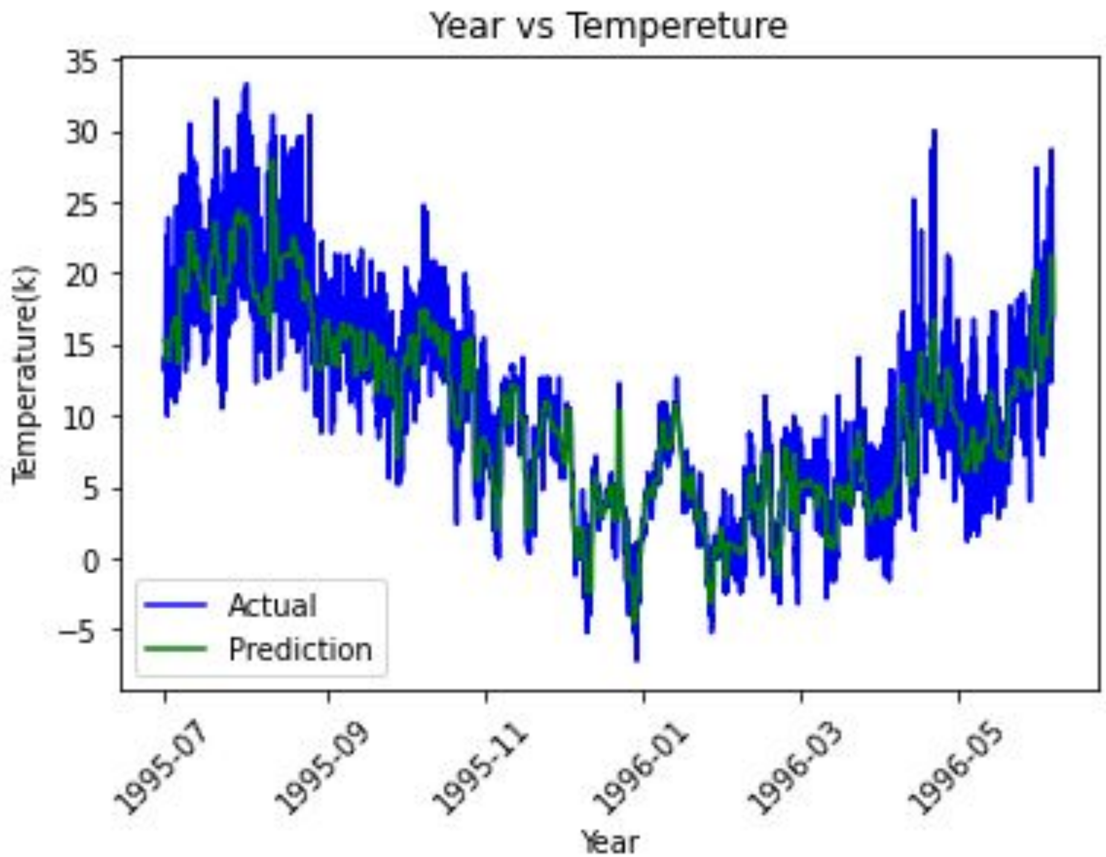
- Cambridge Computer Laboratory Rooftop Weather from 6/30/1995 3:30:00 PM To 11/12/2020 2:00 AM.

Cambridge Weather	Temp in C
Count	439229.000000
Mean	10.68411
Std	6.5817
Min	-40.00
Max	40.00

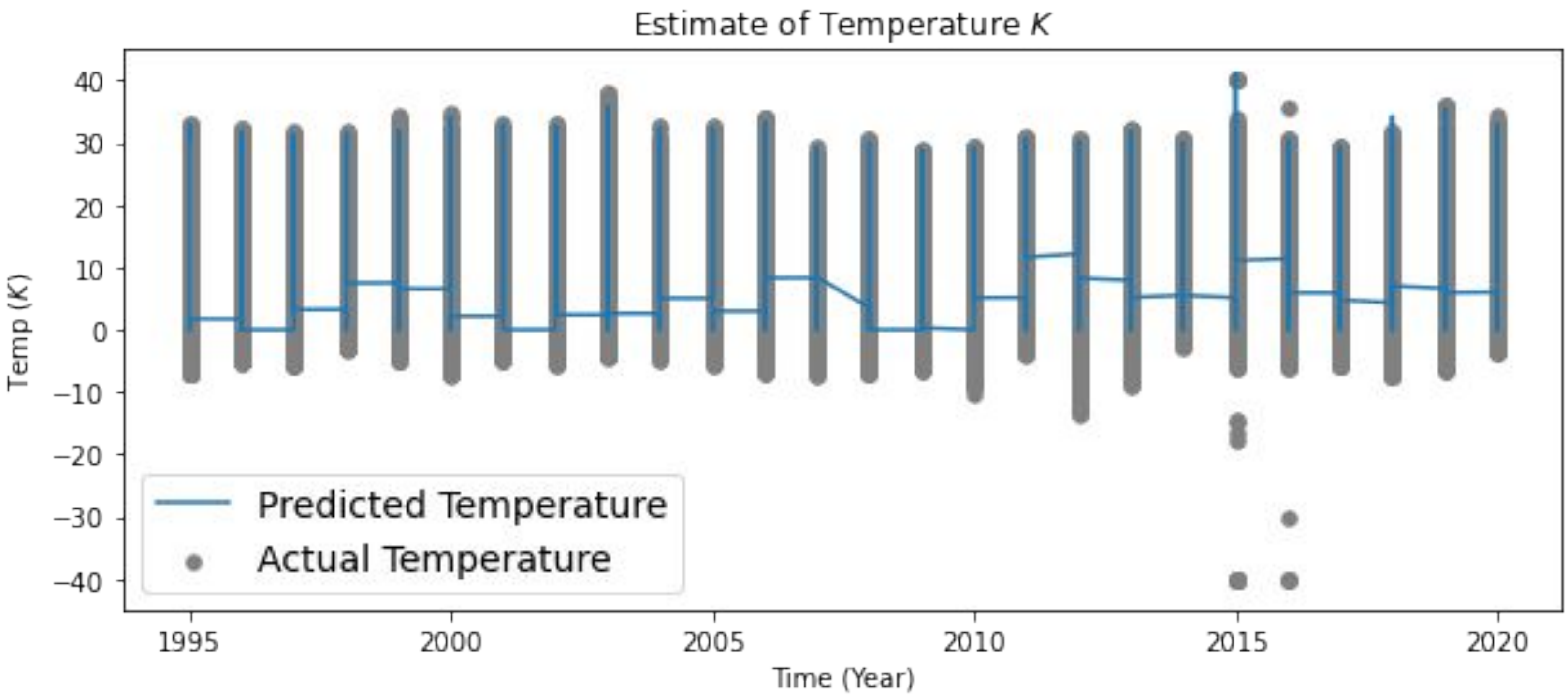
SVM



Year vs Tempereture
Temperature(k)
40
30
20
10
0
−10
−20
−30
−40
Actual
Prediction
1995
2000
2005
2010
2015
2020

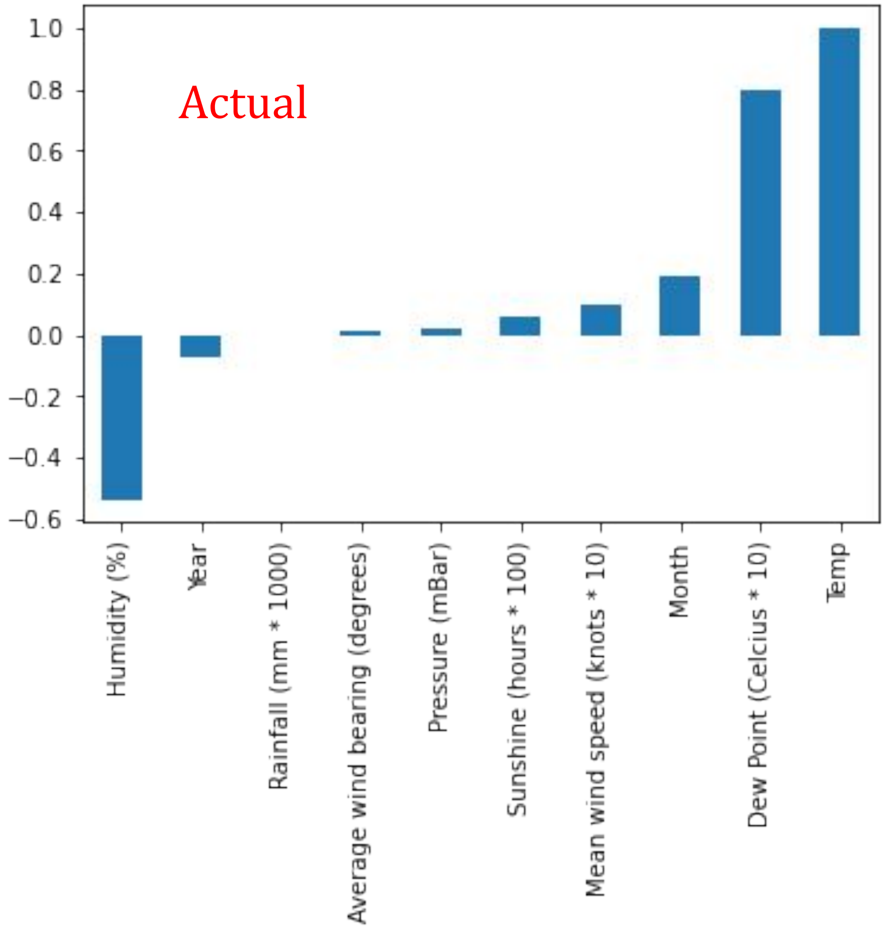

Year vs Tempereture
Temperature(k)
Year
35
30
25
20
15
10
5
0
−5
1995-07
1995-09
1995-11
1996-01
1996-03
1996-05
Actual
Prediction

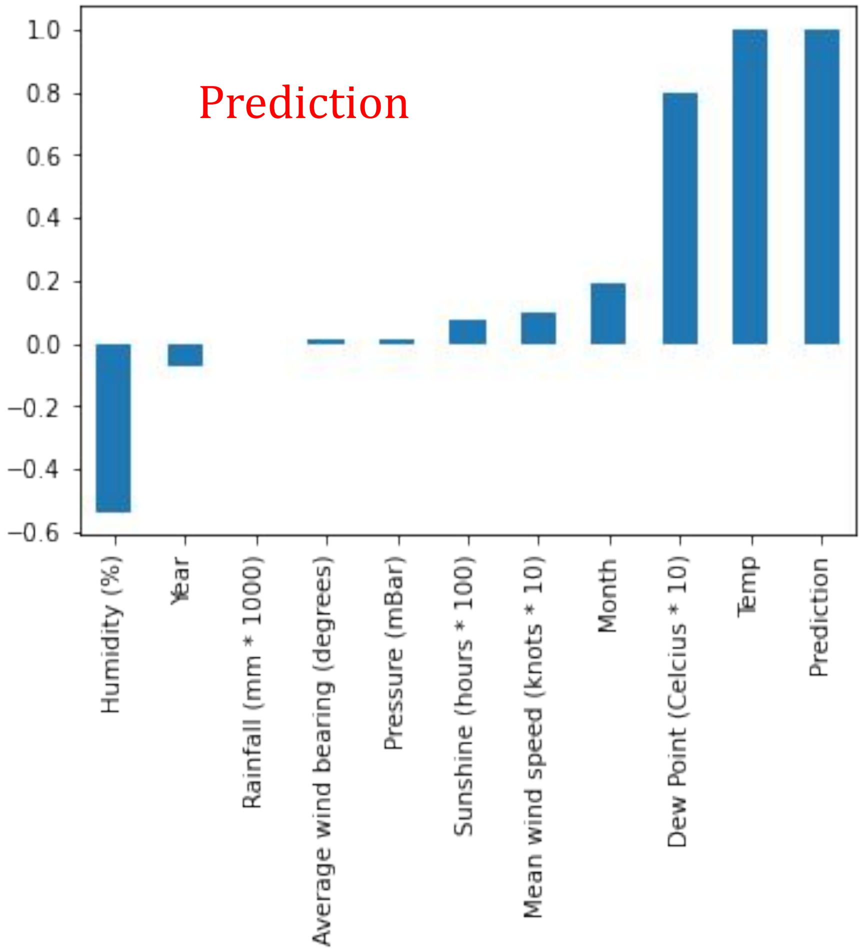
ANN



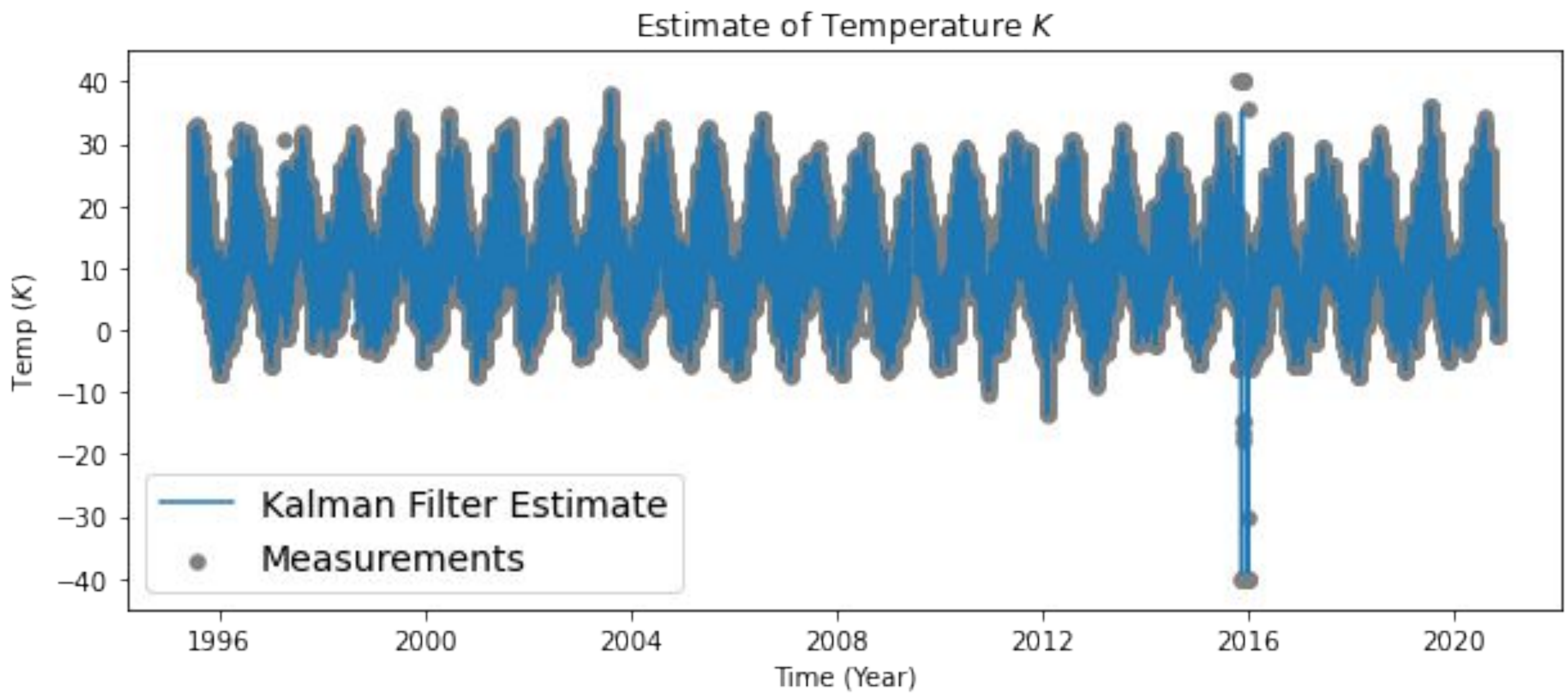
Estimate of Temperature K
Temp (K)
40
30
20
10
0
−10
−20
−30
−40
1995
2000
2005
2010
2015
2020
Time (Year)
Predicted Temperature
Actual Temperature

Correlations

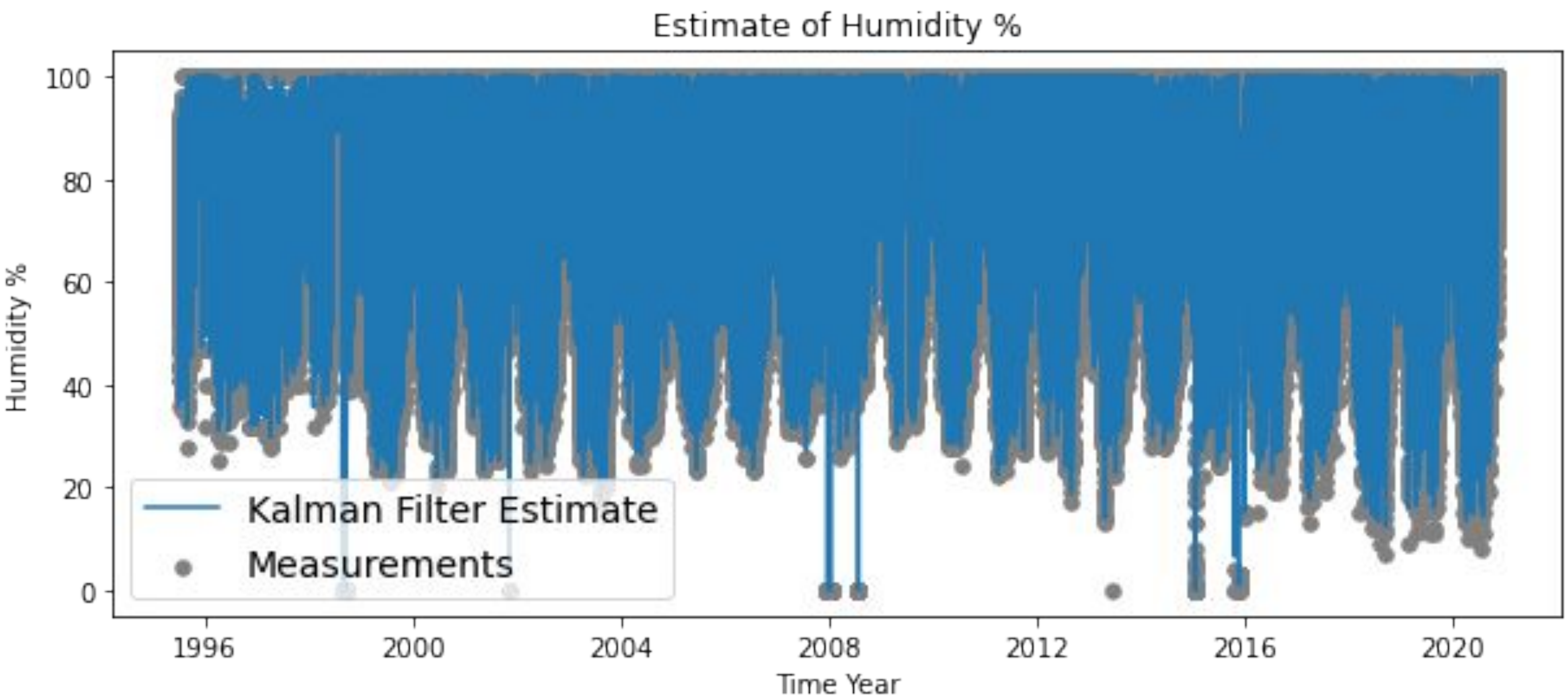




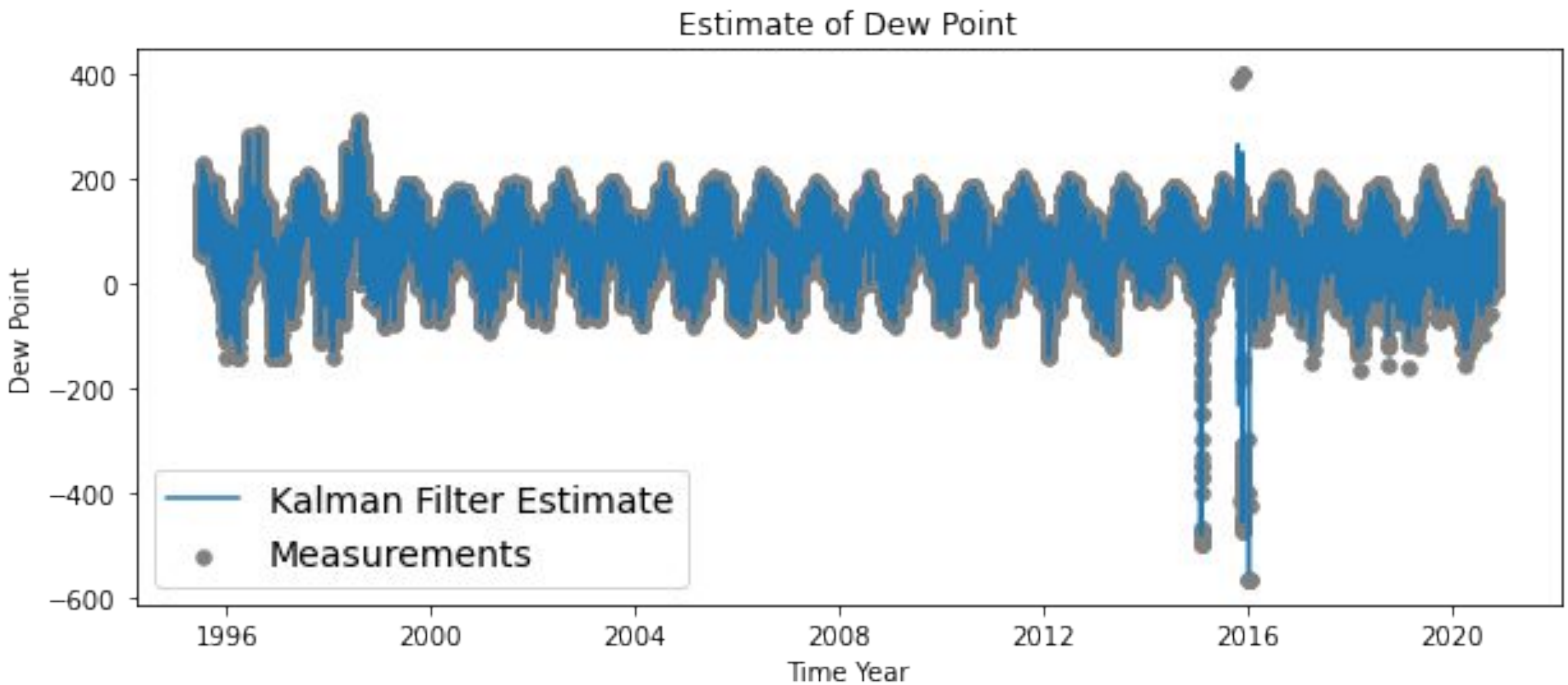
Kalman Filter Results



Kalman Filter Results



Kalman Filter Results



Variable	Variance	RMSE
Temprtature	43.31891699676156	0.349734195177556
Humidity	291.97791472927383	1.8771254817626482
Dew Point	2938.772391505238	3.7496620423068725

Comparison

Data Source	Time Range	# of Data	RMSE	Computation Cost	ML Technique
Cambridge Computer Laboratory Rooftop Weather	6/30/1995 3:30:00 PM To 11/12/2020 2:00	N/A	0.34973419	10 Mins	Kalman Filter
Cambridge Computer Laboratory Rooftop Weather	6/30/1995 3:30:00 PM To 11/12/2020 2:00	Total=439230 Training=315360 Test=123870 Train/Test=2.54	0.49763	8 Mins	ANN
Cambridge Computer Laboratory Rooftop Weather	6/30/1995 3:30:00 PM To 11/12/2020 2:00	Total=439230 Training=315360 Test=123870 Train/Test=2.54	1.857	4.2 Hours	LSTM
Cambridge Computer Laboratory Rooftop Weather	6/30/1995 3:30:00 PM To 11/12/2020 2:00	Total=439230 Training=315360 Test=123870 Train/Test=2.54	3.102	2.8 Hours	SVM

Summary and Research Gap

Most of the research developed in this area (monthly, daily and hourly) are focused on ANN strategies (57%) in comparison with the other widely used strategy SVM (43%). However, it is possible to see that, when SVM and ANN were compared, in most cases, SVM reported a better performance compared with classical ANN-based strategies.

Thank You.

Thank You