Price-Aware Deep Learning for Electricity Markets

Vladimir Dvorkin[†] and Ferdinando Fioretto[‡]

[†]Massachusetts Institute of Technology

[‡]University of Virginia

NeurIPS Workshop on Tackling Climate Change with Machine Learning

December 16, 2023

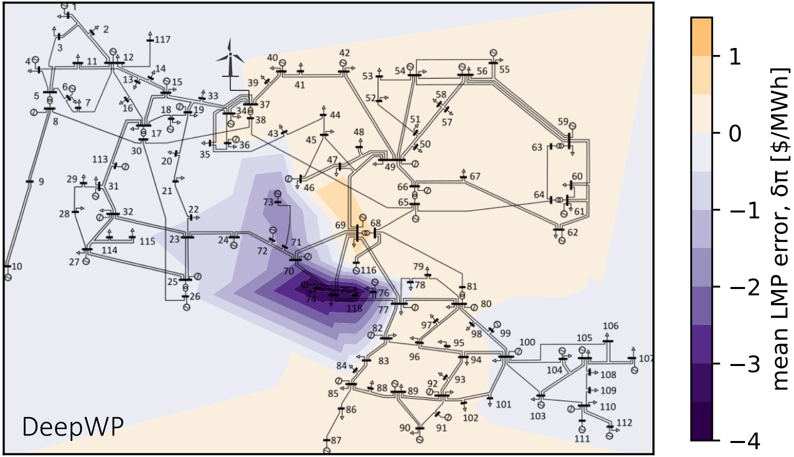
MARIE CURIE ACTIONS

IBERDROLA

What makes wind power commodity so special?

- As of 2022, the share of electricity generation from wind energy sources worldwide constitutes 7.3%.
- Electricity is priced at a *forecast* of variable and uncertain wind power generation, i.e., before the actual realization of wind power is known.
- As a result, forecast errors translate into price errors via electricity market-clearing optimization.
- Although a non-dominant generation resource, it exposes the entire electricity trading to errors

Forecast errors propagate into price errors



Forecast errors from a single wind power plant propagate into locational marginal price (LMP) errors across the IEEE 118-Bus RTS. Many buses demonstrate near zero errors, but electricity at certain buses is systematically over- or under-priced.

Electricity market-clearing optimization

$$\begin{aligned} \underset{\underline{p}\leqslant p\leqslant \overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p && \textit{conventional generator dispatch cost} \\ \text{subject to} \quad & \mathbb{1}^\top(p+\widehat{w}-d)=0 \;:\; \widehat{\lambda}_b, && \textit{power balance condition} \\ & |F(p+\widehat{w}-d)| \leqslant \overline{f} \;:\; \widehat{\lambda}_{\overline{f}}, \widehat{\lambda}_{\underline{f}}, && \textit{power flow limits} \end{aligned}$$

Location marginal prices (LMPs) are derived from the dual solution:

$$\pi(\widehat{w}) = \underbrace{\widehat{\lambda}_b \cdot \mathbb{1}}_{\text{uniform price}} - \underbrace{F^\top(\widehat{\lambda}_{\overline{f}} - \widehat{\lambda}_{\underline{f}})}_{\text{adjustment due to congestion}}$$

which are unique w.r.t forecast $\widehat{w}$ under reasonable assumptions!

The **LMP error** is then defined as:

$$\delta\pi = \pi(\widehat{w}) - \pi(w)$$

i.e., the distance between LMPs induced on the forecast ($\widehat{w}$) and actual realization (w) of wind power.

Electricity market-clearing optimization

$$\begin{aligned} \underset{\underline{p} \leqslant p \leqslant \overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p && \textit{conventional generator dispatch cost} \\ \text{subject to} \quad & \mathbb{1}^\top (p + \widehat{w} - d) = 0 \;:\; \widehat{\lambda}_b, && \textit{power balance condition} \\ & |F(p + \widehat{w} - d)| \leqslant \overline{f} \;:\; \widehat{\lambda}_{\overline{f}}, \widehat{\lambda}_{\underline{f}}, && \textit{power flow limits} \end{aligned}$$

Location marginal prices (LMPs) are derived from the dual solution:

$$\pi(\widehat{w}) = \underbrace{\widehat{\lambda}_b \cdot \mathbb{1}}_{\text{uniform price}} - \underbrace{F^\top (\widehat{\lambda}_{\overline{f}} - \widehat{\lambda}_{\underline{f}})}_{\text{adjustment due to congestion}}$$

which are unique w.r.t forecast $\widehat{w}$ under reasonable assumptions!

The **LMP error** is then defined as:

$$\delta\pi = \pi(\widehat{w}) - \pi(w)$$

i.e., the distance between LMPs induced on the forecast ($\widehat{w}$) and actual realization (w) of wind power.

Electricity market-clearing optimization

$$\begin{aligned} \underset{\underline{p} \leqslant p \leqslant \overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p && \textit{conventional generator dispatch cost} \\ \text{subject to} \quad & \mathbb{1}^\top (p + \widehat{w} - d) = 0 \; : \; \widehat{\lambda}_b, && \textit{power balance condition} \\ & |F(p + \widehat{w} - d)| \leqslant \overline{f} \; : \; \widehat{\lambda}_{\overline{f}}, \widehat{\lambda}_{\underline{f}}, && \textit{power flow limits} \end{aligned}$$

Location marginal prices (LMPs) are derived from the dual solution:

$$\pi(\widehat{w}) = \underbrace{\widehat{\lambda}_b \cdot \mathbb{1}}_{\text{uniform price}} - \underbrace{F^\top(\widehat{\lambda}_{\overline{f}} - \widehat{\lambda}_{\underline{f}})}_{\text{adjustment due to congestion}}$$

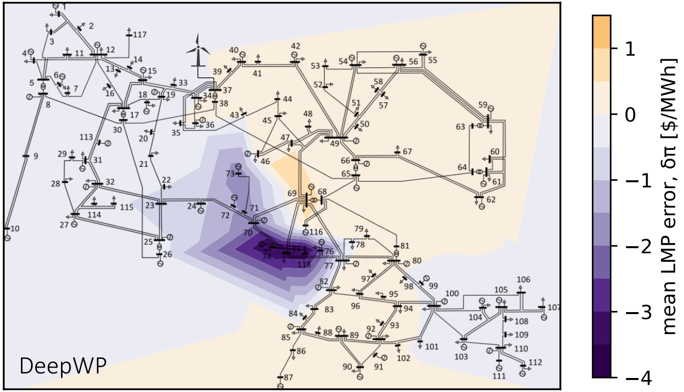
which are unique w.r.t forecast $\widehat{w}$ under reasonable assumptions!

The **LMP error** is then defined as:

$$\delta\pi = \pi(\widehat{w}) - \pi(w)$$

i.e., the distance between LMPs induced on the forecast ($\widehat{w}$) and actual realization (w) of wind power.

Disparities of LMP errors



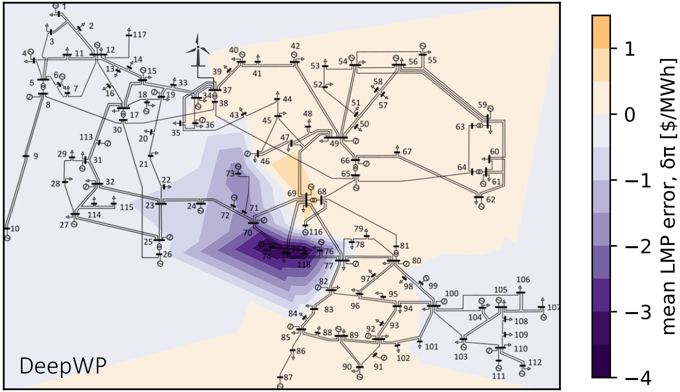
Two properties of LMP errors (informally):

Property #1: Spatial disparity of LMP errors due to congestion

Property #2: Reference bus has the smallest error in the network

Notion of **$\alpha-$fairness**:

$$\alpha = \max_{i\in 1,\dots,n} \|\mathbb{E}[\|\delta\pi_i\|] - \mathbb{E}[\|\delta\pi_{\text{ref}}\|]\|$$



Two properties of LMP errors (informally):

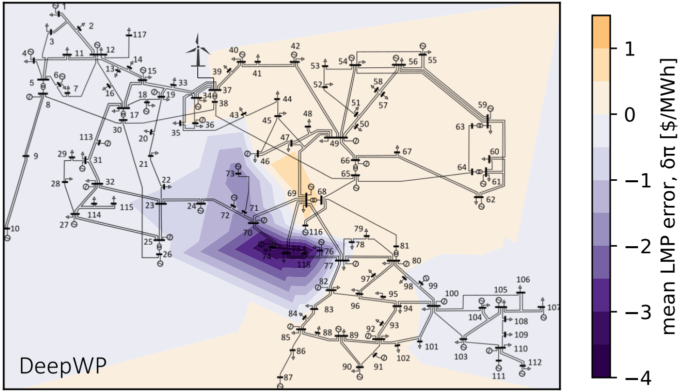
Property #1: Spatial disparity of LMP errors due to congestion

Property #2: Reference bus has the smallest error in the network

Notion of $\alpha-$**fairness**:

$$\alpha = \max_{i \in 1,\dots,n} \left\| \mathbb{E}[\|\delta\pi_i\|] - \mathbb{E}[\|\delta\pi_{\text{ref}}\|] \right\|$$

Disparities of LMP errors



Two properties of LMP errors (informally):

Property #1: Spatial disparity of LMP errors due to congestion

Property #2: Reference bus has the smallest error in the network

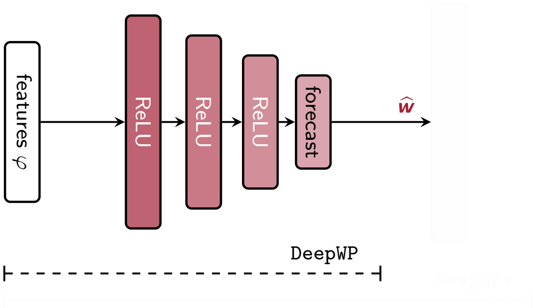
Notion of $\alpha-$**fairness**:

$$\alpha = \max_{i \in 1,\dots,n} \left\| \mathbb{E}[\|\delta\pi_i\|] - \mathbb{E}[\|\delta\pi_{\text{ref}}\|] \right\|$$

Price-awareness for wind power forecast

- Dataset $\{(\varphi_1, w_1), \ldots, (\varphi_m, w_m)\}$ of wind power records, with features φ and measurements w
- Two deep learning architectures `DeepWP` and `DeepWP+` for wind power forecasting:

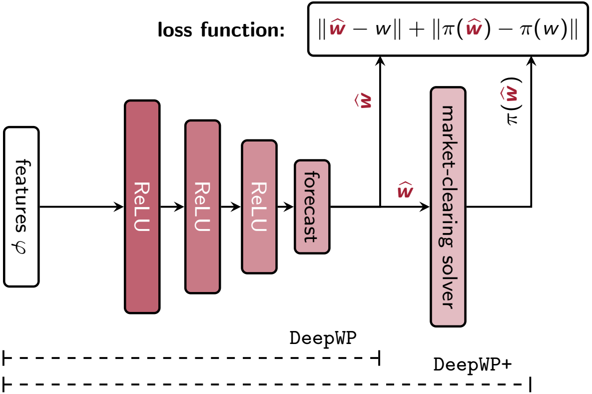
loss function: $\|\widehat{w} - w\|$



- DeepWP+ informs wind power predictions about the downstream pricing errors

Price-awareness for wind power forecast

- Dataset $\{(\varphi_1, w_1), \ldots, (\varphi_m, w_m)\}$ of wind power records, with features φ and measurements w
- Two deep learning architectures `DeepWP` and `DeepWP+` for wind power forecasting:



- DeepWP+ informs wind power predictions about the downstream pricing errors

Market clearing as an optimization layer

Market-clearing optimization $\Longrightarrow$ **Equivalent primal form** $\Longrightarrow$ **Equivalent dual form**

$$\begin{aligned} \underset{\underline{p}\leqslant p\leqslant\overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p \\ \text{subject to} \quad & \mathbb{1}^\top(p + \widehat{w} - d) = 0 \\ & |F(p + \widehat{w} - d)| \leqslant \overline{f} \end{aligned}$$

$$\begin{aligned} \underset{\underline{p}\leqslant p\leqslant\overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p \\ \text{subject to} \quad & Ap \geqslant b(\widehat{w}) \; : \; \lambda \end{aligned}$$

$$\underset{\lambda\geqslant 0}{\text{maximize}} \quad \left(AC^{-1}c + b(\widehat{w})\right)^\top \lambda - \lambda^\top A C^{-1} A^\top \lambda$$

large constrained optimization | only inequality constraints | only non-negativity constraints

Market clearing as an optimization layer

Market-clearing optimization $\Longrightarrow$ **Equivalent primal form** $\Longrightarrow$ **Equivalent dual form**

$$\begin{aligned} \underset{\underline{p}\leqslant p\leqslant \overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p \\ \text{subject to} \quad & \mathbb{1}^\top (p + \widehat{w} - d) = 0 \\ & |F(p + \widehat{w} - d)| \leqslant \overline{f} \end{aligned}$$

$$\begin{aligned} \underset{\underline{p}\leqslant p\leqslant \overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p \\ \text{subject to} \quad & Ap \geqslant b(\widehat{w}) \; : \; \lambda \end{aligned}$$

$$\begin{aligned} \underset{\lambda \geqslant 0}{\text{maximize}} \quad & \left(AC^{-1}c + b(\widehat{w})\right)^\top \lambda \\ & - \lambda^\top A C^{-1} A^\top \lambda \end{aligned}$$

large constrained optimization

only inequality constraints

only non-negativity constraints

Market clearing as an optimization layer

Market-clearing optimization $\Longrightarrow$ **Equivalent primal form** $\Longrightarrow$ **Equivalent dual form**

$$\begin{aligned} \underset{\underline{p}\leqslant p\leqslant \overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p \\ \text{subject to} \quad & \mathbb{1}^\top (p + \widehat{w} - d) = 0 \\ & |F(p + \widehat{w} - d)| \leqslant \overline{f} \end{aligned}$$

large constrained optimization

$$\begin{aligned} \underset{\underline{p}\leqslant p\leqslant \overline{p}}{\text{minimize}} \quad & p^\top C p + c^\top p \\ \text{subject to} \quad & Ap \geqslant b(\widehat{w}) \;:\; \lambda \end{aligned}$$

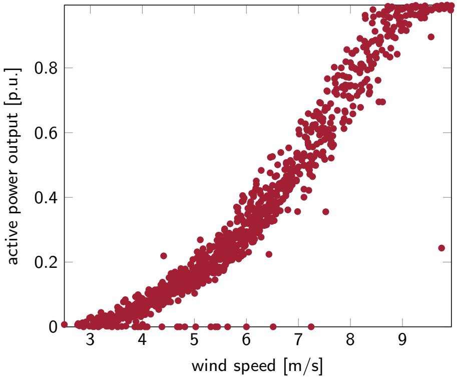
only inequality constraints

$$\begin{aligned} \underset{\lambda\geqslant \mathbb{0}}{\text{maximize}} \quad & \left(AC^{-1}c + b(\widehat{w})\right)^\top \lambda \\ & - \lambda^\top A C^{-1} A^\top \lambda \end{aligned}$$

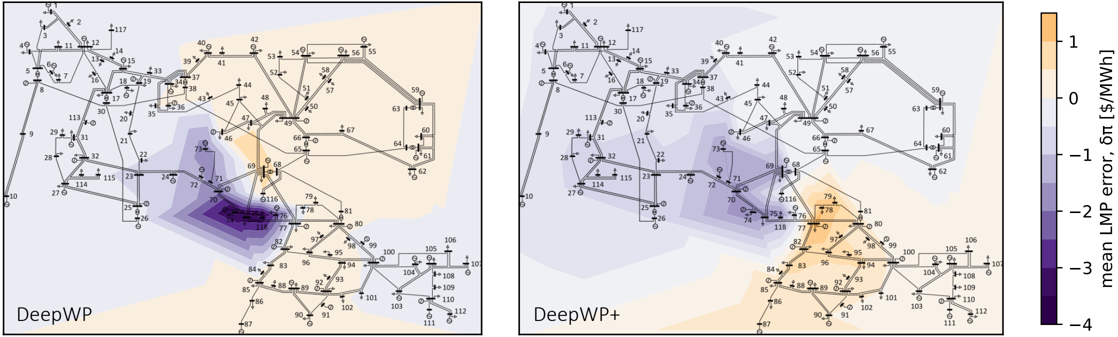
only non-negativity constraints

Numerical experiments

- Standard `PowerModels.jl` test cases
- 1,000 wind power records from a real turbine:
  - Active power output
  - Wind speed and direction
  - Blade pitch angle
- `DeepWP` has 4 hidden layers with 30 neurons each. `DeepWP+` additionally includes an opt. layer
- ADAM optimizer with varying learning rate

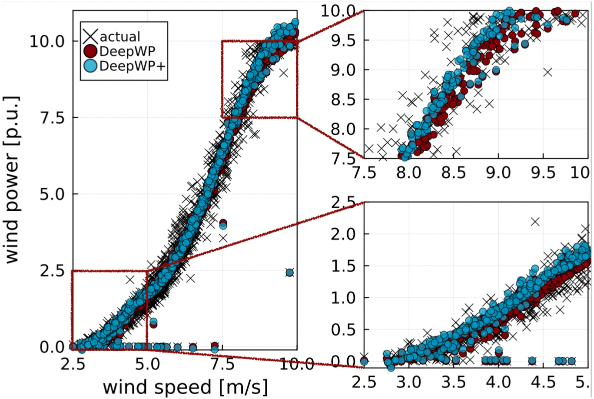


IEEE 118-bus system



$\texttt{DeepWP:}$ Forecast error minimization yields $\delta\pi \in [-4, 1]$ \$/MWh

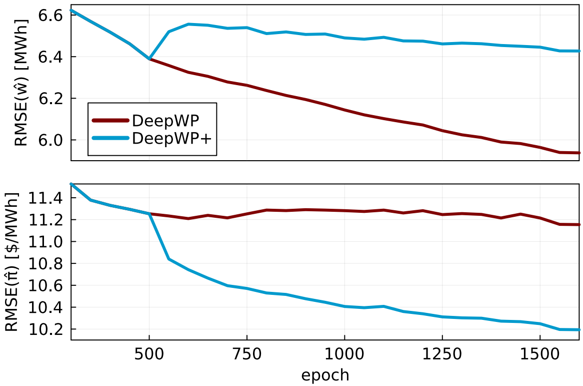
$\texttt{DeepWP+:}$ Price error minimization yields $\delta\pi \in [-1, 1]$ \$/MWh



`DeepWP`: Minimizes the average forecast deviation

`DeepWP+`: Intentionally over-predicts in certain range of wind speeds

Bias of `DeepWP+` model



- `DeepWP+` training starts at iteration 500 using a pre-trained `DeepWP` model
- $\mathtt{RMSE}(\widehat{w})$ and $\mathtt{RMSE}(\widehat{\pi})$ are conflicting objectives which are kept in balance

Underlying trade-offs between forecast errors, price errors, and fairness

case	DeepWP				DeepWP+							
	RMSE($\widehat{w}$)	RMSE($\widehat{\pi}$)	CVaR($\widehat{\pi}$)	$\alpha-$value	RMSE($\widehat{w}$)		RMSE($\widehat{\pi}$)		CVaR($\widehat{\pi}$)		$\alpha-$value	
	MWh	\$/MWh	\$/MWh	\$/MWh	MWh	gain	\$/MWh	gain	\$/MWh	gain	\$/MWh	gain
14_ieee	0.35	0.62	1.52	0	0.35	+0.6%	0.61	**−0.6%**	1.50	−0.8%	0	—
57_ieee	2.31	11.03	34.64	32.08	2.60	+11.2%	10.72	**−2.9%**	33.59	−3.1%	30.92	−3.8%
24_ieee	4.08	8.62	37.70	27.48	4.51	+9.6%	8.33	**−3.5%**	36.35	−3.7%	26.26	−4.6%
39_epri	5.94	11.15	31.21	17.53	6.43	+7.6%	10.19	**−9.4%**	28.02	−11.4%	15.84	−10.7%
73_ieee	4.02	5.12	16.21	32.83	5.51	+26.9%	4.24	**−20.8%**	13.41	−20.9%	26.63	−23.3%
118_ieee	2.29	3.59	11.32	17.91	2.60	+12.1%	2.88	**−24.7%**	9.06	−25.0%	14.09	−27.2%

- Worst-case improvement exceeds that of the average case
- Price error reduction and fairness improves with the size of the network

Conclusions

- Erronouse nature of ML leads to decision errors and algorithmic unfairness
- No need to re-design pricing algorithms to improve fairness
- It is sufficient to provide informed inputs (e.g., forecast)

Thank you for your attention!

Conclusions

- Erronouse nature of ML leads to decision errors and algorithmic unfairness
- No need to re-design pricing algorithms to improve fairness
- It is sufficient to provide informed inputs (e.g., forecast)

Thank you for your attention!