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
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 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

\[
\begin{align*}
\text{minimize} \quad & p^T C p + c^T p \\ 
\text{subject to} \quad & \mathbf{1}^T (p + \hat{w} - d) = 0 : \hat{\lambda}_b, \\
& |F(p + \hat{w} - d)| \leq \bar{\tau} : \hat{\lambda}_\tau, \hat{\lambda}_f,
\end{align*}
\]

\text{conventional generator dispatch cost} \quad \text{power balance condition} \quad \text{power flow limits}

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

\[
\pi(\hat{w}) = \hat{\lambda}_b \cdot \mathbf{1} - F^T (\hat{\lambda}_\tau - \hat{\lambda}_f)
\]

uniform price \quad adjustment due to congestion

which are unique w.r.t forecast \( \hat{w} \) under reasonable assumptions!

The LMP error is then defined as:

\[
\delta \pi = \pi(\hat{w}) - \pi(w)
\]

i.e., the distance between LMPs induced on the forecast (\( \hat{w} \)) and actual realization (\( w \)) of wind power.
Electricity market-clearing optimization

\[
\begin{align*}
\text{minimize} & \quad p^T Cp + c^T p \\
\text{subject to} & \quad 1^T (p + \hat{w} - d) = 0 : \hat{\lambda}_b, \\
& \quad |F(p + \hat{w} - d)| \leq \bar{F} : \hat{\lambda}_f, \hat{\lambda}_f,
\end{align*}
\]

**conventional generator dispatch cost**

**power balance condition**

**power flow limits**

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

\[
\pi(\hat{w}) = \begin{cases} 
\hat{\lambda}_b \cdot 1 & \text{uniform price} \\
F^T(\hat{\lambda}_f - \hat{\lambda}_f) & \text{adjustment due to congestion}
\end{cases}
\]

which are unique w.r.t. forecast \(\hat{w}\) under reasonable assumptions!

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

\[
\delta \pi = \pi(\hat{w}) - \pi(w)
\]

i.e., the distance between LMPs induced on the forecast \((\hat{w})\) and actual realization \((w)\) of wind power.
Electricity market-clearing optimization

\[
\begin{align*}
\text{minimize} & \quad p^T C p + c^T p & \quad \text{conventional generator dispatch cost} \\
\text{subject to} & \quad 1^T (p + \hat{w} - d) = 0 : \hat{\lambda}_b, & \quad \text{power balance condition} \\
& \quad |F(p + \hat{w} - d)| \leq \bar{f} : \hat{\lambda}_f, & \quad \text{power flow limits}
\end{align*}
\]

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

\[
\pi(\hat{w}) = \hat{\lambda}_b \cdot 1 - F^T (\hat{\lambda}_f - \hat{\lambda}_b)
\]

which are unique w.r.t forecast \(\hat{w}\) under reasonable assumptions!

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

\[
\delta \pi = \pi(\hat{w}) - \pi(w)
\]

i.e., the distance between LMPs induced on the forecast \((\hat{w})\) 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 \( \alpha — fairness \):

\[ \alpha = \max_{i \in \{1, \ldots, n\}} |E[||\delta \pi_i||] - E[||\delta \pi_{ref}||]| \]
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:

$$\alpha = \max_{i \in 1, \ldots, n} \|E[\|\delta \pi_i\|] - E[\|\delta \pi_{ref}\|]\|$$
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:**

\[
\alpha = \max_{i \in 1, \ldots, n} |E[|\delta \pi_i|] - E[|\delta \pi_{\text{ref}}|]|\
\]
Price-awareness for wind power forecast

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

$$\text{loss function: } ||\hat{w} - w||$$

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 \( \varphi \) and measurements \( w \)
- Two deep learning architectures DeepWP and DeepWP+ for wind power forecasting:

\[
\text{loss function: } ||\hat{w} - w|| + ||\pi(\hat{w}) - \pi(w)||
\]

- DeepWP+ informs wind power predictions about the downstream pricing errors
Market clearing as an optimization layer

Market-clearing optimization  \(\implies\) Equivalent primal form  \(\implies\) Equivalent dual form

\[
\begin{align*}
\text{minimize} & & p^T C p + c^T p \\
\text{subject to} & & 1^T (p + \hat{w} - d) = 0 \\
\ & & |F(p + \hat{w} - d)| \leq \bar{f}
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & & p^T C p + c^T p \\
\text{subject to} & & A p \geq b(\hat{w}) : \lambda
\end{align*}
\]

\[
\begin{align*}
\text{maximize} & & \left(AC^{-1} c + b(\hat{w})\right)^T \lambda \\
\ & & - \lambda^T AC^{-1} A^T \lambda
\end{align*}
\]

large constrained optimization  only inequality constraints  only non-negativity constraints
Market clearing as an optimization layer

Market-clearing optimization \(\Rightarrow\) Equivalent primal form \(\Rightarrow\) Equivalent dual form

\[
\begin{align*}
\text{minimize} & \quad p^T C p + c^T p \\
\text{subject to} & \quad 1^T (p + \hat{w} - d) = 0 \\
& \quad |F(p + \hat{w} - d)| \leq \bar{f}
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad p^T C p + c^T p \\
\text{subject to} & \quad A p \geq b(\hat{w}) : \lambda \\
\end{align*}
\]

\[
\begin{align*}
\text{maximize} & \quad \left(AC^{-1}c + b(\hat{w})\right)^T \lambda \\
& \quad - \lambda^T AC^{-1} A^T \lambda
\end{align*}
\]

large constrained optimization only inequality constraints only non-negativity constraints
Market clearing as an optimization layer

Market-clearing optimization $\implies$ Equivalent primal form $\implies$ Equivalent dual form

\[
\begin{align*}
\text{minimize} & \quad p^T Cp + c^T p \\
\text{subject to} & \quad \mathbf{1}^T (p + \hat{w} - d) = 0 \\
& \quad |F(p + \hat{w} - d)| \leq \overline{f}
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad p^T Cp + c^T p \\
\text{subject to} & \quad \mathbf{A} p \geq \mathbf{b}(
\hat{\mathbf{w}}) : \lambda \\
& \quad \lambda \geq 0
\end{align*}
\]

large constrained optimization only inequality constraints only non-negativity constraints
Market clearing as an optimization layer

Market-clearing optimization $\implies$ Equivalent primal form $\implies$ Equivalent dual form

\begin{align*}
\text{minimize} & \quad p^T Cp + c^T p & \quad \text{minimize} & \quad p^T Cp + c^T p & \quad \text{maximize} & \quad \left(AC^{-1} c + b(\hat{w})\right)^T \lambda \\text{subject to} & \quad 1^T (p + \hat{w} - d) = 0 & \quad \text{subject to} & \quad Ap \succeq b(\hat{w}) : \lambda & \quad - \lambda^T AC^{-1} A^T \lambda \\
& \quad |F(p + \hat{w} - d)| \leq \bar{f} &
\end{align*}

large constrained optimization \quad only inequality constraints \quad 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

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

DeepWP+: Price error minimization yields $\delta \pi \in [-1, 1] \ $/MWh
Wind power forecasts

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.
- $\text{RMSE}(\hat{w})$ and $\text{RMSE}(\hat{\tau})$ are conflicting objectives which are kept in balance.
**Underlying trade-offs between forecast errors, price errors, and fairness**

<table>
<thead>
<tr>
<th>Case</th>
<th>DeepWP</th>
<th></th>
<th></th>
<th></th>
<th>DeepWP+</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE((\hat{\omega}))</td>
<td>RMSE((\hat{\pi}))</td>
<td>CVaR((\hat{\pi}))</td>
<td>(\alpha)-value</td>
<td>RMSE((\hat{\omega}))</td>
<td>RMSE((\hat{\pi}))</td>
<td>CVaR((\hat{\pi}))</td>
<td>(\alpha)-value</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MWh</td>
<td>$/MWh</td>
<td>$/MWh</td>
<td>$/MWh</td>
<td>MWh</td>
<td>$/MWh</td>
<td>$/MWh</td>
<td>$/MWh</td>
<td></td>
</tr>
<tr>
<td>14.ieee</td>
<td>0.35</td>
<td>0.62</td>
<td>1.52</td>
<td>0</td>
<td>0.35</td>
<td>0.61</td>
<td>-0.6%</td>
<td>1.50</td>
<td>-0.8%</td>
</tr>
<tr>
<td>57.ieee</td>
<td>2.31</td>
<td>11.03</td>
<td>34.64</td>
<td>32.08</td>
<td>2.60</td>
<td>10.72</td>
<td>-2.9%</td>
<td>33.59</td>
<td>-3.1%</td>
</tr>
<tr>
<td>24.ieee</td>
<td>4.08</td>
<td>8.62</td>
<td>37.70</td>
<td>27.48</td>
<td>4.51</td>
<td>8.33</td>
<td>-3.5%</td>
<td>36.35</td>
<td>-3.7%</td>
</tr>
<tr>
<td>39.epri</td>
<td>5.94</td>
<td>11.15</td>
<td>31.21</td>
<td>17.53</td>
<td>6.43</td>
<td>10.19</td>
<td>-9.4%</td>
<td>28.02</td>
<td>-11.4%</td>
</tr>
<tr>
<td>73.ieee</td>
<td>4.02</td>
<td>5.12</td>
<td>16.21</td>
<td>32.83</td>
<td>5.51</td>
<td>4.24</td>
<td>-20.8%</td>
<td>13.41</td>
<td>-20.9%</td>
</tr>
<tr>
<td>118.ieee</td>
<td>2.29</td>
<td>3.59</td>
<td>11.32</td>
<td>17.91</td>
<td>2.60</td>
<td>2.88</td>
<td>-24.7%</td>
<td>9.06</td>
<td>-25.0%</td>
</tr>
</tbody>
</table>

- **Worst-case improvement exceeds that of the average case**
- **Price error reduction and fairness improves with the size of the network**
Underlying trade-offs between forecast errors, price errors, and fairness

<table>
<thead>
<tr>
<th>case</th>
<th>DeepWP</th>
<th>DeepWP+</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE($\hat{\omega}$)</td>
<td>RMSE($\hat{\pi}$)</td>
<td>CVaR($\hat{\pi}$)</td>
</tr>
<tr>
<td></td>
<td>MWh $/\text{MWh}$</td>
<td>$/\text{MWh}$</td>
<td>$/\text{MWh}$</td>
</tr>
<tr>
<td>14_ieee</td>
<td>0.35 0.62</td>
<td>1.52 0</td>
<td>0</td>
</tr>
<tr>
<td>57_ieee</td>
<td>2.31 11.03</td>
<td>34.64 32.08</td>
<td>2.60 +11.2%</td>
</tr>
<tr>
<td>24_ieee</td>
<td>4.08 8.62</td>
<td>37.70 27.48</td>
<td>4.51 +9.6%</td>
</tr>
<tr>
<td>39_epri</td>
<td>5.94 11.15</td>
<td>31.21 17.53</td>
<td>6.43 +7.6%</td>
</tr>
<tr>
<td>73_ieee</td>
<td>4.02 5.12</td>
<td>16.21 32.83</td>
<td>5.51 +26.9%</td>
</tr>
<tr>
<td>118_ieee</td>
<td>2.29 3.59</td>
<td>11.32 17.91</td>
<td>2.60 +12.1%</td>
</tr>
</tbody>
</table>

- Worst-case improvement exceeds that of the average case
- Price error reduction and fairness improves with the size of the network
Underlying trade-offs between forecast errors, price errors, and fairness

<table>
<thead>
<tr>
<th>case</th>
<th>DeepWP</th>
<th></th>
<th>DeepWP+</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE((\hat{\omega}))</td>
<td>RMSE((\hat{\pi}))</td>
<td>CVaR((\hat{\pi}))</td>
<td>(\alpha)–value</td>
</tr>
<tr>
<td></td>
<td>MWh</td>
<td>$/MWh</td>
<td>$/MWh</td>
<td>MWh</td>
</tr>
<tr>
<td>14.ieee</td>
<td>0.35</td>
<td>0.62</td>
<td>1.52</td>
<td>0</td>
</tr>
<tr>
<td>57.ieee</td>
<td>2.31</td>
<td>11.03</td>
<td>34.64</td>
<td>32.08</td>
</tr>
<tr>
<td>24.ieee</td>
<td>4.08</td>
<td>8.62</td>
<td>37.70</td>
<td>27.48</td>
</tr>
<tr>
<td>39.epri</td>
<td>5.94</td>
<td>11.15</td>
<td>31.21</td>
<td>17.53</td>
</tr>
<tr>
<td>73.ieee</td>
<td>4.02</td>
<td>5.12</td>
<td>16.21</td>
<td>32.83</td>
</tr>
<tr>
<td>118.ieee</td>
<td>2.29</td>
<td>3.59</td>
<td>11.32</td>
<td>17.91</td>
</tr>
</tbody>
</table>

- Worst-case improvement exceeds that of the average case
- Price error reduction and fairness improves with the size of the network
Underlying trade-offs between forecast errors, price errors, and fairness

<table>
<thead>
<tr>
<th>case</th>
<th>DeepWP</th>
<th>DeepWP+</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE($\hat{x}$)</td>
<td>RMSE($\hat{y}$)</td>
<td>CVaR($\hat{y}$)</td>
<td>$\alpha -$value</td>
<td>RMSE($\hat{x}$)</td>
<td>RMSE($\hat{y}$)</td>
<td>CVaR($\hat{y}$)</td>
<td>$\alpha -$value</td>
<td>$\text{UR}$</td>
<td>$\text{gain}$</td>
<td>$\text{gain}$</td>
</tr>
<tr>
<td></td>
<td>MWh</td>
<td>$/\text{MWh}$</td>
<td>$/\text{MWh}$</td>
<td>$/\text{MWh}$</td>
<td>MWh</td>
<td>$/\text{MWh}$</td>
<td>$/\text{MWh}$</td>
<td>$/\text{MWh}$</td>
<td>MWh</td>
<td>$/\text{MWh}$</td>
<td>$/\text{MWh}$</td>
</tr>
<tr>
<td>14.ieee</td>
<td>0.35</td>
<td>0.62</td>
<td>1.52</td>
<td>0</td>
<td>0.35</td>
<td>0.61</td>
<td>-0.6%</td>
<td>1.50</td>
<td>-0.8%</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>57.ieee</td>
<td>2.31</td>
<td>11.03</td>
<td>34.64</td>
<td>32.08</td>
<td>2.60</td>
<td>10.72</td>
<td>-2.9%</td>
<td>33.59</td>
<td>-3.1%</td>
<td>30.92</td>
<td>-3.8%</td>
</tr>
<tr>
<td>24.ieee</td>
<td>4.08</td>
<td>8.62</td>
<td>37.70</td>
<td>27.48</td>
<td>4.51</td>
<td>8.33</td>
<td>-3.5%</td>
<td>36.35</td>
<td>-3.7%</td>
<td>26.26</td>
<td>-4.6%</td>
</tr>
<tr>
<td>39.epri</td>
<td>5.94</td>
<td>11.15</td>
<td>31.21</td>
<td>17.53</td>
<td>6.43</td>
<td>10.19</td>
<td>-9.4%</td>
<td>28.02</td>
<td>-11.4%</td>
<td>15.84</td>
<td>-10.7%</td>
</tr>
<tr>
<td>73.ieee</td>
<td>4.02</td>
<td>5.12</td>
<td>16.21</td>
<td>32.83</td>
<td>5.51</td>
<td>4.24</td>
<td>-20.8%</td>
<td>13.41</td>
<td>-20.9%</td>
<td>26.63</td>
<td>-23.3%</td>
</tr>
<tr>
<td>118.ieee</td>
<td>2.29</td>
<td>3.59</td>
<td>11.32</td>
<td>17.91</td>
<td>2.60</td>
<td>2.88</td>
<td>-24.7%</td>
<td>9.06</td>
<td>-25.0%</td>
<td>14.09</td>
<td>-27.2%</td>
</tr>
</tbody>
</table>

- 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!

Price-Aware Deep Learning for Electricity Markets

Vladimir Dvorkin
Massachusetts Institute of Technology
Cambridge, MA 02109
dvorkin@mit.edu

Ferdinando Fioretto
University of Virginia
Charlottesville, VA 22903
fioretto@virginia.edu