Application of machine learning and deep learning methods for load prediction in institutional buildings

By
Sahar Javadian

Supervisor
Prof. Ursula Eicker

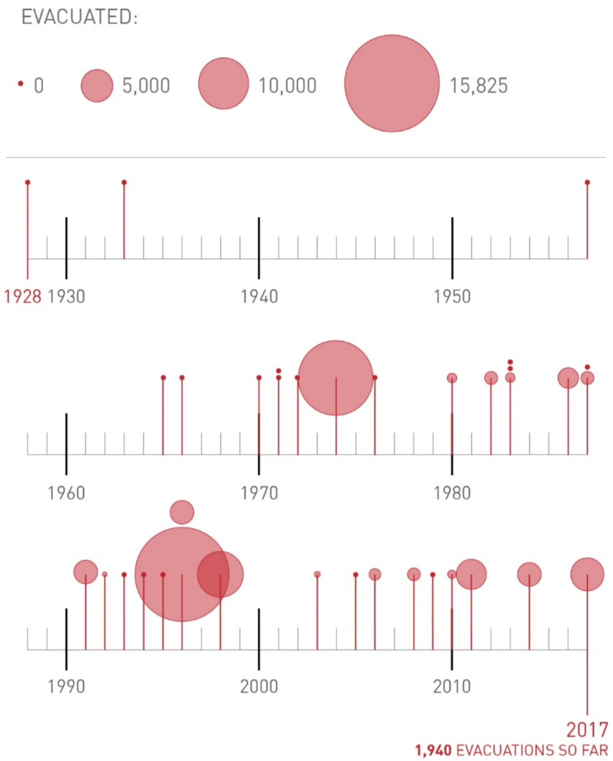
Department of Building, Civil and Environmental engineering

December 21, 2020
Montreal

Outline

Climate change

- History of floods in Quebec

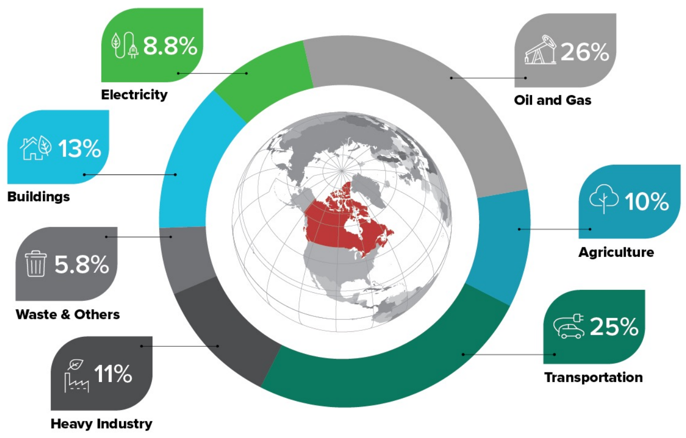


https://www.cbc.ca/news/canada/montreal/quebec-floods-history-timeline-1.4105530

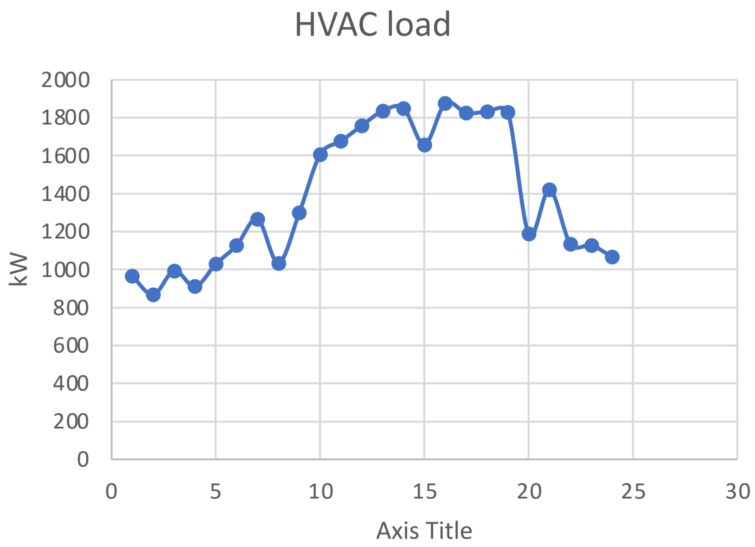
https://www.nationalobserver.com/2020/09/21/opinion/foundations-aim-help-build-better-canada-green-recovery

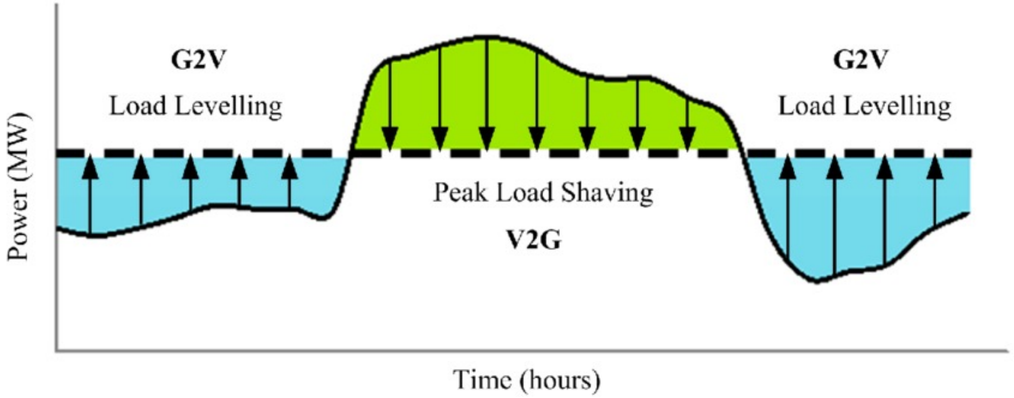
Green house gas emissions

➢ Share different sector for gas emissions



https://www.canada.ca/en/services/environment/weather/climatechange/climate-plan/reduce-emissions.html#shr-pg0





https://www.researchgate.net/figure/The-concept-of-peak-load-shaving-and-load-levelling_fig3_319183694

Reduce energy usage in building sector by:

- Peak load shaving
- use renewable systems

For these two goals we need to analysis the load and to know future load

Machine learning vs Deep learning

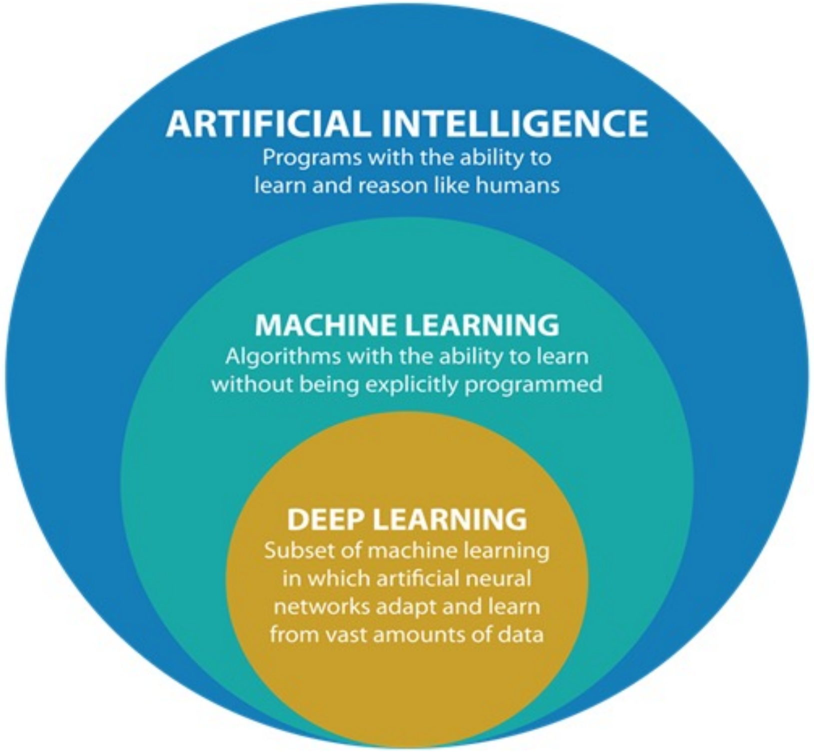
Load prediction by machine learning method:

- Regression models
- Neural networks (NN)
- Support vector machine (SVM)
- decision tree

Deep learning models such as:

- CNN and LSTM

Deep learning



https://datacatchup.com/artificial-intelligence-machine-learning-and-deep-learning/

Literature review

Regression

	Year	authors	Research focus	methods	Results
1	2019	Anand et al	to approximate a factor known as ”energy use per person	Deep neural network and Multiple nonlinear regression	DNN outperformed MNLR.
2	2015	Fumo and Biswas	household energy usage prediction	categories of regression	Regression is easy and efficient model
3	2015	Qiang et al	prediction of cooling load	MLR	MLR has low accuracy in comparison to other models such as ANN
4	2011	Nazih et al	load prediction of one-year time horizon.	linear, polynomial, and exponential	linear and polynomial were at the almost same level.
5	2007	Tso and Yau	predict household energy consumption (Kwh)	regression model, decision tree, and neural network	The effect of each parameter in energy consumption varies in winter and summer analysis.

Literature review

Deep learning methods

	Year	authors	Research focus	methods	Results
1	2019	Wang et al	Next week load prediction	LSTM, MLP, RF	LSTM performed better most of the time
2	2017	Yildiz et al	Forecasting of electrical load	ANN, SVR and MLR	MLR cannot be good as other ML algorithms. the bigger scale prediction was more successful
3	2017	Kong et al	short-term load predication – Comparison of LSTM with other models	LSTM	LSTM showed superiority
4	2016	Marino et al	Forecasting of electrical load	two types of LSTM (standard LSTM and sequence to sequence LSTM)	LSTM performed poorly on the one-minute dataset, but S2S LSTM outperformed on both datasets
5	2011	Feilat and Bouzguenda	peak load of each month	NN and Linear regression	NN method performed better than linear regression

Understanding from literature and Existing gaps

Understanding from literature:

- Prediction of daily or monthly peak demand or load profiles.
- comparing predictive models' performance based on accuracy, error, and running time.
- hybrid models.

- MLR is a fast and easy method but suffer from low accuracy.
- Advantages of using deep learning methods for time series forecasting.

- Application of deep learning on real case study
- Application of load analytics in real-world cases

Objectives

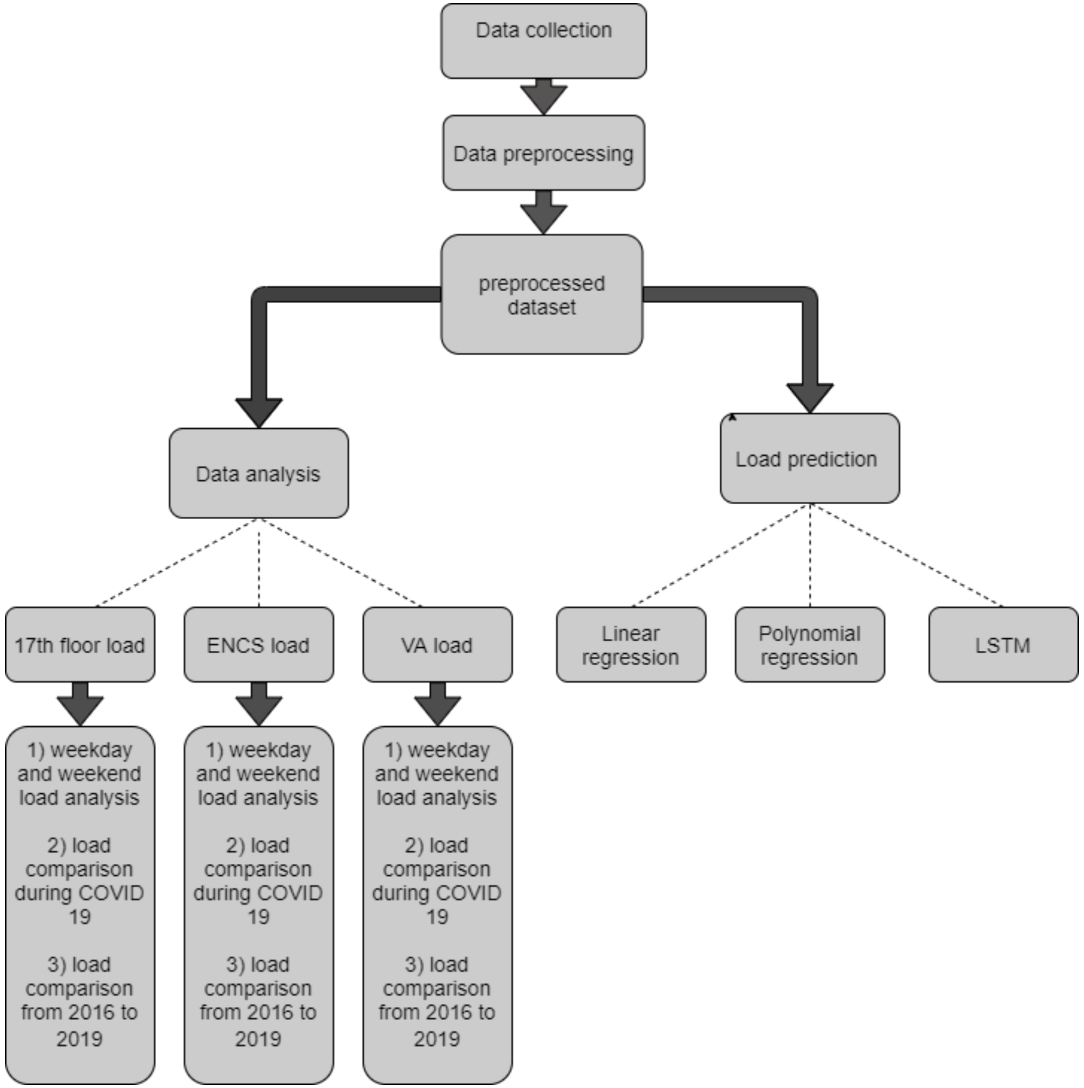
1. To analyze the energy consumption of EV building as base study for load prediction
2. To study the effects of weather variables on load prediction
3. To forecast electrical load for different time horizons with Long Short term memory (LSTM) model
4. To analyze the performance of LSTM on different load types
5. To Provide energy efficient suggestions based on load analysis and load prediction

Research tasks

Load analysis:
- Extraction of typical daily load profile
- Load variation during COVID 19 period
- Load analysis between different years (2015- 2019)

Load prediction:
- Influencing factors on load
- Model evaluation on detecting unusual consumption
- Effects of load types on load forecasting

Models used
- Linear regression *MLR and Polynomial regression
- LSTM



Case study – EV building

Heating:
- Natural gas boilers
- Electrical boiler

Cooling:
- Coolers
- Natural ventilation

ENCS:
- 17 floors
- Mechanical and chemical laboratories on 12^{th} – 16^{th} floors

VA:
- 12 floors
- Mechanical room is on the 12th floor.

https://imtl.org/montreal/building/Concordia-engenring.php

Available data set

	Points in data set	Name of attribute	Unit	Time resolution
EV dataset	Point_1	EV electric boiler	kW	15 minutes
	Point_3	17^{th} floor transformer	kW	15 minutes
	Point_4	ENCS transformer	kW	15 minutes
	Point_5	VA transformer	kW	15 minutes

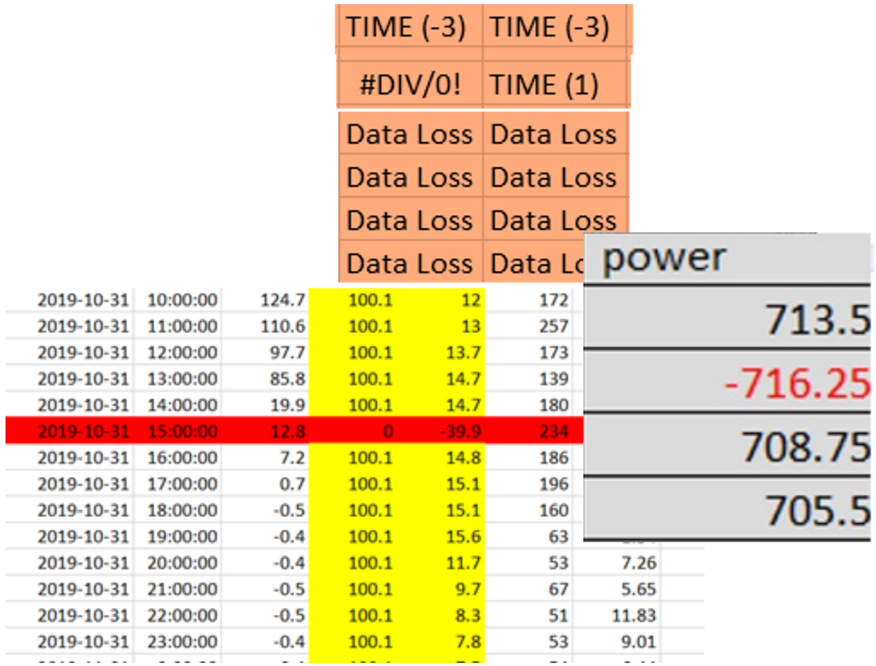
	Features name	Unit	Time resolution
Weather dataset	Solar Radiation	W/m^2	hourly
	Humidity	%	hourly
	Temperature	$^{\circ}c$	hourly
	Wind direction	Degrees from North	hourly
	Wind velocity	m/s	hourly

- EV weather station

Draft Final Report for Environment Canada - Concordia University Project

Data preprocessing

1. Replace string with NaN
2. Duplicate index
3. Complete data-frame
4. Removing not related columns
5. Remove NaN
6. Wrong sign values
7. Wrong values
8. Correct data type
9. Outlier detection
10. Convert to hourly interval
11. Add extra features
12. Data integration



TIME (-3)	TIME (-3)
#DIV/0!	TIME (1)
Data Loss	Data Loss
Data Loss	Data Loss
Data Loss	Data Loss
Data Loss	Data Lo

2019-10-31	10:00:00	124.7	100.1	12	172	
2019-10-31	11:00:00	110.6	100.1	13	257	
2019-10-31	12:00:00	97.7	100.1	13.7	173	
2019-10-31	13:00:00	85.8	100.1	14.7	139	
2019-10-31	14:00:00	19.9	100.1	14.7	180	
2019-10-31	15:00:00	12.8	0	-39.9	234	
2019-10-31	16:00:00	7.2	100.1	14.8	186	
2019-10-31	17:00:00	0.7	100.1	15.1	196	
2019-10-31	18:00:00	-0.5	100.1	15.1	160	
2019-10-31	19:00:00	-0.4	100.1	15.6	63	
2019-10-31	20:00:00	-0.4	100.1	11.7	53	7.26
2019-10-31	21:00:00	-0.5	100.1	9.7	67	5.65
2019-10-31	22:00:00	-0.5	100.1	8.3	51	11.83
2019-10-31	23:00:00	-0.4	100.1	7.8	53	9.01

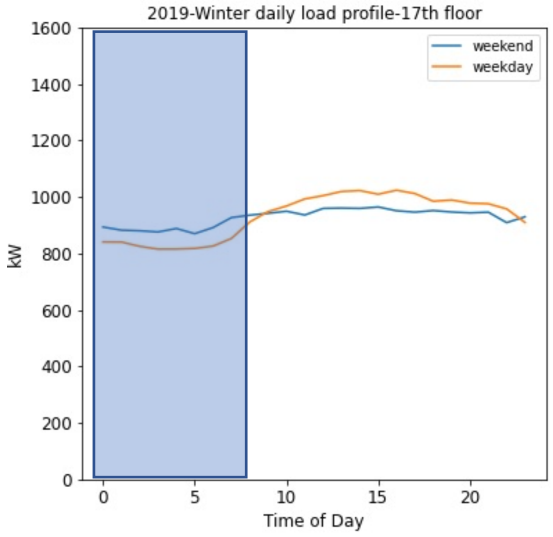
power
713.5
-716.25
708.75
705.5

- Data analysis

Load analysis of HVAC - 17^{th} floor load

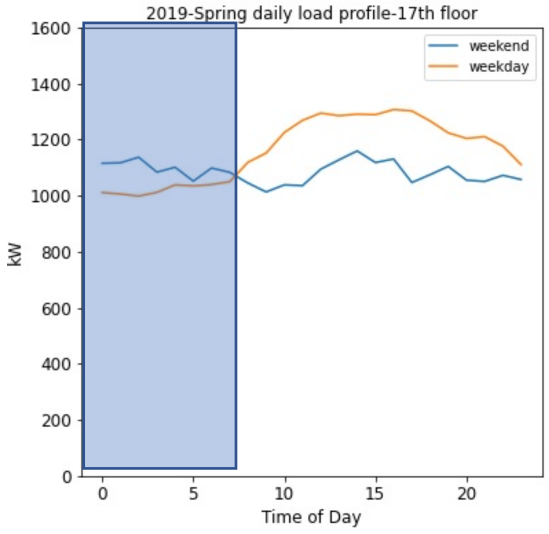
- Office hours and non-office hours are defined as: Office Time: 8 AM -6.59 PM and Non-Office Time: 7 PM – 7:59 PM

Winter



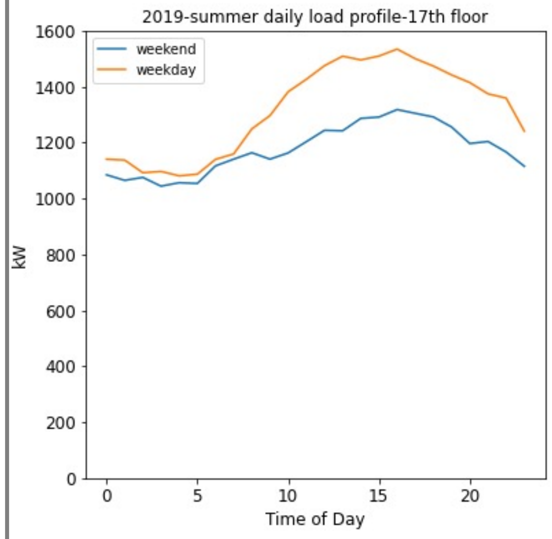
	Office load in kW	Non-Office load in kW	Reduction
Weekday	991	881	11.1 %
Weekend	951	907	4.6%

Spring



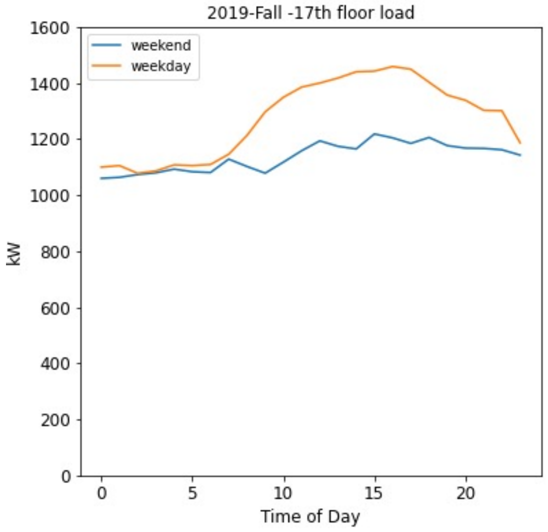
	Office load in kW	Non-Office load in kW	Reduction
Weekday	1254	1086	13.4%
Weekend	1080	1087	-0.6%

Summer



	Office load in kW	Non-Office load in kW	Reduction
Weekday	1441	1213	15.8%
Weekend	1241	1121	9.7%

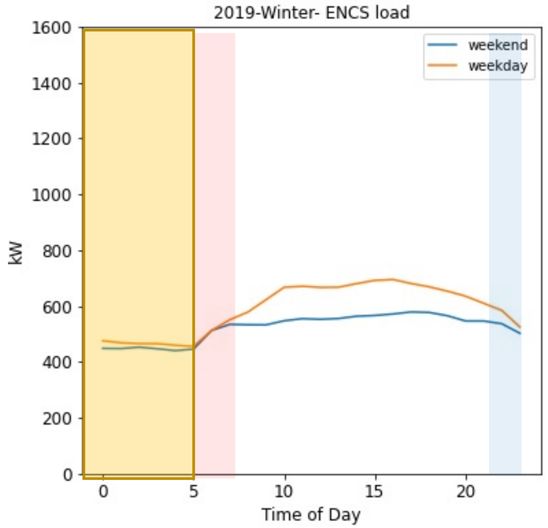
Autumn



	Office load in kW	Non-Office load in kW	Reduction
Weekday	1387	1179	15%
Weekend	1164	1114	4.3%

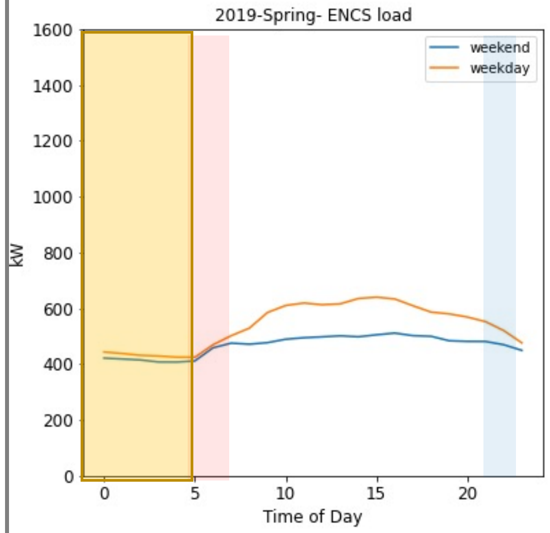
Load analysis - ENCS load

Winter



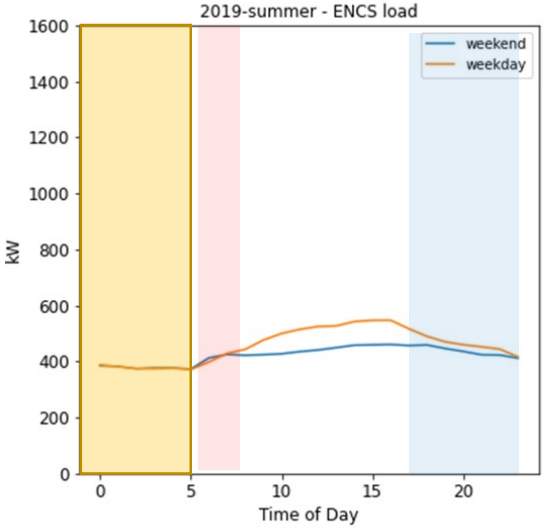
	Office load in kW	Non-Office load in kW	Reduction
Weekday	663	528	20.4%
Weekend	558	495	11.3%

Spring



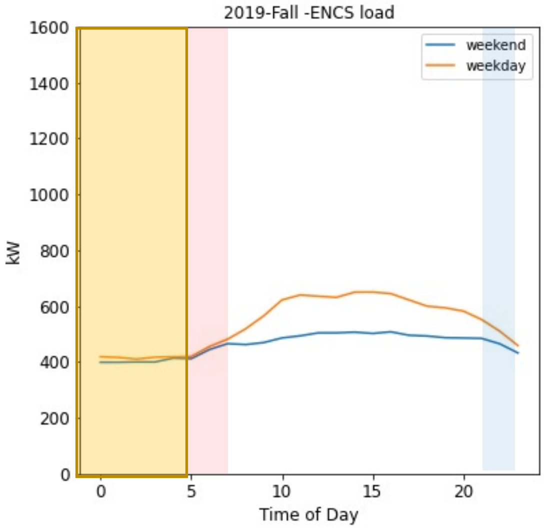
	Office load in kW	Non-Office load in kW	Reduction
Weekday	607	482	20.6%
Weekend	496	445	10.3%

Summer



	Office load in kW	Non-Office load in kW	Reduction
Weekday	512	411	19.7%
Weekend	445	404	9.2%

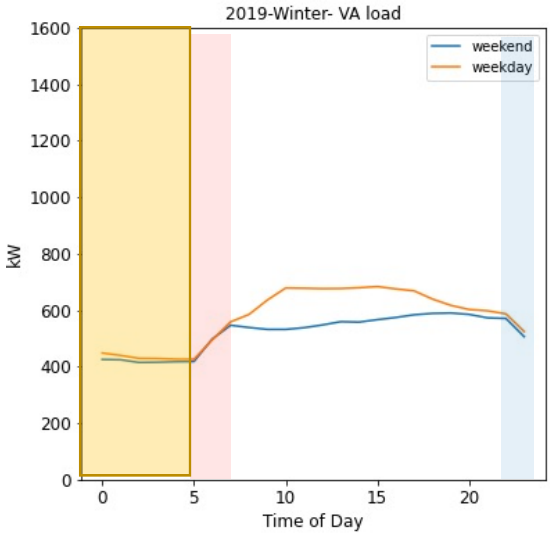
Autumn



	Office load in kW	Non-Office load in kW	Reduction
Weekday	616	472	23.4%
Weekend	494	438	11.3%

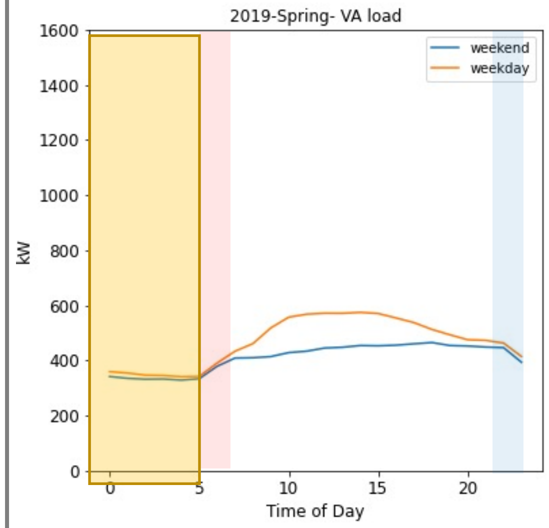
Load analysis - VA load

Winter



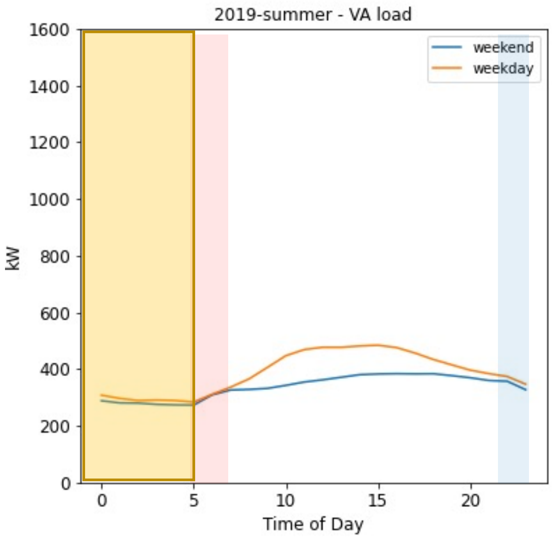
	Office load in kW	Non-Office load in kW	Reduction
Weekday	662	507	23.4%
Weekend	557	492	11.7%

Spring



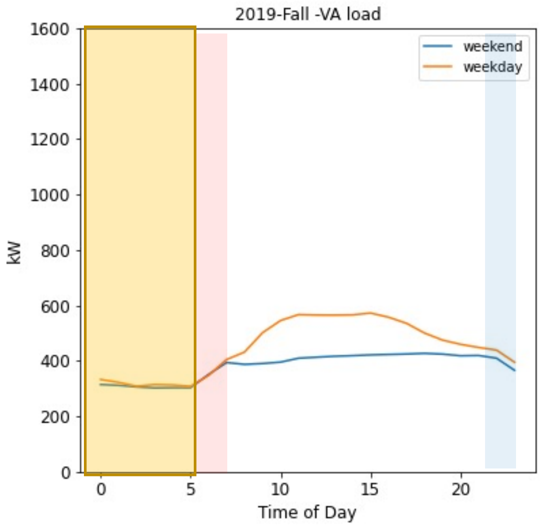
	Office load in kW	Non-Office load in kW	Reduction
Weekday	545	403	26.1%
Weekend	443	384	13.3%

Summer



	Office load in kW	Non-Office load in kW	Reduction
Weekday	453	333	26.5%
Weekend	365	316	13.4%

Autumn



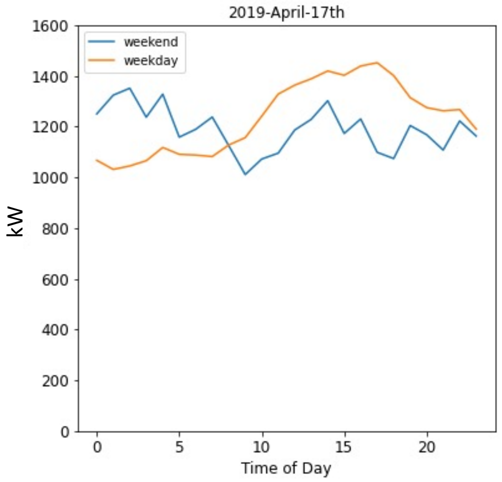
	Office load in kW	Non-Office load in kW	Reduction
Weekday	537	35	30.2%
Weekend	412	356	13.6%

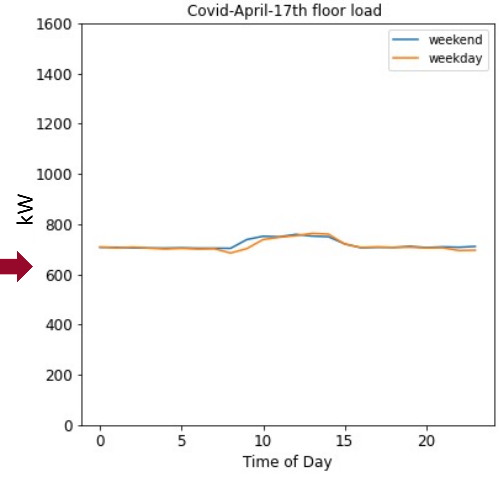
17th floor load comparison – COVID 19

2019

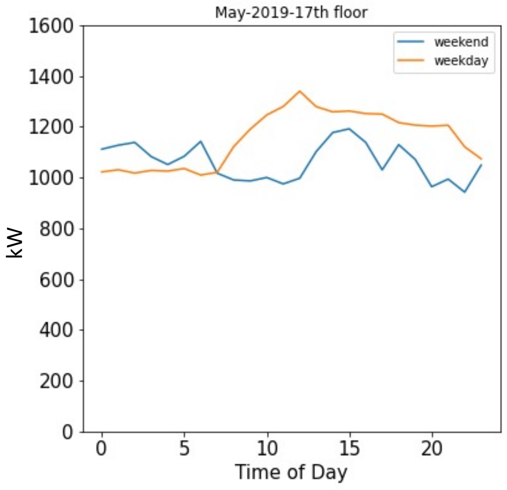
2020

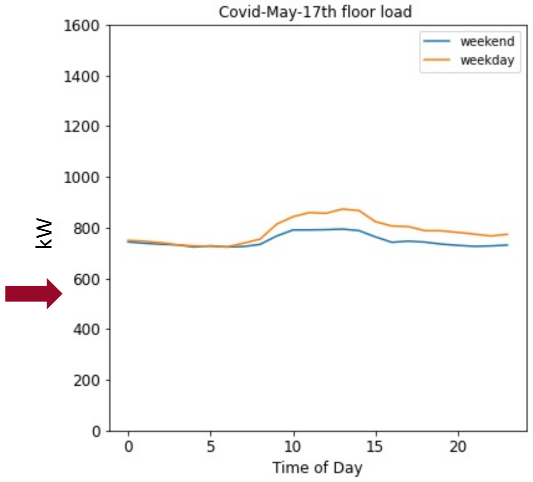
April





May



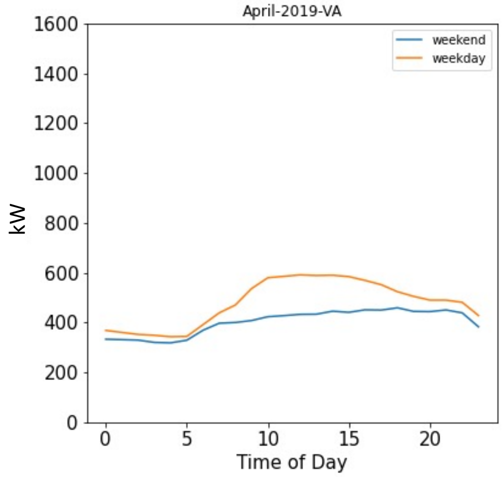


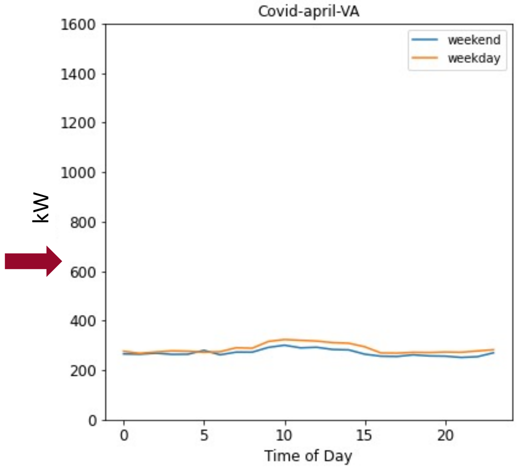
VA load comparison-COVID 19

2019

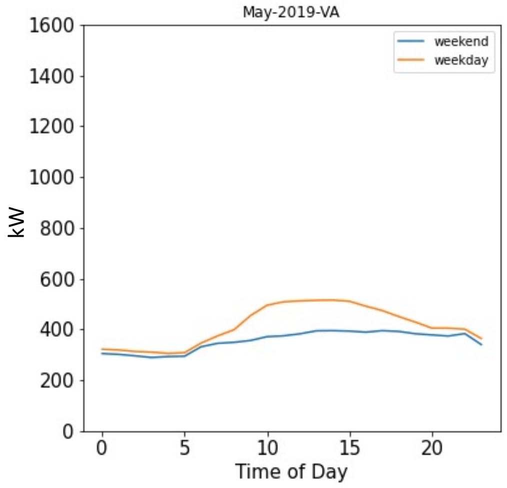
2020

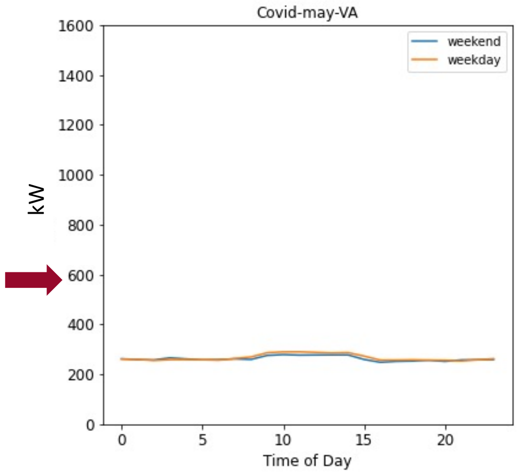
April





May



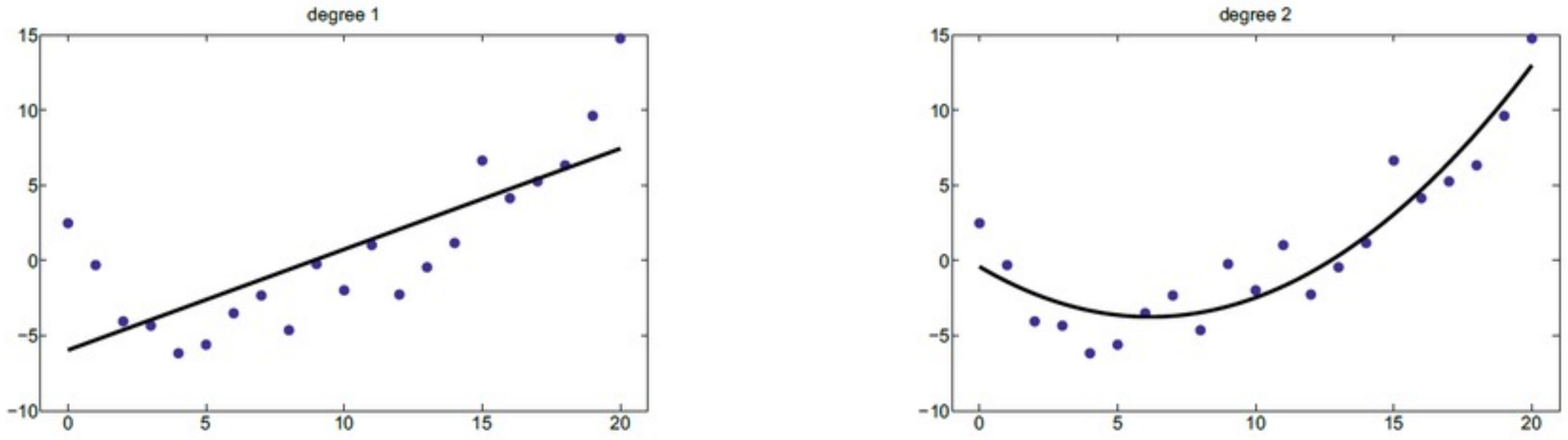


- Load prediction

Linear and Polynomial regression

Linear Regression: $\hat{y}_i = \beta_0 + \beta_1 * x_{i1} + \beta_2 * x_{i2} + ...\beta_j * x_{ij} + \varepsilon$

Polynomial regression: $\hat{y}_i = \beta_0 + \beta_1 * x_{i1} + \beta_2 * x_{i2}^2 + \beta_3 * x_{i3}^3 + ...\beta * x_{ij}^j$

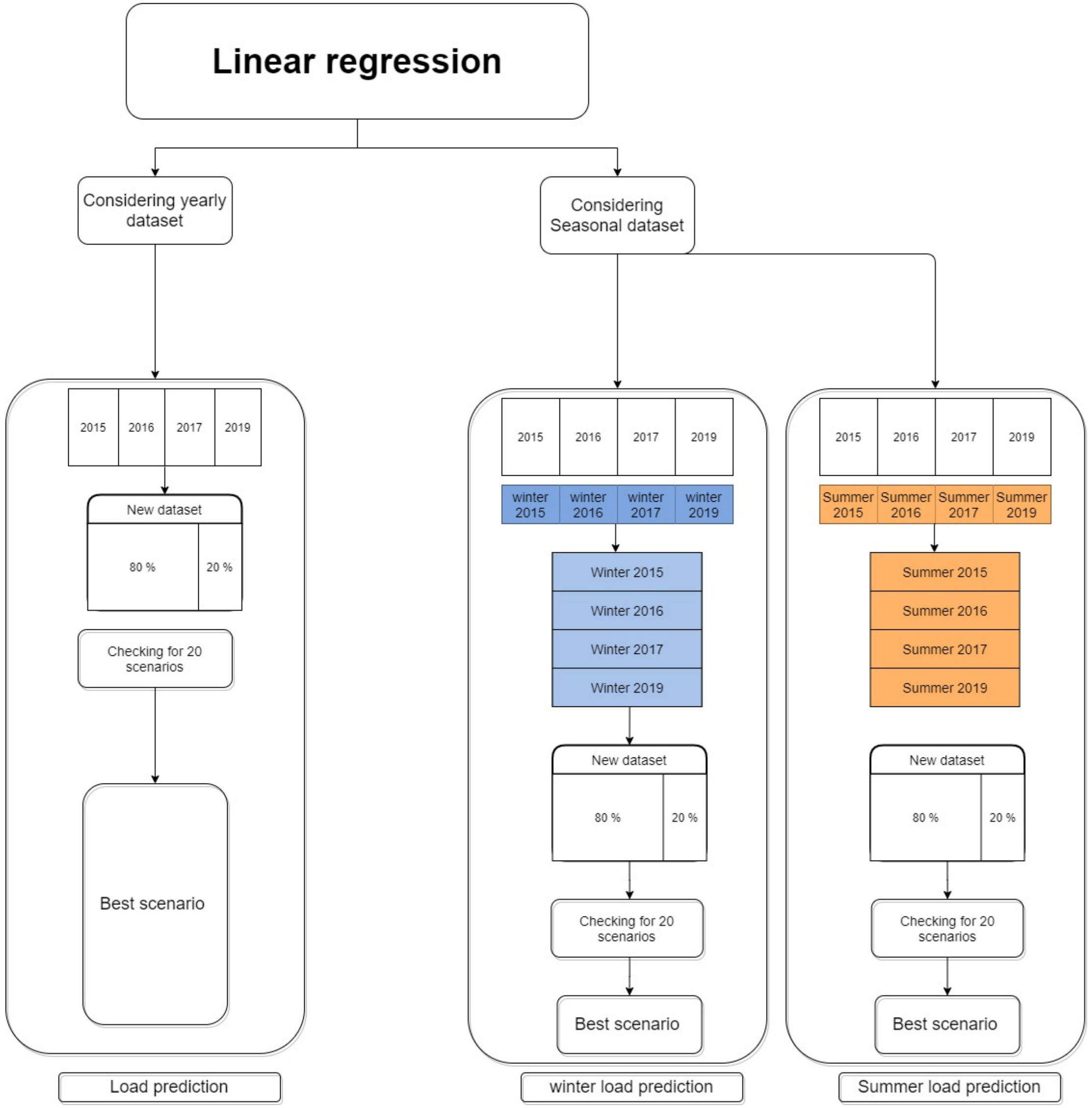


(a) (b)

https://www.researchgate.net/figure/Part-a-Represents-linear-regression-on-a-one-dimensional-data-17-Part-b_fig2_320609829

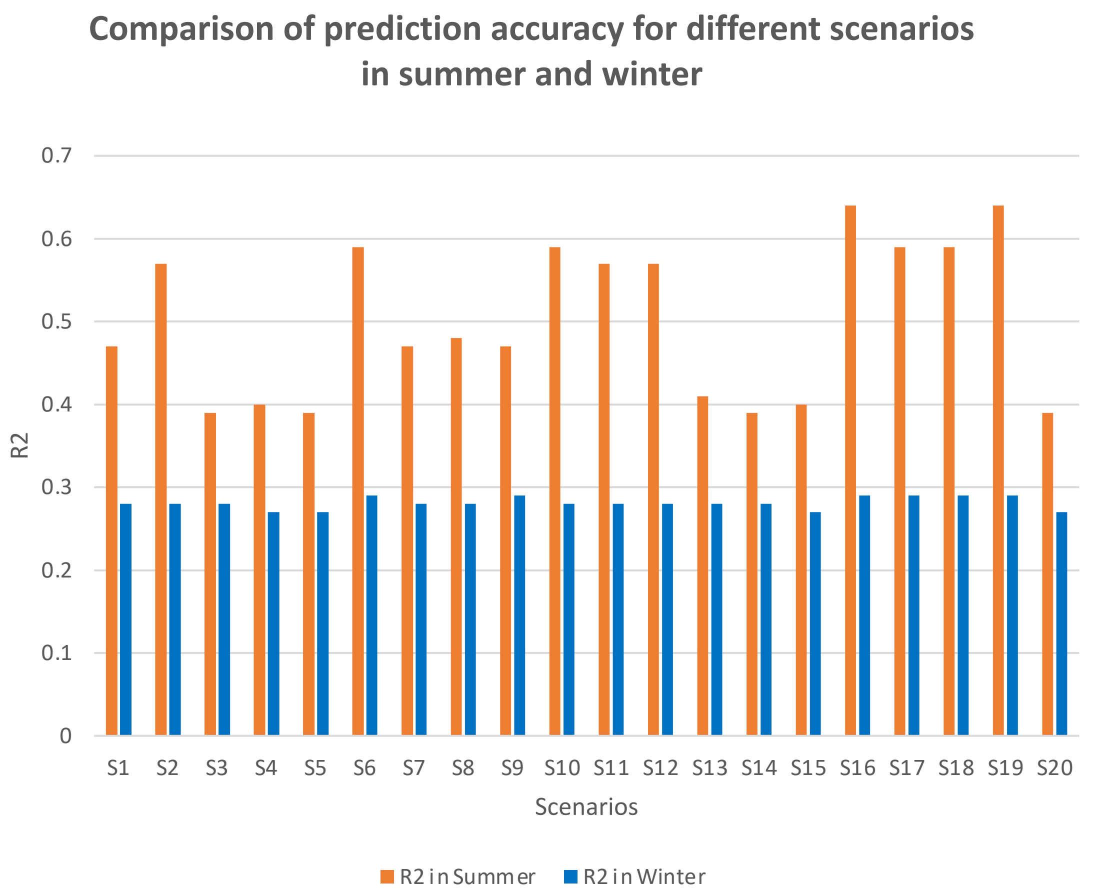
Scenarios

	Calendar data	Solar Radiation	Temperature	Relative humidity	Wind direction	Wind velocity
S1	*	*				
S2	*		*			
S3	*			*		
S4	*				*	
S5	*					*
S6	*	*	*			
S7	*	*		*		
S8	*	*			*	
S9	*	*				*
S10	*		*	*		
S11	*		*		*	
S12	*		*			*
S13	*			*	*	
S14	*			*		*
S15	*				*	*
S16	*	*	*	*		
S17	*	*	*		*	
S18	*	*	*			*
S19	*	*	*	*	*	*
S20	*					

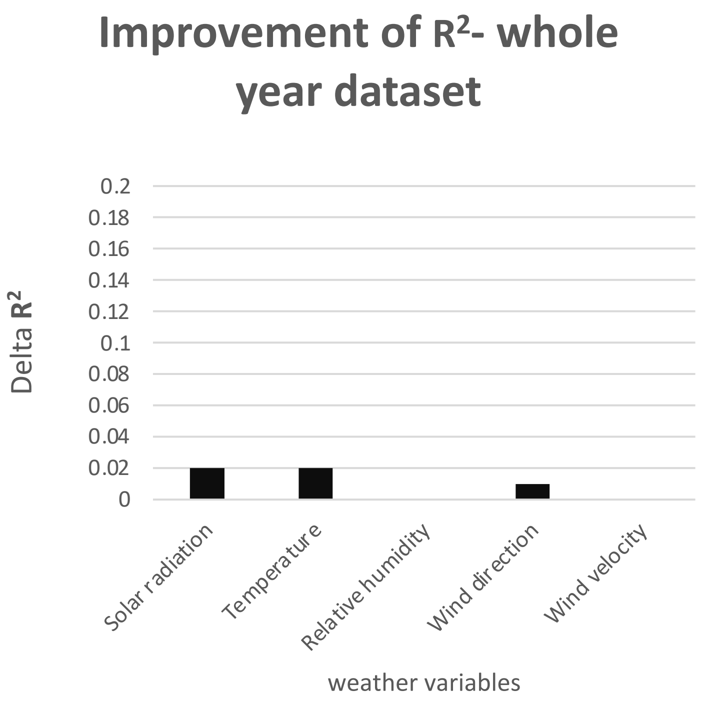


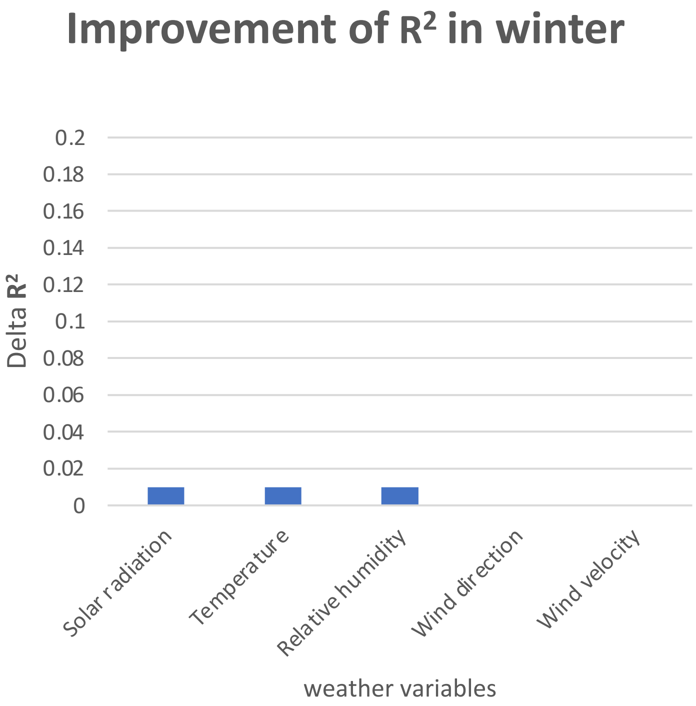
Comparison of different scenarios for summer and winter prediction

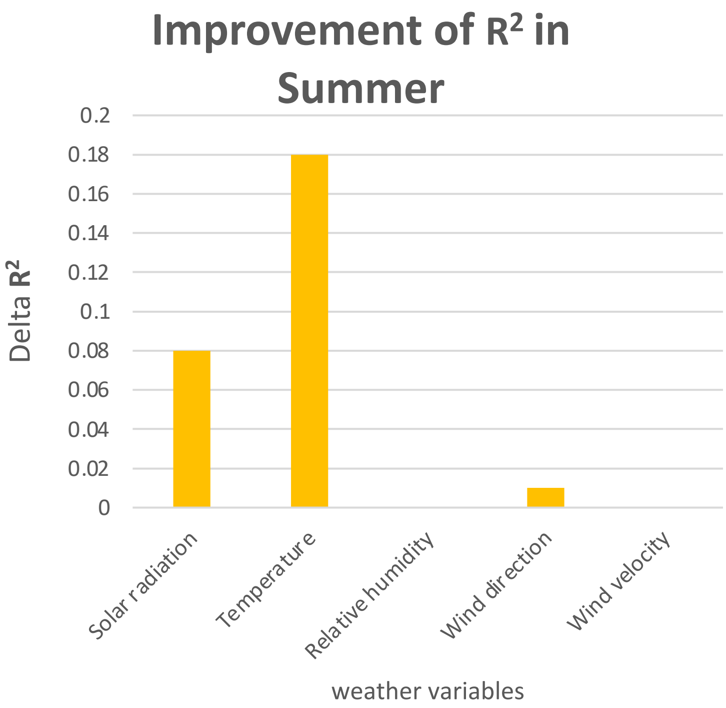
Linear regression scenarios	Summer		Winter	
	R^2 in Summer	MAPE	R^2 in Winter	MAPE
S1	0.47	12.74	0.28	24.52
S2	0.57	11.36	0.28	24.62
S3	0.39	13.58	0.28	24.64
S4	0.4	13.48	0.27	24.68
S5	0.39	13.59	0.27	24.63
S6	0.59	10.94	0.29	24.5
S7	0.47	12.74	0.28	24.51
S8	0.48	12.67	0.28	24.51
S9	0.47	12.74	0.29	24.47
S10	0.59	11.14	0.28	24.61
S11	0.57	11.37	0.28	24.62
S12	0.57	11.35	0.28	24.58
S13	0.41	13.46	0.28	24.64
S14	0.39	13.57	0.28	24.59
S15	0.4	13.48	0.27	24.63
S16	0.64	10.46	0.29	24.5
S17	0.59	10.94	0.29	24.5
S18	0.59	10.94	0.29	24.46
S19	**0.64**	**10.47**	**0.29**	**24.46**
S20	0.39	13.59	0.27	24.68



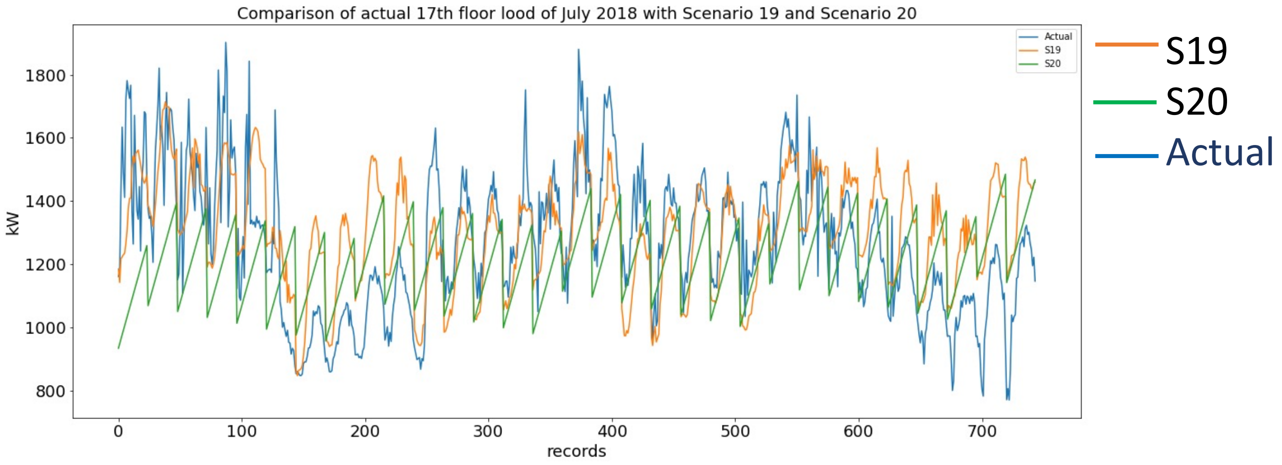
Effect of weather parameters on load prediction





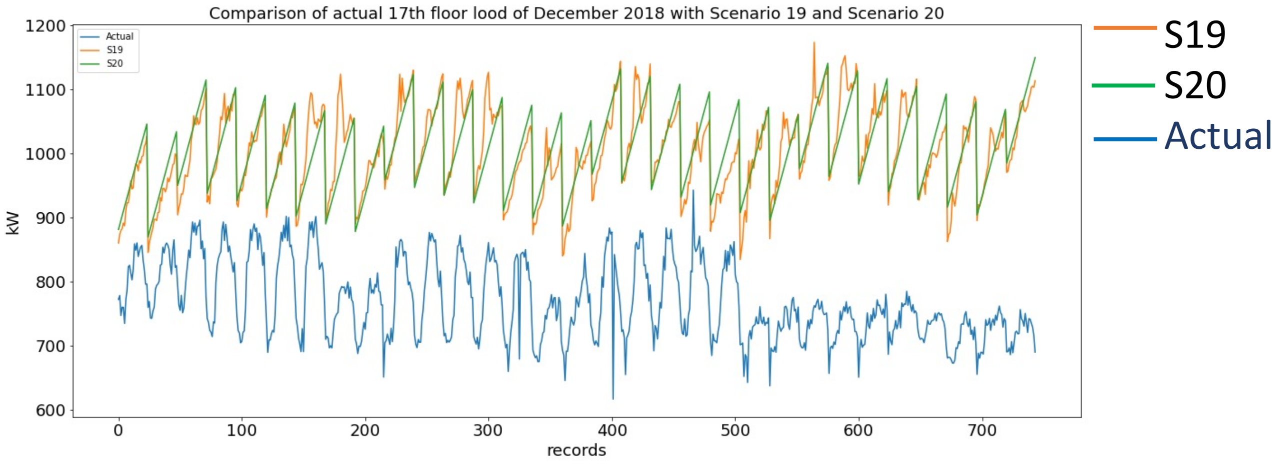


Actual vs predicated load by linear regression



Comparison of July 2018 with Linear regression:

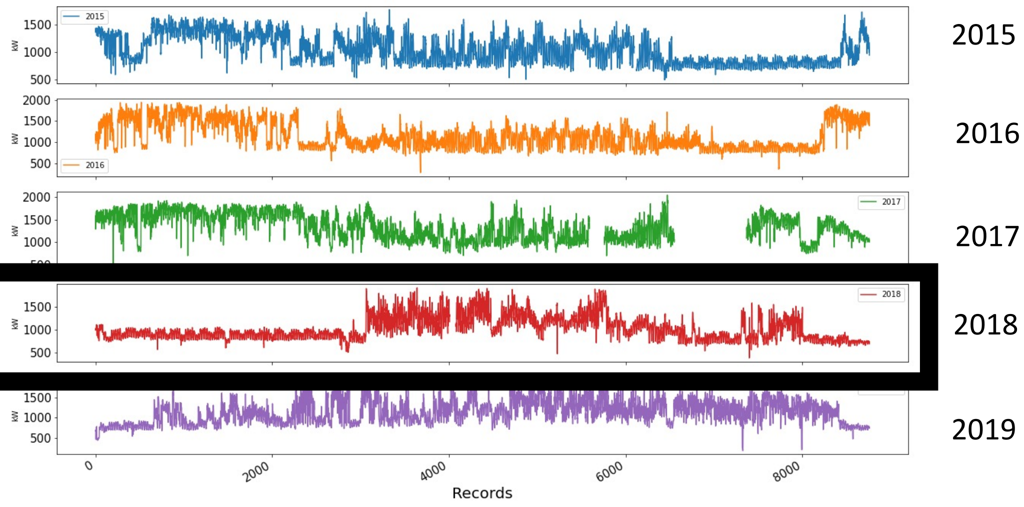
- The model is trained based on summer months of years (2015, 2016 and 2017)



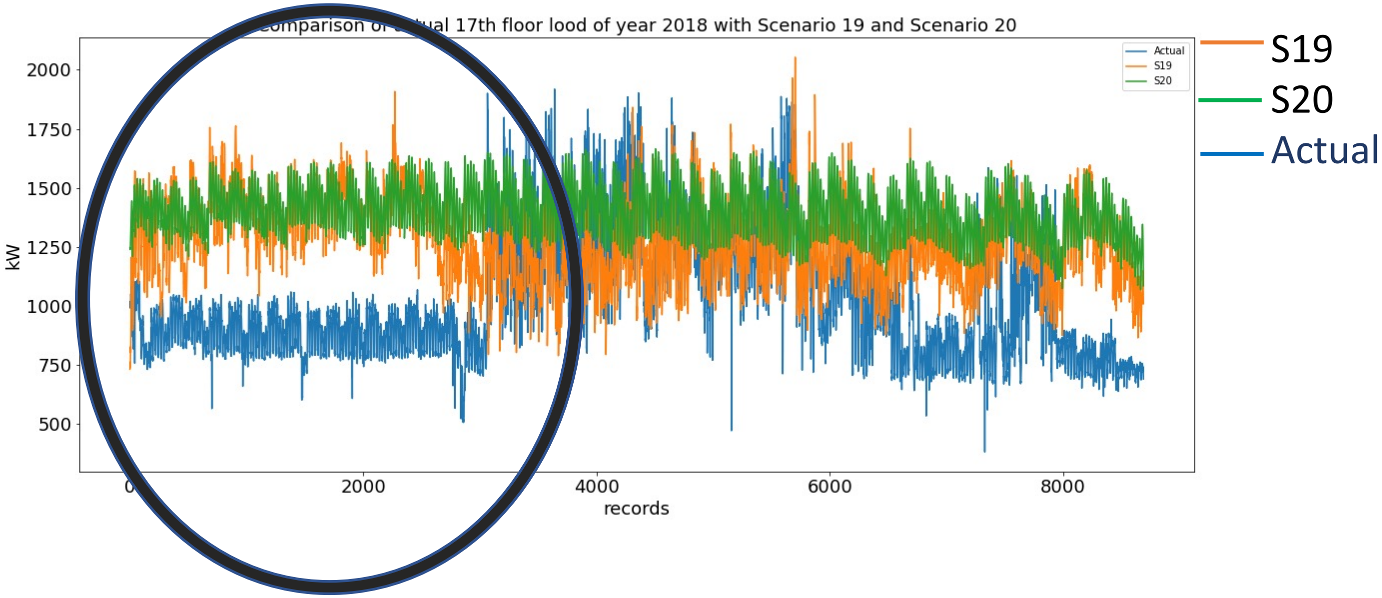
Comparison of December 2018 with Linear regression:

- The model is trained based on winter months of years (2015,2016 and 2017)

Load from 17th floor (2015 – 2019)

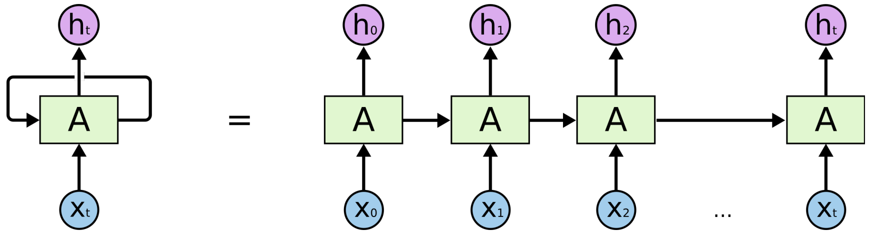


Performance of polynomial regression on 2018

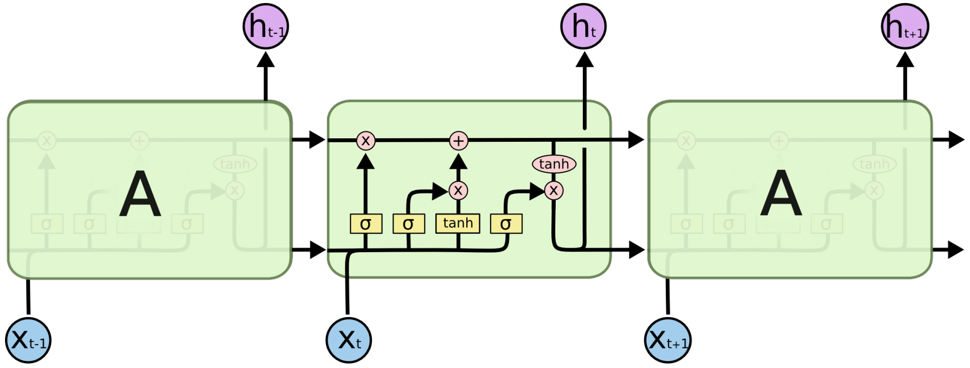


LSTM

RNN



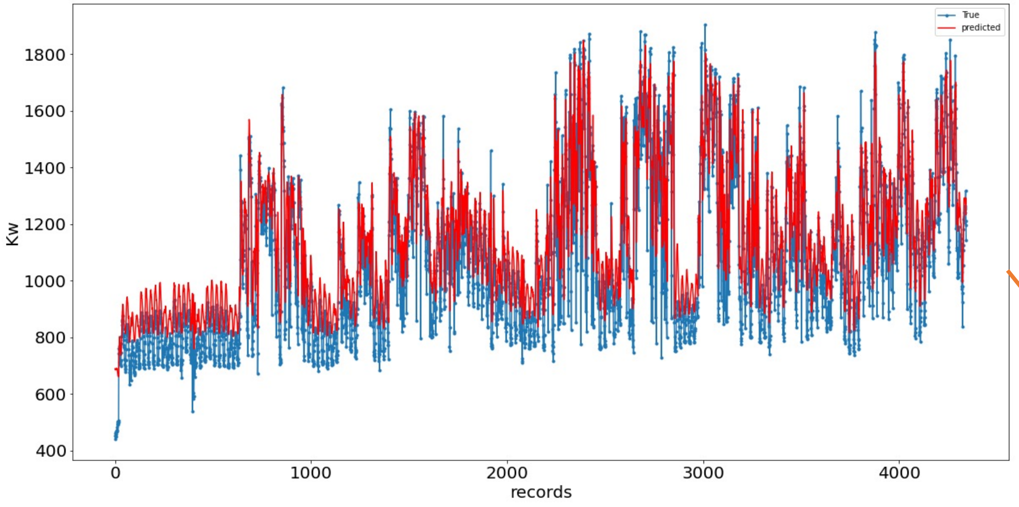
LSTM

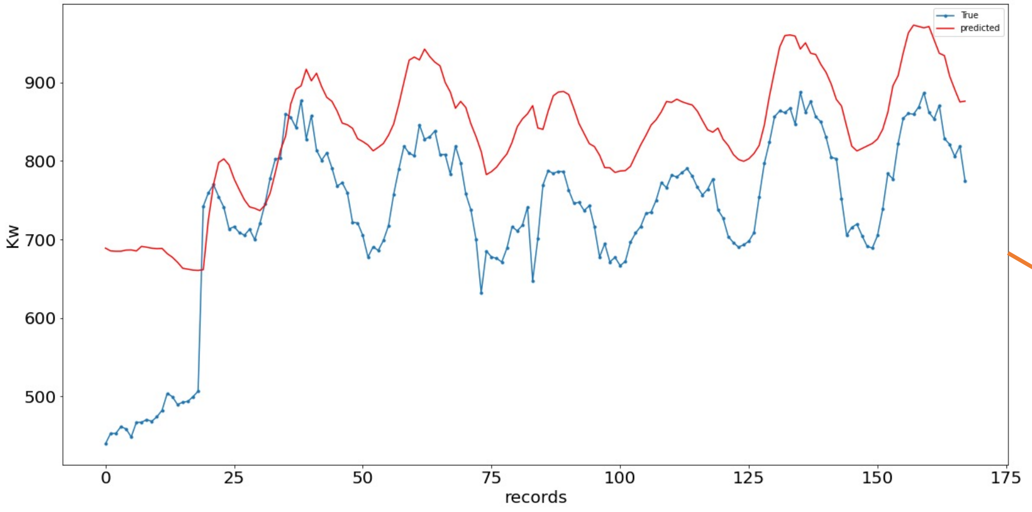


Proposed LSTM hyperparameters

LSTM model properties	
Number of hidden layer	2
Neurons in each layer	50
Window size	24
Learning rate	0.0001
Activation function	relu
Optimizer	adam
epoch	30
Batch size	64
Shuffle	False

LSTM performance for different time horizons





2015-2016-2017 | 2019

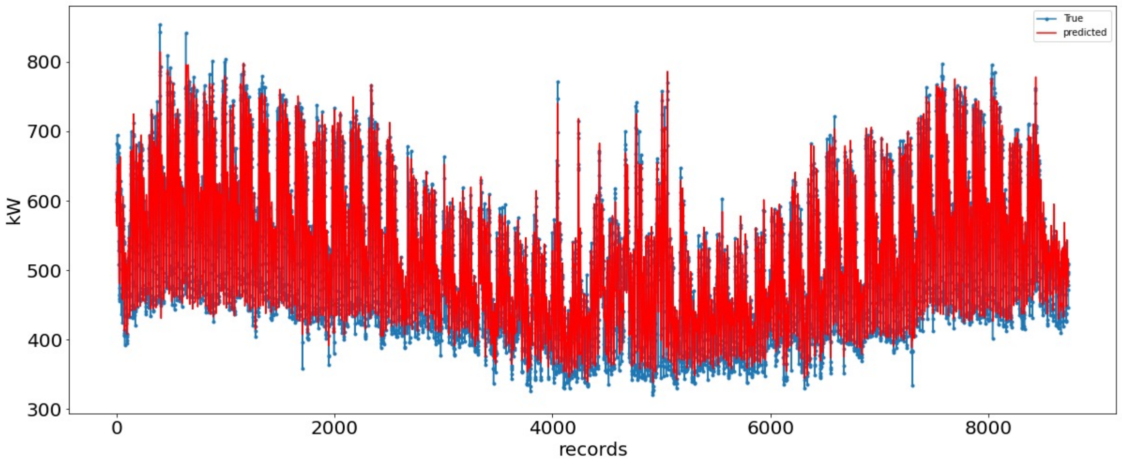
2015-2016-2017 | 2019

2015-2016-2017 | 2019

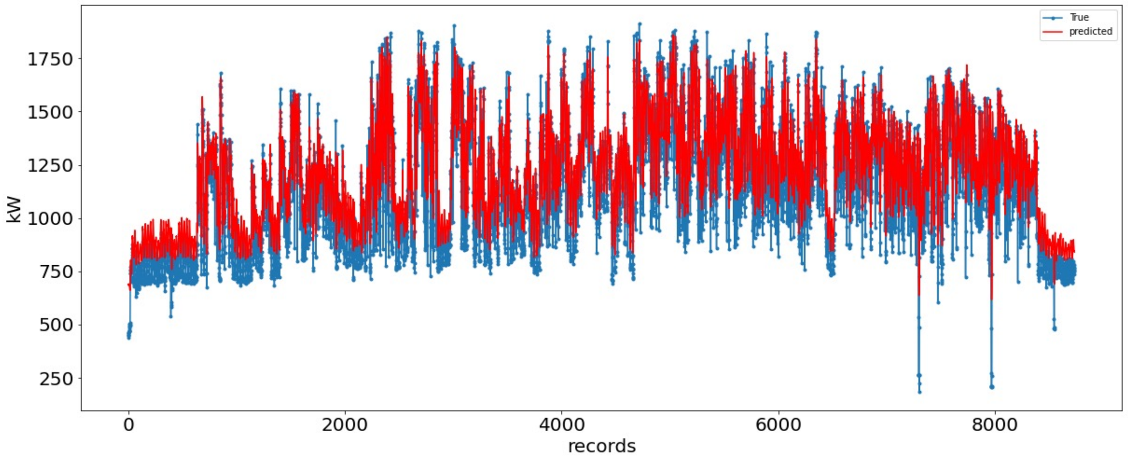
Time horizons	R2	MAPE
Full year	0.75	10.97
6 -month ahead	0.74	11.66
1 month ahead	0.53	13.5

Time horizons	R2	MAPE
2- week ahead	-0.35	14.33
1 -week ahead	-0.12	15.76
1 -day ahead	-1.64	35.64

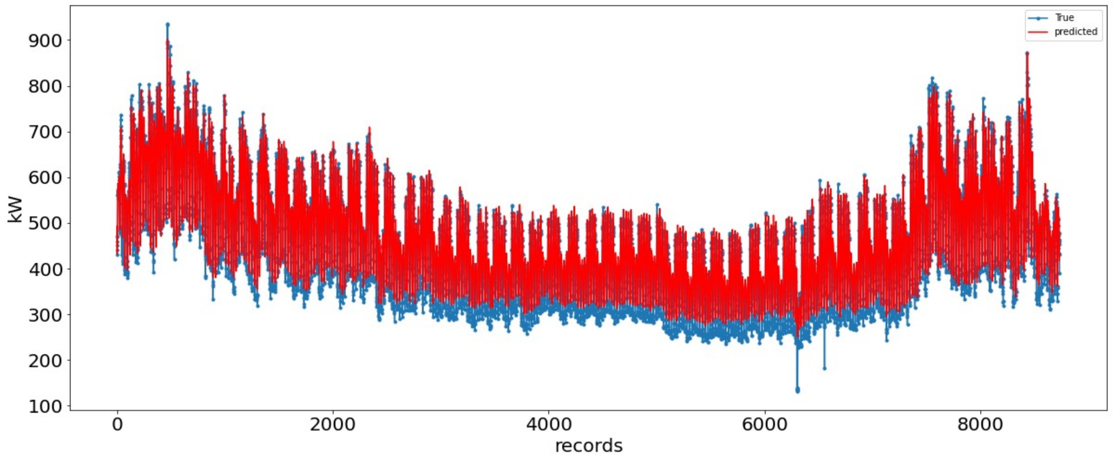
ENCS load forecasting with LSTM - 2019



HVAC load forecasting with LSTM - 2019



VA load forecasting with LSTM - 2019



Comparison of performance for three transformers

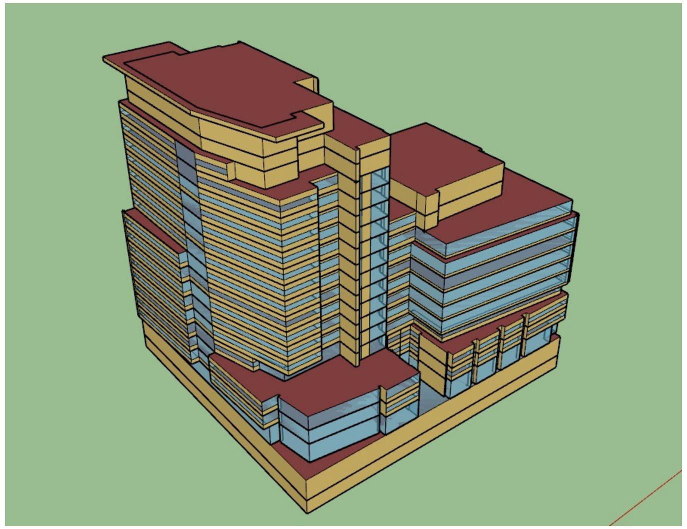
Load types	R^2	MAPE	mse
ENCS	0.92	4.35	818.3
VA	0.93	6.87	1146.4
17^{th} floor load	0.75	10.97	19812.66

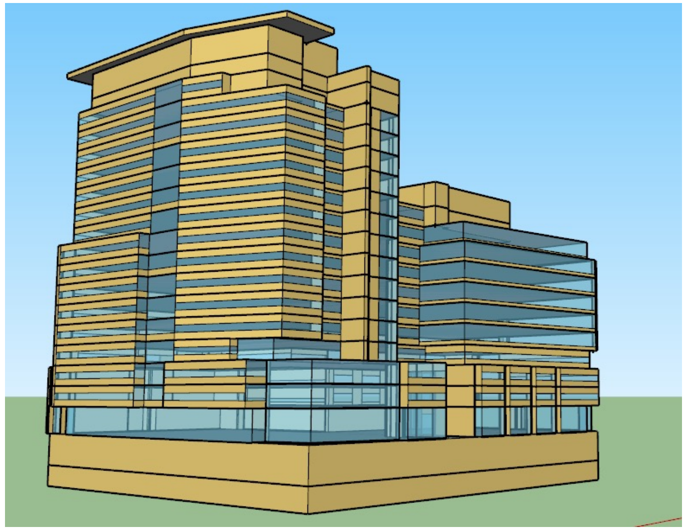
Conclusion

- The pattern of HVAC load and plug loads was extracted and studied. Reflecting schedule settings and behavioral patterns
- The reduction of load during COVID 19 was 42% for HVAC load in April (weekday)
- Temperature and solar radiation are two most affecting weather factors in summer responsible for 18 % and 8% improvement of accuracy.
- The best predictive model is the one considering calendar data with all weather columns. In summer S19 provided 64% accuracy and in winter 29%.
- The prediction model was able to capture unusual consumption.
- The performance of LSTM on ENCS and VA load was higher comparing to HVAC load. More that 90% for ENCS and VA , while HVAC load of 17^{th} floor got 75% accuracy.
- The accuracy of forecasting decreases as the test set size reduce from one year to one month .
- Negative R2 was referring to unusual load data in January.

Future works

- Integration of python codes as simulation blocks to CERC urban energy modeling workflow
- Using 3D Model as input for energy simulation software





Thank you for listening