# Microgrids Introduction to probabilistic forecasting



Jonathan Dumas jdumas@uliege.be

Learning objectives

Through this lecture, it is aimed for the students to be able to:

- Produce **probabilistic** forecasts;
- Perform verification of probabilistic forecasts

### **Residential energy supplier**

Summary

1. Reminder

2. Probabilistic forecasts

3. Verification of probabilistic forecasts

#### Reminder: forecast for decision making

Forecasting is a natural first step to *decision-making* 



Key parameters for a microgrid to forecast:

Generation: PV, Wind Power, Hydraulic Power, etc

Load: office, industrial, residential, etc

Prices: electricity, gas, (futures, day ahead, intraday, imbalances).

**Reminder: point forecast definition** 

A forecast is an **estimate** for time **t** + **k**, conditional to information up to time t.

A point forecast informs of the conditional expectation of power generation.



Reminder: use a strategy to assess forecasts

Several strategies to assess forecasts:

- splitting the dataset into a learning and a validation sets;
- k-cross validation: k pairs or random learning and validation sets.





Reminder: use quantitative metrics

**Bias** or Nbias, for the normalized version:

$$\mathbf{bias}(k) = \frac{1}{T} \sum_{t=1}^{T} \epsilon_{t+k|t}$$

Mean Absolute Error (MAE) or NMAE, for the normalized version:

$$\mathbf{MAE}(k) = \frac{1}{T} \sum_{t=1}^{T} |\epsilon_{t+k|t}|$$

Root Mean Squared Error (RMSE) or NRMSE, for the normalized version:

$$\mathbf{RMSE}(k) = \left[\frac{1}{T} \sum_{t=1}^{T} \epsilon_{t+k|t}^2\right]^{1/2}$$

### **Residential energy supplier**

Summary

1. Reminder

- 2. Probabilistic forecasts
- 3. Verification of probabilistic forecasts

**Probabilistic forecasting** 

The various types of probabilistic forecasts range, from **quantile** to **density** forecasts, **prediction** intervals, and **scenarios**.



#### **Quantile forecast definition**

A **quantile** forecast is to be seen as a probabilistic **threshold** for power generation.



- F the cumulative distribution function (CDF)

**Prediction interval definition** 

A **prediction** interval is an **interval** within which power generation may lie, with a certain probability.



#### Predictive density definition

A predictive density fully describes the **probabilistic distribution** of power generation for every lead time.

$$Y_{t+k} \approx \hat{F}_{t+k|t}$$

with:

- F the cumulative distribution function for Yt+k



#### Trajectories (scenarios) definition

Trajectories are equally-likely **samples** of multivariate predictive densities for power generation (in time and/or space).





PV quantile forecasts computed at 12:00 for the next day along with corresponding observations (Pm in red) and point forecasts (dad 12 in black).

The model are a **feed-forward neural network** (MLP) on the left figure and a **long short term memory neural network** (LSTM) on the right figure.

### **Residential energy supplier**

Summary

1. Reminder

2. Probabilistic forecasts

3. Verification of probabilistic forecasts

Case study: PV parking rooftops from Liège university

PV installation of 466.4 kWp



<u>https://www.uliege.be/cms/c\_7726266/fr/2500-m-de-panneaux-photovoltaiques-bientot-en-fonction-sur-le-campus-du-sart-tilman</u>

Case study: PV forecasting model

The forecasting model g is a **feed-forward neural network**:

- with one hidden layer;
- weather forecasts of **solar** irradiation and **air** temperature as inputs;
- the output layer is composed of **96** neurons (96 time steps);
- it is implemented in **python** using Tensorflow library.

#### **Evaluation methodology**

Forecasting is about being able to **predict future events**, in new situations not only explain what happen in the past.

One need to verify forecasts on data that has not been used for the modelling!

Several strategies:

- splitting the dataset into a learning and a validation sets;
- k-cross validation: k pairs or random learning and validation sets.





**Visual inspection** 

Visual inspection allows you to develop **susbtantial insight** on forecast **quality**.

This comprises a **qualitative analysis** only.

What do you think of these two? Are they good or bad?



Attribute of probabilistic forecast quality

How do you want your forecasts?

- Reliable? (also referred to as "probabilistic calibration")
- **Sharp**? (i.e., informative)

- **Skilled**? (all-round performance, and of higher quality than some benchmark)

Probabilistic calibration (reliability)

Calibration is about respecting the probabilistic contract:

- quantile forecast with a nominal level q = 0.5, one expect than the observations are to be less than the forecast 50% of the times;
- prediction interval with a nominal coverage of 90%, one expect than the observations are to be covered by this prediction 90% of the times

To do it in practice, we take a frequentist approach... we simply count!

#### **Reliability diagrams**

The calibration assessment can be summarized in reliability diagrams.

Predictive densities composed by quantile forecasts with nominal levels {0.05, 0.1,...,0.45, 0.55,...,0.9, 0.95}.

Quantile forecasts are evaluated one by one, and their empirical levels are reported vs. their nominal levels

The **closest** to the **diagonal**, the better!



#### Sharpness

Sharpness is about the **concentration** of probability.

A perfect probabilistic forecast gives a probability of 100% on a single value!

Consequently, a sharpness assessment boils down to evaluating how tight the predictive densities are...

For a given interval forecast

$$\hat{I}_{t+k|t}^{(\alpha)} = \left[ \hat{y}_{t+k|t}^{(q=\alpha/2)}, \hat{y}_{t+k|t}^{(q=1-\alpha/2)} \right]$$

The width is

Average over the validation set

 $\delta^{(lpha)}_{t,k}$ 

 $\mathbf{T}$ 

$$\delta_{t,k}^{(\alpha)} = \hat{y}_{t+k|t}^{(q=1-\alpha/2)} - \hat{y}_{t+k|t}^{(q=\alpha/2)} \qquad \qquad \delta^{(\alpha)}(k) = \frac{1}{T} \sum_{t=1}^{T} \sum_{t=1}^{T} \frac{1}{T} \sum$$

#### Sharpness

Predictive densities are composed by interval forecasts with nominal coverage rates = 0.1, 0.2,..., 0.9.

The interval width increases with the lead time, reflecting higher forecast uncertainty



#### **Overall skill assessment**

The skill of probabilistic forecasts can be assessed by scores, like MAE and RMSE for the deterministic forecasts.

The most common skill score for predictive densities is the **Continuous Ranked Probability Score (CRPS).** 

$$\mathbf{CRPS}_{t,k} = \int_{x} [\hat{F}_{t+k|t}(x) - \mathbf{1}(x \ge y_{t+k})]^{2} dx$$

$$\mathbf{CRPS}(k) = \frac{1}{T} \sum_{t=1}^{T} \mathbf{CRPS}_{t,k}$$

$$\mathbf{CRPS} \text{ and MAE (for deterministic forecasts) can be directly compared.}$$

$$\mathbf{CRPS}_{t,k} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{CRPS}_{t,k}$$

power generation [p.u.]

CRPS example on the case study



**CRPS** per lead time, from 11-cross validation, of three forecasting models for PV quantile forecasts.

#### Conclusion: forecast for decision making

Forecasting is a natural first step to *decision-making* 



Key parameters for a microgrid to forecast:

Generation: PV, Wind Power, Hydraulic Power, etc

Load: office, industrial, residential, etc

Prices: electricity, gas, (futures, day ahead, intraday, imbalances).

**Conclusion: Probabilistic forecasting** 

The various types of probabilistic forecasts range, from **quantile** to **density** forecasts, **prediction** intervals, and **scenarios**.



**Conclusion: Probabilistic forecasting** 

The various types of probabilistic forecasts range, from **quantile** to **density** forecasts, **prediction** intervals, and **scenarios**.



Conclusion: attribute of probabilistic forecast quality

How do you want your forecasts?

- **Reliable**? (also referred to as "probabilistic calibration") -> reliability diagrams

- **Sharp**? (i.e., informative) -> width

- **Skilled**? (all-round performance, and of higher quality than some benchmark) -> CRPS

Conclusion: use a strategy to assess forecasts

Several strategies to assess forecasts:

- splitting the dataset into a learning and a validation sets;
- k-cross validation: k pairs or random learning and validation sets.





#### References

Morales, Juan M., et al. Integrating renewables in electricity markets: operational problems. Vol. 205. Springer Science & Business Media, 2013.

Free online chapter: <u>http://pierrepinson.com/31761/Literature/</u> <u>reninmarkets-chap2.pdf</u>

Online lessons from P. Pinson:

https://energy-markets-school.dk/summer-school-2019/

http://pierrepinson.com/index.php/teaching/

The end, to be continued ...