

Hydro-Québec: Predicting the Hourly Ontario Energy Price in the Medium and Long Term

AUGUST 27, 2020

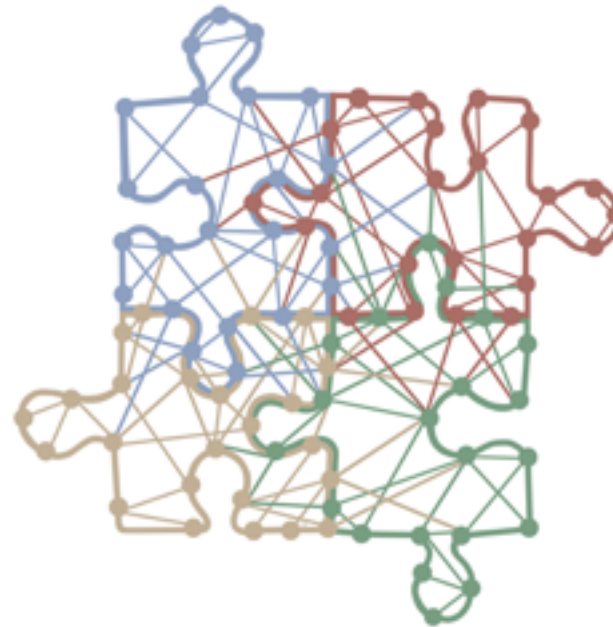
Outline

I. DATASET & GOAL



https://en.wikipedia.org/wiki/Electrical_grid

II. MACHINE LEARNING APPROACH



<https://behavioralscientist.org/scaling-nudges-machine-learning/>

III. MODELING APPROACH



<https://www.dataversity.net/data-modeling-in-the-machine-learning-era/>

Team members

SUPERVISORS: DR. HUAXIONG HUANG

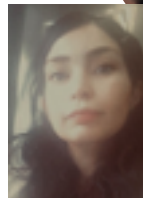
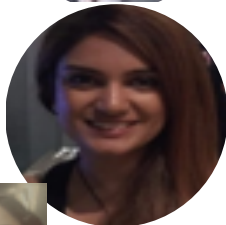


DR. YI YANG



MODELING GROUP

- *Abdoul Haki Maoude*
- *Andrew Day*
- *Ismael Assani*
- *Jingjing Zhang*
- *Maxence Premont*
- *Gita Gonoody*
- *Mozhgan Saeidi*
- *Qi Guo*



MACHINE LEARNING GROUP

- *Pan Liu*
- *Arka Mukherjee*
- *Cédric Poutré*
- *Rémi Galarneau-Vincent*
- *Prabodh Wankhede*
- *Soheila Samiee*
- *Xingwei Yang*

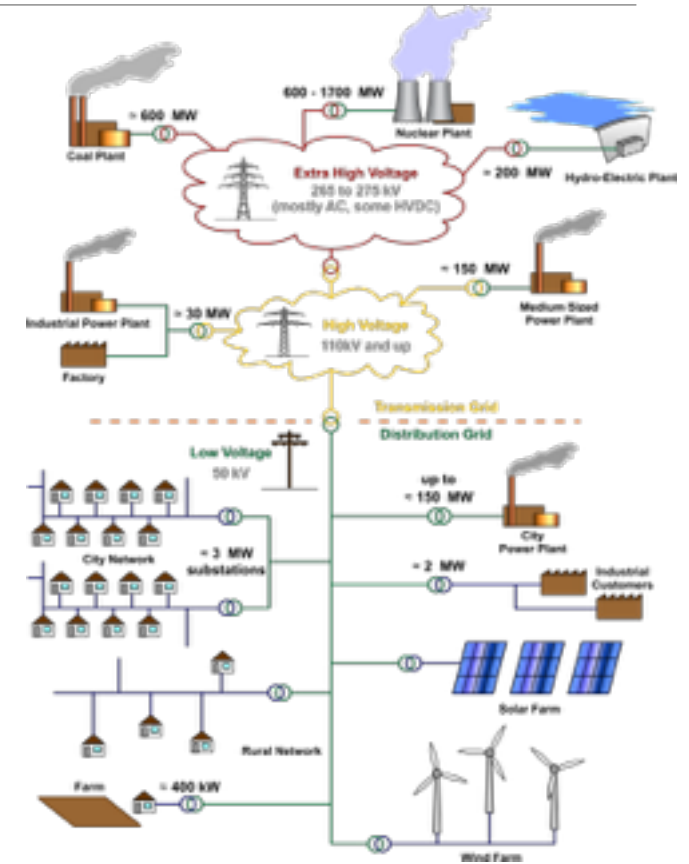


I. Dataset and Goal

I. Dataset and Goal

Goal

- Prediction of the Ontario Energy Price
 - Medium and long-term Periods (18 months)
 - For sales planning
- Ontario Market:
 - A Difficult market to predict:
 - Many fixed price supply contracts
 - 12 % coming from wind-based resources → intermittent
 - A lot of uncertainty in demand



https://en.wikipedia.org/wiki/Electrical_grid

I. Dataset and Goal

Dataset

Available Data sets:

- Predicted weekly data (18 month predictions): 2015 - 2020
- Historical hourly data: 2017 - 2020

	HOEP	Bruce PD	East PD	Essa PD	Niagara PD	NorthEast PD	NorthWest PD	Ottawa PD	SouthWest PD	Toronto PD	...	Expected Hydro Output
count	1473.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	...	1737.000000
mean	16.592956	86.197850	1302.368587	1197.743535	587.803402	1275.523675	471.597209	1228.498074	4012.024750	7388.431608	...	2409.75806
std	9.340326	22.155787	160.801969	167.620619	82.616598	177.474115	74.234597	180.238151	347.982958	843.344809	...	371.01938
min	-1.624762	47.730869	570.300757	762.932920	385.170000	856.255972	279.708361	872.306650	3132.000000	5652.000000	...	0.000000

Goal: Prediction of price from predicted parameters in weekly data

❖ Test:

- Three 18-months prediction files

II. Machine Learning Approach

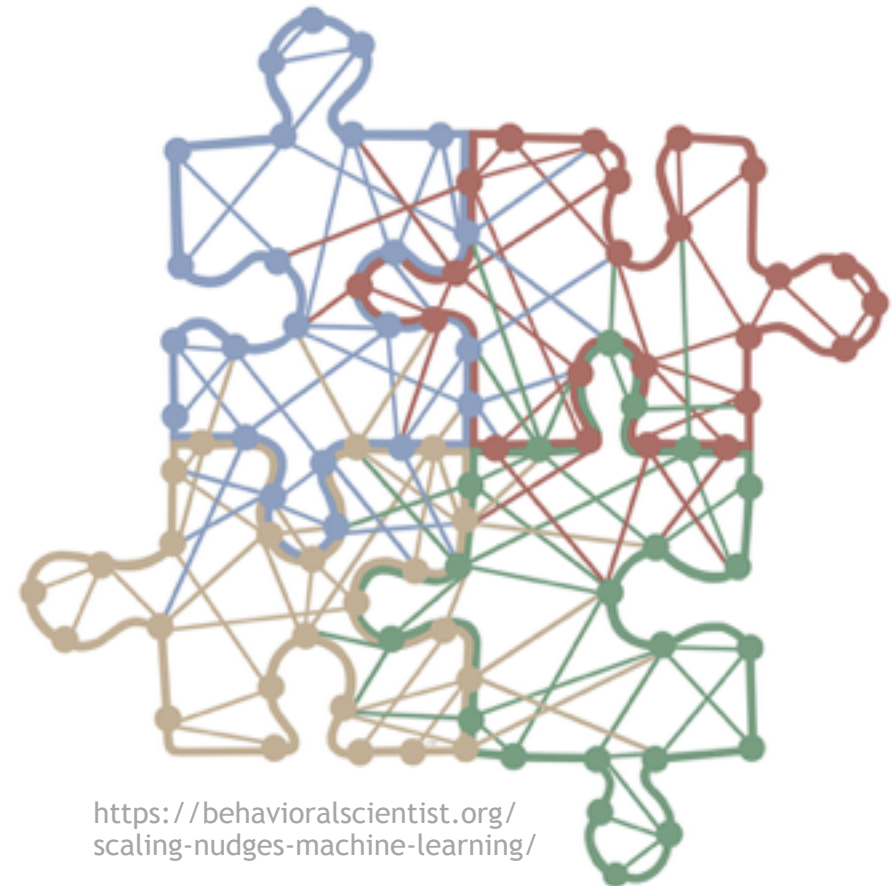
- Classical Machine Learning
- Deep Learning

II. Machine Learning Approach

General overview

➤ Methods:

1. Classical machine learning (CML)
2. Deep learning

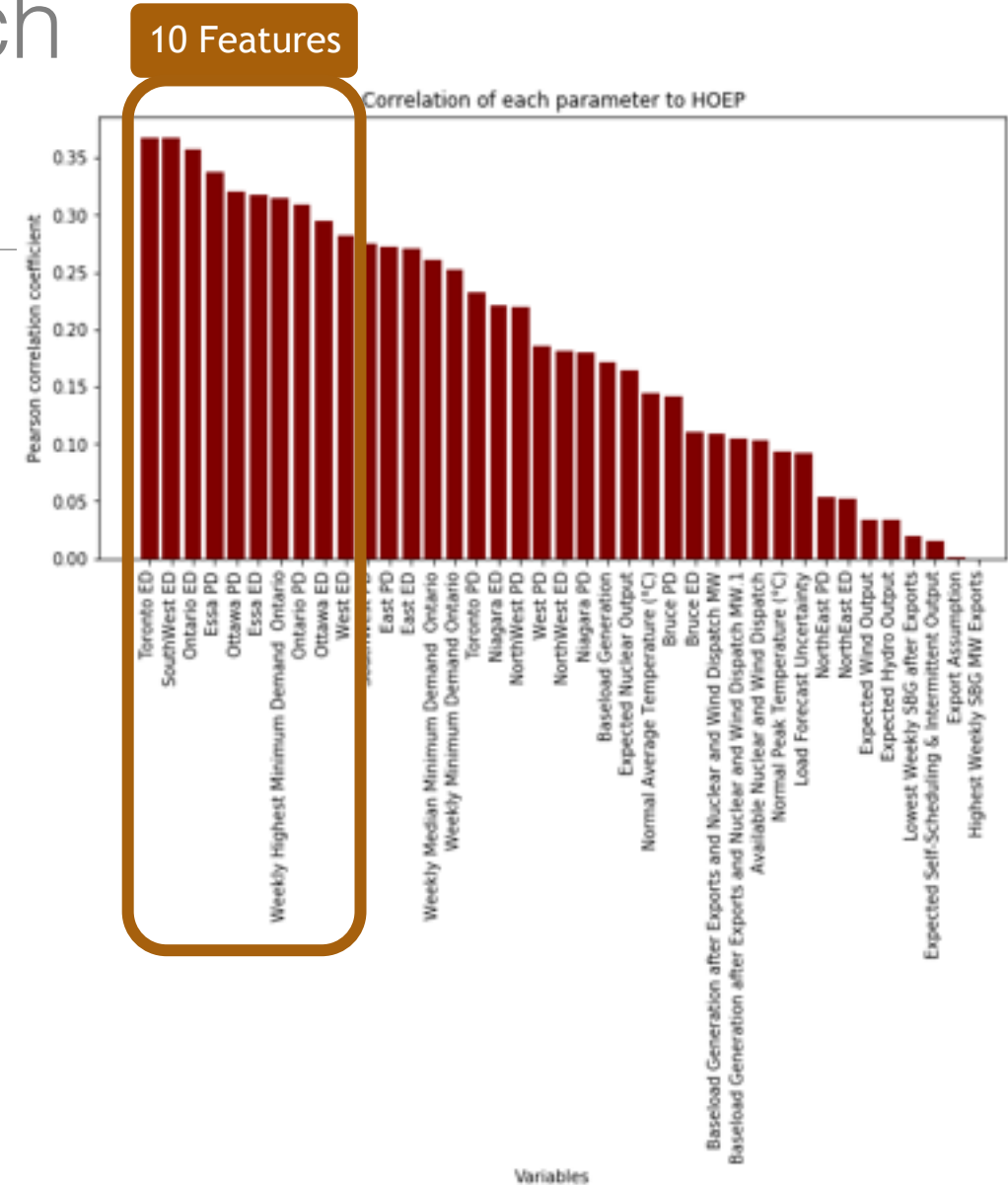


<https://behavioralscientist.org/scaling-nudges-machine-learning/>

II. Machine Learning Approach

CML Method: Features

- Features to be used
 1. Pearson Correlation

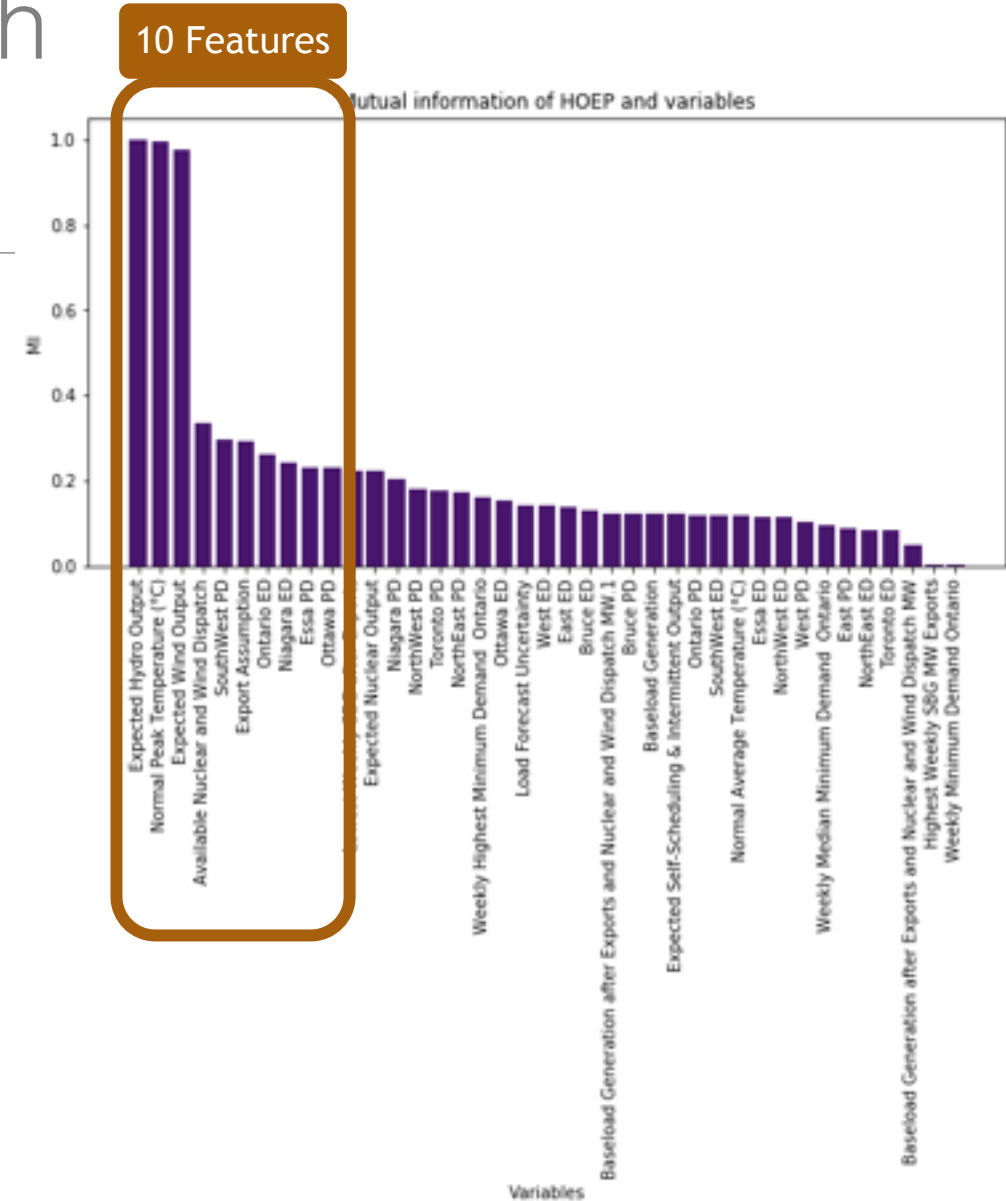


II. Machine Learning Approach

CML Method: Features

➤ Features to be used

1. Pearson Correlation
2. Mutual information



II. Machine Learning Approach

CML Method: Features

➤ Features to be used

1. Pearson Correlation
2. Mutual information
3. Non-zero coefficients from Lasso Regression (all survived feature with $\alpha = 0.4$)

'Bruce PD' 'NorthEast PD' 'NorthWest PD' 'Load Forecast Uncertainty' 'Essa ED'
'NorthEast ED' 'SouthWest ED' 'Toronto ED' 'Baseload Generation after Exports and
Nuclear and Wind Dispatch MW.1' 'Lowest Weekly SBG after Exports'

II. Machine Learning Approach

CML Method: Features

➤ Features to be used

1. Pearson Correlation
2. Mutual information
3. Non-zero coefficients from Lasso Regression

■ All features grouped together:

- 24 unique features:

'Toronto ED', 'SouthWest ED', 'Ontario ED', 'Essa PD', 'Ottawa PD', 'Essa ED', 'Weekly Highest Minimum Demand Ontario', 'Ontario PD', 'Ottawa ED', 'West ED', 'Expected Hydro Output', 'Normal Peak Temperature (°C)', 'Expected Wind Output', 'Available Nuclear and Wind Dispatch', 'SouthWest PD', 'Export Assumption', 'Niagara ED', 'Bruce PD', 'NorthEast PD', 'NorthWest PD', 'Load Forecast Uncertainty', 'NorthEast ED', 'Baseload Generation after Exports and Nuclear and Wind Dispatch MW.1', 'Lowest Weekly SBG after Exports'

II. Machine Learning Approach

CML Method: Regression

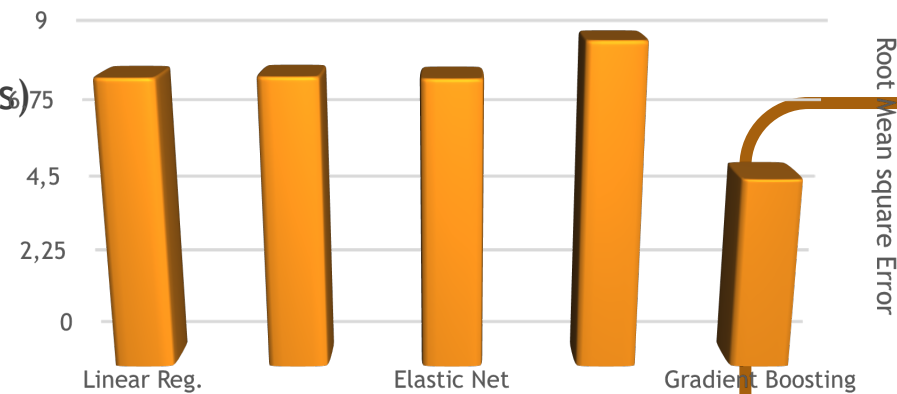
➤ Algorithms:

1. Linear Regression
2. Lasso
3. Elastic Net
4. Neural Network (MLP with two hidden layers)
5. Gradient Boosting

➤ Validation:

- 25% of Training data

Validation Error

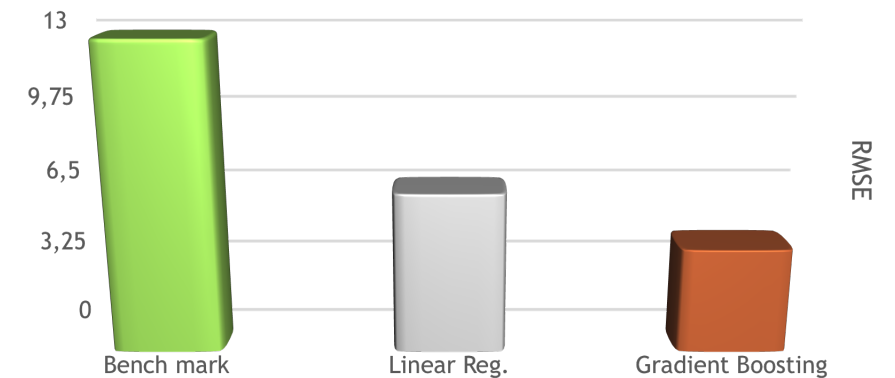
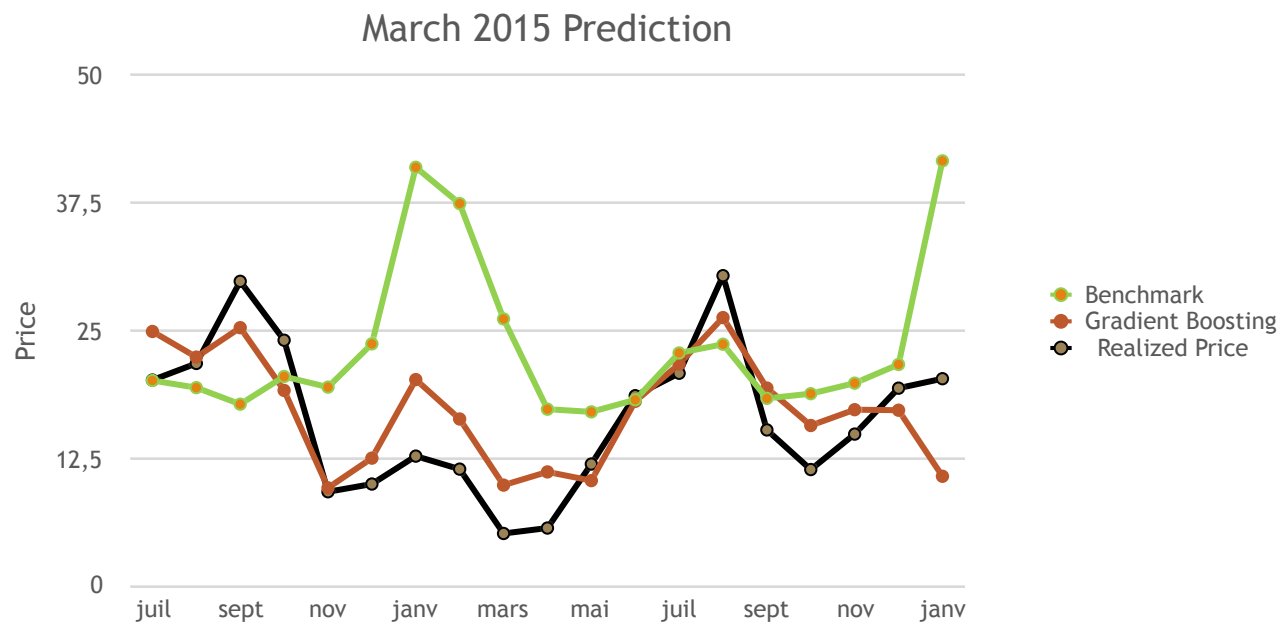


	Linear Reg.	Lasso	Elastic Net	Neural Net.	Gradient Boosting
rMSE (price)	7.95	7.98	7.93	8.91	5.15

II. Machine Learning Approach

CML Results

- Goal: predicting the **real price** better than the **benchmark (FWD HOEP)**
- Best approach based on validation data: **Gradient Boosting**



	Benchmark	Linear Reg.	Gradient Boosting
rMSE	12.85	6.71	4.38

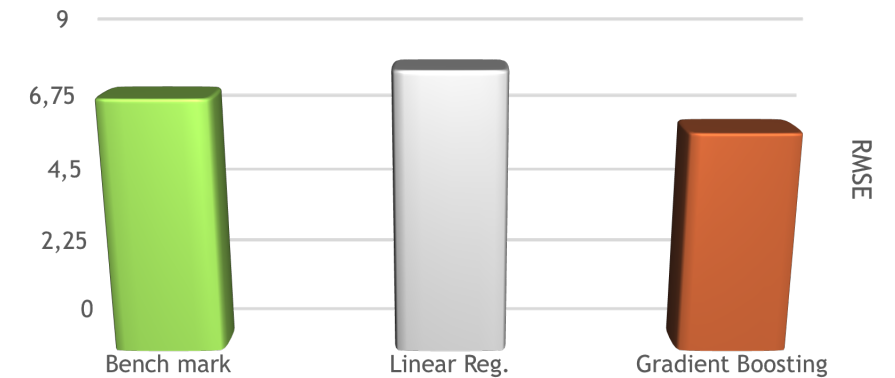
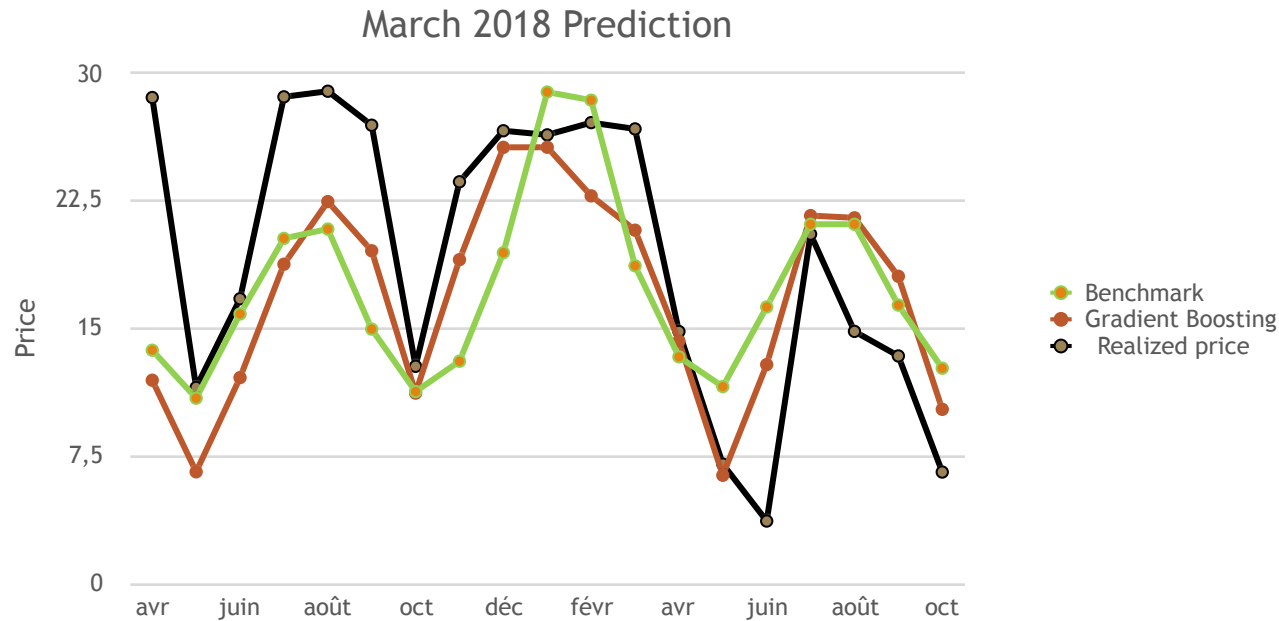
II. Machine Learning Approach

CML Results

➤ Goal: predicting the **real price** better than the **benchmark (FWD HOEP)**

➤ Best approach based on validation data: **Gradient Boosting**

Error in predicting real price (18 month period)

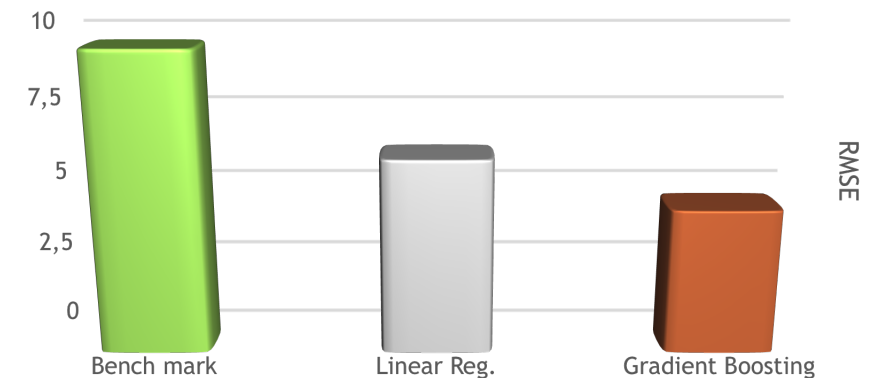
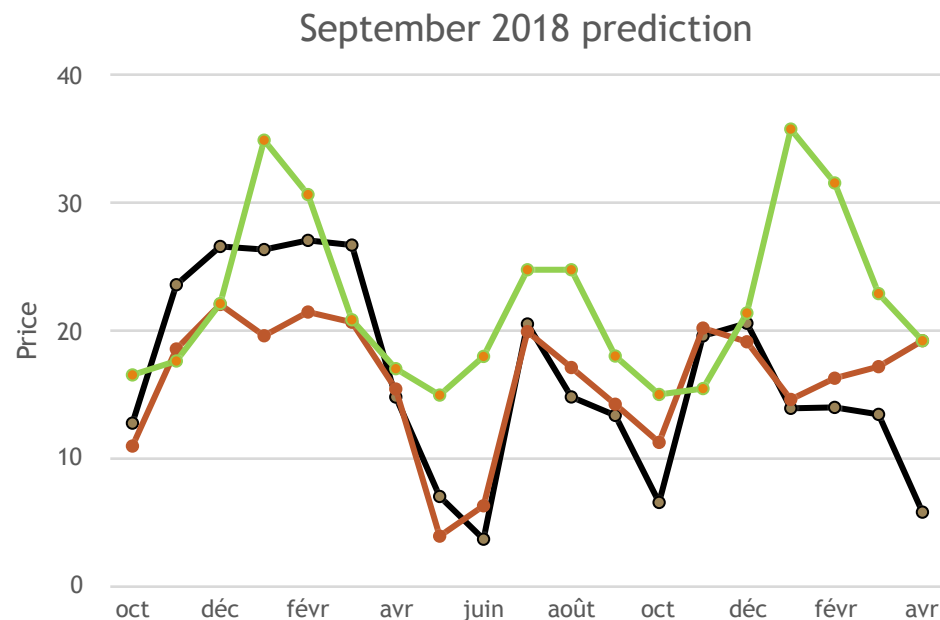


	Benchmark	Linear Reg.	Gradient Boosting
rMSE	7.27	8.04	6.33

II. Machine Learning Approach

CML Results

- Goal: predicting the **real price** better than the **benchmark (FWD HOEP)**
Error in predicting real price (18 month period)
- Best approach based on validation data: **Gradient Boosting**



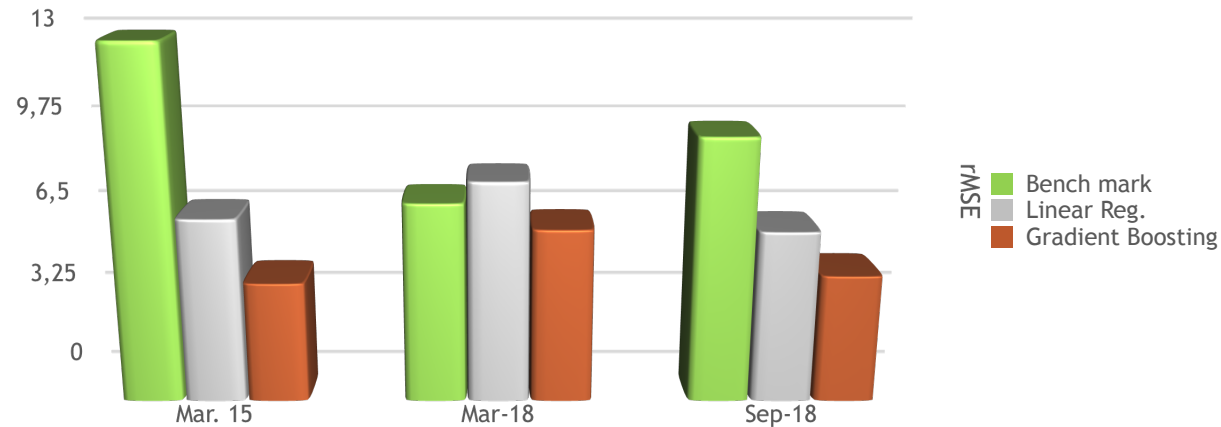
	Benchmark	Linear Reg.	Gradient Boosting
rMSE	9.60	6.25	4.65

II. Machine Learning Approach

CML Results

- Goal: predicting the **real price** better than the **benchmark (FWD HOEP)**

Summary of Error in Predicting Price



18-month Prediction file

	Benchmark	Linear Reg.	Gradient Boosting
Mar. 15	12.85	6.71	4.38
Mar. 18	7.27	8.04	6.33
Sep. 18	9.6	6.25	4.65

II. Machine Learning Approach

Data

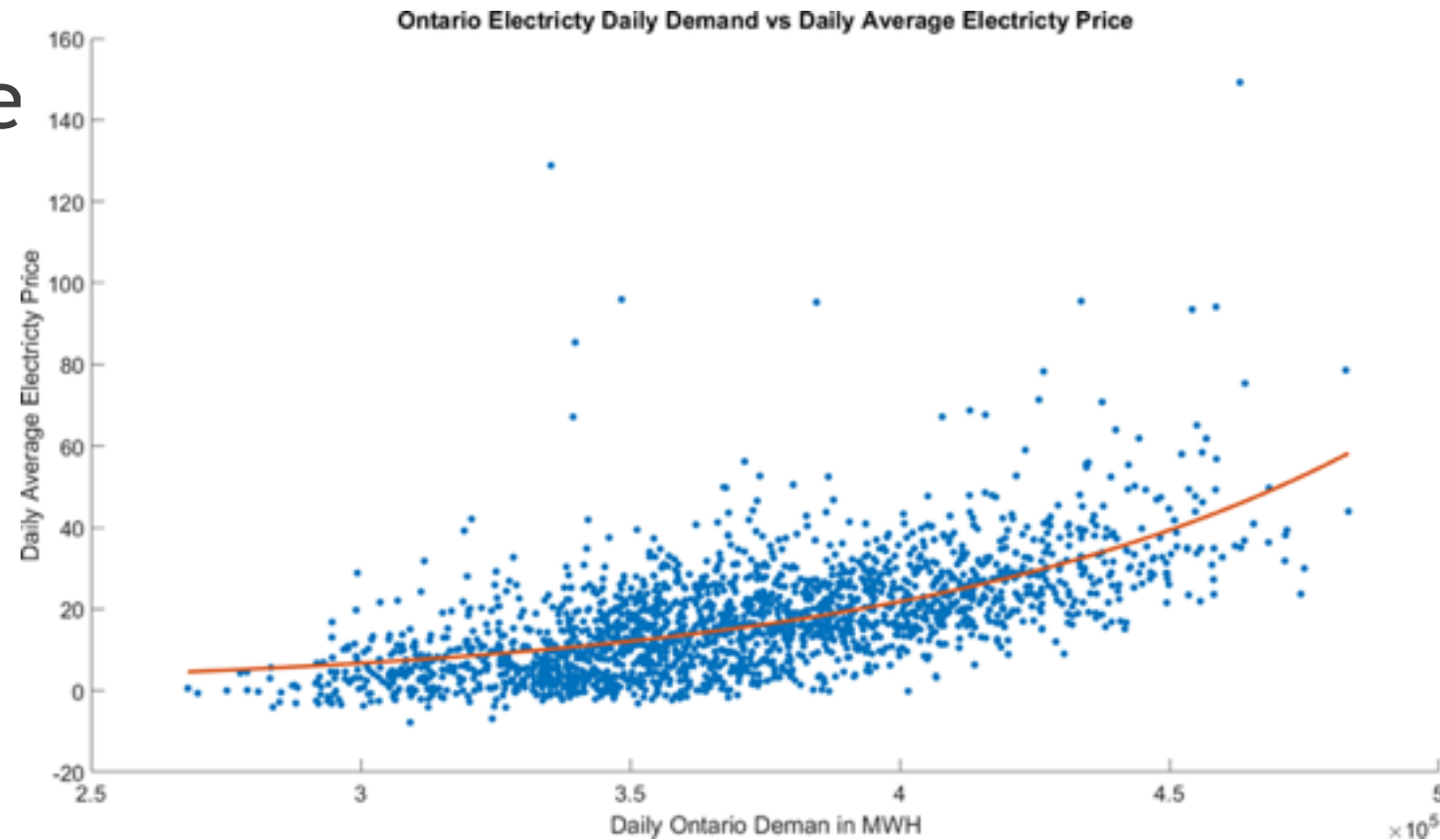
- We consider the IESO hourly realized dataset from 2015/01/01 to 2020/07/01 as well as the weekly outlook dataset provided by Hydro-Quebec.
- From the hourly dataset we aggregate the hourly realized features to generate a daily realized dataset. We also use hourly historical data to generate future hourly outlook for features.
- We use only a few fundamental features that we deem important behind the pricing process.

Simple NL model: $HOEP = f(\text{Demand})$

- To capture the relationship between the price and the Ontario demand we use

$$P_t = \exp(b_0 + b_1 \text{Demand}_t)$$

and minimize the MSE.



II. Machine Learning Approach

NL Methodology

- To forecast the electricity price we input inside the NLR model the forecasted demand found in the weekly dataset .
- Since this method is simple, we use the 2015 and 2016 sample as the training sample to estimate the parameters of the model. The rest of the sample (2017-2020) is the test sample.

HOEP RMSE for the Benchmark and the NL model

Realized vs. Prediction vs. Benchmark (2018-03-01 and 2018-09-01)

	Quartes	Benchmark	NLR Model
In-sampl e	2015-03-01	14.14	9.01
	2015-06-01	12.20	9.45
	2015-09-01	13.88	8.88
	2015-12-01	11.25	7.63
	2016-03-01	13.76	5.76
	2016-06-01	6.15	5.53
	2016-09-01	13.47	5.05
	2016-12-01	10.75	6.37
Out-of-sampl e	2017-03-01	8.35	7.67
	2017-06-01	6.54	7.43
	2017-09-01	6.74	7.84
	2017-12-01	7.18	8.34
	2018-03-01	7.33	8.49
	2018-06-01	6.66	8.16
	2018-09-01	9.34	5.55
	2018-12-01	14.51	5.18
	2019-03-01	14.16	4.82
	2019-06-01	11.14	5.21
	2019-09-01	11.38	5.24
	2019-12-01	12.00	6.23
	2020-03-01	3.99	3.49



II. Machine Learning Approach

General overview

➤ Methods:

1. Classical machine learning (CML)
2. **Deep learning (DL)**
 - Deep learning from hourly data
 - Deep learning from weekly data



<https://behavioralscientist.org/scaling-nudges-machine-learning/>

II. Machine Learning Approach

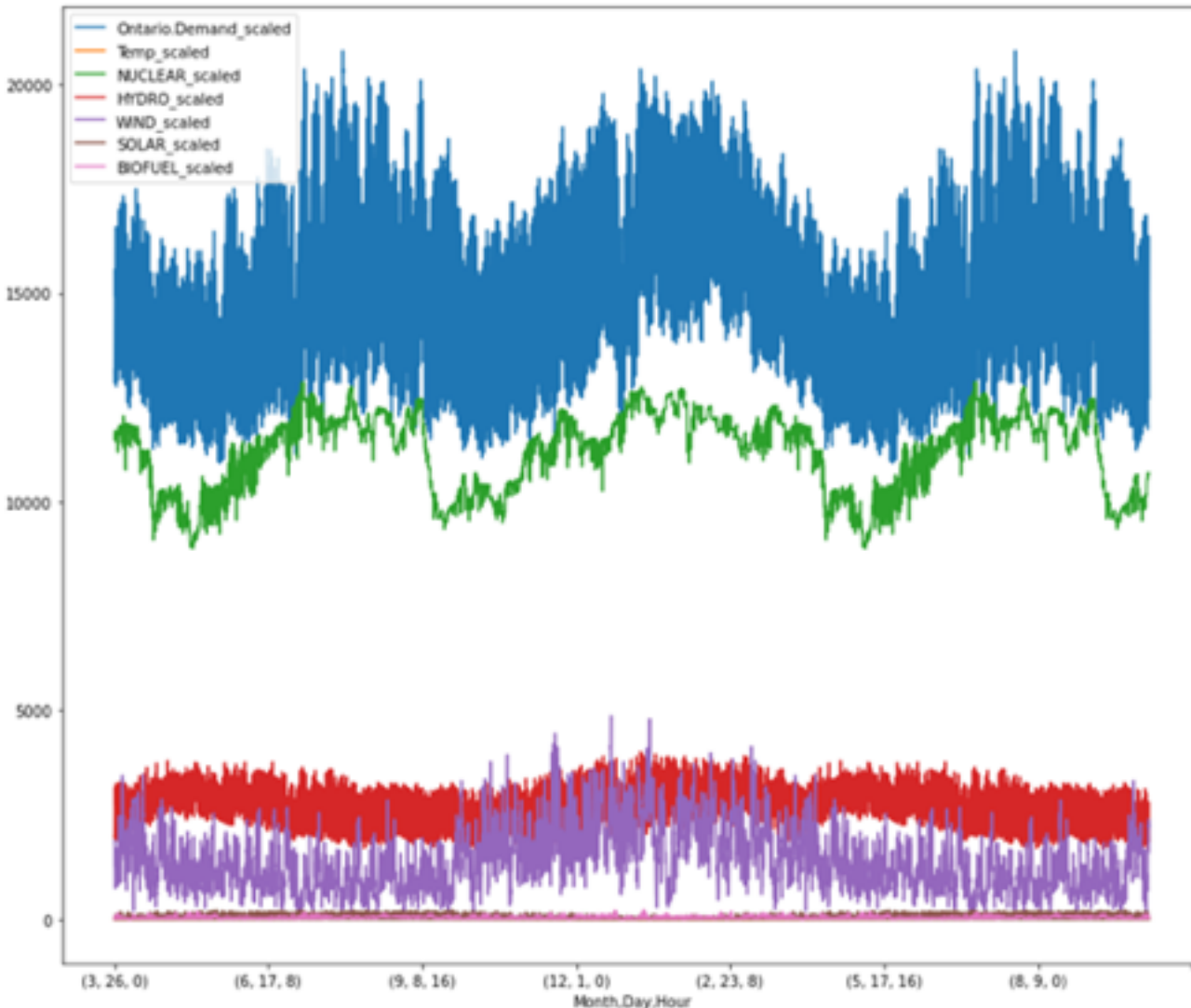
Deep Learning Models

- These were experiments to see if we could **extract valuable information from the realized hourly data** set, instead of the weekly predicted data set (less data).
- Models are trained to learn the pricing mechanism:

$$HOEP_{hour_T} = f(Features_{hour_T}, Features_{hour_T-1}, \dots)$$

- We have trained different deep learning models such as **MLP**, **RNN** (vanilla and stacked) and **LSTM** (vanilla and stacked) models.
- Hyperparameters were tuned using the Tree-structured Parzen Estimator Approach.
- The predictions for each scenario needed **hourly predicted features**, but the given data was on a weekly basis. How to predict hourly features?

Outlook-adjusted historical average - 03012018



Hourly Predicted Features

- 1. Naïve approach: repeating the weekly features for every hour in the week
 - 2. Time series approach: tried Vector Autoregression (VAR) model, etc. Not able to reasonably predict high frequency (hourly) data in a long term. Results abandoned.
 - 3. **Outlook-adjusted historical average:** using average realized hourly historical features adjusted on the Hydro's weekly outlook features.
- > For DL models: the **best** results were obtained with **LSTMs** with the **3rd hourly data generation approach**

DL Model Comparison

- We present two of the most important scenarios as determined by Hydro-Québec:

Model	RMSE
Benchmark	7.33
LSTM	7.45

Model	RMSE
Benchmark	9.34
LSTM	7.90



Realized vs. Benchmark vs. LSTM vs. MLP (2018-03-01 and 2018-09-01)

II. Machine Learning Approach

General overview

➤ Methods:

1. Classical machine learning (CML)
2. **Deep learning (DL)**
 - Deep learning from hourly data
 - **Deep learning from weekly data**



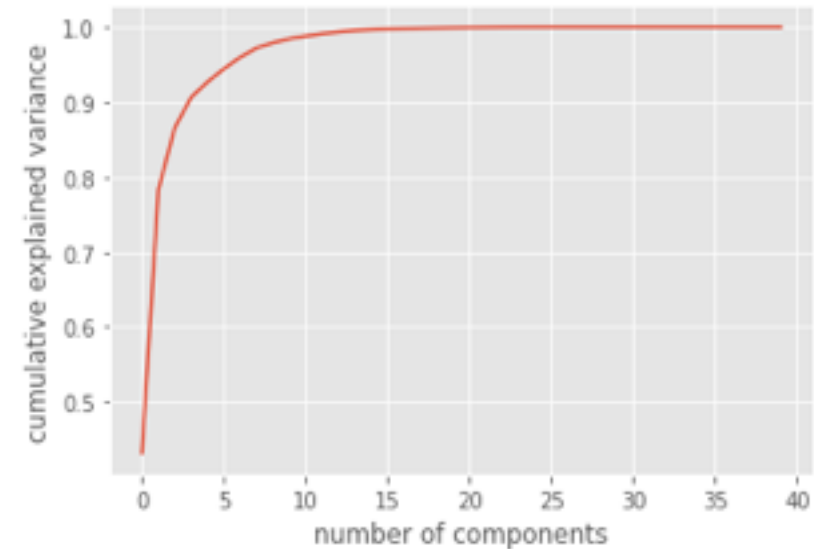
<https://behavioralscientist.org/scaling-nudges-machine-learning/>

II. Machine Learning Approach

DL Method (LSTM Model): Features

➤ Features to be used

1. Augmented Dickey Fuller Test - 9 features are non-stationary but there is weak evidence to reject null hypothesis. Hence, PCA can be applied.
2. Principal Component Analysis
 - 12 unique features that explain 99% of cumulative variance
['Ontario PD', 'Toronto PD', 'Expected Self-Scheduling & Intermittent Output', 'Expected Wind Output', 'Baseload Generation after Exports and Nuclear and Wind Dispatch MW', 'Export Assumption', 'Expected Hydro Output', 'Weekly Median Minimum Demand Ontario', 'Available Nuclear and Wind Dispatch', 'Lowest Weekly SBG after Exports', 'Weekly Highest Minimum Demand Ontario', 'Load Forecast Uncertainty']



II. Machine Learning Approach

DL Method (LSTM Model): Parameters

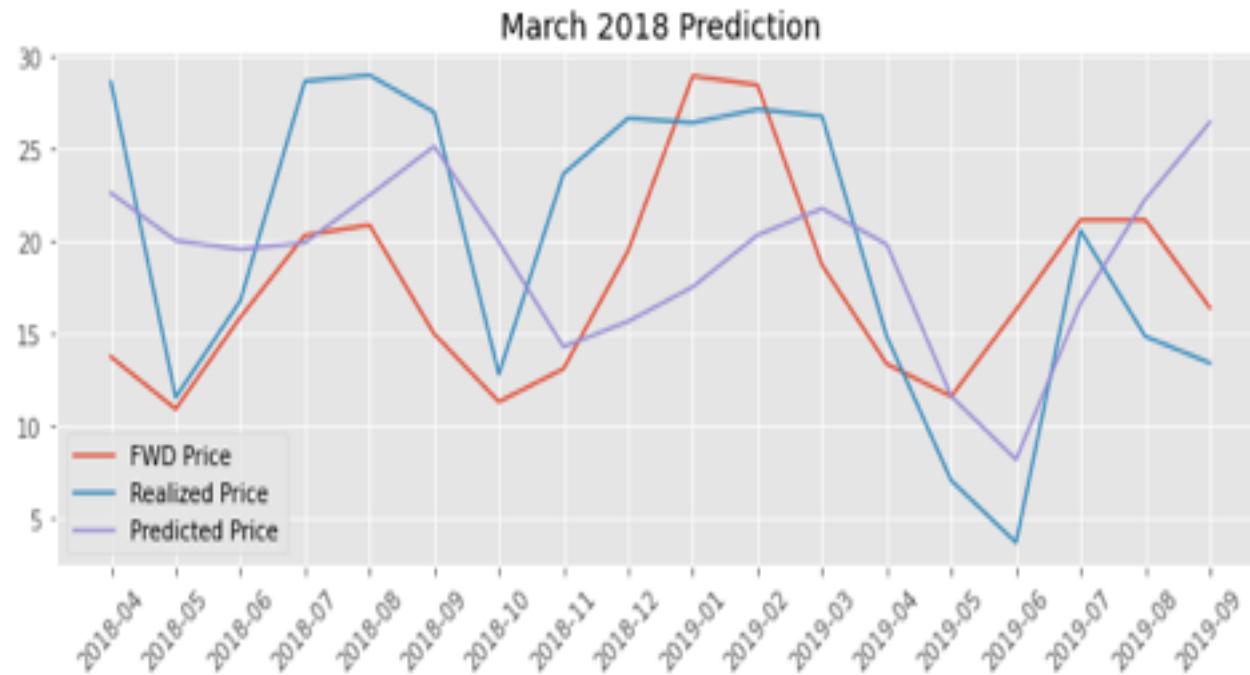
➤ Value of hyperparameters

1. Learning Rate - 0.01
2. L2 Regularization - 0.003
3. Input Sequence Length - 52
4. Stacked Layers - 2
5. Gradient Clipping Value - 2.5

II. Machine Learning Approach

DL Result: Stacked LSTM

- Goal: predicting the **real price** better than the **benchmark (FWD HOEP)**



	Benchmark	LSTM
RMSE	7.33	7.27

II. Machine Learning Approach

DL Result: Stacked LSTM

- Goal: predicting the **real price** better than the **benchmark (FWD HOEP)**



	Benchmark	LSTM
RMSE	9.34	7.49

III. Modeling Approach

Models

Results

III. Modeling Approach

General overview

➤ Method:

1. **Daily Historical Regression**
2. Weekly Regression
3. Time series method



<https://www.dataversity.net/data-modeling-in-the-machine-learning-era/>

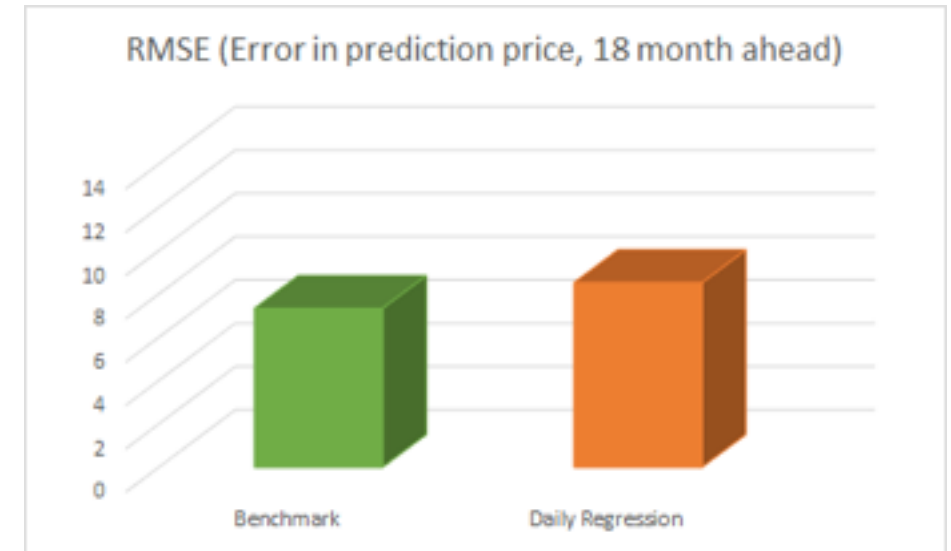
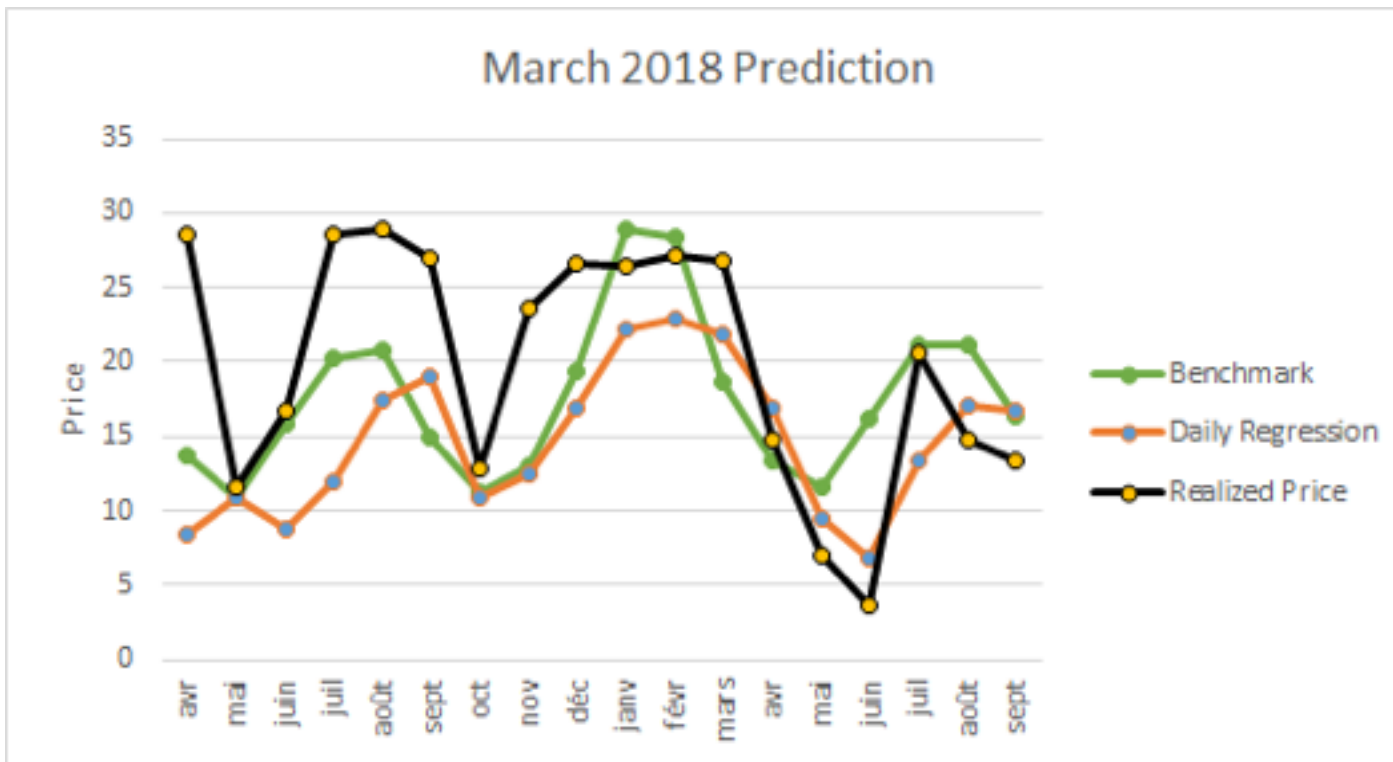
III. Modeling Approach

Daily Historical Regression

- Model $P_t = \beta_0 + \beta_1 Temp_t + \beta_2 Demand_t + \beta_3 WH_t + \beta_4 Saison_t$
- P_t is the HOEP at day t
 - $Temp_t$ is the *temperature* at day t
 - $Demand_t$ is Demand at day t
 - WH_t is whether the day t is a weekend or holiday
 - $Saison_t$ is whether the day t is a specific saison "winter", "spring", "Autumn" or "Summer".
-
- *Daily data for training the model*
 - *Weekly prevision only need weekly data for $Temp_t$ and $Demand_t$*

III. Modeling Approach

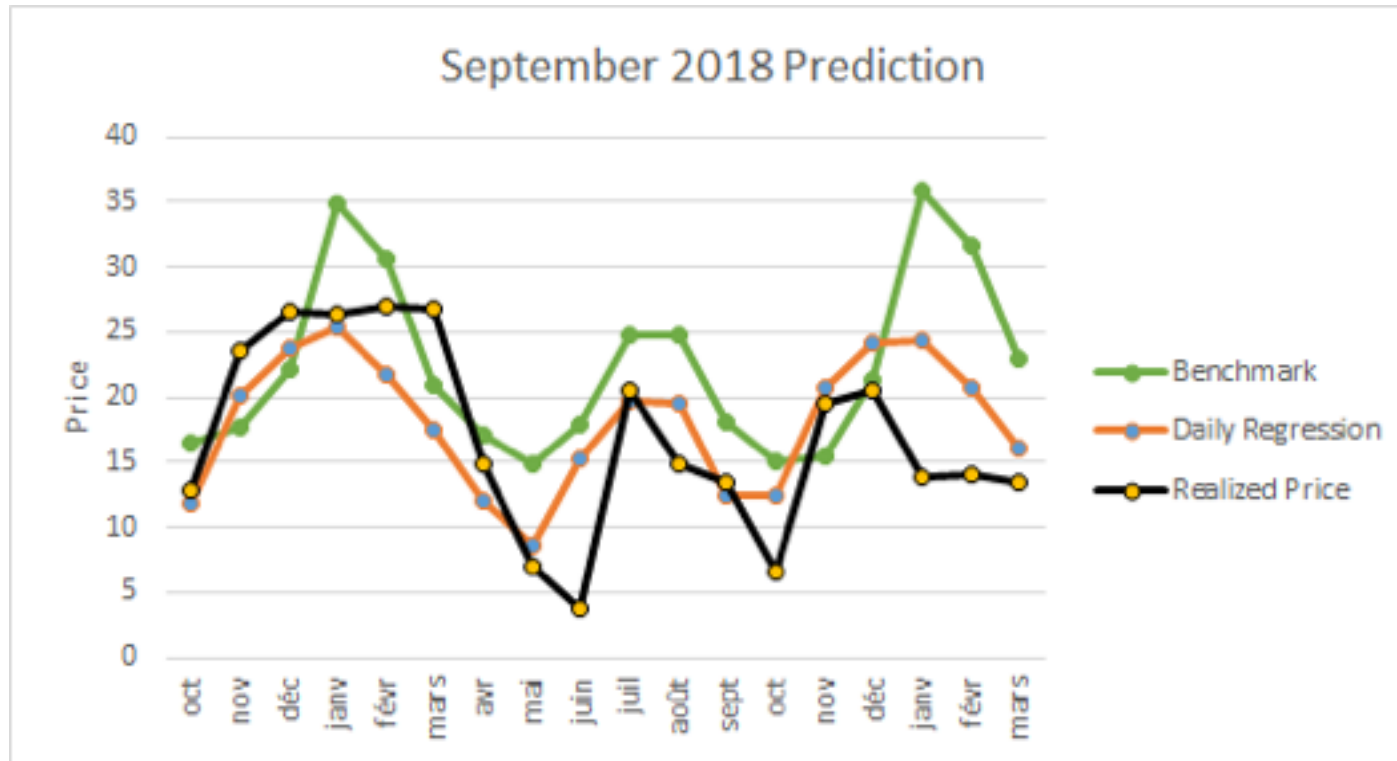
Daily Historical Regression : Results



	Benchmark	Daily Reg.
rMSE	7.33	8.52

III. Modeling Approach

Daily Historical Regression : Results



	Benchmark	Daily Reg.
rMSE	9.34	5.34

III. Modeling Approach

General overview

➤ Method:

1. Daily Historical Regression
2. **Weekly Regression**
3. Time series method



<https://www.dataversity.net/data-modeling-in-the-machine-learning-era/>

III. Modeling Approach

General overview

- Use **weekly data** (aggregation from hourly data)
- **Stepwise method** to select the variables to include in the model
- On **historical** data
 - Training period: 2017, January 2nd to 2019 March 31st
 - Validation sample: 2019 April 1st to 2020 July 5th
- On **forecast** data
 - Test samples: 2015-06 and 2018-03 and 2018-09
 - Get monthly forecast from weekly forecast by averaging (mean) the weekly forecast for corresponding month

III. Modeling Approach

Model specification

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2.386e+01	1.896e+01	-1.258	0.211274	
NUCLEAR_mean	-6.918e-03	7.567e-04	-9.142	8.14e-15	***
WIND_mean	-6.514e-03	1.512e-03	-4.308	3.89e-05	***
`Temp` (^C)_mean`	-4.025e-01	2.025e-01	-1.988	0.049610	*
Northwest_max	4.359e-02	3.211e-02	1.358	0.177625	
East_max	1.291e-02	6.135e-03	2.104	0.037944	*
Toronto_max	-2.934e-03	1.454e-03	-2.018	0.046322	*
Northeast_sum	1.064e-04	6.762e-05	1.573	0.118910	
Northwest_sum	-5.106e-04	2.151e-04	-2.374	0.019530	*
Ottawa_sum	2.051e-04	5.361e-05	3.825	0.000229	***
Bruce_sum	2.953e-04	1.272e-04	2.321	0.022330	*
Southwest_sum	2.321e-04	4.200e-05	5.527	2.65e-07	***
HYDRO_min	-8.632e-03	1.308e-03	-6.600	2.06e-09	***
`Temp` (^C)_mean2`	-9.307e-02	2.679e-02	-3.474	0.000761	***
meanTempLT11	-2.978e+00	6.363e-01	-4.680	9.11e-06	***
meanTempLT11_2	2.429e-01	5.688e-02	4.270	4.50e-05	***
meanTempLTm5	2.078e+00	8.483e-01	2.450	0.016040	*
meanTempLT57	1.503e+00	3.920e-01	3.833	0.000222	***

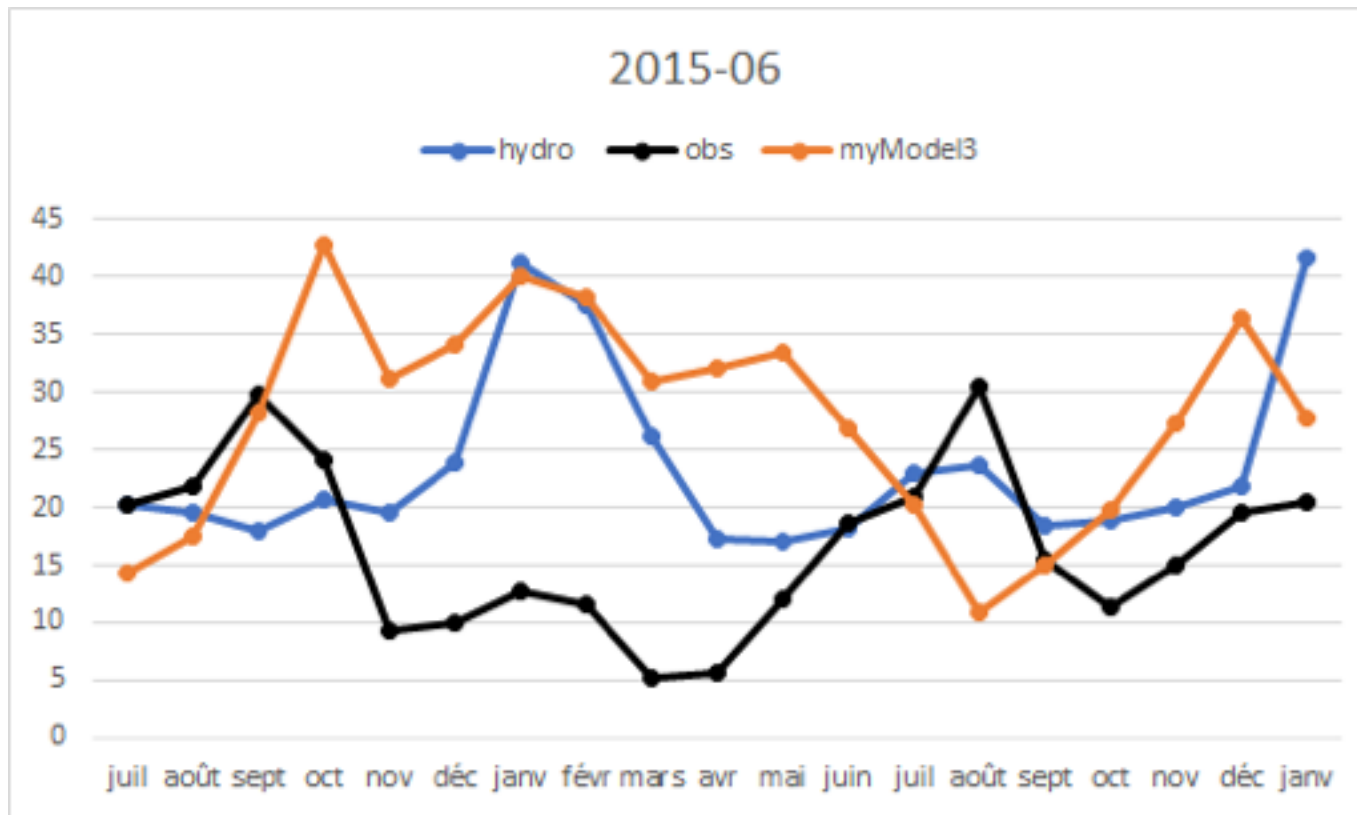
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.307 on 99 degrees of freedom
Multiple R-squared: 0.8415, Adjusted R-squared: 0.8142
F-statistic: 30.91 on 17 and 99 DF, p-value: < 2.2e-16

- meanTempLT are variables created based on average temperature variable to take into account the specific relationship with the price depending on the temperature
- Temp (c)_mean2 the square of the average temperature

III. Modeling Approach

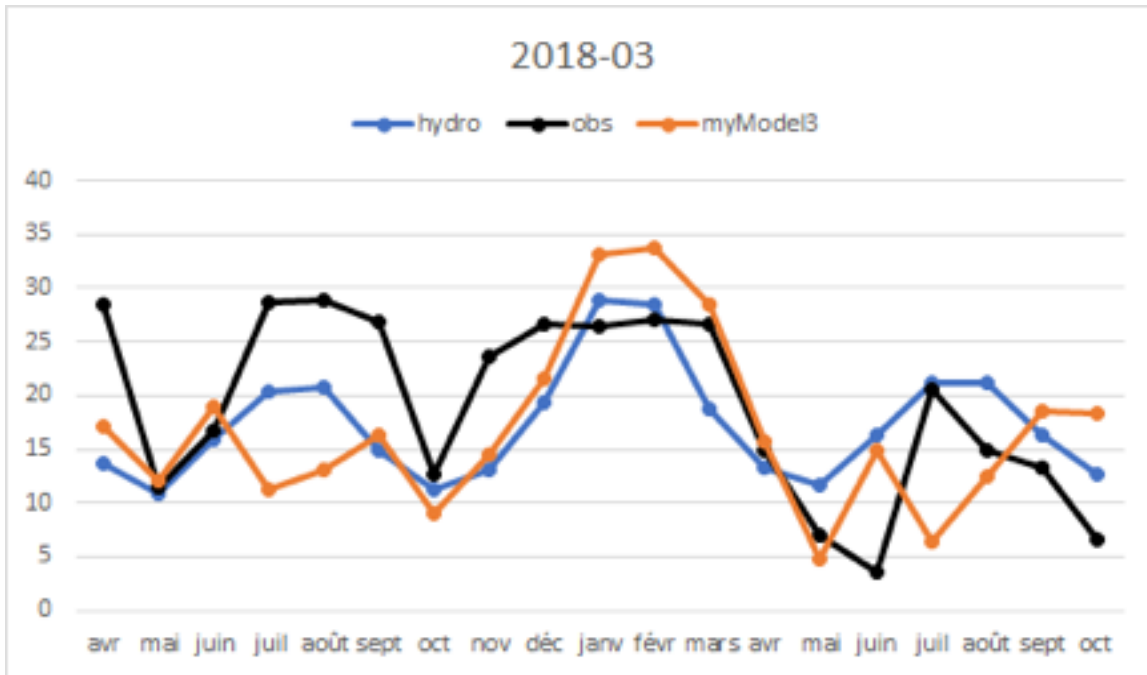
Results - 2015



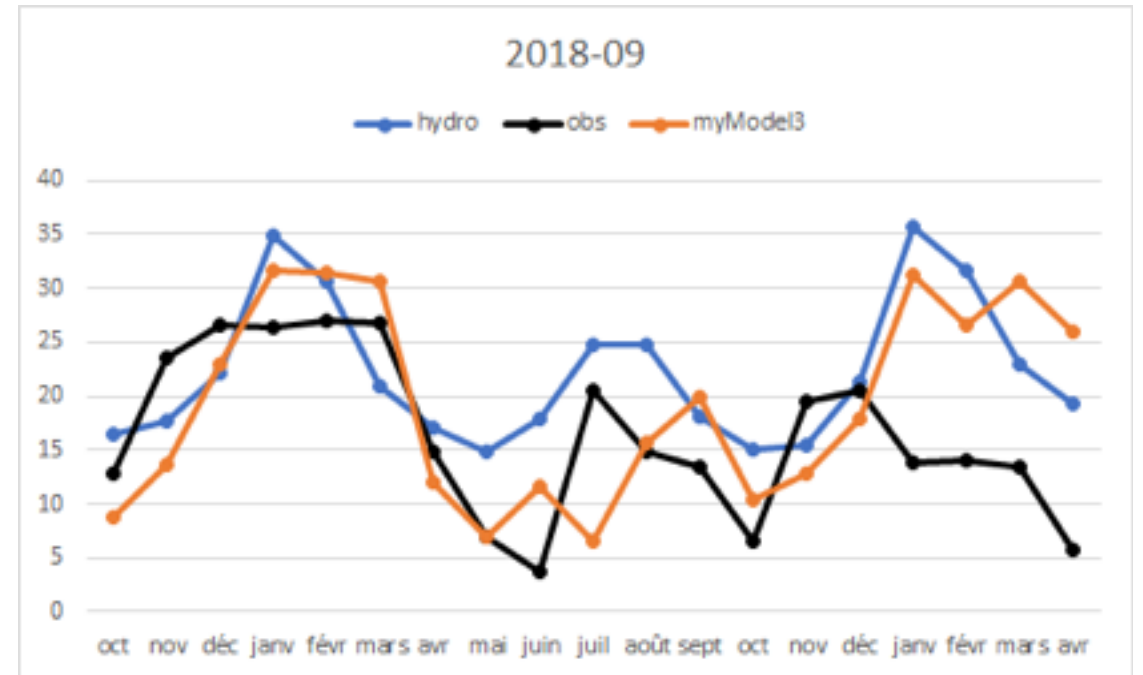
	Hydro model	Our model
RMSE	12.85	17.39

III. Modeling Approach

Results - 2018



	Hydro model	Our model
RMSE	7.27	8.96



	Hydro model	Our model
RMSE	9.60	9.55

III. Modeling Approach

General overview

➤ Method:

1. Daily Historical Regression
2. Weekly Regression
3. Time series method



<https://www.dataversity.net/data-modeling-in-the-machine-learning-era/>

III. Modeling Approach

Time series method:

- ARMA model $P_t = \alpha + \beta_1 P_{t-1} + \beta_2 P_{t-2} + \dots + \beta_p P_{t-p} + \gamma_1 \epsilon_{t-1} + \gamma_2 \epsilon_{t-2} + \dots + \gamma_q \epsilon_{t-q}$
 P_t : is the HOEP at time point t
 ϵ_t : is a sequence of white noise
- ARIMA model:
times and apply the ARMA model.
- SARIMA model: Consider the ARIMA method with the seasonal effect.
- ARX model: AutoRegressive models with exogenous variables

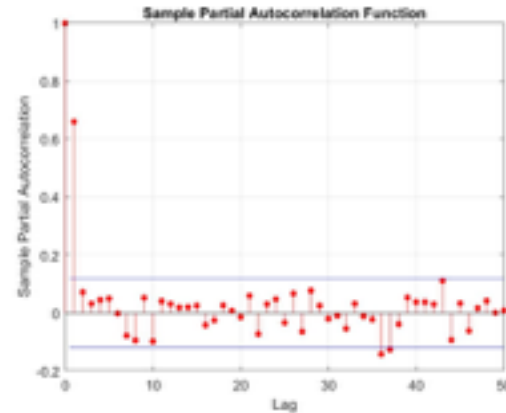
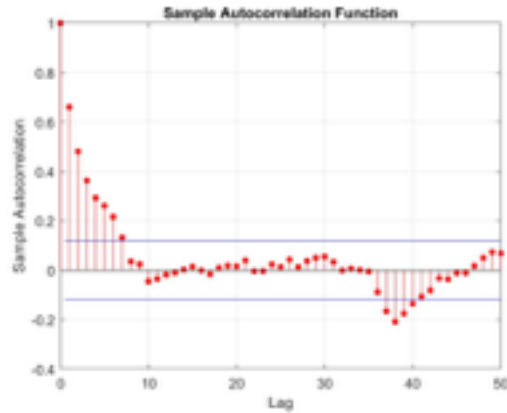
$$P_{t+1} = \sum_{i=1}^{n_a} a_i P_{t-i+1} + \sum_{i=1}^{n_b} b_i u_{t-i+1}$$

P_t : is the HOEP at time point t

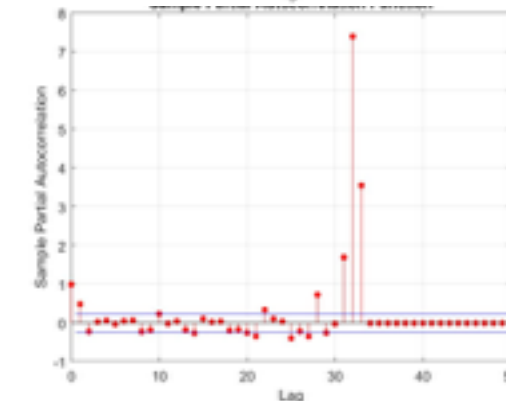
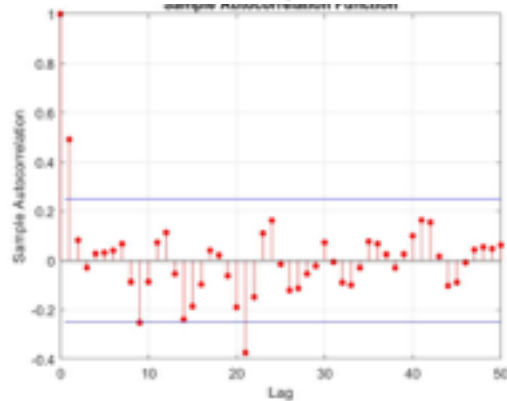
u_t : is a sequence of the input features

III. Modeling Approach

Time series method:



- ACF and PACF for weekly data.
- Nonstationary.
- No seasonal effect.

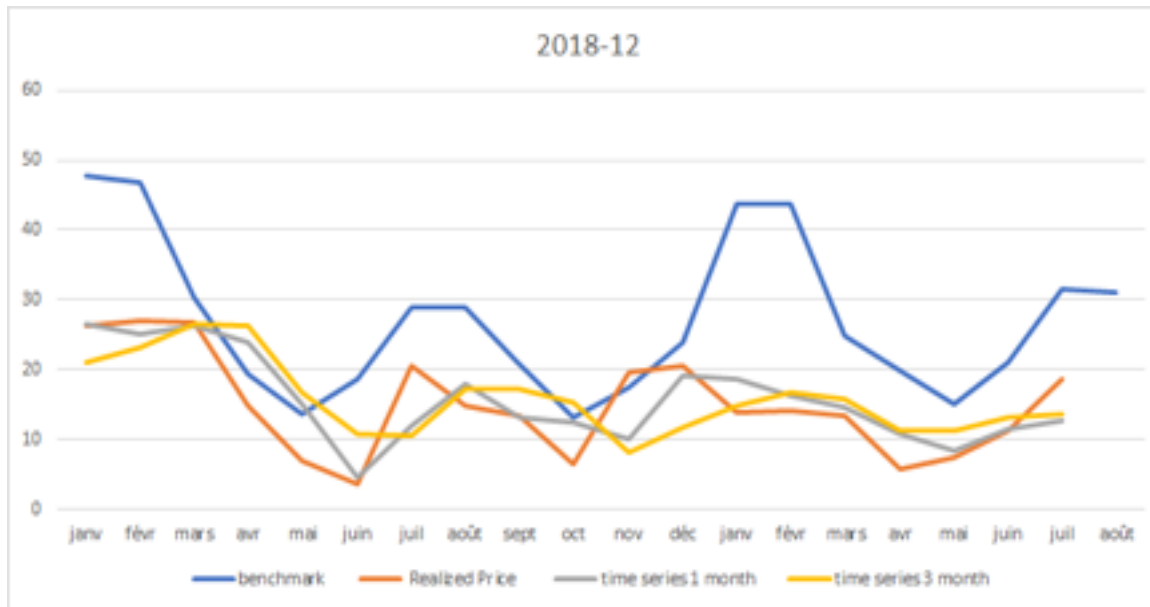


- ACF and PACF for monthly data.
- Nonstationary.
- Have seasonal effect.

III. Modeling Approach

Time series method:

Monthly Forecast by applying weekly data



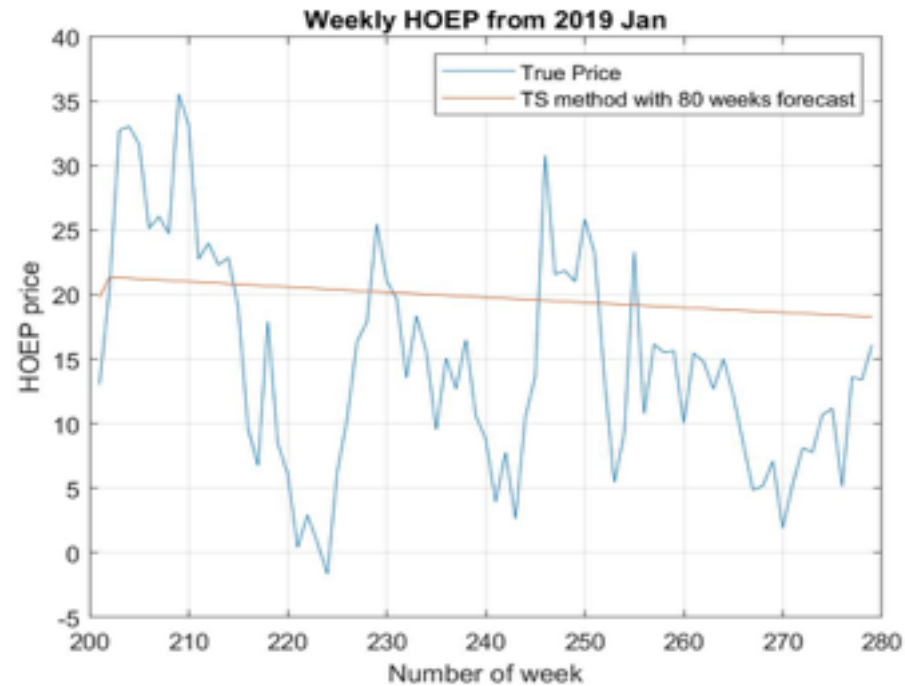
- We apply the ARIMA (1,1,2)
- Using the rolling method to compute the 1 month and 3 month's forecast.

	Benchmark	TS 1 month	TS 3 months
RMSE	14.4	4.88	6.57

III. Modeling Approach

Time series method:

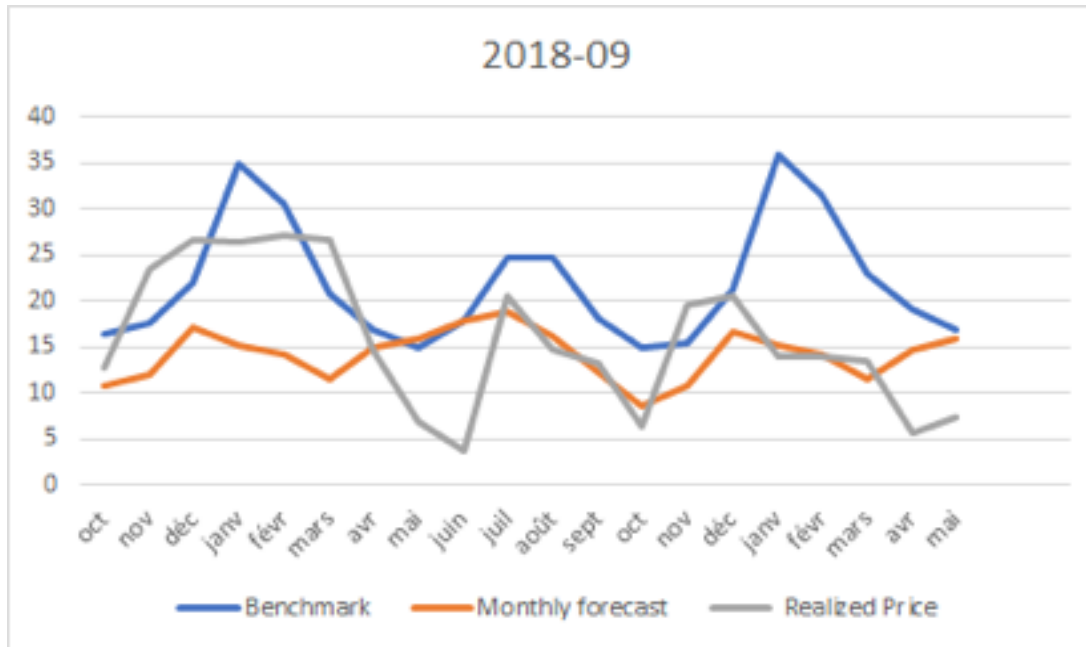
- Weekly forecast with long prediction horizon
- Time series will be a smooth line and fail



III. Modeling Approach

Time series method:

Monthly Forecast by applying monthly data



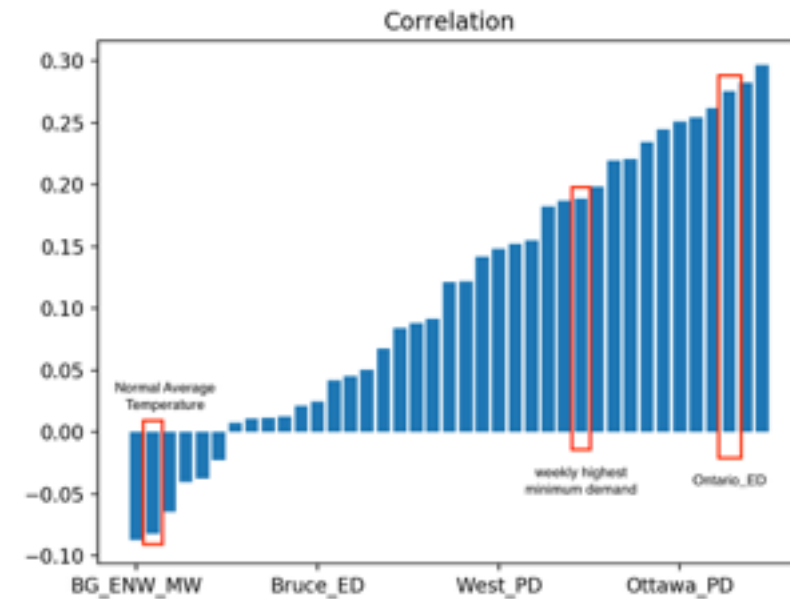
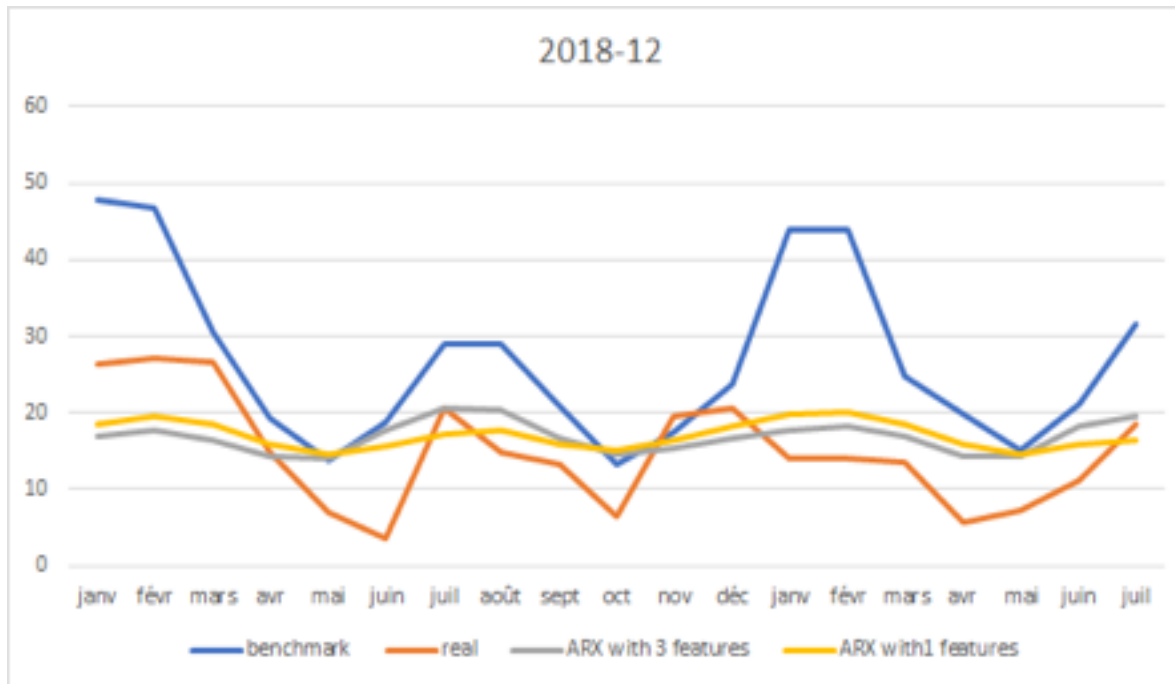
- We apply the SARIMA (3,1,2) (0,1,1)₁₂ for the monthly data.
- Apply the monthly data 2015-03 to 2018-09 to train and next 20 months for the forecast.

	Benchmark	TS monthly
RMSE	9.6	8.03

III. Modeling Approach

Time series method:

ARX by applying weekly data :



	Benchmark	ARX with 3 features	ARX with 1 features
RMSE	14.4	6.83	6.39

Summary

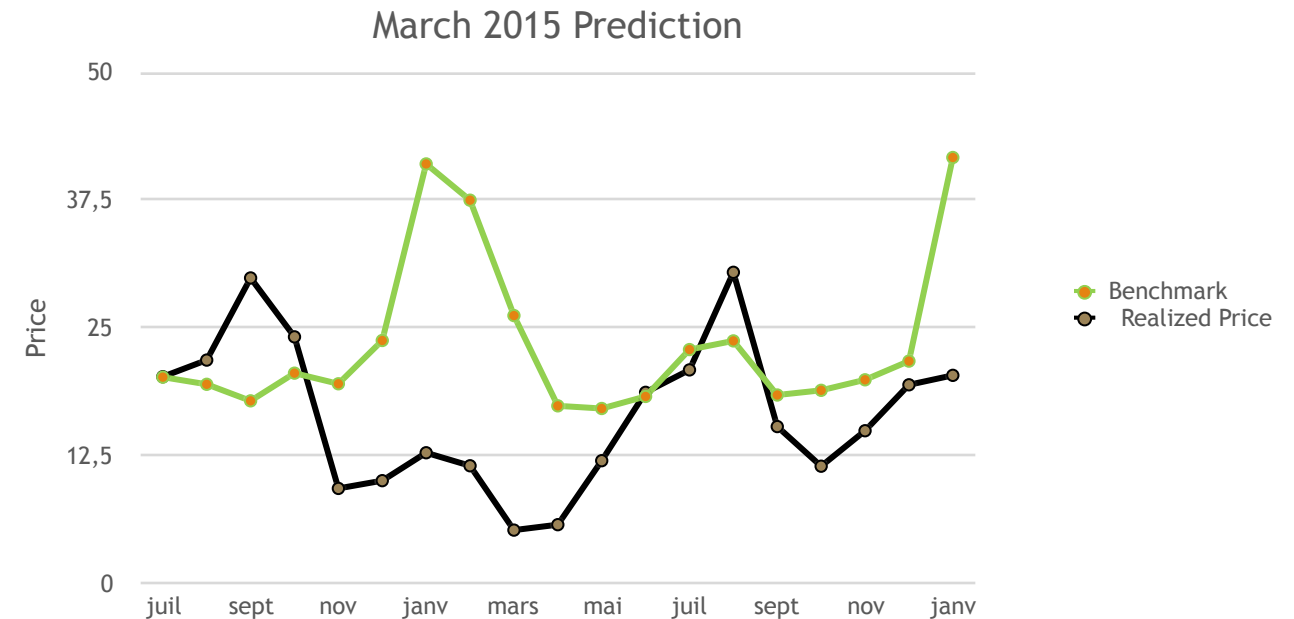
➤ **Goal:** Ontario energy price prediction better than **benchmark**

➤ **Approaches:**

✓ Classical Machine Learning

✓ Deep Learning

✓ Modeling



Thank you
