Machine Learning techniques for Electricity Price Forecasting

Thesis Defence

Léonard Tschora

Pr. Siegfried Nijssen	Université Catholique de Louvain
Pr. Massih-Reza Amini	Université Grenoble Alpes
Dr. Charlotte Laclau	Télécom Paris
Pr. Elisa Fromont	Université Rennes 1
Pr. Stéphane Canu	INSA Rouen
Dr. Erwan Pierre	elmy
Pr. Céline Robardet	INSA Lyon
Pr. Marc Plantevit	EPITA

17^{th} Jan, 2024

UNIVERSITÉ DE LYON INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

elmy

La fée electricité, Raoul Dufy 1937

Balancing consumption and generation

Consumption

Balancing consumption and generation

Consumption

Generation

Balancing consumption and generation

Consumption

Generation

$$=$$

How can suppliers and consumers agree on a common price?

Balancing consumption and generation

Consumption

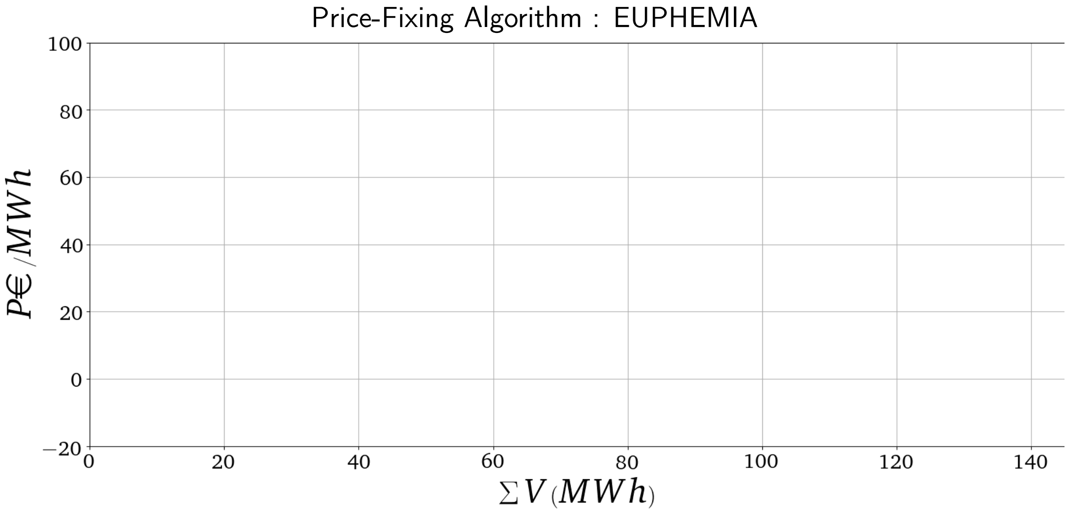
=

Generation

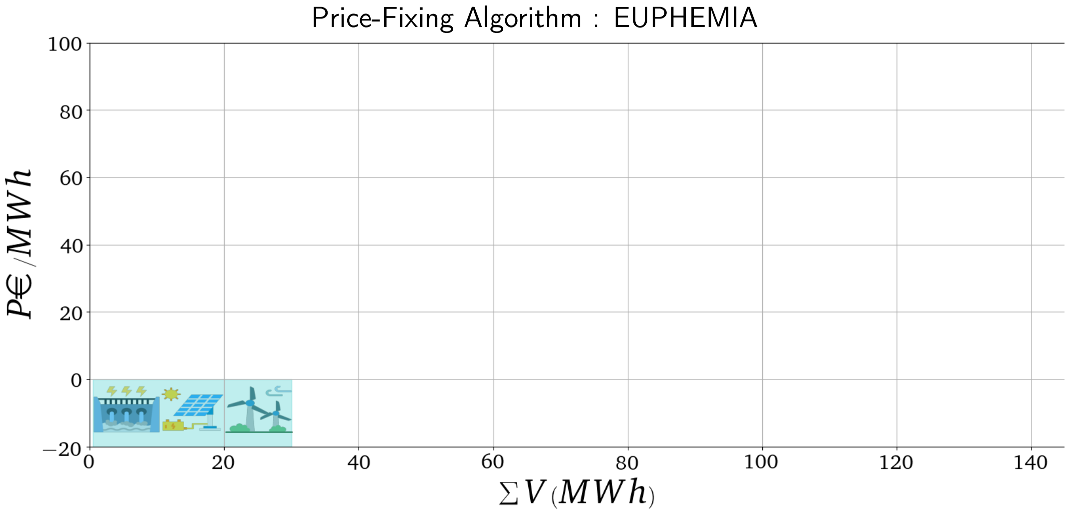
How can suppliers and consumers agree on a common price?

They use a Price-Fixing Algorithm

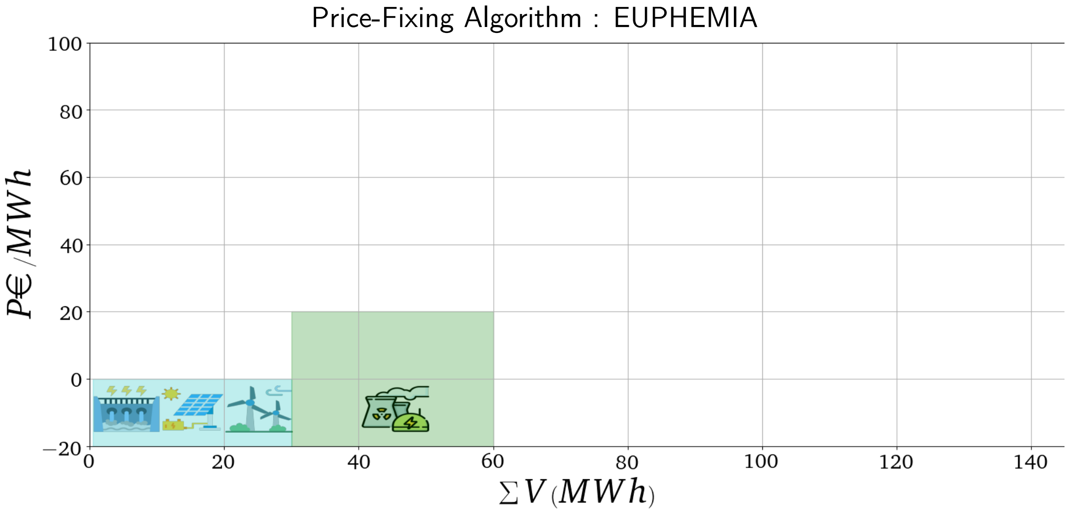
How does the Price-Fixing Algorithm works?



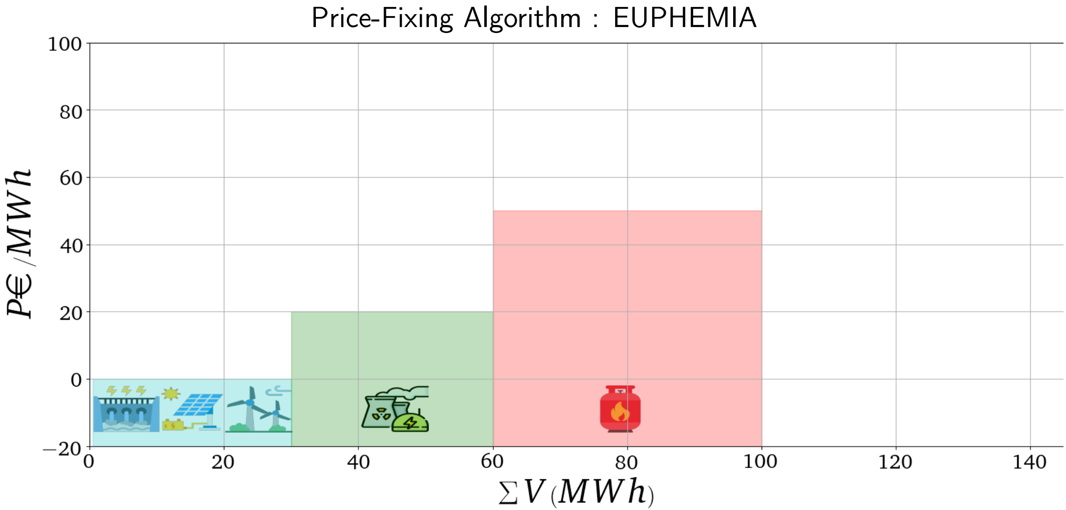
How does the Price-Fixing Algorithm works?



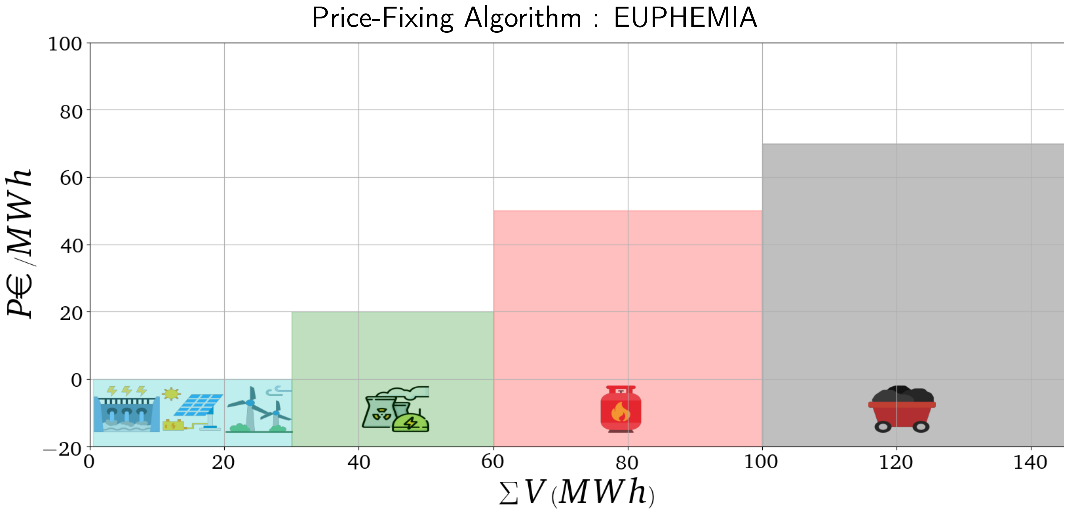
How does the Price-Fixing Algorithm works?



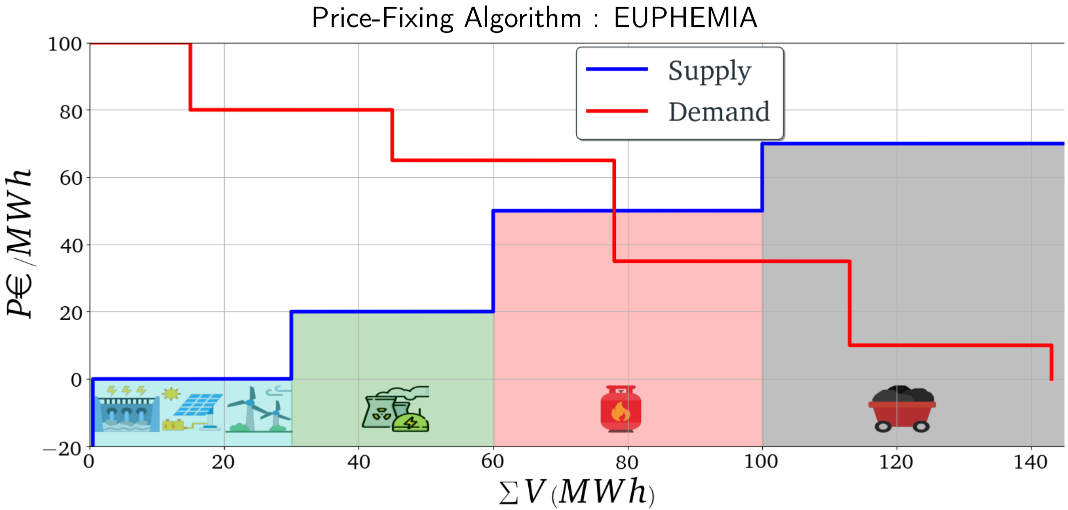
How does the Price-Fixing Algorithm works?



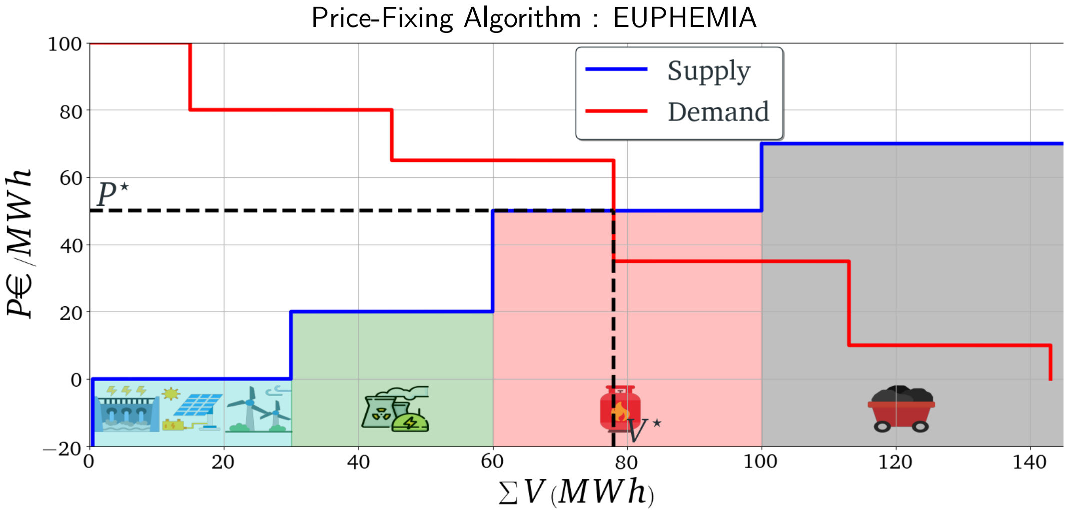
How does the Price-Fixing Algorithm works?



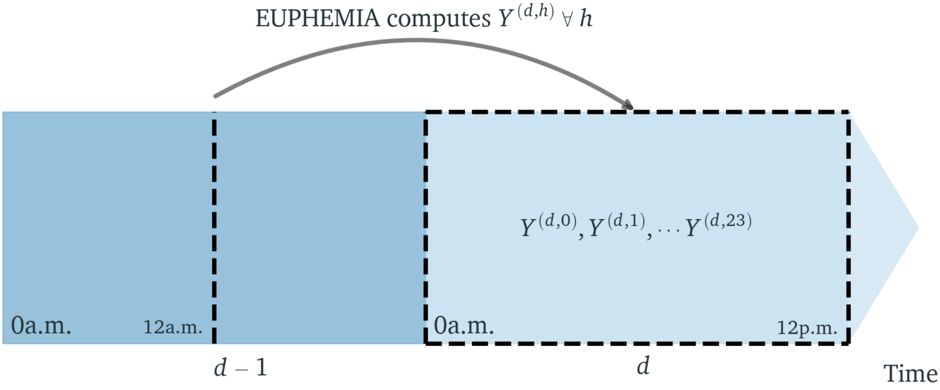
How does the Price-Fixing Algorithm works?



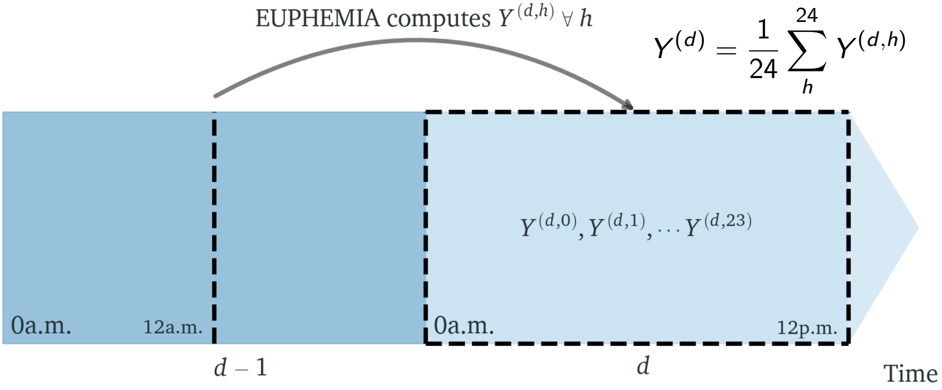
How does the Price-Fixing Algorithm works?



The Day-Ahead Prices Y are computed everyday at 12a.m.

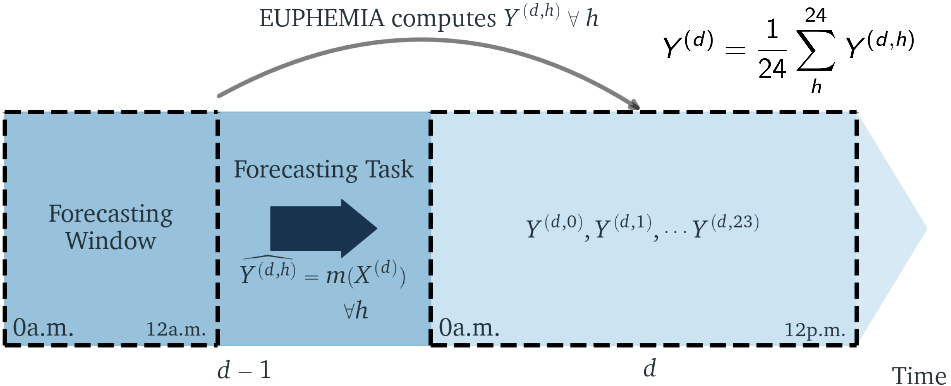


The Day-Ahead Prices Y are computed everyday at 12a.m.

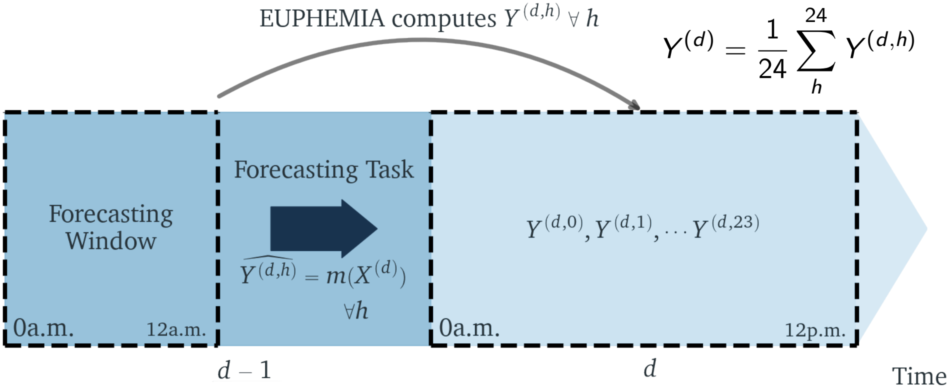


The Day-Ahead Prices Y are computed everyday at 12a.m.

$$Y^{(d)} = \frac{1}{24}\sum_{h}^{24} Y^{(d,h)}$$

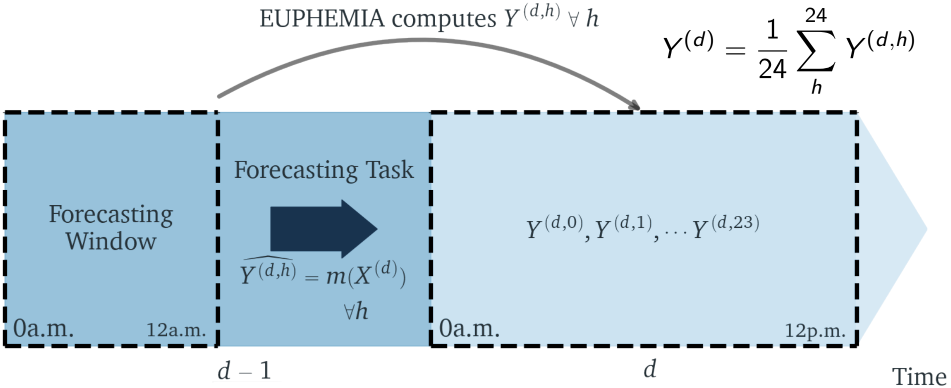


The Day-Ahead Prices Y are computed everyday at 12a.m.



Consumption
Programmable Generation
Renewables Generation

The Day-Ahead Prices Y are computed everyday at 12a.m.

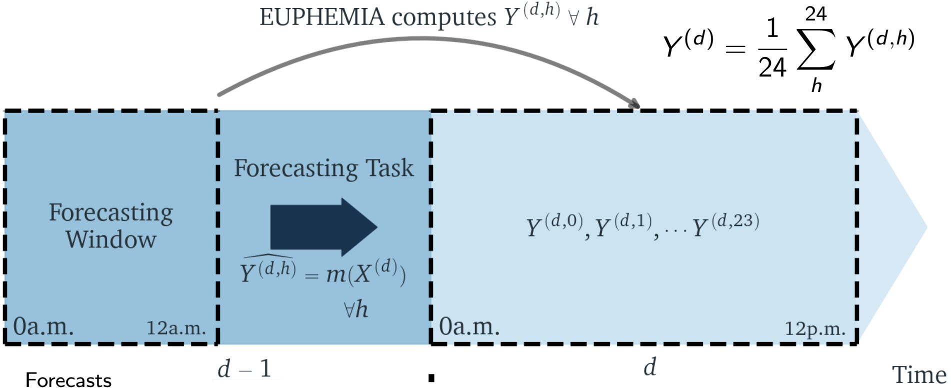


Consumption
Programmable Generation
Renewables Generation

Not available at forecasting time!

The Day-Ahead Prices Y are computed everyday at 12a.m.

$$Y^{(d)} = \frac{1}{24}\sum_{h}^{24} Y^{(d,h)}$$



$X =$ Forecasts

$C^{(d,h)}, C^{(d-1,h)}, C^{(d-7,h)}$

$G^{(d,h)}, G^{(d-1,h)}, G^{(d-7,h)}$

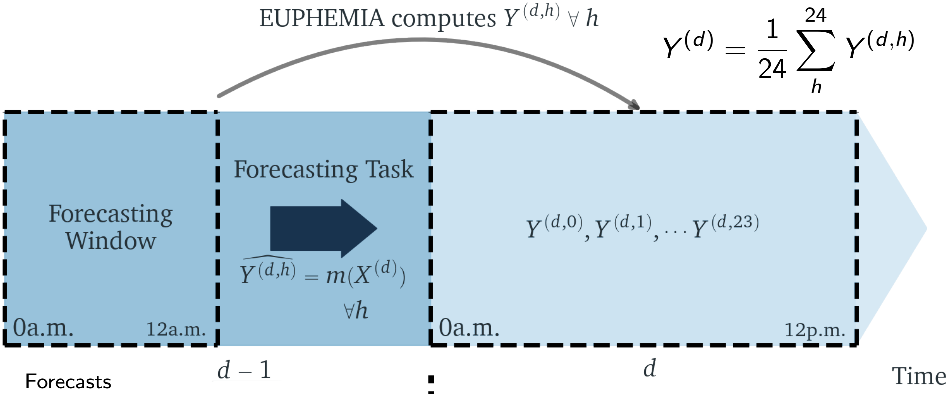
$R^{(d,h)}, R^{(d-1,h)}, R^{(d-7,h)}$

Consumption
Programmable Generation
Renewables Generation

Not available at forecasting time!

The Day-Ahead Prices Y are computed everyday at 12a.m.

$$Y^{(d)} = \frac{1}{24}\sum_{h}^{24} Y^{(d,h)}$$



$X =$ Forecasts
- $C^{(d,h)}, C^{(d-1,h)}, C^{(d-7,h)}$
- $G^{(d,h)}, G^{(d-1,h)}, G^{(d-7,h)}$
- $R^{(d,h)}, R^{(d-1,h)}, R^{(d-7,h)}$

Past Prices
- $Y^{(d-1,h)}, Y^{(d-2,h)}, Y^{(d-3,h)}, Y^{(d-7,h)}$

Consumption
Programmable Generation
Renewables Generation

Not available at forecasting time!

Electricity Price Forecasts usage : The Islander project

islander

ACCELERATING THE DECARBONIZATION OF ISLANDS' ENERGY SYSTEMS

This project has received funding from the European Union's Horizon 2020 research and innovation under grant agreement No 957669

elmy

The island of Borkum, Germany

Electricity Price Forecasts usage : The Islander project

islander

ACCELERATING THE DECARBONIZATION OF ISLANDS' ENERGY SYSTEMS

This project has received funding from the European Union's Horizon 2020 research and innovation under grant agreement No 957669

elmy

The island of Borkum, Germany

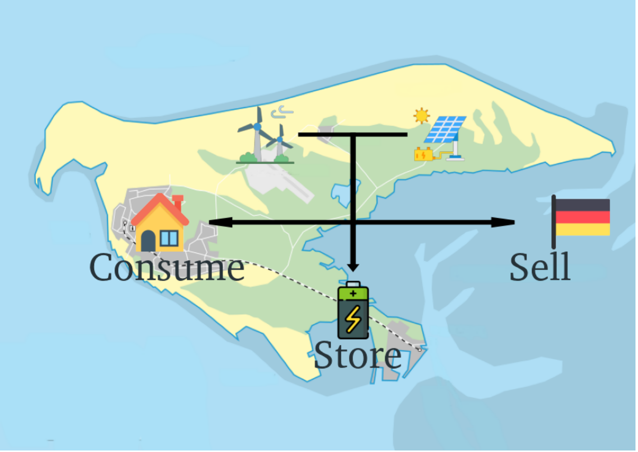
Electricity Price Forecasts usage : The Islander project

islander
ACCELERATING THE DECARBONIZATION OF ISLANDS' ENERGY SYSTEMS

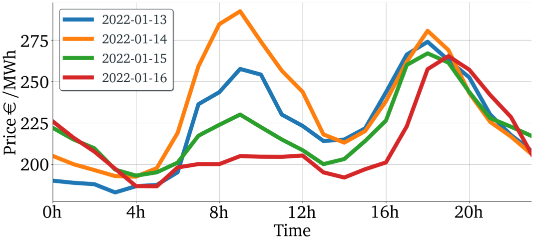
This project has received funding from the European Union's Horizon 2020 research and innovation under grant agreement No 957669

elmy

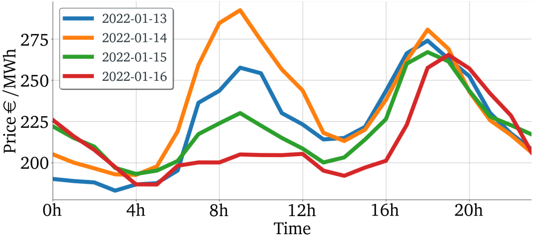
The island of Borkum, Germany

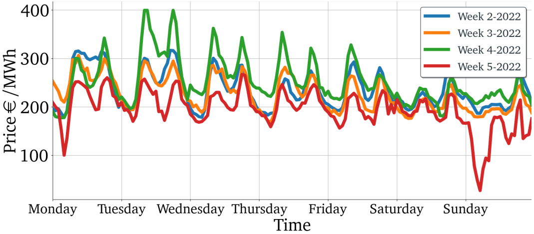


Model performance VS user confidence

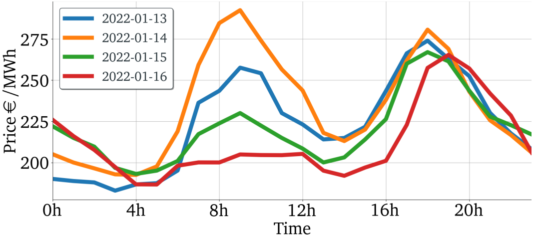


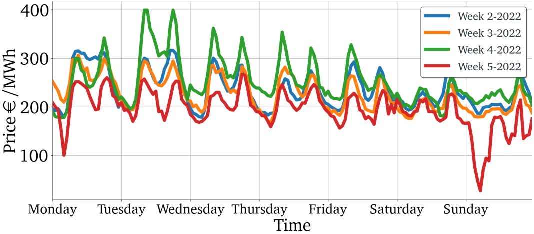
Model performance VS user confidence

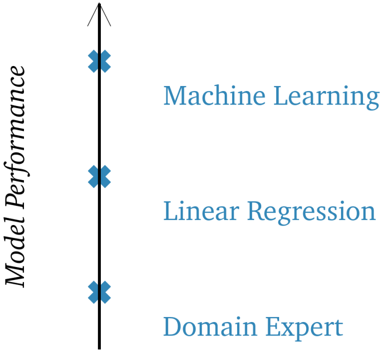




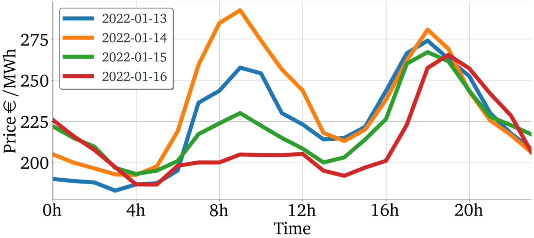
Model performance VS user confidence

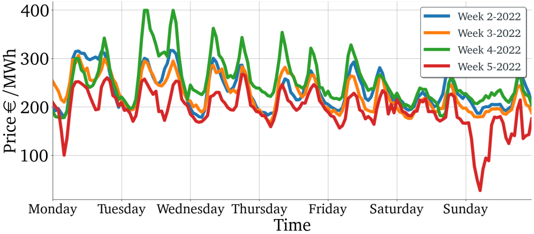


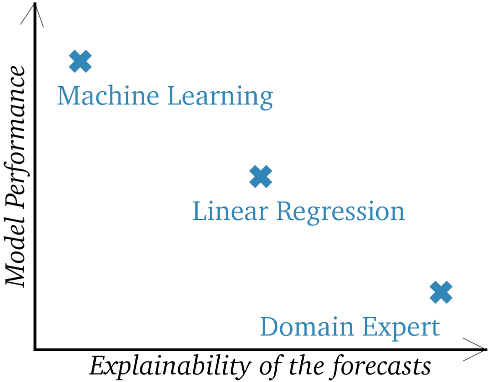




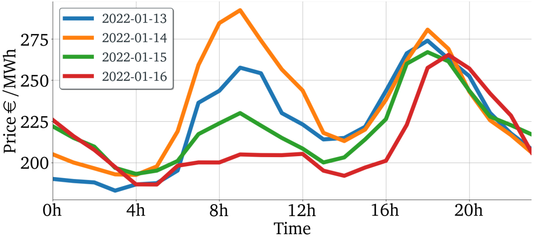
Model performance VS user confidence

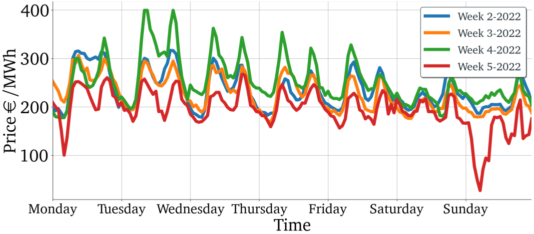




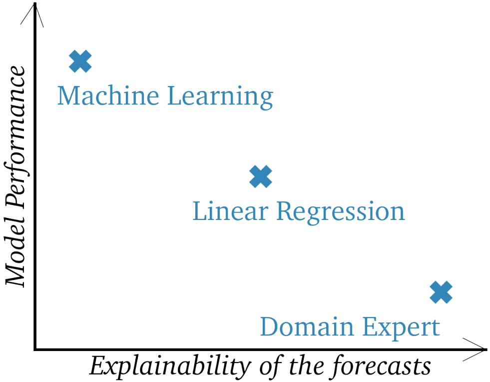


Model performance VS user confidence

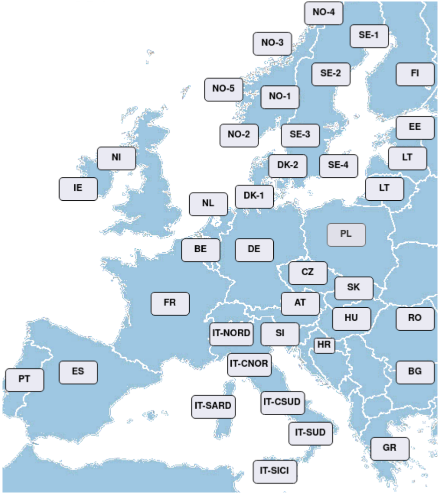




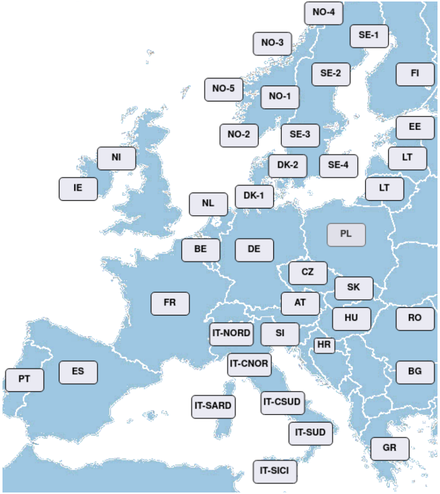
Necessity to explain the forecasts!



Constraining Energy Flow Exchanges in the European Network



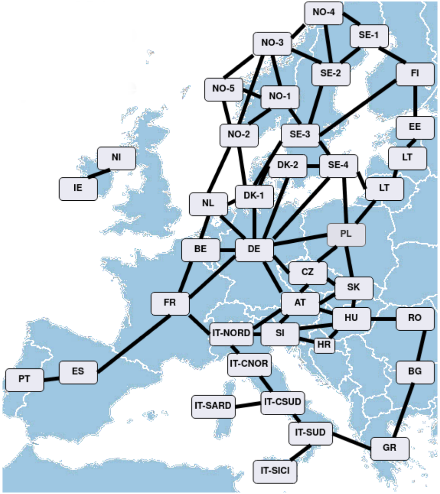
Constraining Energy Flow Exchanges in the European Network



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

How can we forecast the prices of all markets simultaneously?

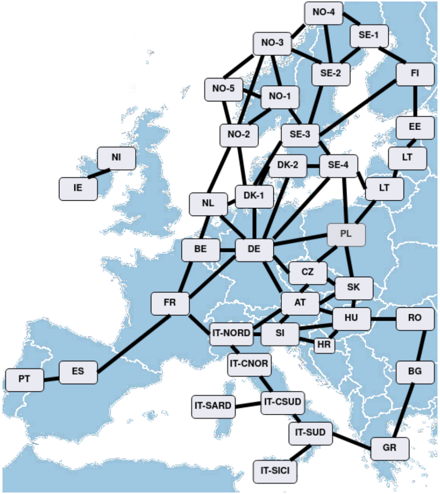
Constraining Energy Flow Exchanges in the European Network



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

How can we forecast the prices of all markets simultaneously?

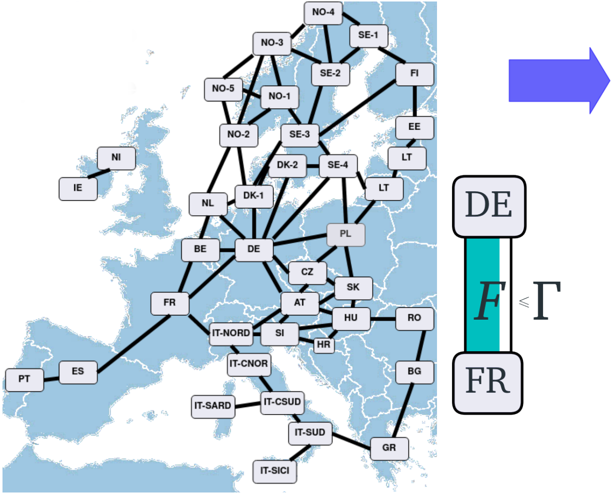
Constraining Energy Flow Exchanges in the European Network



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

How can we forecast the prices of all markets simultaneously?

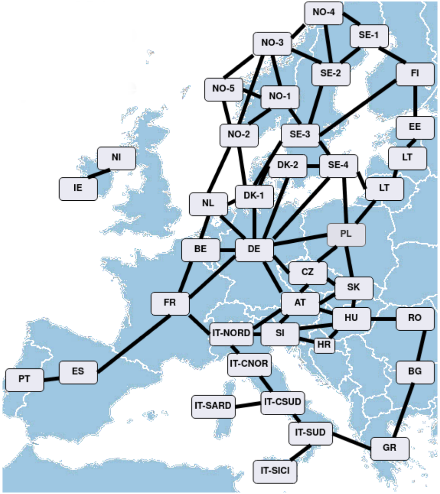
Constraining Energy Flow Exchanges in the European Network



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

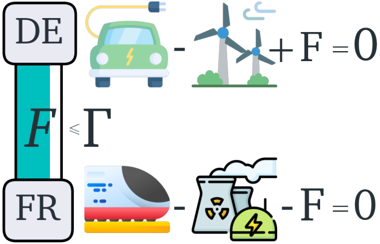
How can we forecast the prices of all markets simultaneously?

Constraining Energy Flow Exchanges in the European Network

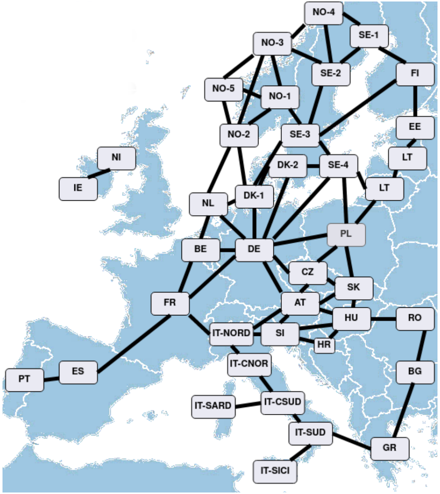


$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

How can we forecast the prices of all markets simultaneously?

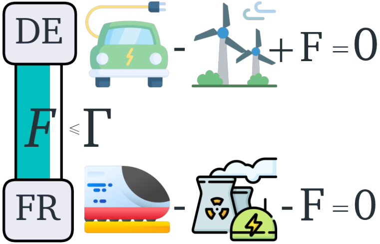


Constraining Energy Flow Exchanges in the European Network



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

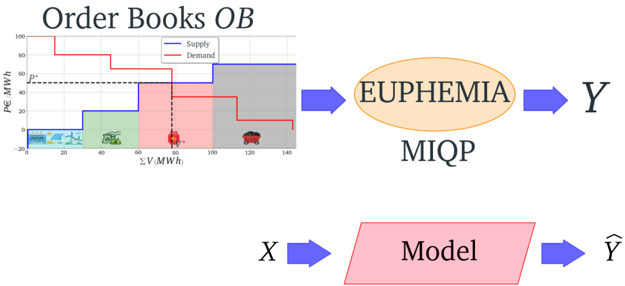
How can we forecast the prices of all markets simultaneously?



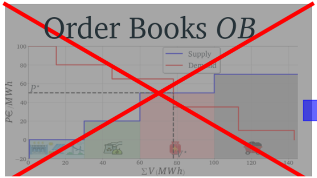
How can we consider the constrained energy flows while forecasting prices?

The Price-Fixing Algorithm

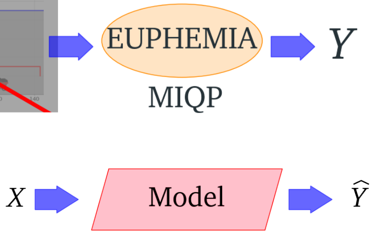
The Price-Fixing Algorithm



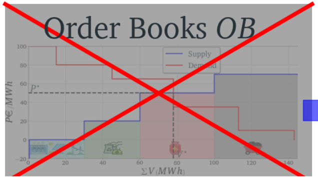
The Price-Fixing Algorithm



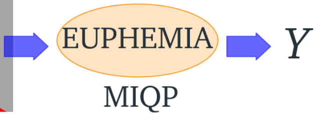
Order Books are not available before price fixation!

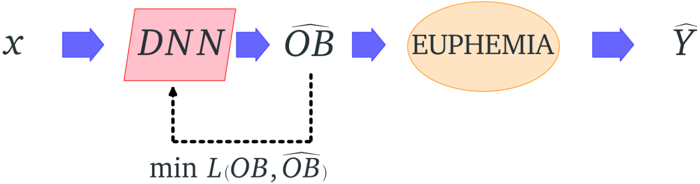


The Price-Fixing Algorithm

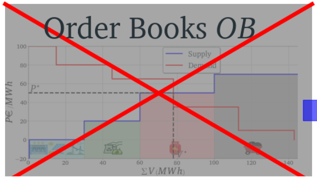


Order Books are not available before price fixation!

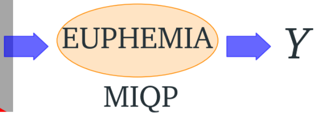


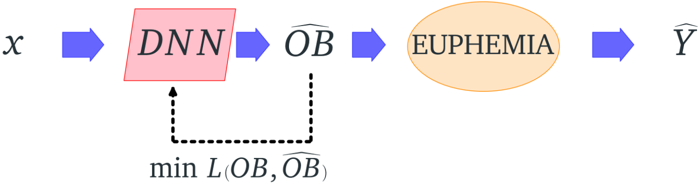


The Price-Fixing Algorithm

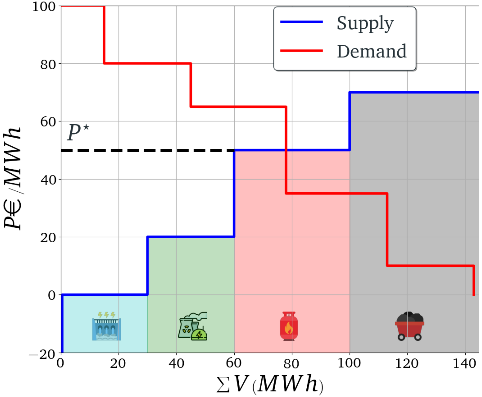


Order Books are not available before price fixation!

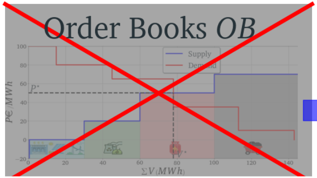




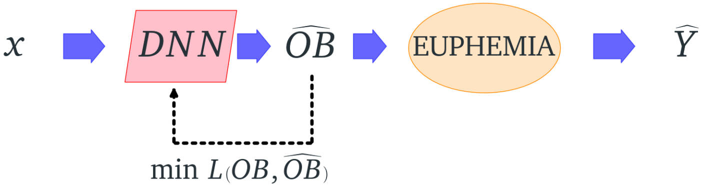
$$\min L(OB, \widehat{OB}) \neq \min L(Y, \widehat{Y})$$



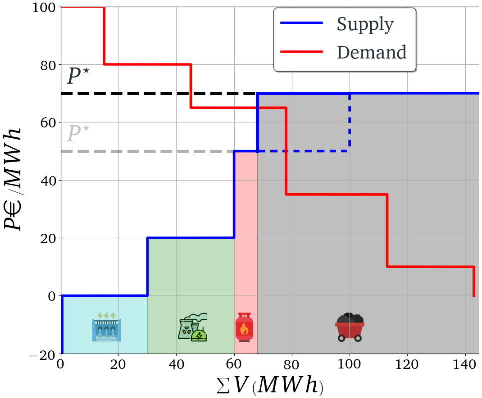
The Price-Fixing Algorithm



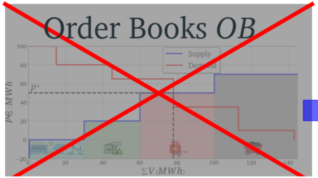
Order Books are not available before price fixation!



$$\min L(OB, \widehat{OB}) \neq \min L(Y, \widehat{Y})$$

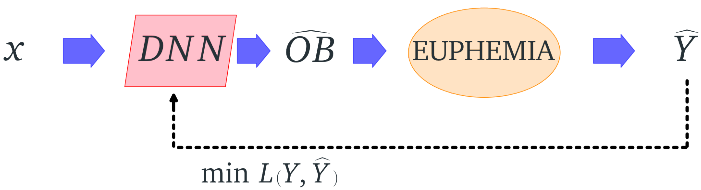


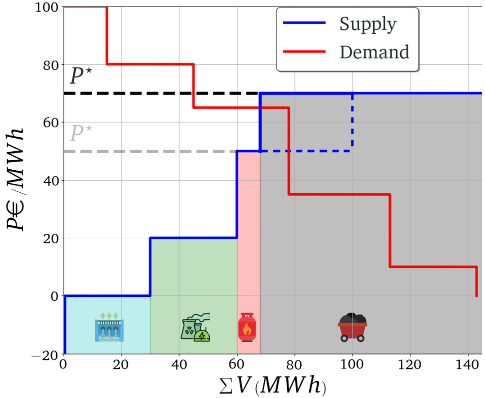
The Price-Fixing Algorithm



Order Books are not available before price fixation!

$$\min L(OB, \widehat{OB}) \neq \min L(Y, \widehat{Y})$$





How can we minimize the Price Forecasting Error?

Plan

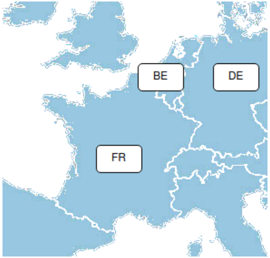
Extending the State-of-the-Art benchmark

J. Lago *et. al* **Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark**, Applied Energy, 2021

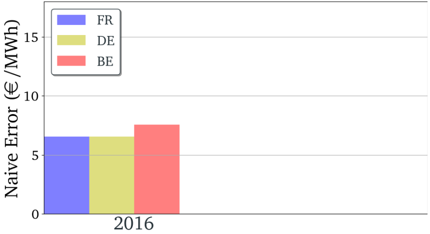
Models:

- Deep Neural Network

Features



Test Period



$\hat{Y}^{(d,h)}_{naive} =$

$Y^{(d-7,h)}$ if d is a week-end

$Y^{(d-1,h)}$ otherwise

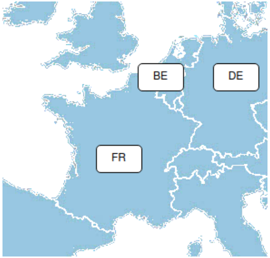
Extending the State-of-the-Art benchmark

J. Lago *et. al* **Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark**, Applied Energy, 2021

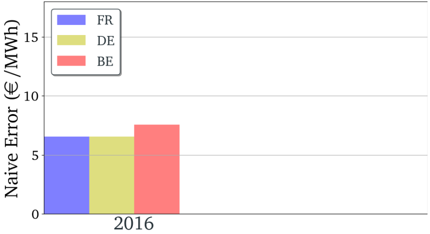
Models:

- Deep Neural Network
- **Convolutional Neural Network**
- **Random Forest**
- **Support Vector (Chain, Multi)**

Features



Test Period



$\widehat{Y}^{(d,h)}_{naive} =$

$Y^{(d-7,h)}$ if d is a week-end

$Y^{(d-1,h)}$ otherwise

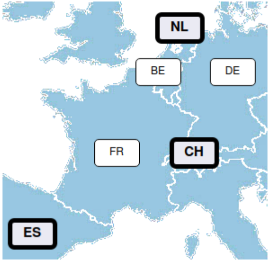
Extending the State-of-the-Art benchmark

J. Lago *et. al* **Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark**, Applied Energy, 2021

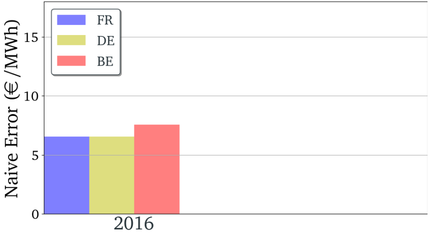
Models:

- Deep Neural Network
- **Convolutional Neural Network**
- **Random Forest**
- **Support Vector (Chain, Multi)**

Features



Test Period



$\widehat{Y}^{(d,h)}_{naive} =$

$Y^{(d-7,h)}$ if d is a week-end

$Y^{(d-1,h)}$ otherwise

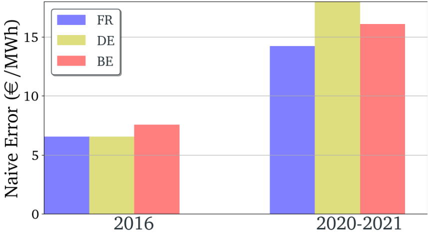
Extending the State-of-the-Art benchmark

J. Lago *et. al* **Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark**, Applied Energy, 2021

Models:

- Deep Neural Network
- **Convolutional Neural Network**
- **Random Forest**
- **Support Vector (Chain, Multi)**

Test Period

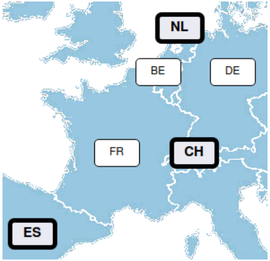


$\widehat{Y}^{(d,h)}_{naive} =$

$Y^{(d-7,h)}$ if d is a week-end

$Y^{(d-1,h)}$ otherwise

Features



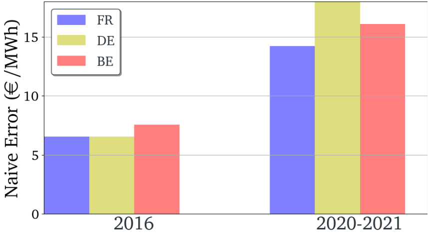
Extending the State-of-the-Art benchmark

J. Lago *et. al* **Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark**, Applied Energy, 2021

Models:

- Deep Neural Network
- **Convolutional Neural Network**
- **Random Forest**
- **Support Vector (Chain, Multi)**

Test Period

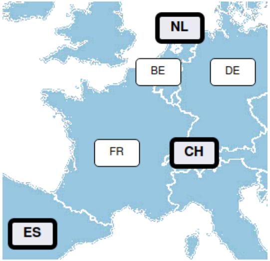


$\widehat{Y}_{naive}^{(d,h)} =$

$Y^{(d-7,h)}$ if d is a week-end

$Y^{(d-1,h)}$ otherwise

Features



Multi-market forecasting model M

$$\boldsymbol{X}_{FR,DE,BE}^{(d)} \rightarrow \boldsymbol{M} \rightarrow [\boldsymbol{Y}_{FR}^{(d)}, \boldsymbol{Y}_{DE}^{(d)}, \boldsymbol{Y}_{BE}^{(d)}]$$

How well do the model perform?

Recalibration = the model is retrained before each prediction

$$RMAE(Y, \widehat{Y}) = \frac{MAE(Y, \widehat{Y})}{MAE(Y, \widehat{Y}_{naive})} \in [0, 1] \qquad \widehat{Y}_{naive}^{(d,h)} = \begin{cases} Y^{(d-1,h)} & \text{if } d \text{ is a week day} \\ Y^{(d-7,h)} & \text{otherwise} \end{cases}$$

How well do the model perform?

Recalibration = the model is retrained before each prediction

$$RMAE(Y, \widehat{Y}) = \frac{MAE(Y, \widehat{Y})}{MAE(Y, \widehat{Y}_{naive})} \in [0, 1] \qquad \widehat{Y}^{(d,h)}_{naive} = \begin{cases} Y^{(d-1,h)} & \text{if } d \text{ is a week day} \\ Y^{(d-7,h)} & \text{otherwise} \end{cases}$$

Markets	Datasets	CNN	RF	SVR Chain	SVR Multi	SOTA DNN
FR	SOTA	0.64	0.71	**0.60**	0.61	0.62
	Enriched	0.69	0.61	**0.54**	0.55	0.58
	Multi-Market	0.59	0.64	**0.55**	**0.55**	0.57
	[2020-2021]	0.73	0.66	0.48	**0.46**	0.56
DE	SOTA	0.50	0.57	**0.45**	**0.45**	**0.45**
	Enriched	0.44	0.51	**0.43**	0.45	**0.43**
	Multi-Market	**0.45**	0.57	**0.45**	**0.45**	**0.45**
	[2020-2021]	0.47	0.58	0.46	0.48	**0.42**
BE	SOTA	0.73	0.74	0.73	**0.71**	**0.71**
	Enriched	0.70	0.74	**0.69**	0.70	0.72
	Multi-Market	0.68	0.75	**0.67**	**0.67**	**0.67**
	[2020-2021]	0.88	0.76	**0.58**	0.59	0.73

How well do the model perform?

Recalibration = the model is retrained before each prediction

$$RMAE(Y, \widehat{Y}) = \frac{MAE(Y, \widehat{Y})}{MAE(Y, \widehat{Y}_{naive})} \in [0, 1] \qquad \widehat{Y}^{(d,h)}_{naive} = \begin{cases} Y^{(d-1,h)} & \text{if } d \text{ is a week day} \\ Y^{(d-7,h)} & \text{otherwise} \end{cases}$$

Markets	Datasets	CNN	RF	SVR Chain	SVR Multi	SOTA DNN
FR	SOTA	0.64	**0.71**	**0.60**	**0.61**	**0.62**
	Enriched	0.69	**0.61**	**0.54**	**0.55**	**0.58**
	Multi-Market	0.59	0.64	0.55	0.55	0.57
	[2020-2021]	0.73	0.66	0.48	0.46	0.56
DE	SOTA	**0.50**	**0.57**	**0.45**	0.45	**0.45**
	Enriched	**0.44**	**0.51**	**0.43**	0.45	**0.43**
	Multi-Market	0.45	0.57	0.45	0.45	0.45
	[2020-2021]	0.47	0.58	0.46	0.48	0.42
BE	SOTA	**0.73**	0.74	**0.73**	**0.71**	0.71
	Enriched	**0.70**	0.74	**0.69**	**0.70**	0.72
	Multi-Market	0.68	0.75	0.67	0.67	0.67
	[2020-2021]	0.88	0.76	0.58	0.59	0.73

Improvement (FR: SOTA → Enriched)

Improvement (DE: SOTA → Enriched)

Improvement (BE: SOTA → Enriched)

How well do the model perform?

Recalibration = the model is retrained before each prediction

$$RMAE(Y, \widehat{Y}) = \frac{MAE(Y, \widehat{Y})}{MAE(Y, \widehat{Y}_{naive})} \in [0, 1] \qquad \widehat{Y}^{(d,h)}_{naive} = \begin{cases} Y^{(d-1,h)} & \text{if } d \text{ is a week day} \\ Y^{(d-7,h)} & \text{otherwise} \end{cases}$$

Markets	Datasets	CNN	RF	SVR Chain	SVR Multi	SOTA DNN
FR	SOTA	0.64	0.71	0.60	0.61	0.62
	Enriched	**0.69**	0.61	0.54	0.55	**0.58**
	Multi-Market	0.59	0.64	0.55	0.55	0.57
	[2020-2021]	0.73	0.66	0.48	0.46	0.56
DE	SOTA	0.50	0.57	0.45	0.45	0.45
	Enriched	0.44	0.51	0.43	0.45	0.43
	Multi-Market	0.45	0.57	0.45	0.45	0.45
	[2020-2021]	0.47	0.58	0.46	0.48	0.42
BE	SOTA	0.73	0.74	0.73	0.71	0.71
	Enriched	**0.70**	0.74	**0.69**	**0.70**	**0.72**
	Multi-Market	0.68	0.75	0.67	0.67	0.67
	[2020-2021]	0.88	0.76	0.58	0.59	0.73

FR: Partial improvement

DE: No improvement

BE: Improvement

How well do the model perform?

Recalibration = the model is retrained before each prediction

$$RMAE(Y,\widehat{Y}) = \frac{MAE(Y,\widehat{Y})}{MAE(Y,\widehat{Y}_{naive})} \in [0,1] \qquad \widehat{Y}^{(d,h)}_{naive} = \begin{cases} Y^{(d-1,h)} & \text{if } d \text{ is a week day} \\ Y^{(d-7,h)} & \text{otherwise} \end{cases}$$

Markets	Datasets	CNN	RF	SVR Chain	SVR Multi	SOTA DNN
FR	SOTA	0.64	0.71	0.60	0.61	0.62
	Enriched	0.69	0.61	0.54	0.55	0.58
	Multi-Market	0.59	0.64	**0.55**	**0.55**	**0.57**
	[2020-2021]	0.73	0.66	**0.48**	**0.46**	**0.56**
DE	SOTA	0.50	0.57	0.45	0.45	0.45
	Enriched	0.44	0.51	0.43	0.45	0.43
	Multi-Market	0.45	0.57	0.45	0.45	**0.45**
	[2020-2021]	0.47	0.58	0.46	0.48	**0.42**
BE	SOTA	0.73	0.74	0.73	0.71	0.71
	Enriched	0.70	0.74	0.69	0.70	0.72
	Multi-Market	0.68	0.75	**0.67**	**0.67**	0.67
	[2020-2021]	0.88	0.76	**0.58**	**0.59**	0.73

FR: Improvement

DE: Partial Improvement

BE: Improvement

Explaining the DDN's German price forecasts using Shap Values

S. Lundberg *et al.* **A Unified Approach to Interpreting Model Predictions.**, NIPS 2017

$$\widehat{Y}^{(d,h)} = \sum_{f,l,h'} \Phi_{f,l,h'}^{(d,h)}$$

Explaining the DDN's German price forecasts using Shap Values

S. Lundberg *et al.* **A Unified Approach to Interpreting Model Predictions.**, NIPS 2017

$$\widehat{Y}^{(d,h)} = \sum_{f,l,h'} \Phi_{f,l,h'}^{(d,h)}$$

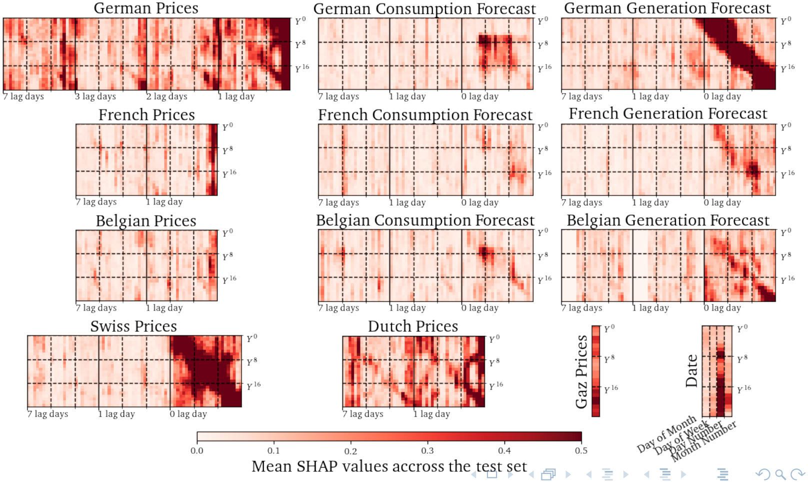
$$\bar{\Phi}_{f,l,h'}^{(h)} = \frac{1}{n_d} \sum_{d} \Phi_{f,l,h'}^{(d,h)}$$

Explaining the DDN's German price forecasts using Shap Values

S. Lundberg *et al.* **A Unified Approach to Interpreting Model Predictions.**, NIPS 2017

$$\widehat{Y}^{(d,h)} = \sum_{f,l,h'} \Phi_{f,l,h'}^{(d,h)}$$

$$\bar{\Phi}_{f,l,h'}^{(h)} = \frac{1}{n_d} \sum_{d} \Phi_{f,l,h'}^{(d,h)}$$



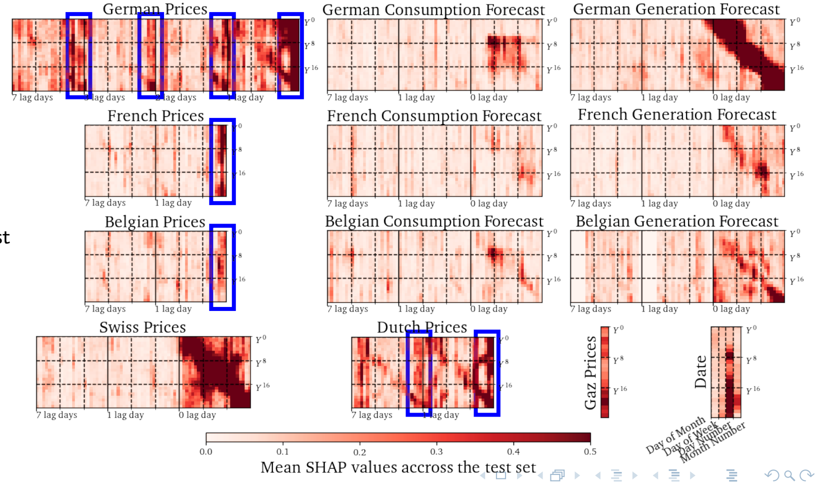
Explaining the DDN's German price forecasts using Shap Values

S. Lundberg *et al.* **A Unified Approach to Interpreting Model Predictions.**, NIPS 2017

$$\widehat{Y}^{(d,h)} = \sum_{f,l,h'} \Phi^{(d,h)}_{f,l,h'}$$

$$\bar{\Phi}^{(h)}_{f,l,h'} = \frac{1}{n_d} \sum_{d} \Phi^{(d,h)}_{f,l,h'}$$

- Vertical lines for end-of-the-day Past Prices



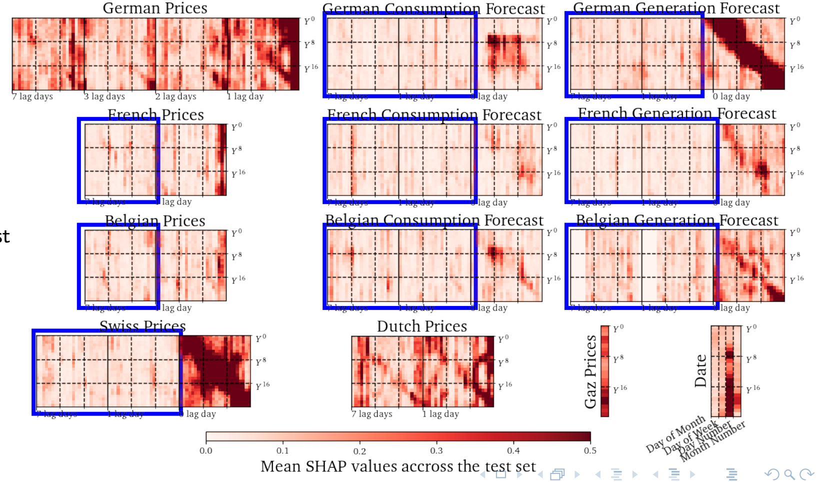
Explaining the DDN's German price forecasts using Shap Values

S. Lundberg *et al.* **A Unified Approach to Interpreting Model Predictions.**, NIPS 2017

$$\widehat{Y}^{(d,h)} = \sum_{f,l,h'} \Phi_{f,l,h'}^{(d,h)}$$

$$\bar{\Phi}_{f,l,h'}^{(h)} = \frac{1}{n_d} \sum_{d} \Phi_{f,l,h'}^{(d,h)}$$

- Vertical lines for end-of-the-day Past Prices
- Past Features are not important

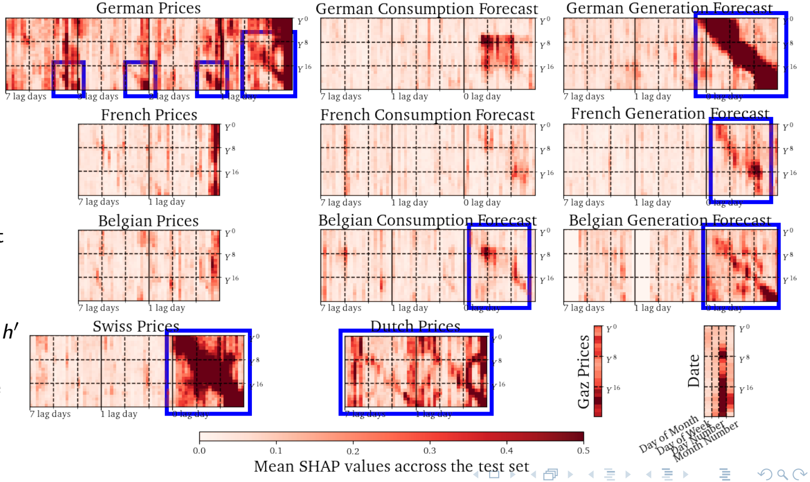


Explaining the DDN's German price forecasts using Shap Values

S. Lundberg *et al.* **A Unified Approach to Interpreting Model Predictions.**, NIPS 2017

$$\widehat{Y}^{(d,h)} = \sum_{f,l,h'} \Phi_{f,l,h'}^{(d,h)}$$

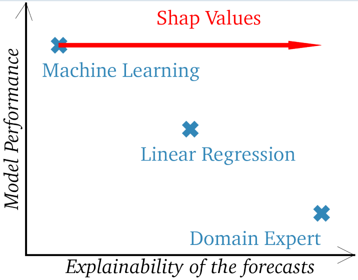
$$\bar{\Phi}_{f,l,h'}^{(h)} = \frac{1}{n_d} \sum_{d} \Phi_{f,l,h'}^{(d,h)}$$



- Vertical lines for end-of-the-day Past Prices
- Past Features are not important
- Diagonals: $\bar{\Phi}_{f,l,h'}^{(h)}$ is high when $h = h'$
  Importance of Generation Forecasts and Foreign Prices (Switzerland, the Netherlands)

Synthesis

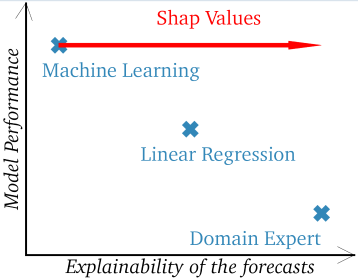
Using a **SVR** or a **DNN** combined with **Shap Values** bridges the gap between forecasts explainability and model performance



Synthesis

Using a **SVR** or a **DNN** combined with **Shap Values** bridges the gap between forecasts explainability and model performance

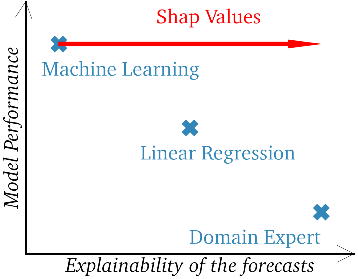
Electricity price forecasting on the day-ahead market using machine learning Applied Energy 313 (2022)



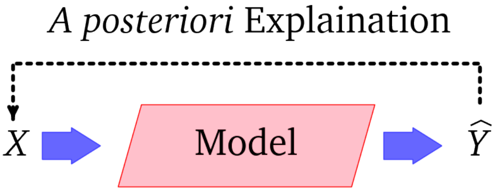
Synthesis

Using a **SVR** or a **DNN** combined with **Shap Values** bridges the gap between forecasts explainability and model performance

Electricity price forecasting on the day-ahead market using machine learning Applied Energy 313 (2022)



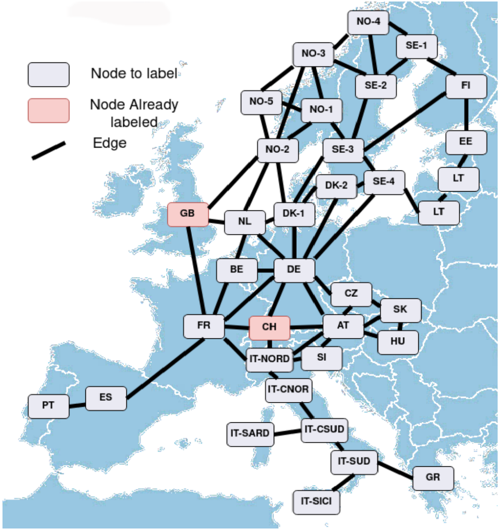
Shap Values link predictions with **Domain-Knowledge**



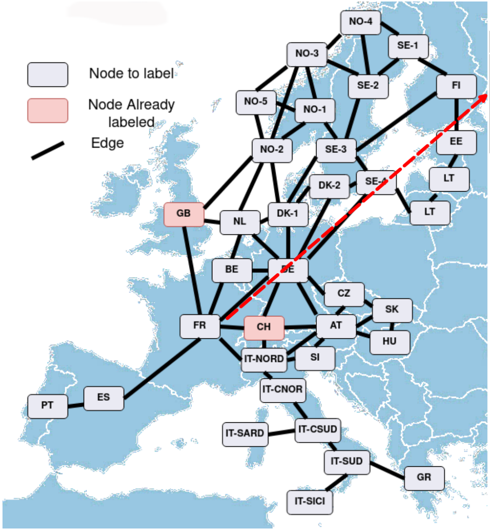
Plan

1. Introduction
2. Explaining the Forecasts
3. Optimize-then-Predict approach
4. A differentiable Optimization Approach
5. Conclusion

Modeling the European network as a Graph

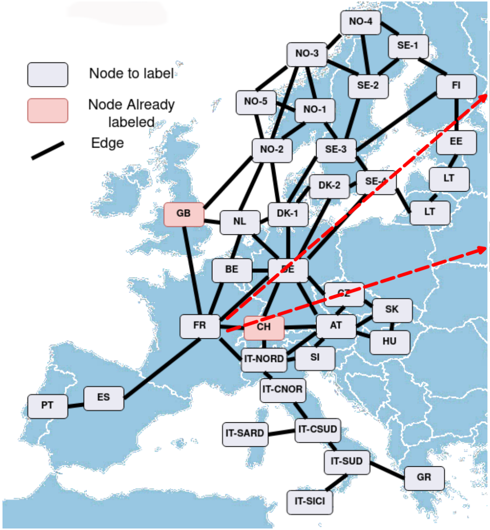


Modeling the European network as a Graph



Attributes	
$\mathbf{Y}_{FR}^{(d-1)}$	Past prices
$\mathbf{C}_{FR}^{(d)}$	Consumption Forecast
$\mathbf{G}_{FR}^{(d)}$	Programmable Generation Forecast
$\mathbf{R}_{FR}^{(d)}$	Renewables Generation Forecast

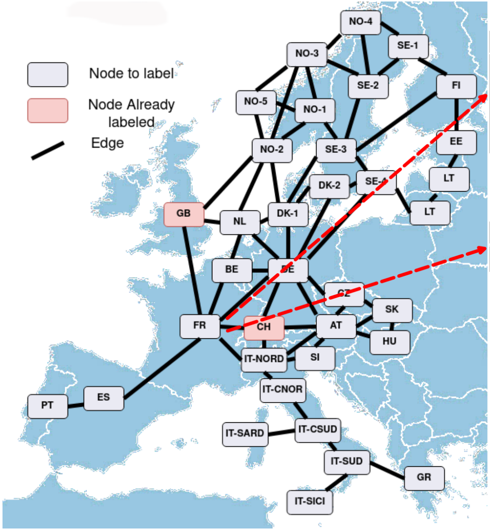
Modeling the European network as a Graph



Attributes	
$\mathbf{Y}_{FR}^{(d-1)}$	Past prices
$\mathbf{C}_{FR}^{(d)}$	Consumption Forecast
$\mathbf{G}_{FR}^{(d)}$	Programmable Generation Forecast
$\mathbf{R}_{FR}^{(d)}$	Renewables Generation Forecast

Labels	
$\mathbf{Y}_{FR}^{(d)}$	Day-Ahead Prices

Modeling the European network as a Graph



Attributes	
$\mathbf{Y}_{FR}^{(d-1)}$	Past prices
$\mathbf{C}_{FR}^{(d)}$	Consumption Forecast
$\mathbf{G}_{FR}^{(d)}$	Programmable Generation Forecast
$\mathbf{R}_{FR}^{(d)}$	Renewables Generation Forecast

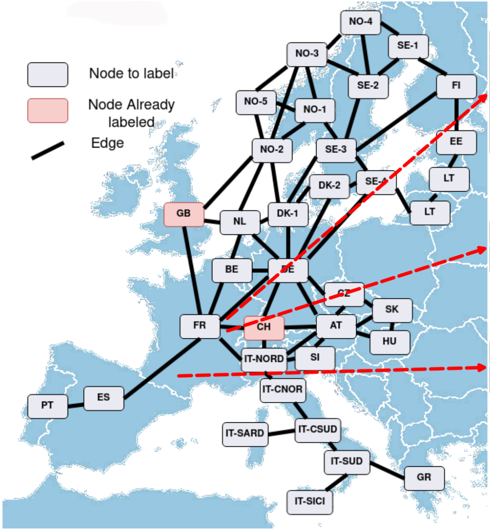
Labels	
$\mathbf{Y}_{FR}^{(d)}$	Day-Ahead Prices

Node-Labeling task

Graph Neural Network

$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

Modeling the European network as a Graph



Attributes	
$\mathbf{Y}_{FR}^{(d-1)}$	Past prices
$\mathbf{C}_{FR}^{(d)}$	Consumption Forecast
$\mathbf{G}_{FR}^{(d)}$	Programmable Generation Forecast
$\mathbf{R}_{FR}^{(d)}$	Renewables Generation Forecast

Labels	
$\mathbf{Y}_{FR}^{(d)}$	Day-Ahead Prices

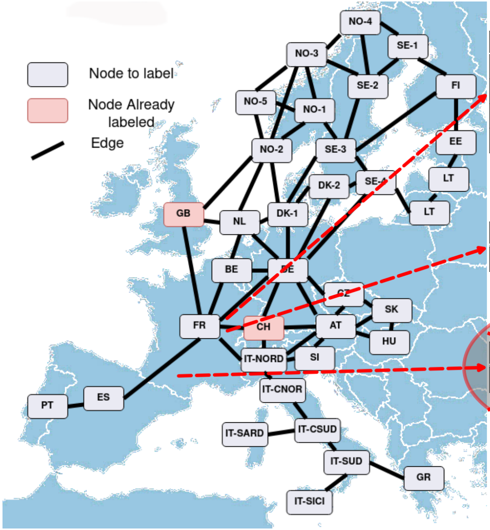
Attributes	
$\mathbf{F}_{FR,ES}^{(d)}$	Cross-Market Flows

Node-Labeling task

Graph Neural Network

$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

Modeling the European network as a Graph



Attributes	
$\mathbf{Y}_{FR}^{(d-1)}$	Past prices
$\mathbf{C}_{FR}^{(d)}$	Consumption Forecast
$\mathbf{G}_{FR}^{(d)}$	Programmable Generation Forecast
$\mathbf{R}_{FR}^{(d)}$	Renewables Generation Forecast

Labels	
$\mathbf{Y}_{FR}^{(d)}$	Day-Ahead Prices

Attributes	
$\mathbf{F}_{FR,ES}^{(d)}$	Cross-Market Flows

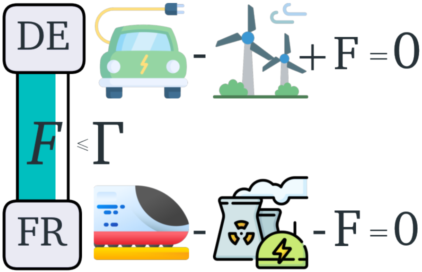
Node-Labeling task

Graph Neural Network

$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

F is not available, but we can approximate it using the maximal capacity Γ

Estimating the flows using an Optimization problem

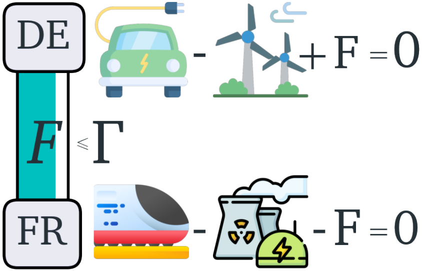


Estimating the flows using an Optimization problem

Attributes	
$\mathbf{Y}^{(d-1)}$	Past prices
C	Consumption Forecast
G	Programmable Generation Forecast
R	Renewables Generation Forecast

$$Fs = \arg\max_{F_{z,z'}} \sum_{z,z'} F_{z,z'}\left(Y_{z'}^{(d-1)} - Y_{z}^{(d-1)}\right)$$

$$\text{under const.} \begin{cases} G_z + R_z - C_z + \sum_{z'} F_{z',z} - \sum_{z'} F_{z,z'} = 0 & \forall z \\ F_{z,z'} \leq \Gamma_{z,z'} & \forall z, z' \end{cases}$$



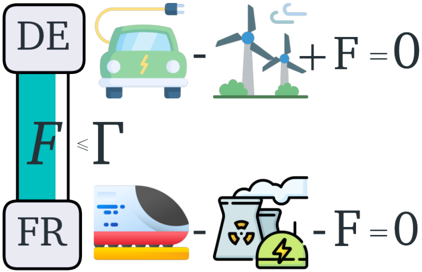
Estimating the flows using an Optimization problem

Attributes	
$\mathbf{Y}^{(d-1)}$	Past prices
C	Consumption Forecast
G	Programmable Generation Forecast
R	Renewables Generation Forecast

$$Fs = \arg\max_{F_{z,z'}} \sum_{z,z'} F_{z,z'}\left(Y_{z'}^{(d-1)} - Y_{z}^{(d-1)}\right)$$

$$\textit{under const.} \begin{cases} G_z + R_z - C_z + \sum_{z'} F_{z',z} - \sum_{z'} F_{z,z'} = 0 & \forall z \\ F_{z,z'} \leq \Gamma_{z,z'} & \forall z, z' \end{cases}$$

Impossible to enforce using forecasts **G**, **R**, **C**



Estimating the flows using an Optimization problem

Attributes	
$\mathbf{Y}^{(d-1)}$	Past prices
C	Consumption Forecast
G	Programmable Generation Forecast
R	Renewables Generation Forecast

$$Fs = \arg\max_{F_{z,z'}} \sum_{z,z'} F_{z,z'}(Y_{z'}^{(d-1)} - Y_z^{(d-1)})$$

$$\textit{under const.} \begin{cases} G_z + R_z - C_z + \sum_{z'} F_{z',z} - \sum_{z'} F_{z,z'} = 0 & \forall z \\ F_{z,z'} \leq \Gamma_{z,z'} & \forall z, z' \end{cases}$$

Impossible to enforce using forecasts **G, R, C**

Using Programmable Generation **E** as an optimization variable

Flin

$$\arg\max_{F_{z,z'},\, E_z} \sum_{z,z'} F_{z,z'}(P_{z'} - P_z)$$

$$\begin{cases} E_z + R_z - C_z + \sum_{z'} F_{z,z'} - \sum_{z'} F_{z',z} = 0 & \forall z \\ 0 \leq F_{z,z'} \leq A_{z,z'} & \forall z, z' \\ 0 \leq E_z \leq V_z & \forall z \end{cases}$$

Estimating the flows using an Optimization problem

Attributes	
$\mathbf{Y}^{(d-1)}$	Past prices
C	Consumption Forecast
G	Programmable Generation Forecast
R	Renewables Generation Forecast

$$Fs = \underset{F_{z,z'}}{\arg\max} \sum_{z,z'} F_{z,z'}\left(Y_{z'}^{(d-1)} - Y_{z}^{(d-1)}\right)$$

$$\text{under const.} \begin{cases} G_z + R_z - C_z + \sum_{z'} F_{z',z} - \sum_{z'} F_{z,z'} = 0 & \forall z \\ F_{z,z'} \leq \Gamma_{z,z'} & \forall z, z' \end{cases}$$

Impossible to enforce using forecasts **G**, **R**, **C**

Using Programmable Generation **E** as an optimization variable

Flin

$$\underset{F_{z,z'},\, E_z}{\arg\max} \sum_{z,z'} F_{z,z'}(P_{z'} - P_z)$$

$$\begin{cases} E_z + R_z - C_z + \sum_{z'} F_{z,z'} - \sum_{z'} F_{z',z} = 0 & \forall z \\ 0 \leq F_{z,z'} \leq A_{z,z'} & \forall z, z' \\ 0 \leq E_z \leq V_z & \forall z \end{cases}$$

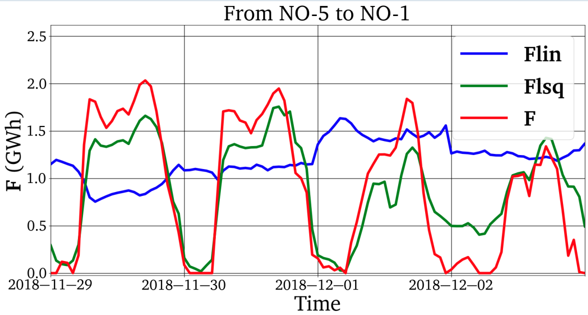
Penalize deviation from the **Energy Balance**

Flsq

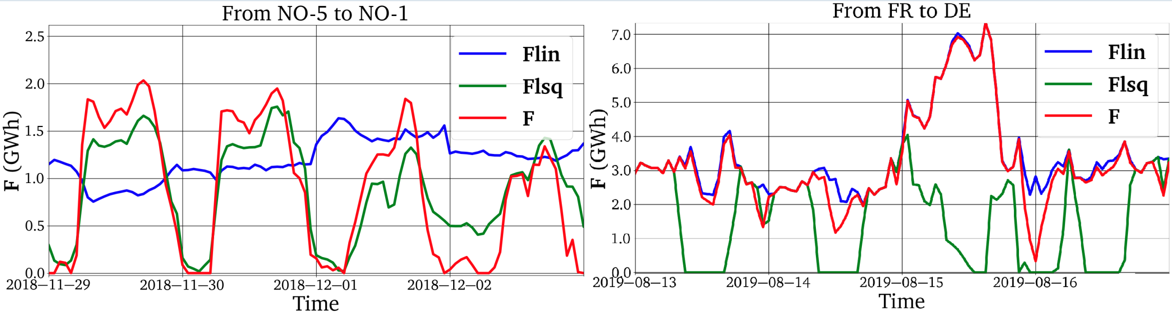
$$\underset{F_{z,z'}}{\arg\min} \sum_{z} \left(R_z + G_z - C_z + \sum_{z'} F_{z',z} - \sum_{z'} F_{z,z'} \right)^2$$

$$u.c \;\; 0 \leq F_{z,z'} \leq A_{z,z'} \; \forall_{z,z'}$$

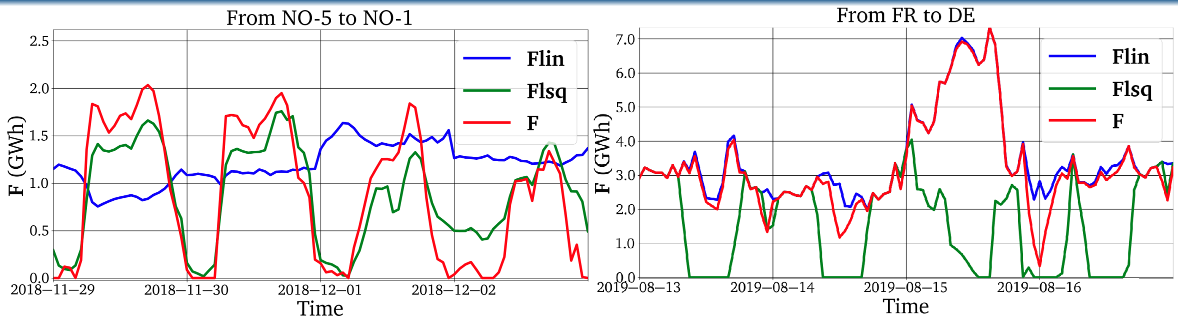
Combining the Flow estimates



Combining the Flow estimates

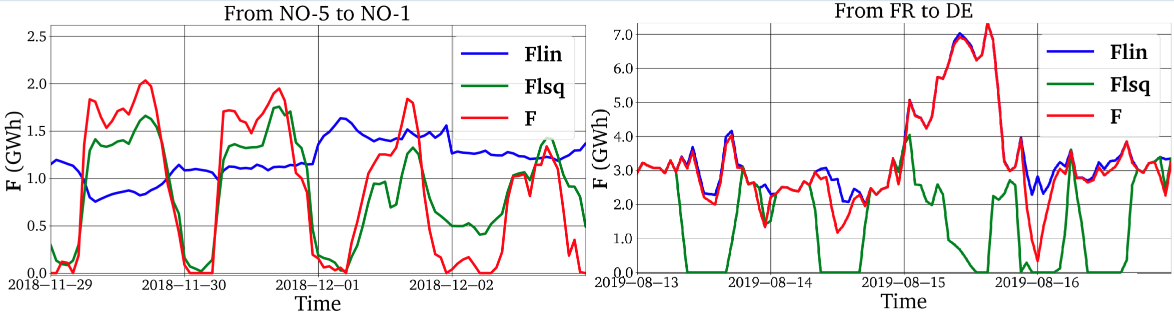


Combining the Flow estimates



$$L^{(t)}(z, z') = |\mathbf{F}^{(t)}_{z,z'} - \mathbf{Flsq}^{(t)}_{z,z'}| - |\mathbf{F}^{(t)}_{z,z'} - \mathbf{Flin}^{(t)}_{z,z'}|$$

Combining the Flow estimates

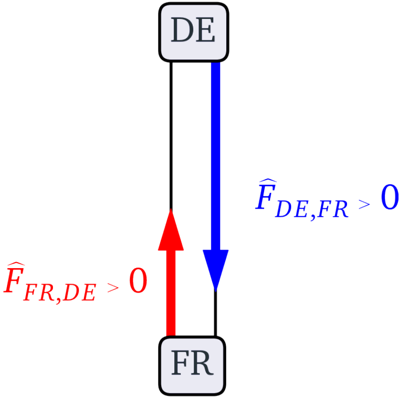


$$L^{(t)}(z,z') = |\mathbf{F}^{(t)}_{z,z'} - \mathbf{Flsq}^{(t)}_{z,z'}| - |\mathbf{F}^{(t)}_{z,z'} - \mathbf{Flin}^{(t)}_{z,z'}|$$

$$\mathbf{Fcmb} = \begin{cases} \mathbf{Flin} & \text{if } L^{(t)}(z,z') > 0 \\ \mathbf{Flsq} & \text{otherwise} \end{cases}$$

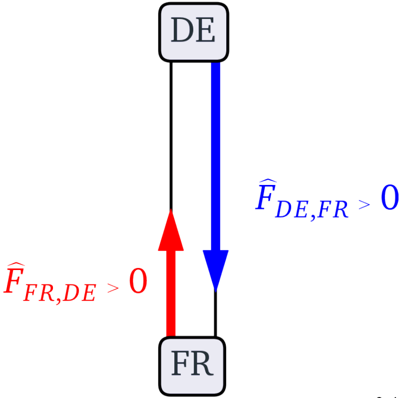
Enforcing One-sided flows

Bilateral flows



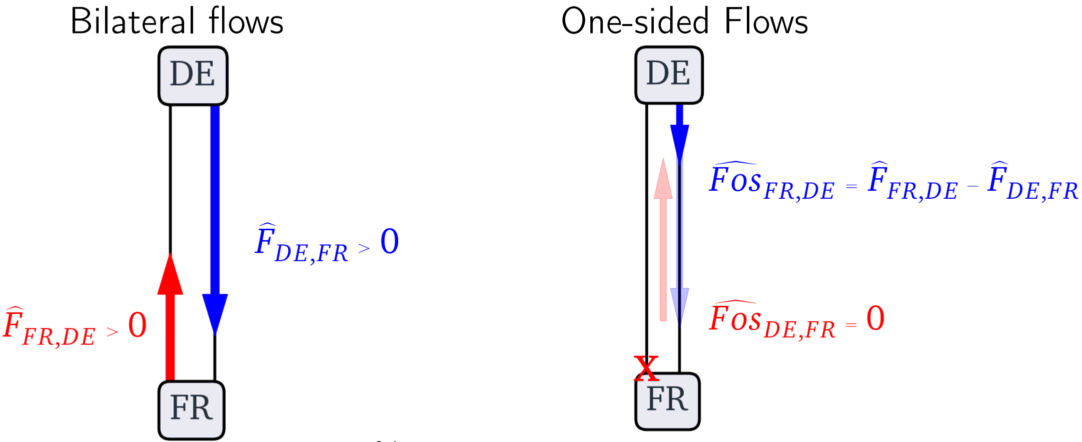
Enforcing One-sided flows

Bilateral flows



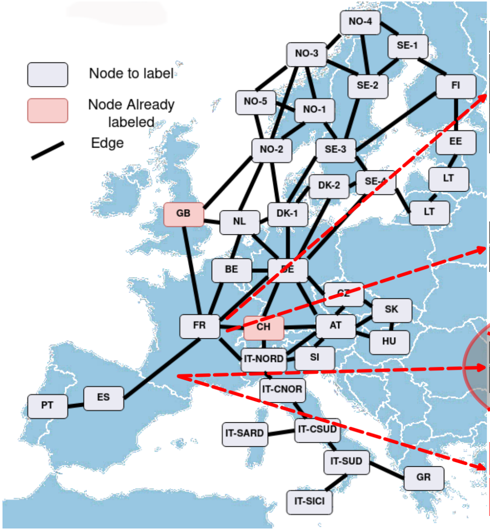
If a connection is 75 % one sided in the dataset, we always apply One-Sideness to its flows

Enforcing One-sided flows



If a connection is 75 % one sided in the dataset, we always apply One-Sideness to its flows

An Optimize-then Predict approach



Attributes	
$\mathbf{Y}_{FR}^{(d-1)}$	Past prices
$\mathbf{C}_{FR}^{(d)}$	Consumption Forecast
$\mathbf{G}_{FR}^{(d)}$	Programmable Generation Forecast
$\mathbf{R}_{FR}^{(d)}$	Renewables Generation Forecast

Labels	
$\mathbf{Y}_{FR}^{(d)}$	Day-Ahead Prices

→ Graph Neural Network →

$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

Attributes	
$\mathbf{F}_{FR,ES}^{(d)}$	Cross-Market Flows

Attributes	
$\widehat{\mathbf{F}}_{FR,ES}^{(d)}$	Estimated Flows

← Optimization Problem

How does the flow estimates improve price forecast quality?

$\widehat{F}$	SMAPE	
Γ		DNN
Flin		DNN
Flsq		DNN
Fcmb		DNN
Fos		DNN
Γ		CNN
Flin		CNN
Flsq		CNN
Fcmb		CNN
Fos		CNN
Γ		GNN
Flin		GNN
Flsq		GNN
Fcmb		GNN
Fos		GNN

How does the flow estimates improve price forecast quality?

	$\widehat{F}$	SMAPE
DNN	Γ	29.76
	Flin	28.51
	Flsq	28.26
	Fcmb	28.38
	Fos	28.84
CNN	Γ	32.17
	Flin	32.23
	Flsq	32.01
	Fcmb	31.81
	Fos	31.87
GNN	Γ	24.59
	Flin	24.6
	Flsq	24.6
	Fcmb	24.46
	Fos	24.52

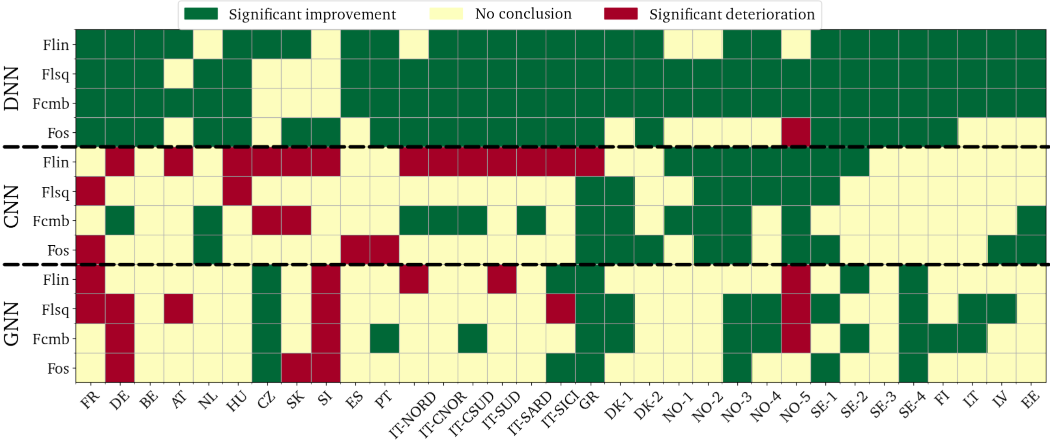
How does the flow estimates improve price forecast quality?

$\hat{F}$	SMAPE	
Γ	29.76	DNN
Flin	28.51	DNN
Flsq	28.26	DNN
Fcmb	28.38	DNN
Fos	28.84	DNN
Γ	32.17	CNN
Flin	32.23	CNN
Flsq	32.01	CNN
Fcmb	31.81	CNN
Fos	31.87	CNN
Γ	**24.59**	GNN
Flin	**24.6**	GNN
Flsq	**24.6**	GNN
Fcmb	**24.46**	GNN
Fos	**24.52**	GNN

How does the flow estimates improve price forecast quality?

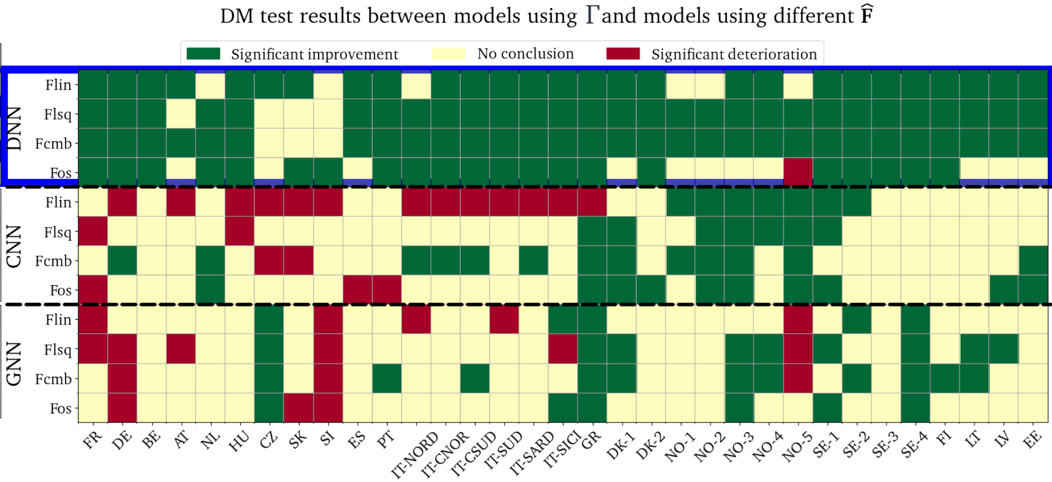
$\widehat{F}$	SMAPE
Γ	29.76
Flin	28.51
Flsq	28.26
Fcmb	28.38
Fos	28.84
Γ	32.17
Flin	32.23
Flsq	32.01
Fcmb	31.81
Fos	31.87
Γ	24.59
Flin	24.6
Flsq	24.6
Fcmb	24.46
Fos	24.52

DM test results between models using Γ and models using different $\widehat{\mathbf{F}}$



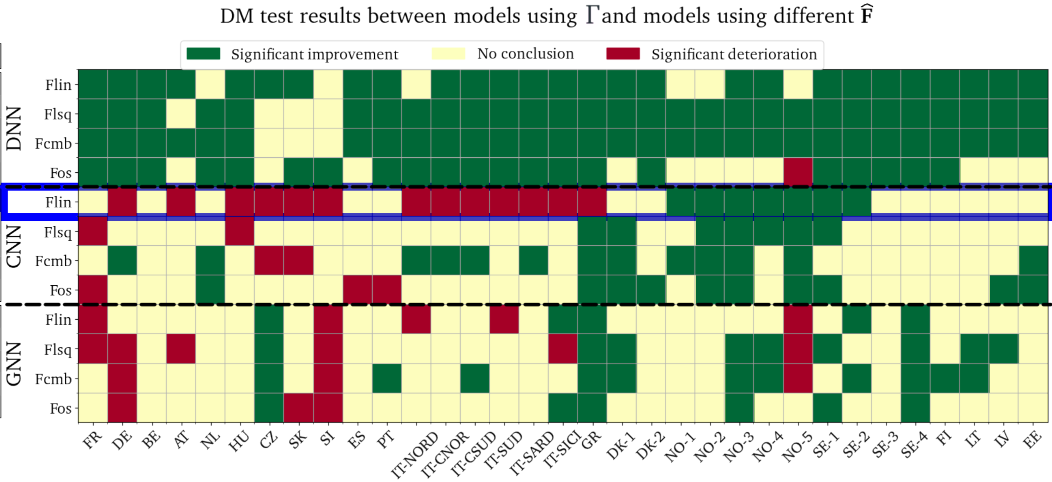
How does the flow estimates improve price forecast quality?

$\widehat{F}$	SMAPE
Γ	29.76
Flin	28.51
Flsq	28.26
Fcmb	28.38
Fos	28.84
Γ	32.17
Flin	32.23
Flsq	32.01
Fcmb	31.81
Fos	31.87
Γ	24.59
Flin	24.6
Flsq	24.6
Fcmb	24.46
Fos	24.52



How does the flow estimates improve price forecast quality?

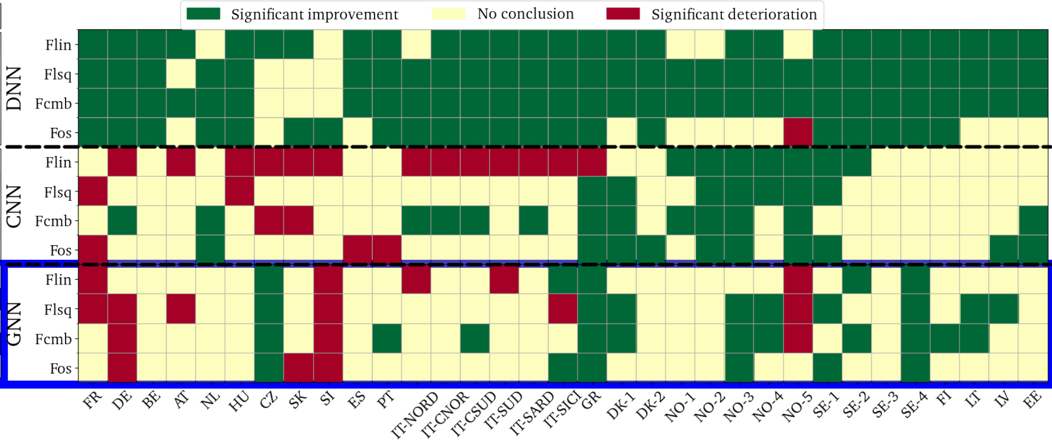
$\hat{F}$	SMAPE
Γ	29.76
Flin	28.51
Flsq	28.26
Fcmb	28.38
Fos	28.84
Γ	32.17
Flin	32.23
Flsq	32.01
Fcmb	31.81
Fos	31.87
Γ	24.59
Flin	24.6
Flsq	24.6
Fcmb	24.46
Fos	24.52



How does the flow estimates improve price forecast quality?

$\widehat{F}$	SMAPE
Γ	29.76
Flin	28.51
Flsq	28.26
Fcmb	28.38
Fos	28.84
Γ	32.17
Flin	32.23
Flsq	32.01
Fcmb	31.81
Fos	31.87
Γ	24.59
Flin	24.6
Flsq	24.6
Fcmb	24.46
Fos	24.52

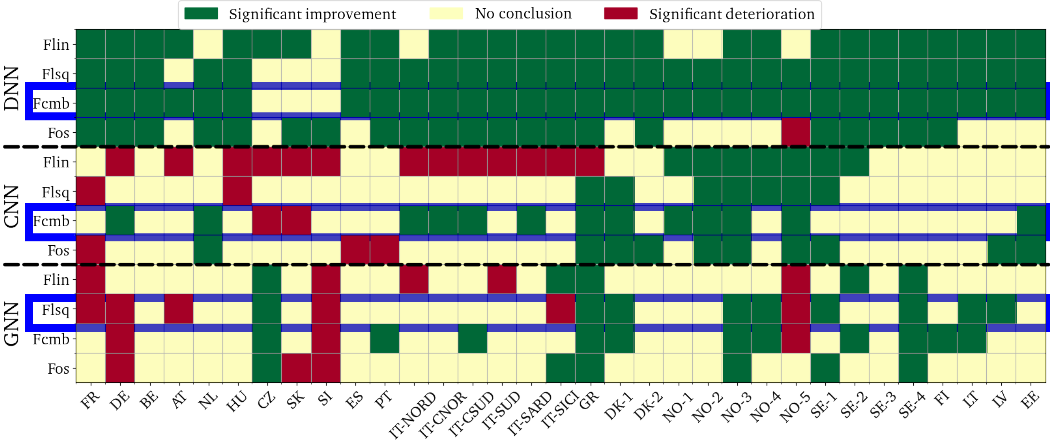
DM test results between models using Γ and models using different $\widehat{F}$



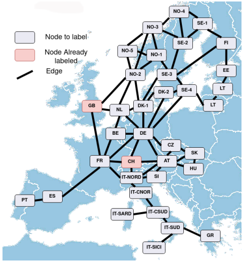
How does the flow estimates improve price forecast quality?

$\hat{F}$	SMAPE
Γ	29.76
Flin	28.51
Flsq	28.26
Fcmb	28.38
Fos	28.84
Γ	32.17
Flin	32.23
Flsq	32.01
Fcmb	31.81
Fos	31.87
Γ	24.59
Flin	24.6
Flsq	24.6
Fcmb	24.46
Fos	24.52

DM test results between models using Γ and models using different $\hat{\mathbf{F}}$



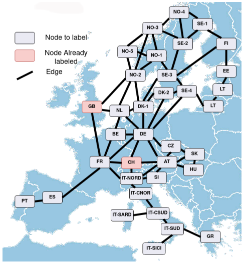
Synthesis



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

Multi-Market Forecasting using a **Graph Network Constraints** considered using **Flow Estimation Problems**

Synthesis

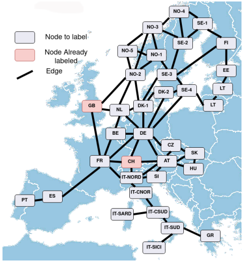


$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

Multi-Market Forecasting using a **Graph Network Constraints** considered using **Flow Estimation Problems**

Forecasting Electricity Prices: An Optimize Then Predict-Based Approach, IDA 2023.

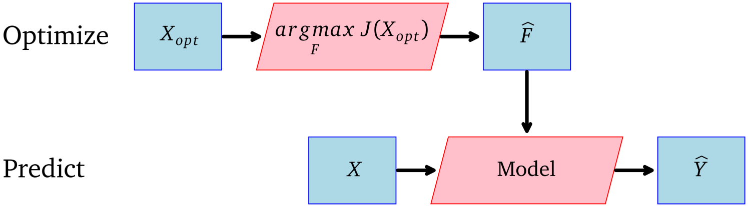
Synthesis



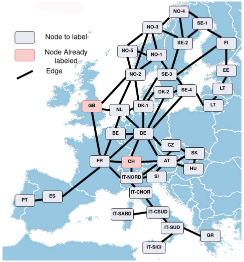
$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

Multi-Market Forecasting using a **Graph Network Constraints** considered using **Flow Estimation Problems**

Forecasting Electricity Prices: An Optimize Then Predict-Based Approach, IDA 2023.



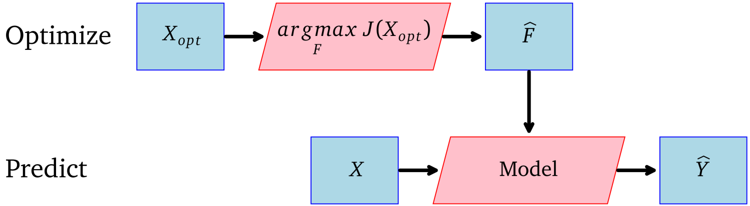
Synthesis



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

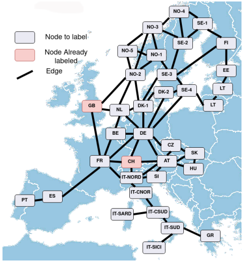
Multi-Market Forecasting using a **Graph Network Constraints** considered using **Flow Estimation Problems**

Forecasting Electricity Prices: An Optimize Then Predict-Based Approach, IDA 2023.



Domain-Knowledge integrated in the input

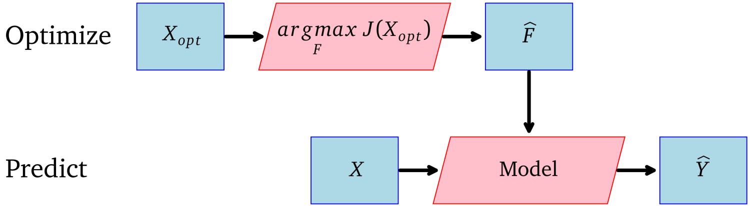
Synthesis



$$\begin{cases} \widehat{Y}_{FR} \\ \widehat{Y}_{ES} \\ \dots \\ \widehat{Y}_{GR} \end{cases}$$

Multi-Market Forecasting using a **Graph Network Constraints** considered using **Flow Estimation Problems**

Forecasting Electricity Prices: An Optimize Then Predict-Based Approach, IDA 2023.

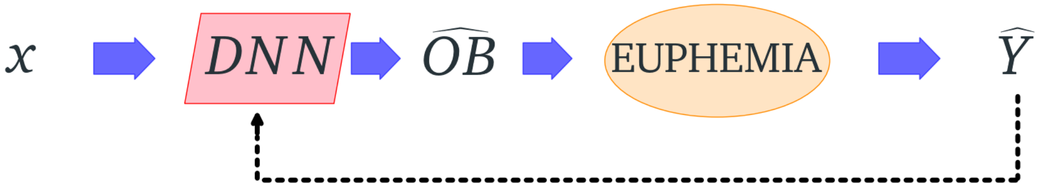


Domain-Knowledge integrated in the input

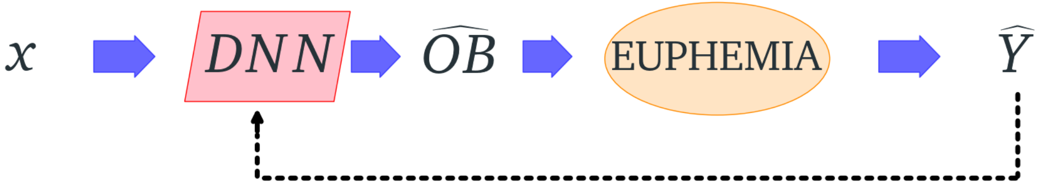
But this is done *A priori*!

Plan

A Differentiable Optimization framework for EPF

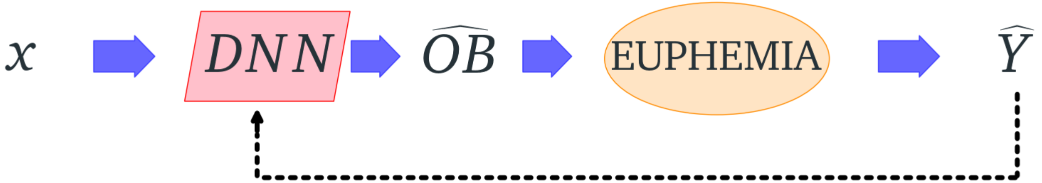


A Differentiable Optimization framework for EPF



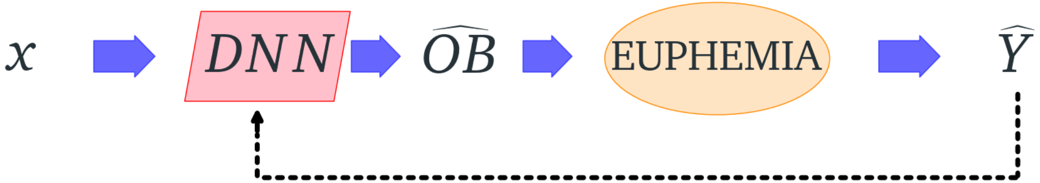
$$\frac{\partial L(Y,\widehat{Y})}{\partial x} \propto \frac{\partial L(Y,\widehat{Y})}{\partial \widehat{Y}} \times \frac{\partial \widehat{Y}}{\partial \widehat{OB}} \times \frac{\partial \widehat{OB}}{\partial x}$$

A Differentiable Optimization framework for EPF



$$\frac{\partial L(Y,\widehat{Y})}{\partial x} \propto \frac{\partial L(Y,\widehat{Y})}{\partial \widehat{Y}} \times \frac{\partial \widehat{Y}}{\partial \widehat{OB}} \times \frac{\partial \widehat{OB}}{\partial x}$$

A Differentiable Optimization framework for EPF

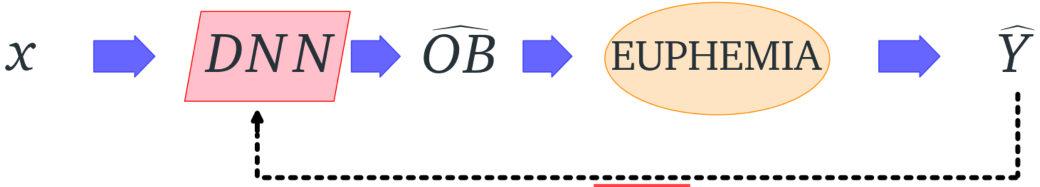


$$\frac{\partial L(Y,\widehat{Y})}{\partial x} \propto \frac{\partial L(Y,\widehat{Y})}{\partial \widehat{Y}} \times \boxed{\frac{\partial \widehat{Y}}{\partial \widehat{OB}}} \times \frac{\partial \widehat{OB}}{\partial x}$$

Differentiable Optimization

Amos B, Kolter JZ. **Optnet: Differentiable optimization as a layer in neural networks.**, ICML 2017

A Differentiable Optimization framework for EPF



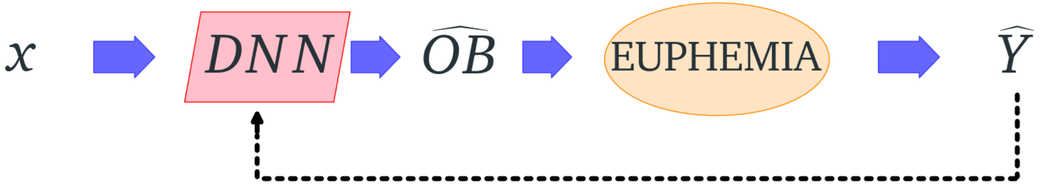
$$\frac{\partial L(Y,\widehat{Y})}{\partial x} \propto \frac{\partial L(Y,\widehat{Y})}{\partial \widehat{Y}} \times \boxed{\frac{\partial \widehat{Y}}{\partial \widehat{OB}}} \times \frac{\partial \widehat{OB}}{\partial x}$$

Differentiable Optimization

Amos B, Kolter JZ. **Optnet: Differentiable optimization as a layer in neural networks.**, ICML 2017

$|OB|$ = thousands of Orders per hour

A Differentiable Optimization framework for EPF



$$\frac{\partial L(Y,\widehat{Y})}{\partial x} \propto \frac{\partial L(Y,\widehat{Y})}{\partial \widehat{Y}} \times \boxed{\frac{\partial \widehat{Y}}{\partial \widehat{OB}}} \times \frac{\partial \widehat{OB}}{\partial x}$$

Differentiable Optimization

Amos B, Kolter JZ. **Optnet: Differentiable optimization as a layer in neural networks.**, ICML 2017

$|OB|$ = thousands of Orders per hour

We have to define **EUPHEMIA**'s forward and backward pass

Formalizing Euphemia's forward pass on 1 market

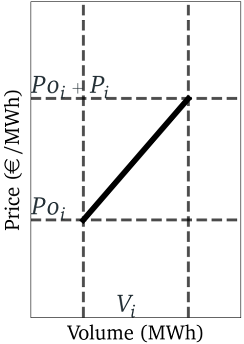
$$\max_{A \in [0,1]^n} \text{Social Welfare}(A, OB)$$

$$u.c. \text{ Energy Balance}(A, OB) = 0$$

Formalizing Euphemia's forward pass on 1 market

$$\max_{A \in [0,1]^n} \text{Social Welfare}(A, OB)$$

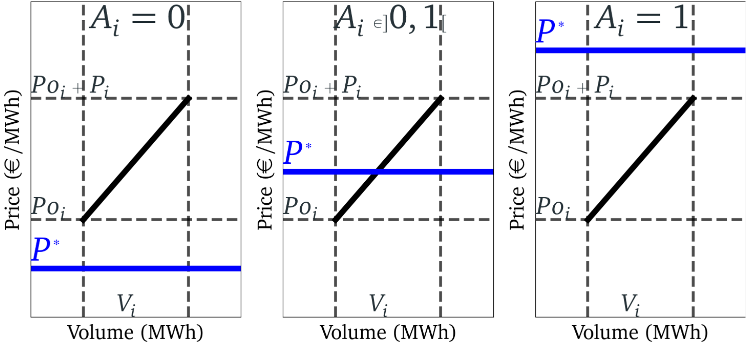
$$u.c. \ \text{Energy Balance}(A, OB) = 0$$



Formalizing Euphemia's forward pass on 1 market

$$\max_{A \in [0,1]^n} \text{Social Welfare}(A, OB)$$

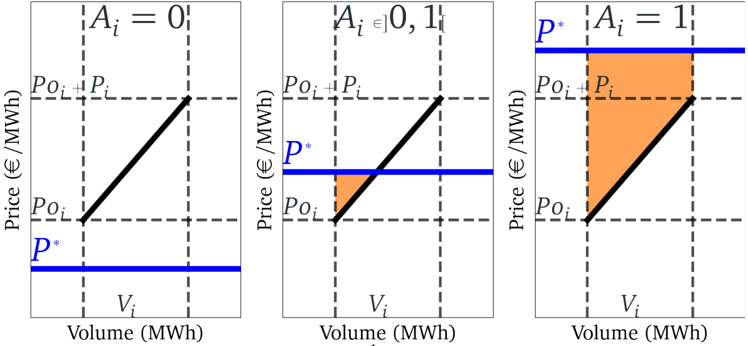
$$u.c. \text{ Energy Balance}(A, OB) = 0$$



Formalizing Euphemia's forward pass on 1 market

$$\max_{A \in [0,1]^n} \text{Social Welfare}(A, OB)$$

$$u.c. \text{ Energy Balance}(A, OB) = 0$$



$$SW(i) = A_i V_i P^* - \frac{1}{2} A_i^2 V_i P_i - A_i V_i Po_i - \theta_i$$

Formalizing Euphemia's forward pass on 1 market

$$\max_{A \in [0,1]^n} \text{Social Welfare}(A, OB)$$

$$u.c.\ \text{Energy Balance}(A, OB) = 0$$

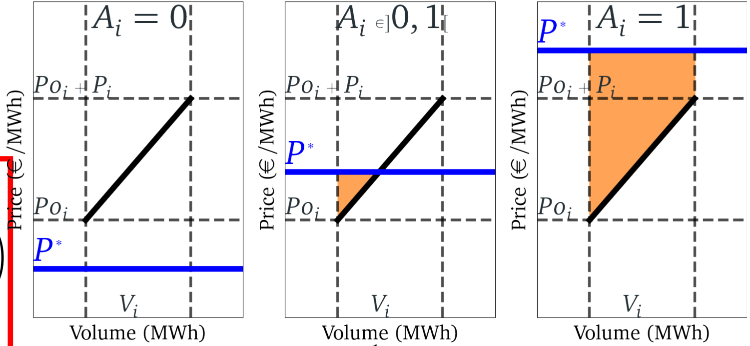
EUPHEMIA

$$\max_{A} \sum_{i \in OB} \left(-\frac{1}{2} A_i^2 V_i P_i - A_i V_i P_{o_i} \right)$$

$$\text{u.c.} \sum_{i \in OB} A_i V_i = 0,$$

$$-A_i \leq 0,$$

$$A_i - 1 \leq 0$$



$$SW(i) = A_i V_i P^* - \frac{1}{2} A_i^2 V_i P_i - A_i V_i Po_i - \theta_i$$

Formalizing Euphemia's forward pass on 1 market

$$\max_{A\in[0,1]^n} \text{Social Welfare}(A, OB)$$

$$u.c. \ \text{Energy Balance}(A, OB) = 0$$

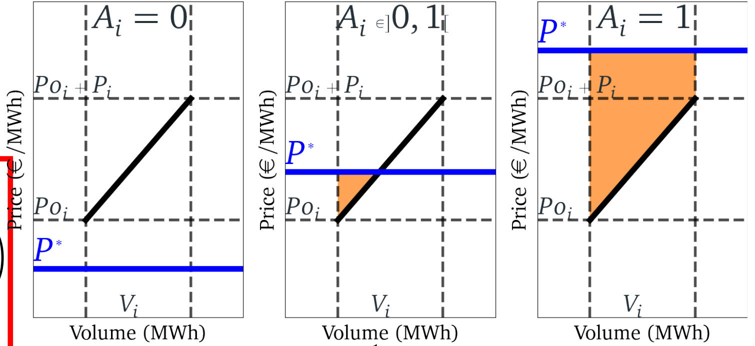
EUPHEMIA

$$\max_{A} \sum_{i\in OB} \left(-\frac{1}{2} A_i^2 V_i P_i - A_i V_i P_{o_i} \right)$$

$$\text{u.c.} \quad \sum_{i\in OB} A_i V_i = 0,$$

$$-A_i \leq 0,$$

$$A_i - 1 \leq 0$$



$$SW(i) = A_i V_i P^* - \frac{1}{2} A_i^2 V_i P_i - A_i V_i Po_i - \theta_i$$

How can we find **Y** while solving **EUPHEMIA**?

Writing the Dual Problem and its derivative

EUPHEMIA

$$
\begin{aligned}
\max_{A} \sum_{i \in \, OB} & \left(-\frac{1}{2} A_i^2 V_i P_i - A_i V_i P_{o_i} \right) \\
\text{u.c.} \sum_{i \in \, OB} & A_i V_i = 0, \\
& - A_i \leq 0, \\
& A_i - 1 \leq 0
\end{aligned}
$$

Writing the Dual Problem and its derivative

EUPHEMIA

$$\max_{A} \sum_{i \in OB} \left(-\frac{1}{2} A_i^2 V_i P_i - A_i V_i P_{o_i} \right)$$

$$\text{u.c.} \sum_{i \in OB} \boldsymbol{A_i V_i} = 0,$$

$$-\boldsymbol{A_i} \leq 0,$$

$$\boldsymbol{A_i} - 1 \leq 0$$

Lagrangian

$$\mathcal{L}(A, \lambda, M, K) = \sum_{i \in OB} \left(-\frac{A_i^2 V_i P_i}{2} - A_i V_i Po_i + \boldsymbol{\lambda A_i V_i} - \boldsymbol{M_i A_i} + \boldsymbol{K_i}(\boldsymbol{A_i} - 1) \right)$$

Writing the Dual Problem and its derivative

Dual Problem

$$\lambda^\star = \min_{\lambda} \sum_{i \in OB} \mathcal{D}_i(\lambda)$$

$$\text{with } \mathcal{D}_i(\lambda) = \begin{cases} (1)\ 0, & \text{if } V_i(Po_i - \lambda) > 0 \\ (2)\ V_i(\lambda - \frac{P_i}{2} - Po_i), & \text{if } V_i(\lambda - P_i - Po_i)) > 0 \\ (3)\ \frac{V_i}{2P_i}(\lambda - Po_i)^2, & \text{if } \lambda \in [Po_i, Po_i + P_i] \end{cases}$$

EUPHEMIA

$$\max_{A} \sum_{i \in OB} \left(-\frac{1}{2} A_i^2 V_i P_i - A_i V_i P_{o_i} \right)$$

$$\text{u.c.} \sum_{i \in OB} \boldsymbol{A_i V_i} = 0,$$

$$-\boldsymbol{A_i} \leq 0,$$

$$\boldsymbol{A_i} - 1 \leq 0$$

Lagrangian

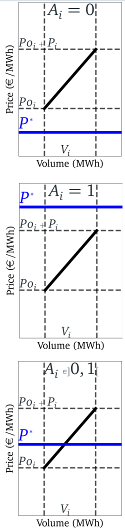
$$\mathcal{L}(A, \lambda, M, K) = \sum_{i \in OB} \left(-\frac{A_i^2 V_i P_i}{2} - A_i V_i Po_i + \boldsymbol{\lambda A_i V_i} - \boldsymbol{M_i A_i} + \boldsymbol{K_i}(\boldsymbol{A_i} - 1) \right)$$

Writing the Dual Problem and its derivative

Dual Problem

$$\lambda^\star = \min_{\lambda} \sum_{i \in OB} \mathcal{D}_i(\lambda)$$

$$\text{with } \mathcal{D}_i(\lambda) = \begin{cases} (1)\ 0, & \text{if } V_i(Po_i - \lambda) > 0 \\ (2)\ V_i(\lambda - \frac{P_i}{2} - Po_i), & \text{if } V_i(\lambda - P_i - Po_i)) > 0 \\ (3)\ \frac{V_i}{2P_i}(\lambda - Po_i)^2, & \text{if } \lambda \in [Po_i, Po_i + P_i] \end{cases}$$



Lagrangian

$$\mathcal{L}(A, \lambda, M, K) = \sum_{i \in OB} \left(-\frac{A_i^2 V_i P_i}{2} - A_i V_i Po_i + \lambda A_i V_i - M_i A_i + K_i(A_i - 1) \right)$$

Writing the Dual Problem and its derivative

Dual Problem

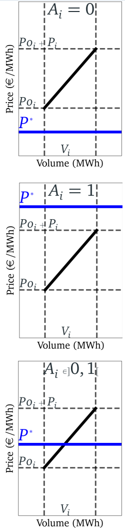
$$\lambda^\star = \min_{\lambda} \sum_{i \in OB} \mathcal{D}_i(\lambda)$$

$$\text{with } \mathcal{D}_i(\lambda) = \begin{cases} (1)\ 0, & \text{if } V_i(Po_i - \lambda) > 0 \\ (2)\ V_i(\lambda - \frac{P_i}{2} - Po_i), & \text{if } V_i(\lambda - P_i - Po_i)) > 0 \\ (3)\ \frac{V_i}{2P_i}(\lambda - Po_i)^2, & \text{if } \lambda \in [Po_i, Po_i + P_i] \end{cases}$$

$\lambda^\star$ is the Day-Ahead Price!

Lagrangian

$$\mathcal{L}(A, \lambda, M, K) = \sum_{i \in OB} \left(-\frac{A_i^2 V_i P_i}{2} - A_i V_i Po_i + \lambda A_i V_i - M_i A_i + K_i(A_i - 1) \right)$$

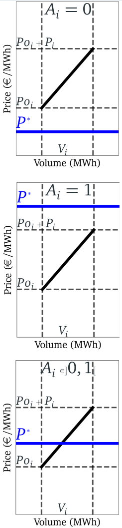


Writing the Dual Problem and its derivative

Dual Problem

$$\lambda^\star = \min_{\lambda} \sum_{i \in OB} \mathcal{D}_i(\lambda)$$

$$\text{with } \mathcal{D}_i(\lambda) = \begin{cases} (1)\ 0, & \text{if } V_i(Po_i - \lambda) > 0 \\ (2)\ V_i(\lambda - \frac{P_i}{2} - Po_i), & \text{if } V_i(\lambda - P_i - Po_i)) > 0 \\ (3)\ \frac{V_i}{2P_i}(\lambda - Po_i)^2, & \text{if } \lambda \in [Po_i, Po_i + P_i] \end{cases}$$



$$\mathcal{D}'(\lambda^\star) = 0$$

$\lambda^\star$ is the Day-Ahead Price!

Lagrangian

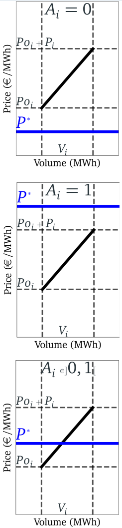
$$\mathcal{L}(A, \lambda, M, K) = \sum_{i \in OB} \left(-\frac{A_i^2 V_i P_i}{2} - A_i V_i Po_i + \lambda A_i V_i - M_i A_i + K_i(A_i - 1) \right)$$

Writing the Dual Problem and its derivative

Dual Problem

$$\lambda^\star = \min_{\lambda} \sum_{i \in OB} \mathcal{D}_i(\lambda)$$

$$\text{with } \mathcal{D}_i(\lambda) = \begin{cases} (1)\ 0, & \text{if } V_i(Po_i - \lambda) > 0 \\ (2)\ V_i(\lambda - \frac{P_i}{2} - Po_i), & \text{if } V_i(\lambda - P_i - Po_i)) > 0 \\ (3)\ \frac{V_i}{2P_i}(\lambda - Po_i)^2, & \text{if } \lambda \in [Po_i, Po_i + P_i] \end{cases}$$



$$\mathcal{D}'(\lambda^\star) = 0$$

$$H(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 \text{ if } x \geq 0 \end{cases}$$

$$x_i = V_i(\lambda - Po_i)$$

$$y_i = V_i(\lambda - Po_i - P_i)$$

$\lambda^\star$ is the Day-Ahead Price!

Lagrangian

$$\mathcal{L}(A, \lambda, M, K) = \sum_{i \in OB} \left(-\frac{A_i^2 V_i P_i}{2} - A_i V_i Po_i + \lambda A_i V_i - M_i A_i + K_i(A_i - 1) \right)$$

Writing the Dual Problem and its derivative

Dual Problem

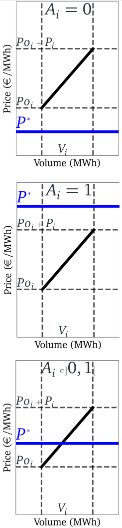
$$\lambda^\star = \min_\lambda \sum_{i \in OB} \mathcal{D}_i(\lambda)$$

$$\text{with } \mathcal{D}_i(\lambda) = \begin{cases} (1)\ 0, & \text{if } V_i(Po_i - \lambda) > 0 \\ (2)\ V_i(\lambda - \frac{P_i}{2} - Po_i), & \text{if } V_i(\lambda - P_i - Po_i)) > 0 \\ (3)\ \frac{V_i}{2P_i}(\lambda - Po_i)^2, & \text{if } \lambda \in [Po_i, Po_i + P_i] \end{cases}$$

$\lambda^\star$ is the Day-Ahead Price!

Lagrangian

$$\mathcal{L}(A, \lambda, M, K) = \sum_{i \in OB} \left(-\frac{A_i^2 V_i P_i}{2} - A_i V_i Po_i + \lambda A_i V_i - M_i A_i + K_i(A_i - 1) \right)$$



$$\mathcal{D}'(\lambda^\star) = 0$$

$$\mathcal{D}'(\lambda) = \sum_i \frac{x_i H(x_i) - y_i H(y_i)}{P_i}$$

$$H(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 \text{ if } x \geq 0 \end{cases}$$

$$x_i = V_i(\lambda - Po_i)$$

$$y_i = V_i(\lambda - Po_i - P_i)$$

Solving $\mathcal{D}'(\lambda^\star) = 0$ using a Differentiable Dichotomic search

Algorithm 1 Differentiable dichotomic search.

lb ← -500€/MWh
ub ← 3000€/MWh
found ←**False**
while (found = **False**) and $(ub - lb > 2 * 0.01)$ **do**
  $\lambda \leftarrow \frac{ub+lb}{2}$
  $\mathcal{D}'_k \leftarrow \mathcal{D}'(\lambda)$
  found ← $\mathcal{D}'_k = 0$
  $ub \leftarrow ub - H(\mathcal{D}'_k) * (ub - \lambda)$
  $lb \leftarrow \lambda - H(\mathcal{D}'_k) * (\lambda - lb)$
end while

Solving $\mathcal{D}'(\lambda^\star) = 0$ using a Differentiable Dichotomic search

Algorithm 1 Differentiable dichotomic search.

lb ← -500€/MWh
ub ← 3000€/MWh
found ← **False**
while (found = **False**) and $(ub - lb > 2 * 0.01)$ **do**
  $\lambda \leftarrow \frac{ub+lb}{2}$
  $\mathcal{D}'_k \leftarrow \mathcal{D}'(\lambda)$
  found $\leftarrow \mathcal{D}'_k = 0$
  $ub \leftarrow ub - H(\mathcal{D}'_k) * (ub - \lambda)$
  $lb \leftarrow \lambda - H(\mathcal{D}'_k) * (\lambda - lb)$
end while

$$\frac{\partial \widehat{Y}}{\partial \widehat{OB}} = \sum_m \nabla_m \frac{\partial m}{\partial \widehat{OB}} \ \forall\ m \text{ used to compute } \widehat{Y}$$

Solving $\mathcal{D}'(\lambda^\star) = 0$ using a Differentiable Dichotomic search

Algorithm 1 Differentiable dichotomic search.

```
lb ← -500€/MWh
ub ← 3000€/MWh
found ← False
while (found = False) and (ub − lb > 2 ∗ 0.01) do
    λ ← (ub+lb)/2
    D'_k ← D'(λ)
    found ← D'_k = 0
    ub ← ub − H(D'_k) ∗ (ub − λ)
    lb ← λ − H(D'_k) ∗ (λ − lb)
end while
```

$$\frac{\partial \widehat{Y}}{\partial \widehat{OB}} = \sum_m \nabla_m \frac{\partial m}{\partial \widehat{OB}} \; \forall \; m \text{ used to compute } \widehat{Y}$$

$$= \sum_{k=1}^{N-1} \nabla_{\mathcal{D}'_k} \frac{\partial \mathcal{D}'_k}{\partial \widehat{OB}},$$

Solving $\mathcal{D}'(\lambda^\star) = 0$ using a Differentiable Dichotomic search

Algorithm 1 Differentiable dichotomic search.

> lb ← -500€/MWh
> ub ← 3000€/MWh
> found ← **False**
> **while** (found = **False**) and $(ub - lb > 2 * 0.01)$ **do**
> $\lambda \leftarrow \frac{ub+lb}{2}$
> $\mathcal{D}'_k \leftarrow \mathcal{D}'(\lambda)$
>   found $\leftarrow \mathcal{D}'_k = 0$
> $ub \leftarrow ub - H(\mathcal{D}'_k) * (ub - \lambda)$
> $lb \leftarrow \lambda - H(\mathcal{D}'_k) * (\lambda - lb)$
> **end while**

Accumulate $\frac{\partial \mathcal{D}'_k}{\partial \widehat{OB}}$ during each step k

$$\frac{\partial \widehat{Y}}{\partial \widehat{OB}} = \sum_m \nabla_m \frac{\partial m}{\partial \widehat{OB}} \ \forall\ m \text{ used to compute } \widehat{Y}$$

$$= \sum_{k=1}^{N-1} \nabla_{\mathcal{D}'_k} \frac{\partial \mathcal{D}'_k}{\partial \widehat{OB}},$$

Solving $\mathcal{D}'(\lambda^\star) = 0$ using a Differentiable Dichotomic search

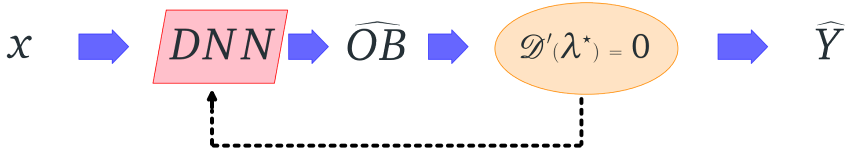
Algorithm 1 Differentiable dichotomic search.

lb ← -500€/MWh
ub ← 3000€/MWh
found ←**False**
while (found = **False**) and $(ub - lb > 2 * 0.01)$ **do**
  $\lambda \leftarrow \frac{ub+lb}{2}$
  $\mathcal{D}'_k \leftarrow \mathcal{D}'(\lambda)$
  found ← $\mathcal{D}'_k = 0$
  $ub \leftarrow ub - H(\mathcal{D}'_k) * (ub - \lambda)$
  $lb \leftarrow \lambda - H(\mathcal{D}'_k) * (\lambda - lb)$
end while

Accumulate $\frac{\partial \mathcal{D}'_k}{\partial \widehat{OB}}$ during each step k

$$\frac{\partial \widehat{Y}}{\partial \widehat{OB}} = \sum_m \nabla_m \frac{\partial m}{\partial \widehat{OB}} \;\forall\; m \text{ used to compute } \widehat{Y}$$

$$= \sum_{k=1}^{N-1} \nabla_{\mathcal{D}'_k} \frac{\partial \mathcal{D}'_k}{\partial \widehat{OB}},$$



Solving $\mathcal{D}'(\lambda^\star) = 0$ using a Differentiable Dichotomic search

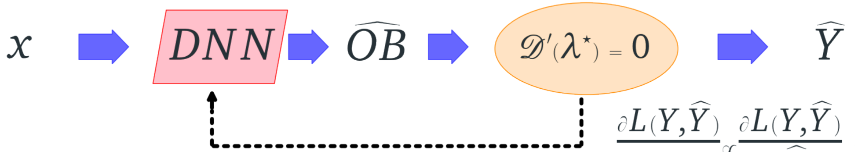
Algorithm 1 Differentiable dichotomic search.

lb ← -500€/MWh
ub ← 3000€/MWh
found ←**False**
while (found = **False**) and $(ub - lb > 2 * 0.01)$ **do**
  $\lambda \leftarrow \frac{ub+lb}{2}$
  $\mathcal{D}'_k \leftarrow \mathcal{D}'(\lambda)$
  found ← $\mathcal{D}'_k = 0$
  $ub \leftarrow ub - H(\mathcal{D}'_k) * (ub - \lambda)$
  $lb \leftarrow \lambda - H(\mathcal{D}'_k) * (\lambda - lb)$
end while

Accumulate $\frac{\partial \mathcal{D}'_k}{\partial \widehat{OB}}$ during each step k

$$\frac{\partial \widehat{Y}}{\partial \widehat{OB}} = \sum_m \nabla_m \frac{\partial m}{\partial \widehat{OB}} \; \forall \; m \text{ used to compute } \widehat{Y}$$

$$= \sum_{k=1}^{N-1} \nabla_{\mathcal{D}'_k} \frac{\partial \mathcal{D}'_k}{\partial \widehat{OB}},$$



$$\frac{\partial L(Y,\widehat{Y})}{\partial x} \propto \frac{\partial L(Y,\widehat{Y})}{\partial \widehat{Y}} \times \frac{\partial \widehat{Y}}{\partial \widehat{OB}} \times \frac{\partial \widehat{OB}}{\partial x}$$

Evaluation of the Differentiable Optimization DO approach on the 2019 period

Market	Model			
BE				
DE				
FR				
NL				

Evaluation of the Differentiable Optimization DO approach on the 2019 period

Market	Model			
BE	DNN			
DE	DNN			
FR	DNN			
NL	DNN			

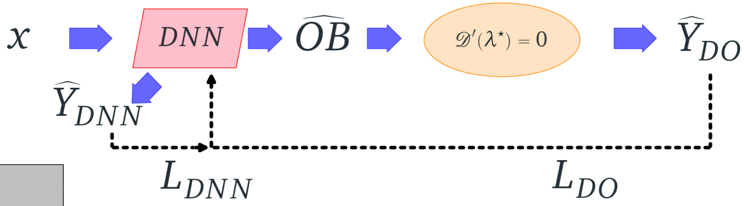
Evaluation of the Differentiable Optimization DO approach on the 2019 period

Market	Model			
BE	DNN			
	DO			
DE	DNN			
	DO			
FR	DNN			
	DO			
NL	DNN			
	DO			

Evaluation of the Differentiable Optimization DO approach on the 2019 period

Market	Model			
BE	DNN			
	DO			
	DNN + DO			
DE	DNN			
	DO			
	DNN + DO			
FR	DNN			
	DO			
	DNN + DO			
NL	DNN			
	DO			
	DNN + DO			

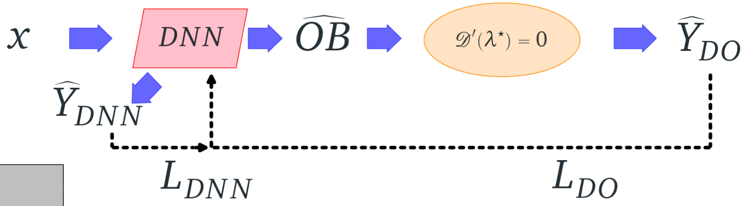
Evaluation of the Differentiable Optimization DO approach on the 2019 period



Market	Model			
BE	DNN			
	DO			
	DNN + DO			
DE	DNN			
	DO			
	DNN + DO			
FR	DNN			
	DO			
	DNN + DO			
NL	DNN			
	DO			
	DNN + DO			

Evaluation of the Differentiable Optimization DO approach on the 2019 period

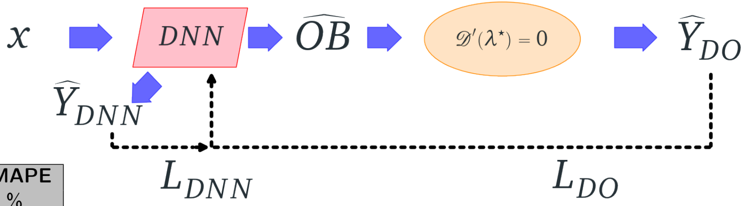
$$L_{DNN+DO} = \frac{1}{2} L_{DNN} + \frac{1}{2} L_{DO}$$



Market	Model			
BE	DNN			
	DO			
	DNN + DO			
DE	DNN			
	DO			
	DNN + DO			
FR	DNN			
	DO			
	DNN + DO			
NL	DNN			
	DO			
	DNN + DO			

Evaluation of the Differentiable Optimization DO approach on the 2019 period

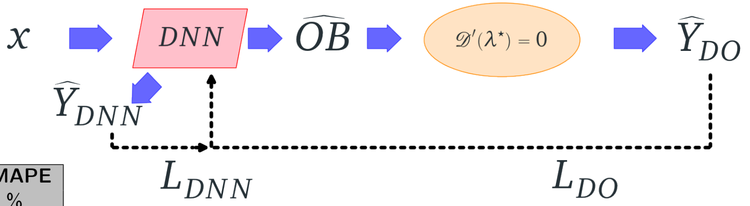
$$L_{DNN+DO} = \frac{1}{2} L_{DNN} + \frac{1}{2} L_{DO}$$



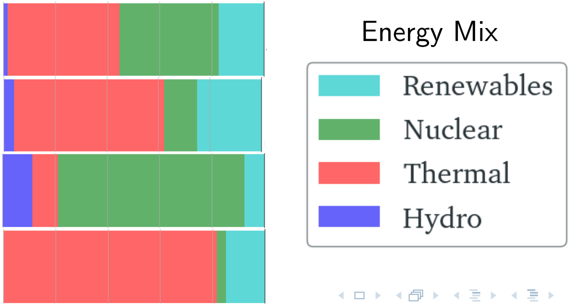
Market	Model	MAE €/MWh	RMAE	SMAPE %
BE	DNN	7.74	0.941	21.27
	DO	7.27	0.884	19.73
	DNN + DO	**6.28**	**0.763**	**17.28**
DE	DNN	7.28	0.778	29.83
	DO	9.01	0.958	29.87
	DNN + DO	**6.99**	**0.745**	**25.97**
FR	DNN	**4.54**	**0.653**	**15.5**
	DO	6.47	0.93	20.31
	DNN + DO	5.3	0.759	16.2
NL	DNN	6.32	1.057	18.84
	DO	6.53	1.092	16.47
	DNN + DO	**5.22**	**0.874**	**13.4**

Evaluation of the Differentiable Optimization DO approach on the 2019 period

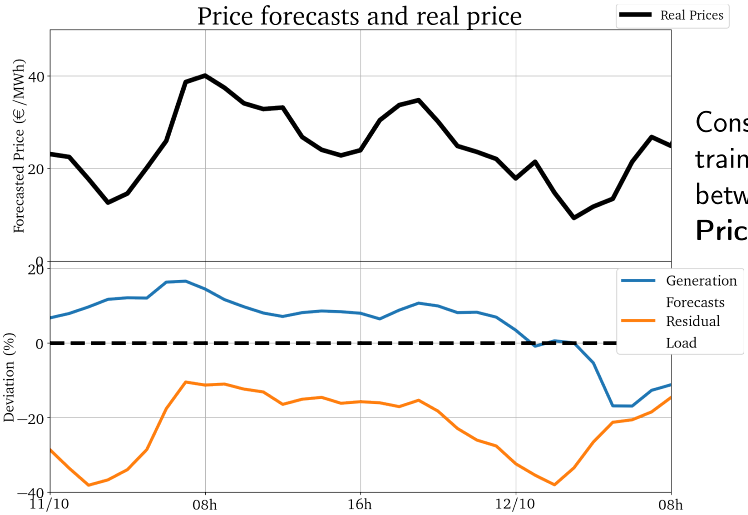
$$L_{DNN+DO} = \frac{1}{2} L_{DNN} + \frac{1}{2} L_{DO}$$



Market	Model	MAE €/MWh	RMAE	SMAPE %
BE	DNN	7.74	0.941	21.27
	DO	7.27	0.884	19.73
	DNN + DO	**6.28**	**0.763**	**17.28**
DE	DNN	7.28	0.778	29.83
	DO	9.01	0.958	29.87
	DNN + DO	**6.99**	**0.745**	**25.97**
FR	DNN	**4.54**	**0.653**	**15.5**
	DO	6.47	0.93	20.31
	DNN + DO	5.3	0.759	16.2
NL	DNN	6.32	1.057	18.84
	DO	6.53	1.092	16.47
	DNN + DO	**5.22**	**0.874**	**13.4**

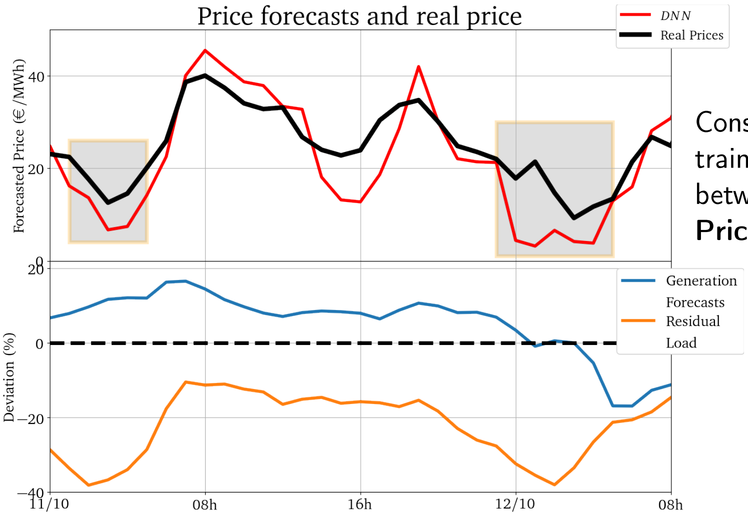


Discussion



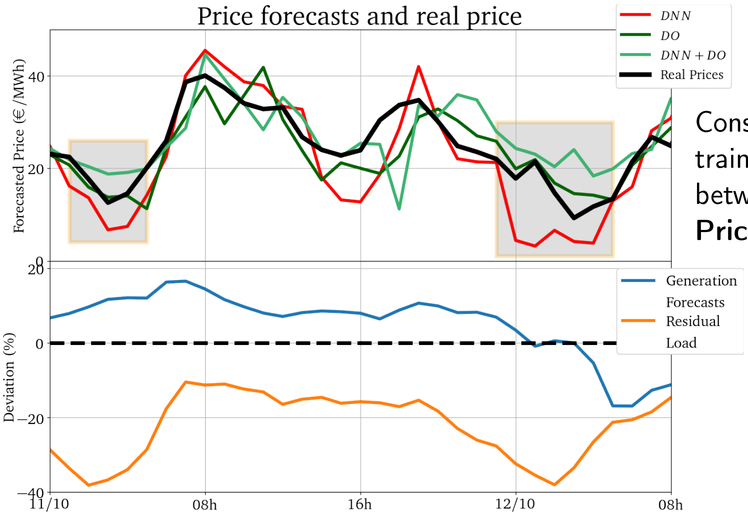
Considering **Domain Knowledge** during training captures the real relationship between **Consumption**, **Generation** and **Prices**.

Discussion



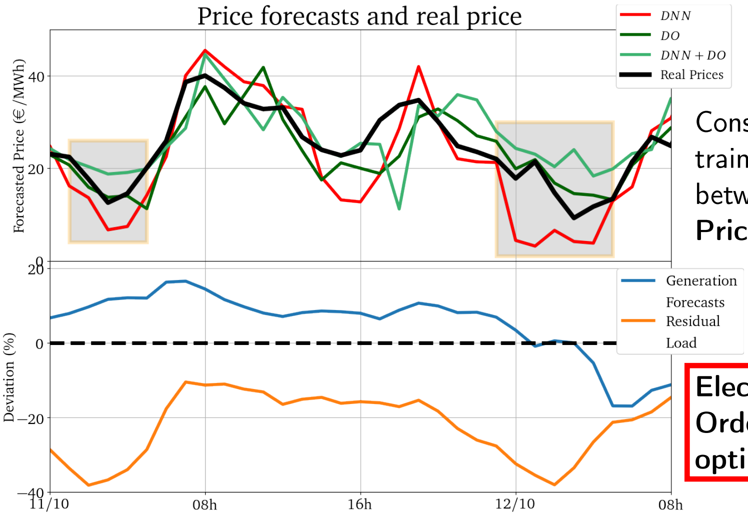
Considering **Domain Knowledge** during training captures the real relationship between **Consumption**, **Generation** and **Prices**.

Discussion



Considering **Domain Knowledge** during training captures the real relationship between **Consumption**, **Generation** and **Prices**.

Discussion



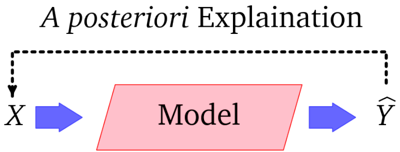
Considering **Domain Knowledge** during training captures the real relationship between **Consumption**, **Generation** and **Prices**.

Electricity Price Forecasting based on Order Books: a differentiable optimization approach, DSAA 2023.

Plan

1. Introduction
2. Explaining the Forecasts
3. Optimize-then-Predict approach
4. A differentiable Optimization Approach
5. Conclusion

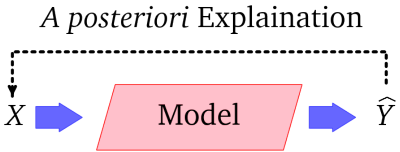
Summary of the Contributions



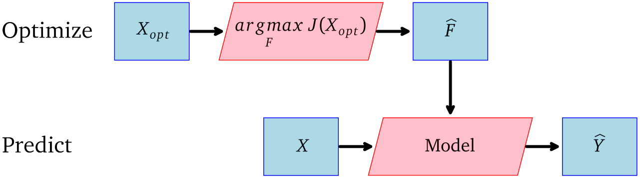
Linking predictions with

Domain-Knowledge

Summary of the Contributions



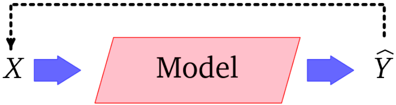
Linking predictions with **Domain-Knowledge**



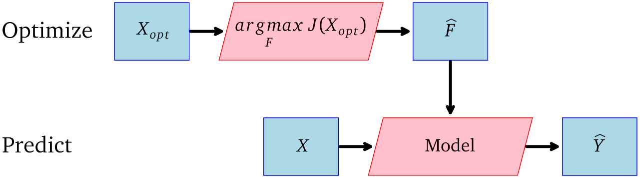
Integrating **Domain-Knowledge** in the input

Summary of the Contributions

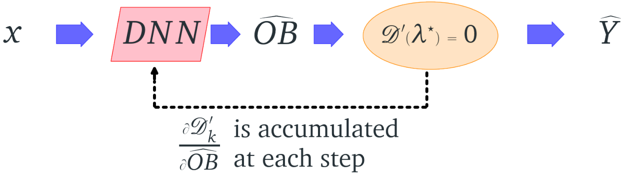
A posteriori Explanation



Linking predictions with **Domain-Knowledge**



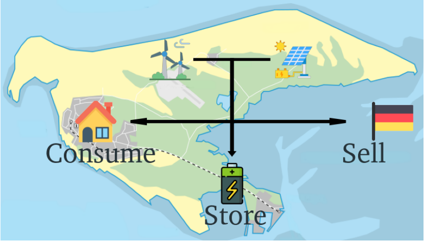
Integrating **Domain-Knowledge** in the input



Considering **Domain Knowledge** during training

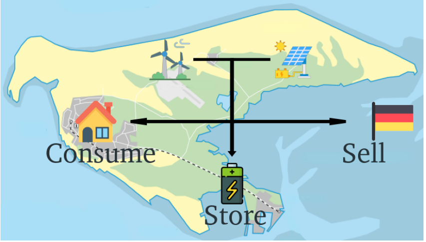
Industrial Impact of the thesis

Germany : Islander Project

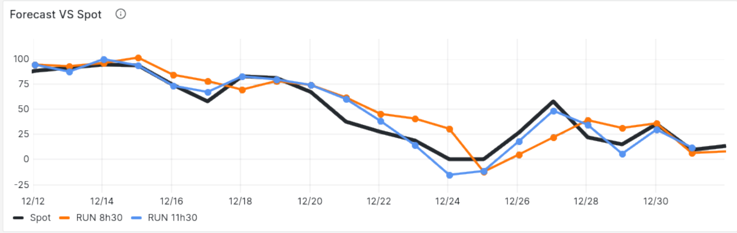


Industrial Impact of the thesis

Germany : Islander Project

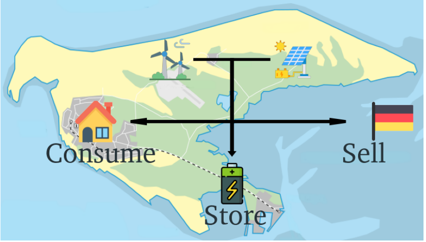


France : Trading on the Day-Ahead Market

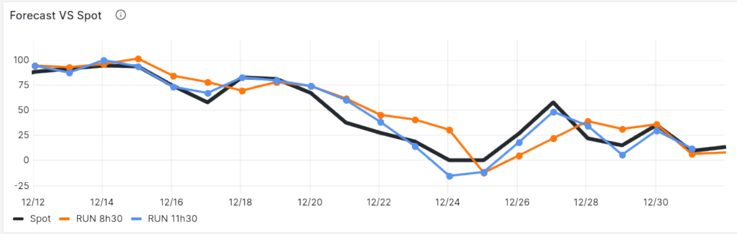


Industrial Impact of the thesis

Germany : Islander Project



France : Trading on the Day-Ahead Market



Minimizing the Task Loss using **Differentiable Optimization**


Thanks
Optimize
Predict
Model
Shap Values
Machine Learning
Regression
Domain Expert
Graph Neural Network
Optimization Problem
Algorithm 1 Differentiable dichotomy search
found ← False
Forecast VS Spot