# Approximating complex functions in mixed integer non-linear optimization: a hybrid data- and knowledge-driven approach

Claudia D'Ambrosio
LIX CNRS, École Polytechnique, Institut Polytechnique de Paris

joint work with

M. Cuesta, M. Durbán, V. Guerrero, R. Spencer Trindade

**Acknowledgments**
Programme Gaspard Monge pour l'Optimisation (PGMO), Integrated Urban Mobility (IUM) chair,
PID2019-104901RB-I00, PDC2022-133359-I00, PID2022-13724OB-I00 and IJC2020-045220-I
(funded by MCIN/AEI/10.13039/501100011033 and the last also supported by European Union "NextGenerationEU/PRTR" funds).

MIP Workshop 2025, Minneapolis, U.S.A.

# Outline

# Outline

# Introduction

- Solving **MINLP problems** faster by replacing **complex functions** with simpler ones (**surrogate**).

$$
\begin{aligned}
\min_{x \in \mathbb{R}^p} g_0(x) & \\
g_m(x) \leq 0 & \qquad \forall m = 1, \dots, \overline{m} \\
x_j \in \mathbb{Z} & \qquad \forall j \in I \subseteq \{1, \dots, p\} \\
\underline{x}_j \leq x_j \leq \overline{x}_j & \qquad \forall j \in \{1, \dots, p\}.
\end{aligned}
$$

$$\boxed{\text{MINLP with complex functions } g(x)} \longrightarrow \boxed{\text{Surrogate MI(NL)P with surrogate functions } \tilde{g}(x)}$$

# Introduction

- Solving **MINLP problems** faster by replacing **complex functions** with simpler ones (**surrogate**).

$$
\begin{aligned}
&\min_{x\in\mathbb{R}^p} \tilde{g}_0(x) && \\
&\tilde{g}_m(x) \le 0 && \forall m = 1,\ldots,\overline{m} \\
&x_j \in \mathbb{Z} && \forall j \in I \subseteq \{1,\ldots,p\} \\
&\underline{x}_j \le x_j \le \overline{x}_j && \forall j \in \{1,\ldots,p\}.
\end{aligned}
$$

**MINLP with complex functions $g(x)$** → **Surrogate MI(NL)P with surrogate functions $\tilde{g}(x)$**

# Knowledge-driven Approaches

MINLP with complex functions $g(x)$ $\xrightarrow{\text{Knowledge-driven approach}}$ Surrogate MI(NL)P with surrogate functions $\tilde{g}(x)$

Rebennack and Kallrath [2015]: PWL approximation of bivariate functions.
Under- or over-estimate.
Guarantee that the approximation error remains within a given tolerance.

Codsi et al. [2025], Duguet and Ngueveu [2022]: PWL approximation of univariate/bivariate functions.
Bounded approximation error while minimizing the number of pieces of the PWL approximation.

Gößet al. [2025]: parabolic approximations of MINLPs.
To find the surrogate function $\rightarrow$ solve MILPs.

# Data-driven Approaches

$$\boxed{\text{MINLP with complex functions } g(x)} \xrightarrow{\text{Data-driven approach}} \boxed{\text{Surrogate MI(NL)P with surrogate functions } \tilde{g}(x)}$$

Sample $g(x) \rightarrow$ learn $\tilde{g}(x)$ with MILP-repres ML $\rightarrow$ solve $\tilde{P}$

- Bertsimas and Öztürk [2023]: hyperplane-based decision-trees
- Bertsimas and Margaritis [2025]: gradient boosted trees, multi layer perceptrons, support vector machine

# Data- vs Knowledge-driven Approaches: **Limitations**

**Data-driven**:

- too simplistic to capture the structure of the original MINLP Geißler et al. [2012]
- lack of interpretability [Goodman and Flaxman, 2017, Rudin et al., 2022]
- difficulty of incorporating expert knowledge [Gambella et al., 2021]

**Knowledge-driven**:

- tailored to specific nonlinear function $\rightarrow$ challenging generalization
- finding the surrogate might be time consuming, e.g., MILP solving Gößet al. [2025]

# Data+Knowledge-driven Approach

Mixed-integer Smoothing Surrogate Optimization with Constraints (**MiSSOC**)

- Data-driven approach where expert knowledge can be integrated
- B-splines $\rightarrow$ piecewise polynomials
- Surrogate problem $\tilde{P}$ is MINLP but more tractable $\rightarrow$ better approx?
- This flexibility comes at the expense of not being able to guarantee an error bound on the entire domain.

# Data+Knowledge-driven Approaches

$$\boxed{\text{MINLP with complex functions } g(x)} \xrightarrow{\text{Data+knowledge}} \boxed{\text{Surrogate MI(NL)P with surrogate functions } \tilde{g}(x)}$$

- **Data+knowledge-driven approach**: extension of [Navarro-García, Guerrero, and Durban, 2023, 2024] to **constrained smooth additive regression models**

- **Surrogate MINLPs**: MINLPs in which non-convexities in the objective function and/or constraints are sum of non-convex univariate functions [D'Ambrosio et al., 2012b, Spencer Trindade et al., 2023].

**Key point**: Find tradeoff between **approximation quality and tractability**.

# Outline

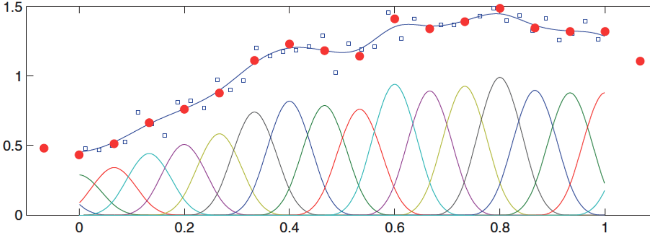
# Outline

# Additive regression models

- Approximate complex multivariate functions of the general MINLPs by simpler surrogate functions, $f_1, f_2, \ldots, f_p$.

$$y = g(x_1, x_2, \ldots, x_p) \approx f_1(x_1) + f_2(x_2) + \ldots + f_p(x_p)$$

- Each $f_j$ is approximated by a linear combination of $B-$splines basis functions.



# Univariate smoothing with $B-$splines

$\{(x_i, y_i)\}$, $i = 1, \ldots, n$, are $n$ realizations in the sample of a function $y_i = g(x_i)$

$k =$ number of intervals in which the domain is split $\rightarrow k + 1$ "internal" knots (equally-spaced and increasing, w.l.o.g.).

Function $g$ is approx as a linear combination of $k + d$ $B-$splines of degree $d$.

Coefficients $\theta_j$, for $j = 1, \ldots, k + d$, are found by solving:

$$\min_{\substack{\theta_j \in \mathbb{R} \\ j=1,\ldots,k+d}} \sum_{i=1}^{n} \left( y_i - \sum_{j=1}^{k+d} \theta_j B_{j,d,\mathbf{t}}(x_i) \right)^2$$

s.t. Sign, monotonicity or shape requirements if needed.

- $B_{j,d,\mathbf{t}}(x_i)$; $j-$th $B-$spline basis function associated evaluated at $x_i$;
- $\theta_j$ is the coefficient associated to the $j-$th basis function

# Data-driven surrogate MINLP

- The complex multivariate function

$$y = g(x) \approx f_1(x_1) + f_2(x_2) + \ldots + f_p(x_p)$$

  is approximated by sums of univariate piecewise polynomial functions.

- Function $g$ appearing in the MINLP is replaced by this approximation.

Why sum of **univariate surrogate function**?

Computational tractability: SC-MINLP can be used to obtain the global solution of such a surrogate MINLP model, as well as other MINLP solvers.

# Outline

# Sequential Convex - MINLP

Sequential Convex - MINLP (SC-MINLP) [D'Ambrosio et al., 2012b, Spencer Trindade et al., 2023] solves MINLPs in which non-convexities in the objective function and/or constraints appear as the sum of non-convex univariate functions faster than standard solvers for MINLPs.

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \sum_{j \in N} c_j x_j \\ \text{subject to} \quad & h(x) \leq 0, \\ & r_i(x) + \sum_{k \in H(i)} f_{ik}(x_k) \leq 0 \quad && i \in M, \\ & l_j \leq x_j \leq u_j && \forall j \in N, \\ & x_j \text{ integer} && \forall j \in J, \end{aligned}$$

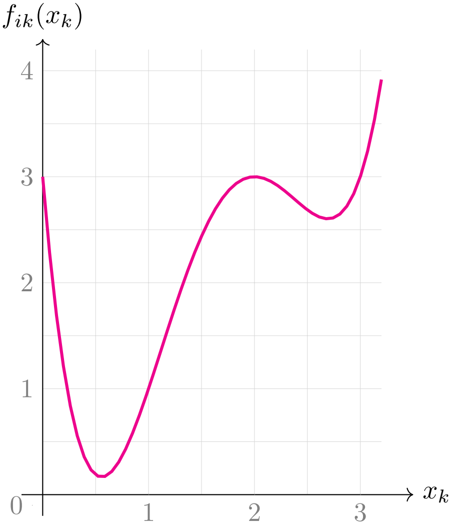
where $N = \{1, \ldots, n\}$, $M = \{1, \ldots, m\}$, $H(i) \subseteq N$, $J \subseteq N$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}^q$ and $r_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ are multivariate convex functions and $f_{ik} : \mathbb{R} \longrightarrow \mathbb{R}$ are non-convex univariate functions, $\forall k \in H(i)$, $\forall i \in M$.

# Sequential Convex - MINLP

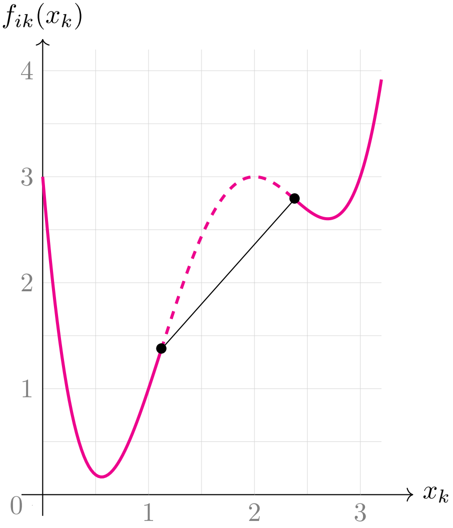
The **SC-MINLP**: iterative technique

- **Lower bound**: solve a **convex MINLP relaxation** of the separable MINLP.
  - Computing a piecewise-convex relaxation of each $f_{ik}$: the concave parts are substituted for linear functions and the convex parts are kept as they are.
- **Upper bound**: solve a **non-convex NLP restriction** of the separable MINLP problem obtained by fixing the variables $x_j \in J$.
  - Locally solving the non-convex NLP restriction an upper bound of the original MINLP problem is obtained.
- **Refinement technique**: improve the quality of the lower bound and thus decrease the gap between the lower and upper bounds.

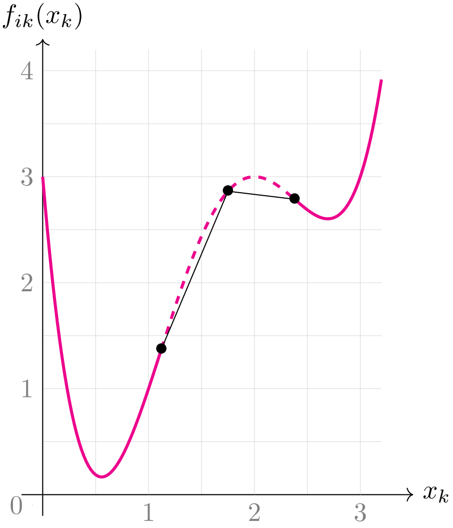
# Sequential Convex - MINLP



# Sequential Convex - MINLP



# Sequential Convex - MINLP



# Outline

# The proposed approach

**Input**: Problem $P$; degree $d$; intervals $k$.

Using random sampling, generate set $\mathcal{T} = \{(x_i, g_0(x_i)),\ i = 1, \ldots, n\}$.
$\tilde{g}_0 \leftarrow$ estimate $g_0$ using data in $\mathcal{T}$.
Generate $\tilde{P}$ using $\tilde{g}_0$.
$\tilde{\mathbf{x}} \leftarrow$ solution of $\tilde{P}$ with SC-MINLP.
$\mathbf{x}^* \leftarrow$ solve the restricted NLP using $\tilde{\mathbf{x}}$ as a warm start.

**Output**: $\tilde{\mathbf{x}}$; $\mathbf{x}^*$.

$\tilde{P}$ has a piecewise polynomial function replacing complex $g_m(x)$ for $m \in C = \{0\}$.

# Outline

# Preliminary Computational Results

Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz processor with 64GB RAM.

**Preliminary test:**

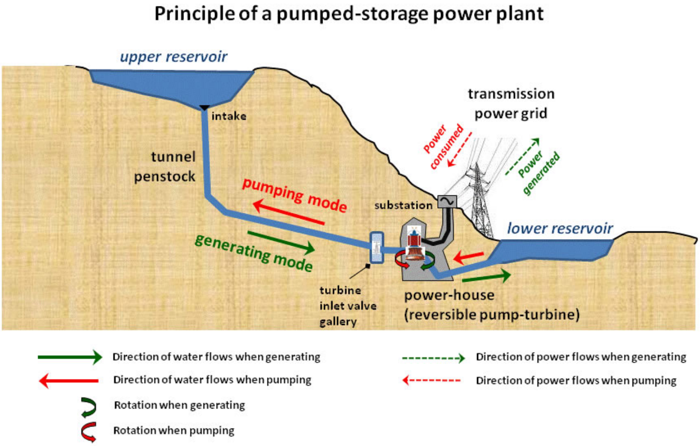
- Number of intervals of the B-spline $k = 10$
- Degree of the polynomials: $d = 3$

Surrogate functions found in **milliseconds**.

# Outline

# Single-reservoir Hydro Unit Commitment Problem

Several units which can operate as turbine/pump or be off.




- Turbine: water flow $q$ goes downhill and produces power $p$.
- Pump: water flow $q$ goes uphill and consumes power $p$.

# Single reservoir HUC, Taktak & D'A. (2017)

**Physical constraints**:

- Water flow balance equations
- Respect allowed operational points: (dis-)continuous, discrete, turbine/pump related
- Forbid of simultaneous pump and turbine mode
- Power production depending on water flow and head effect
- Minimum number of periods to be spent in a status by the unit (minimum starting up/down times)
- spillage
- ...

**Strategic constraints**:

- Ramp up/down bound constraint
- Irrigation requirement/Ecological flows/Water rights
- Load balance equations constraints
- Minimum release of water per period
- Minumum final reservoir level

**Objective Function**:

- **Maximize profit**

# HUC: the complex function

Power function:

$$\varphi(q, v) = \frac{9.81}{1000} \cdot q \cdot \sum_{h=0}^{6} \left( L_h q^h \left( \sum_{k=0}^{6} K_k (1000^k v^k) - \underline{L} - R_0 q^2 \right) \right)$$

$q =$ **water flow** in the unit [m$^3$/s]
$v =$ **water volume** in the basin [m$^3$]

# Hydro Unit Commitment

| | | MINLP | Surrogate $f(v,q)$ | | | Surrogate $f(v,q,v+q)$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | B-splines degree | | | B-splines degree | | |
| | | | 2 | 3 | 4 | 2 | 3 | 4 |
| Baron | Real Objective | 3692.91 | 14538.87 | 14540.09 | 14540.08 | 14517.91 | 14385.71 | 14539.90 |
| | Surrogate Objective | | 14615.62 | 14617.81 | 14617.62 | 14527.46 | 14392.81 | 14546.37 |
| | Time | 600.00 | 156.41 | 413.79 | 446.69 | 600.24 | 600.22 | 600.70 |
| Bonmin | Real Objective | 7285.32 | 14184.88 | 3578.29 | 12996.68 | 14024.72 | 14018.56 | 14023.47 |
| | Surrogate Objective | | 14226.42 | 3384.58 | 13059.56 | 14036.07 | 14026.22 | 14031.38 |
| | Time | 1.11 | 607.74 | 607.79 | 607.37 | 604.90 | 607.91 | 605.15 |
| Couenne | Real Objective | - | 10383.09 | 12697.24 | 11548.79 | - | -906.69 | -19510.34 |
| | Surrogate Objective | | 10394.44 | 12797.97 | 11589.48 | - | -908.53 | -19510.44 |
| | Time | 0.22 | 604.46 | 601.67 | 601.65 | 0.36 | 604.97 | 603.74 |
| SC-MINLP | Real Objective | | 14538.77 | 14540.08 | 14540.05 | 14539.25 | 14540.12 | 14540.10 |
| | Surrogate Objective | | 14615.61 | 14617.79 | 14617.60 | 14545.02 | 14546.98 | 14546.51 |
| | Time | | 1.52 | 1.55 | 1.58 | 600.02 | 600.18 | 600.07 |

# Hydro Unit Commitment

| | | MINLP | Surrogate $f(v,q)$ | | | Surrogate $f(v,q,v+q)$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | B-splines degree | | | B-splines degree | | |
| | | | 2 | 3 | 4 | 2 | 3 | 4 |
| Baron | Real Objetive | 3692.91 | **14538.87** | **14540.09** | **14540.08** | 14517.91 | 14385.71 | 14539.90 |
| | Surrogate Objective | | 14615.62 | 14617.81 | 14617.62 | 14527.46 | 14392.81 | 14546.37 |
| | Time | 600.00 | 156.41 | 413.79 | 446.69 | 600.24 | 600.22 | 600.70 |
| Bonmin | Real Objective | 7285.32 | 14184.88 | 3578.29 | 12996.68 | 14024.72 | 14018.56 | 14023.47 |
| | Surrogate Objective | | 14226.42 | 3384.58 | 13059.56 | 14036.07 | 14026.22 | 14031.38 |
| | Time | 1.11 | 607.74 | 607.79 | 607.37 | 604.90 | 607.91 | 605.15 |
| Couenne | Real Objective | - | 10383.09 | 12697.24 | 11548.79 | - | -906.69 | -19510.34 |
| | Surrogate Objective | | 10394.44 | 12797.97 | 11589.48 | - | -908.53 | -19510.44 |
| | Time | 0.22 | 604.46 | 601.67 | 601.65 | 0.36 | 604.97 | 603.74 |
| SC-MINLP | Real Objective | | **14538.77** | **14540.08** | **14540.05** | 14539.25 | 14540.12 | 14540.10 |
| | Surrogate Objective | | 14615.61 | 14617.79 | 14617.60 | 14545.02 | 14546.98 | 14546.51 |
| | Time | | 1.52 | 1.55 | 1.58 | 600.02 | 600.18 | 600.07 |

Degree 1 (MILP): Time **0.23"**; Surrogate obj **14633.39**; Real obj **14533.96**

# Hydro Unit Commitment

| | | MINLP | Surrogate $f(v,q)$ B-splines degree 2 | 3 | 4 | Surrogate $f(v,q,v+q)$ B-splines degree 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|
| Baron | Real Objetive | 3692.91 | **14538.87** | **14540.09** | **14540.08** | **14517.91** | 14385.71 | **14539.90** |
| | Surrogate Objective | | 14615.62 | 14617.81 | 14617.62 | 14527.46 | 14392.81 | 14546.37 |
| | Time | 600.00 | 156.41 | 413.79 | 446.69 | 600.24 | 600.22 | 600.70 |
| Bonmin | Real Objective | 7285.32 | 14184.88 | 3578.29 | 12996.68 | 14024.72 | 14018.56 | 14023.47 |
| | Surrogate Objective | | 14226.42 | 3384.58 | 13059.56 | 14036.07 | 14026.22 | 14031.38 |
| | Time | 1.11 | 607.74 | 607.79 | 607.37 | 604.90 | 607.91 | 605.15 |
| Couenne | Real Objective | - | 10383.09 | 12697.24 | 11548.79 | - | -906.69 | -19510.34 |
| | Surrogate Objective | | 10394.44 | 12797.97 | 11589.48 | - | -908.53 | -19510.44 |
| | Time | 0.22 | 604.46 | 601.67 | 601.65 | 0.36 | 604.97 | 603.74 |
| SC-MINLP | Real Objective | | **14538.77** | **14540.08** | **14540.05** | **14539.25** | **14540.12** | **14540.10** |
| | Surrogate Objective | | 14615.61 | 14617.79 | 14617.60 | 14545.02 | 14546.98 | 14546.51 |
| | Time | | 1.52 | 1.55 | 1.58 | **600.02** | **600.18** | **600.07** |

# Outline

# Conclusions & Ongoing work

- **Statistical modeling** + **mathematical optimization** to get surrogate MINLPs.
- Think about **approximation quality and tractability** when identifying the surrogate function.
- Tailored and general-purpose **solvers**.

**Ongoing work:**

- More extensive and stable results (test new gurobi).
- Considering black-box functions.
- **Under/over estimation** to obtain lower/upper bounds.

**Future work:**

- How to select **intervals, degree**, or, in general, surrogate function **property**?
- **Dynamically add $B-$spline basis functions** to increase the quality of the approximation in some regions.

# References

D. Bertsimas and G. Margaritis. Global optimization: a machine learning approach. *Journal on Global Optimization*, 2025.

D. Bertsimas and B. Öztürk. Global optimization via optimal decision trees. *Journal on Global Optimization*, 2023.

A. Bhosekar and M. Ierapetritou. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Computers & Chemical Engineering*, 108:250–267, 2018.

Julien Codsi, Sandra Ulrich Ngueveu, and Bernard Gendron. Lina: a faster approach to piecewise linear approximations using corridors and its application to mixed-integer optimization. *Mathematical Programming Computation*, 17(2):265–306, 6 2025. ISSN 1867-2957.

Carl De Boor. *A practical guide to splines*, volume 27. Springer, New York, 1978.

Aloïs Duguet and Sandra Ulrich Ngueveu. Piecewise linearization of bivariate nonlinear functions: Minimizing the number of pieces under a bounded approximation error. In Ivana Ljubić, Francisco Barahona, Santanu S. Dey, and A. Ridha Mahjoub, editors, *Combinatorial Optimization*, pages 117–129, Cham, 2022. Springer International Publishing.

C. D'Ambrosio, J. Lee, and A. Wächter. An algorithmic framework for MINLP with separable non-convexity. In J. Lee and S. Leyffer, editors, *Mixed Integer Nonlinear Programming*, volume 154 of *The IMA Volumes in Mathematics and its Applications*, pages 315–347. Springer, New York, 2012a.

C. D'Ambrosio, J. Lee, and A. Wächter. An algorithmic framework for MINLP with separable non-convexity. In J. Lee and S. Leyffer, editors, *Mixed Integer Nonlinear Programming*, volume 154 of *The IMA Volumes in Mathematics and its Applications*, pages 315–347. Springer, New York, 2012b.

C. D'Ambrosio, A. Frangioni, and C. Gentile. Strengthening the sequential convex MINLP technique by perspective reformulations. *Optimization Letters*, 13:673–684, 2019.

C. Gambella, B. Ghaddar, and J. Naoum-Sawaya. Optimization problems for machine learning: a survey. *Eurpean Journal of Operational Research*, 290(3):807–828, 2021.

# References


B. Geißler, A. Martin, A. Morsi, and L. Schewe. Using piecewise linear functions for solving MINLPs. In J. Lee and S. Leyffer, editors, *Mixed Integer Nonlinear Programming*, volume 154 of *The IMA Volumes in Mathematics and its Applications*, pages 287–314. Springer, New York, 2012.

B. Goodman and S. Flaxman. European union regulations on algorithmic decision-making and a "right to explanation"'. *AI Magazine*, 38(3):50–57, 2017.

Adrian Göß, Robert Burlacu, and Alexander Martin. Parabolic approximation & relaxation for minlp, 2025. URL https://arxiv.org/abs/2407.06143.

B. Grimstad and A. Sandnes. Global optimization with spline constraints: a new branch-and-bound method based on B-splines. *Journal of Global Optimization*, 65:401–439, 2016.

M. Navarro-García, V. Guerrero, and M. Durban. On constrained smoothing and out-of-range prediction using p-splines: a conic optimization approach. *Applied Mathematics and Computation*, 441:127679, 2023.

M. Navarro-García, V. Guerrero, and M. Durban. A mathematical optimization approach to shape-constrained generalized additive models. *Expert Systems with Applications*, 255(C):124654, 2024.

Steffen Rebennack and Josef Kallrath. Continuous piecewise linear delta-approximations for bivariate and multivariate functions. *Journal of Optimization Theory and Applications*, 167(1):102–117, 10 2015.

C. Rudin, C. Chen, Z. Chen, H. Huang, L. Semenova, and C. Zhong. Interpretable machine learning: fundamental principles and 10 grand challenges. *Statistics Surveys*, 1:1–85, 2022.

R. Spencer Trindade, C. D'Ambrosio, A. Frangioni, and C. Gentile. Comparing perspective reformulations for piecewise-convex optimization. *Operations Research Letters*, 51(6):702–708, 2023.

Juan Pablo Vielma, Shabbir Ahmed, and George Nemhauser. Mixed-integer models for nonseparable piecewise-linear optimization: Unifying framework and extensions. *Operations Research*, 58(2):303–315, 2009.

$$
\begin{aligned}
\text{maximize} \quad & \sum_{j\in J}\sum_{t\in T}\Big( \tau\, \Pi_t\, p_{jt} - C_j\, \widetilde{w}_{jt} - (D_j + \Pi_t E_j)\widetilde{z}_{jt}\Big) && \\
\text{subject to} \quad & v_{\bar{t}} = V_{\bar{t}}, && \\
& v_t = v_{t-1} + 3600\,\tau\,(I_t - \sum_{j\in J} q_{jt} - s_t) && t \in T, \\
& -Q^- \le \sum_{j\in J}(q_{jt} - q_{jt-1}) \le +Q^+ && t \in T, \\
& Q_j^-\, u_{jt} + \underline{Q}_j\, g_{jt} \le q_{jt} \le Q_j^-\, u_{jt} + \overline{Q}_j\, g_{jt} && j \in J,\ t \in T, \\
& P_j^-\, u_{jt} + \underline{P}_j\, g_{jt} \le p_{jt} \le P_j^-\, u_{jt} + \overline{P}_j\, g_{jt} && j \in J,\ t \in T, \\
& \sum_{j\in J} q_{jt} + s_t \ge \underline{\Theta} && t \in T, \\
& s_t - \sum_{j\in J}(W_j\, \widetilde{w}_{jt} + Y_j\, \widetilde{z}_{jt}) \ge 0 && t \in T, \\
& g_{jt} + u_{kt} \le 1 && j \in J,\ k \in J,\ t \in T, \\
& g_{jt} - g_{jt-1} - (\widetilde{w}_{jt} - w_{jt}) = 0 && j \in J,\ t \in T, \\
& \widetilde{w}_{jt} + w_{jt} \le 1 && j \in J,\ t \in T, \\
& u_{jt} - u_{jt-1} - (\widetilde{z}_{jt} - z_{jt}) = 0 && j \in J,\ t \in T, \\
& \widetilde{z}_{jt} + z_{jt} \le 1 && j \in J,\ t \in T, \\
& p_{jt} \le 9.81 q_{jt} \sum_{h=0}^{6}\left( L_h q_{jt}^h \left( \sum_{k=0}^{6} K_k v_t^k - \underline{L} - R_0 q_{jt}^2 \right)\right) g_{jt} && j \in J,\ t \in T, \\
& g_{jt},\ u_{jt},\ \widetilde{w}_{jt}, \widetilde{z}_{jt},\ w_{jt},\ z_{jt} \in \{0,1\} && j \in J,\ t \in T, \\
& Q_j^- \le q_{jt} \le \overline{Q}_j,\ \underline{V} \le v_t \le \overline{V},\ P_j^- \le p_{jt} \le \overline{P}_j,\ 0 \le s_t \le \overline{S} && j \in J,\ t \in T.
\end{aligned}
$$