Power Load Forecasting By Electrical Substation

Team: Alain, Alexandre, Charlie, Maksym, Olivier and Sun

Hydro-Québec
(TransÉnergie et Équipement)

August 26th, 2022

Team and schedule

Hydro-Québec

- Alain Marcotte
- Olivier Milon

Students

- Charlie Hébert-Pinard, UQAM
- Maksym Shpakovych, UNILIM
- Zuming Sun, UCalgary

Coordination

- Alexandre Blondin Massé, HQ/UQAM

Schedule

- **Monday**: problem presentation, questions/answers about the domain, retrieval of the dataset
- **Tuesday**: exploration of the dataset, preparation of the dataset, description of some models used at Hydro-Québec
- **Wednesday**: implementation of various models (statistical and machine-learning based), reports about preliminary results
- **Thursday**: improvement of the models, additional reports, preparation of today's presentation

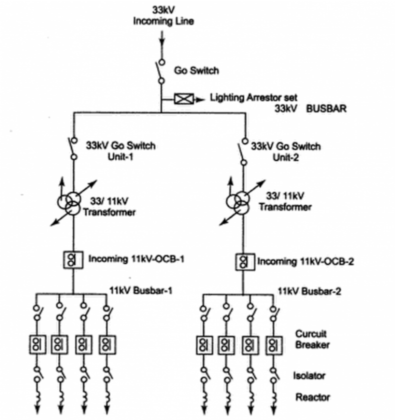
Summary of the submitted problem

Current situation

- « **Besoin québécois** » or **BQ** (in MW): load demand on the whole *connected* Quebec territory
- **Forecast**: typically 4 hours-ahead, 1 day-ahead or 2 days-ahead, with hourly time steps, updated 3 times per hour
- **Input**: *load consumption history* and *meteorological* state (temperature, cloudiness, wind speed, precipitation)

What is needed

Forecasts at the electrical *substation* level



(Source: WatElectrical.com)
A single line diagram (SLD) of a 33kV electrical substation

More details

Ideal result

- **BQ** (in MW) for the complete station
- **Transformers**: active power (in MW) and reactive power (in MVar) for each transformer
- **Bars**: current (in A) for each bar
- **Lines**: current (in A) for phases (3)

Difficulty: dependencies

- Strong **dependency** between those quantities
- Hard to **characterize formally**
- Layouts (or topology) **vary** a lot between substations
- Multiple possible substation **configurations** (or **states**)

Strategy?

1. Start with independent models
2. Chain them in time (for the forecasted horizon)
3. Chain them with respect to the substation topology

Description of the dataset

Input values

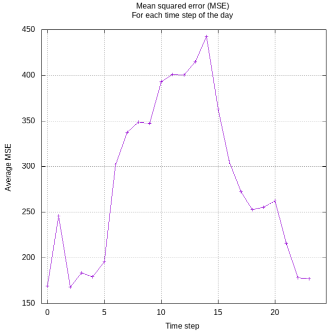
- Observed valued to be forecasted
  - Transformers Power (MW)
  - Bars Current (A)
  - Lines Current (A)
- Configuration setting
  - Breakers state (opened/closed)
- Weather conditions
  - Temperature
  - Wind Speed
  - Cloud Coverage (on a 0-10 scale)
  - Precipitation Type (coded string)

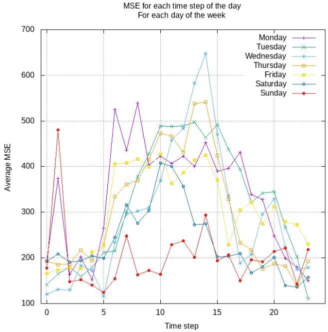
annee	mois	jour	jsem	periode	heure	BQ	Baiss	Tyul	Nyul	Vyul	Precyul
2018	12	31	1	0	24	23686	1709	28	1	9	10
2019	1	1	2	0	1	23675	1734	85	1	7	10
2019	1	1	2	0	2	23614	1744	32	1	9	10
2019	1	1	2	0	3	23463	1785	4	2	3	10
2019	1	1	2	0	4	23423	1744	32	2	10	18
2019	1	1	2	0	5	23556	1790	72	0	4	10
2019	1	1	2	0	6	23829	1811	55	0	5	10
2019	1	1	2	0	7	23981	1834	28	0	1	10
2019	1	1	2	0	8	24117	1837	12	0	1	9
2019	1	1	2	0	9	23860	1842	8	0	8	7
2019	1	1	2	0	10	23631	1895	83	0	9	6
2019	1	1	2	0	11	23458	1924	24	1	10	24
2019	1	1	2	0	12	23756	1858	9	0	6	7
2019	1	1	2	0	13	23669	1917	61	0	2	6
2019	1	1	2	0	14	23355	1834	28	0	8	4
2019	1	1	2	0	15	23628	1782	2	-1	5	4
2019	1	1	2	0	16	25006	1822	92	-3	2	2
2019	1	1	2	0	17	26824	1786	93	-5	2	4
2019	1	1	2	0	18	27260	1761	36	-6	5	2
2019	1	1	2	0	19	27076	1891	1	-8	1	19
2019	1	1	2	0	20	27144	1898	67	-9	1	14
2019	1	1	2	0	21	27100	1855	11	-9	7	2
2019	1	1	2	0	22	26826	1874	5	-10	4	1
2019	1	1	2	0	23	26420	1811	55	-11	1	13
2019	1	1	2	0	24	26274	1780	3	-11	4	1
2019	1	2	3	0	1	26291	1751	89	-12	3	1
2019	1	2	3	0	2	26726	1712	12	-13	1	4
2019	1	2	3	0	3	27135	1696	2	-13	7	3
2019	1	2	3	0	4	27525	1695	8	-13	9	7

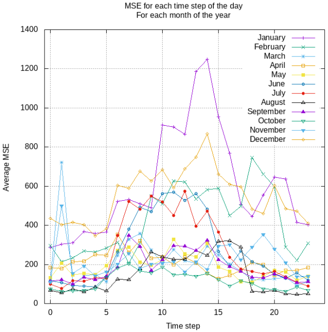
Dataset view

Multiple equation model (baseline)

- **ARMAX-type model**: *auto-regressive*, with *moving average* and exogenous variables
- **Seasonality addressed explicitly**: one model for each time step of the day, one-hot encoding for the week day and linear combination of sine and cosine expressions for the year seasonality







Chained 2-layers neural networks

Definition (Chained regression model)

The chained regression model is the model where the forecast for each consecutive point in a timeseries is done by a separate model, which uses the previous forecast as an input.

The idea is to train the chained model where a nonlinear element (NN) is used instead of a linear (ARMAX) in the chain. The *NN* writes as

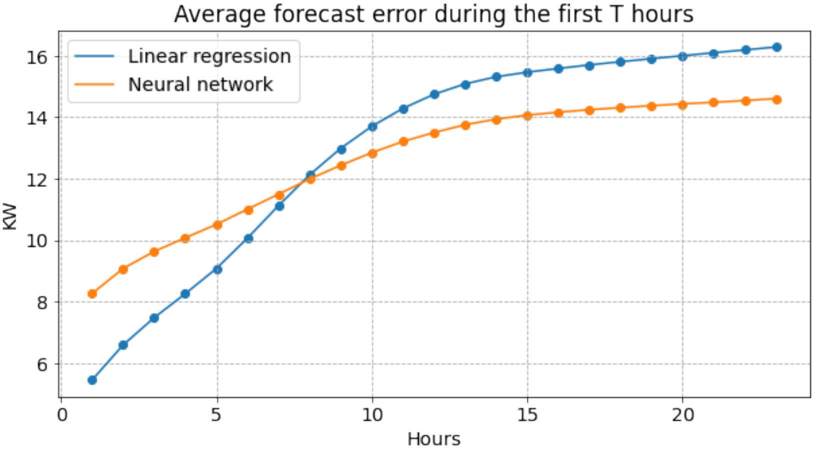
$$NN(x) = w_2^\top g(W_1 x + b_1, \theta) + b_2, \tag{1}$$

where $W_1 \in \mathbb{R}^{64\times n}$, $w_2 \in \mathbb{R}^{64}$, $b_1 \in \mathbb{R}^{64}$, $b_2 \in \mathbb{R}$, $\theta \in \mathbb{R}^{64}$ are trainable parameters and $g(x, \theta) = \{g_1(x_1, \theta_1), \dots, g_m(x_m, \theta_m)\}$ where

$$g_i(x_i, \theta_i) = \begin{cases} x_i & x_i \geq 0 \\ \theta_i x_i & x_i < 0 \end{cases} \quad \text{for } i \in \{1, \dots, m\}, \tag{2}$$

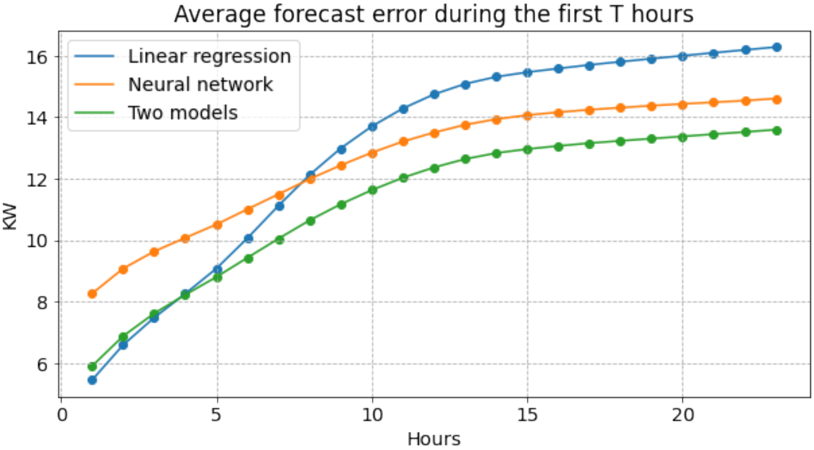
is an m-dimensional parametric ReLU function.

Chained 2-layers neural networks



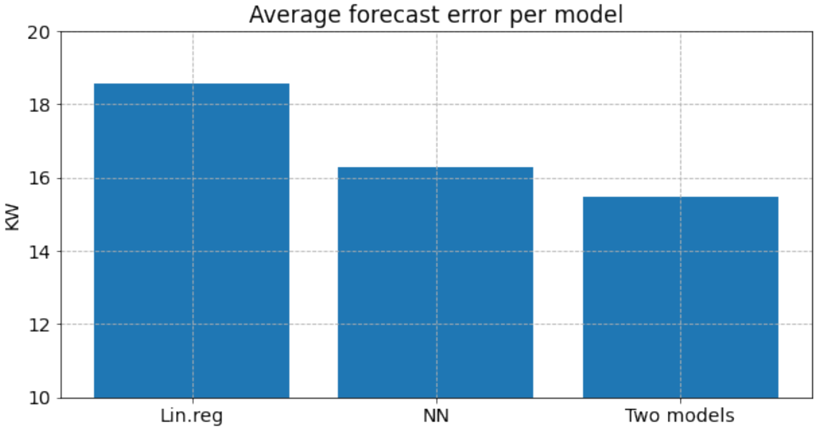
Forecast quality comparison for the first T hours.

Chained 2-layers neural networks combined with a linear model



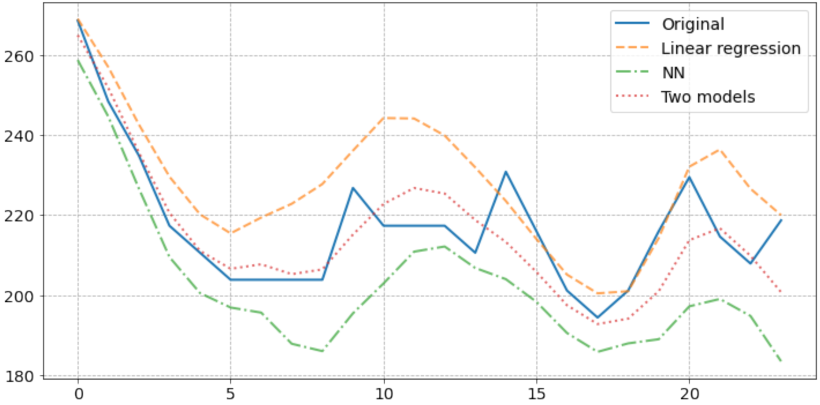
Forecast quality comparison for the first T hours.

Chained 2-layers neural networks combined with a linear model



Forecast quality comparison for all validation set.

Chained 2-layers neural networks combined with a linear model



Forecast for 24 hours example.

Next steps

- To build a global model for all elemnts of the system (transformers, bars, brakers) using the information of the brakers state.
- To implement a parallel training for neural network models in the chain.
- To use a faster optimization algorithm than default stochastic gradient descent.