Development of Advanced Stochastic Unit Commitment Formulation for Management of Uncertainty

C. Lindsay Anderson
M. Gabriela Martínez
Laura Lindley Tupper

August 6, 2014

CERTS
Consortium for Electric Reliability Technology Solutions

Overview

- Background & Objectives
- Current Progress
  - Chance-constrained unit commitment
  - Scenario generation and reduction
- Conclusions & Future Directions

Objective

Development of a solution for SCUC that will be

- Flexible: able to include uncertain renewable resources in a realistic way, integrate with other tools
- Robust: provide optimal (or ϵ-optimal) solutions
- Scalable: applicable to reasonably-sized systems in practical computation time

Objective

Development of a solution for SCUC that will be

- Flexible: able to include uncertain renewable resources in a realistic way, integrate with other tools
- Robust: provide optimal (or ϵ-optimal) solutions
- Scalable: applicable to reasonably-sized systems in practical computation time

Combination of formulation and algorithm implementation

Background

In 2013, we proposed to investigate chance-constrained formulations that will

- ease the requirement for perfectly binding constraints,
- require constraints to be met with some *large* probability
- ensure scalability through *separation of stochastic complexity and system complexity*

Background

In 2013, we proposed to investigate chance-constrained formulations that will

- ease the requirement for perfectly binding constraints,
- require constraints to be met with some *large* probability
- ensure scalability through *separation of stochastic complexity and system complexity*

Replicating the reliability requirements of the power system operator

Goals

These objectives led to two primary goals previously described:

1. Use modestly-sized test cases, to filter approach to most promising
2. Investigate methods for reduction of scenario sets and approximation of excluded scenarios

Goals

These objectives led to two primary goals previously described:

1. Use modestly-sized test cases, to filter approach to most promising → Gabriela Martínez
2. Investigate methods for reduction of scenario sets and approximation of excluded scenarios

Goals

These objectives led to two primary goals previously described:

1. Use modestly-sized test cases, to filter approach to most promising → Gabriela Martínez
2. Investigate methods for reduction of scenario sets and approximation of excluded scenarios→ Laura Tupper

Stochastic Models

Gabriela Martínez

CERTS
Consortium for Electric Reliability Technology Solutions

Stochastic Models

1. Chance-constrained UC Models
   - Quantile-region MIP approximation
   - Quantile-based Relaxation
2. Stochastic Two-stage Models
   - Risk-neutral two-stage UC Model (Expected recourse actions)
   - Risk-averse two-stage UC Model (Conditional value-at-risk recourse actions)

Goals: formulate numerically tractable stochastic models based on data – scenarios that capture statistical properties of renewable generation, and possibly load

Chance-constrained Optimization
Introduction

The general form of the optimization model considered:

- C_G total cost of thermal-generation
- P_G^t total thermal power generation at time t
- $\mathbf{L}^t$ total load
- $\mathbf{P}_r^t$ total renewable generation at time t
- T scheduling horizon

$$\begin{aligned} &\min C_G(P_G), \\ &\text{s.t:} P_G^t \geq \mathbf{L}^t - \mathbf{P}_r^t, t = 1, \ldots, T \\ &\quad P_G \in \mathcal{C}_G. \end{aligned}$$

Set of deterministic decisions $\mathcal{C}_G$ (operational constraints of thermal units). $\mathbf{L} - \mathbf{P}_r \in \mathbb{R}^T$ is a random vector.

Chance-constrained Model

Introduction

The general form of the optimization model considered:

- C_G total cost of thermal-generation
- P_G^t total thermal power generation at time t
- $\mathbf{L}^t$ total load
- $\mathbf{P}_r^t$ total renewable generation at time t
- T scheduling horizon

$$\begin{aligned} &\min C_G(P_G), \\ &\text{s.t:}\mathbb{P}(P_G^t \geq \mathbf{L}^t - \mathbf{P}_r^t, t = 1, \ldots, T) \geq \pi \\ &P_G \in \mathcal{C}_G. \end{aligned}$$

Distribution of data $\mathbf{L} - \mathbf{P}_r$ is represented with scenarios.

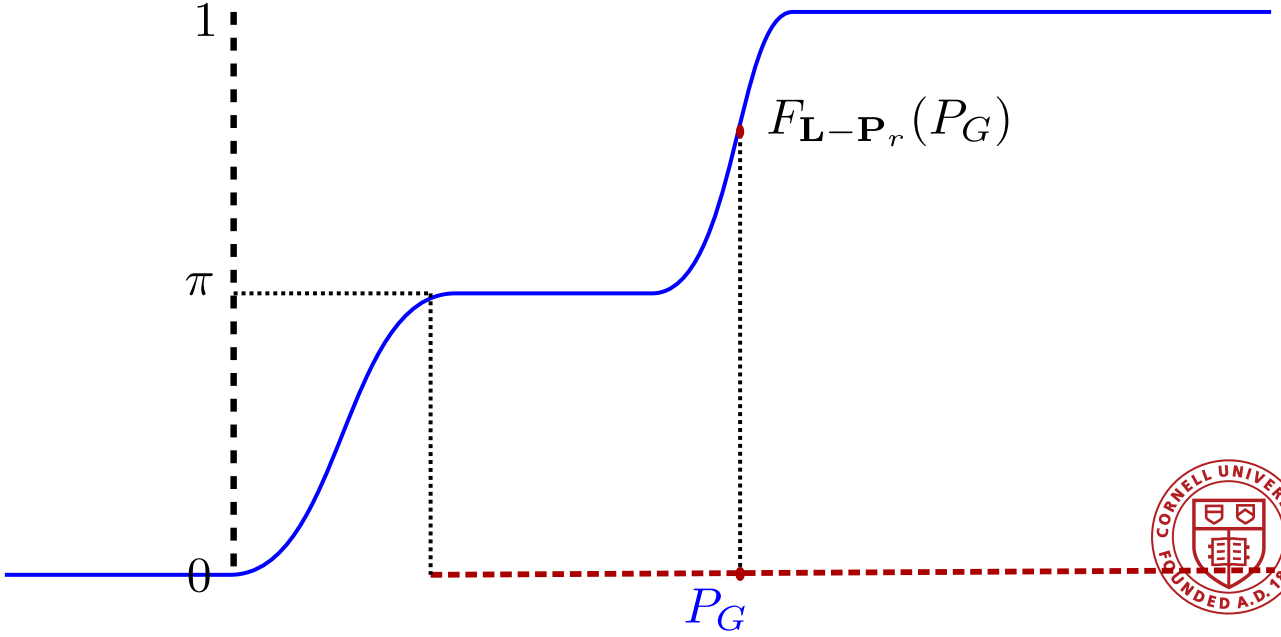
Chance-constrained Model

Introduction

The general form of the optimization model considered:

- C_G total cost of thermal-generation
- P_G^t total thermal power generation at time t
- $\mathbf{L}^t$ total load
- $\mathbf{P}_r^t$ total renewable generation at time t
- T scheduling horizon

$$\begin{aligned} &\min C_G(P_G), \\ &\text{s.t:} F_{\mathbf{L}-\mathbf{P}_r}(P_G) \geq \pi \\ &\quad P_G \in \mathcal{C}_G. \end{aligned}$$



CERTS
Consortium for Electric Reliability Technology Solutions

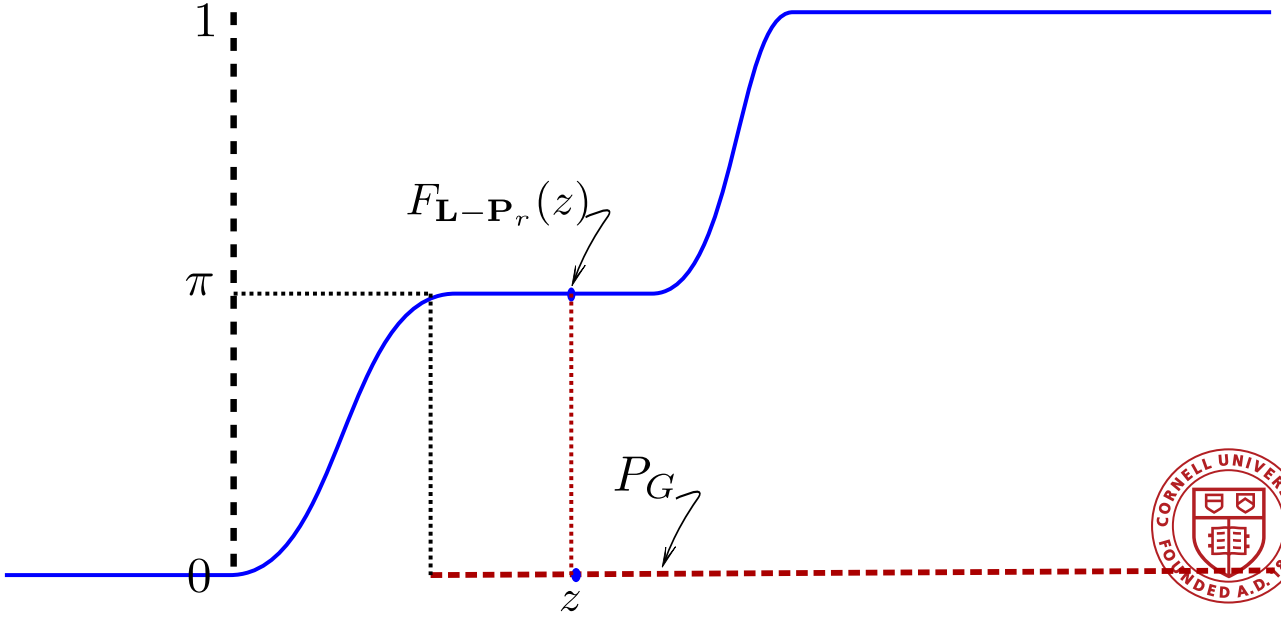
Chance-constrained Model

Introduction

The general form of the optimization model considered The general form of the optimization model considered:

- C_G total cost of thermal-generation
- P_G^t total thermal power generation at time t
- $\mathbf{L}^t$ total load
- $\mathbf{P}_r^t$ total renewable generation at time t
- T scheduling horizon

$$
\begin{aligned}
&\min C_G(P_G),\\
&\text{s.t:} P_G \geq z\\
&\qquad F_{\mathbf{L}-\mathbf{P}_r}(z) \geq \pi\\
&\qquad P_G \in \mathcal{C}_G.
\end{aligned}
$$



Chance-constrained Model

Introduction

The general form of the optimization model considered The general form of the optimization model considered:

- C_G total cost of thermal-generation
- P_G^t total thermal power generation at time t
- $\mathbf{L}^t$ total load
- $\mathbf{P}_r^t$ total renewable generation at time t
- T scheduling horizon

$$\min C_G(P_G),$$
$$\text{s.t:} P_G \geq z$$
$$z \in \mathcal{Z}_\pi$$
$$P_G \in \mathcal{C}_G.$$

Stochastic-Feasible Set: $\mathcal{Z}_\pi := \{z \in \mathbb{R}^T : \mathbb{P}(\mathbf{L} - \mathbf{P}_r \leq z) \geq \pi\}$.

Difficulty: $\mathcal{Z}_\pi$ – handle unknown distribution, and numerical tractability.

CERTS
Consortium for Electric Reliability Technology Solutions

Cornell University Founded A.D. 1865

Chance-constrained Model

Introduction

The general form of the optimization model considered The general form of the optimization model considered:

- C_G total cost of thermal-generation
- P_G^t total thermal power generation at time t
- $\mathbf{L}^t$ total load
- $\mathbf{P}_r^t$ total renewable generation at time t
- T scheduling horizon

$$\min C_G(P_G),$$
$$\text{s.t:} P_G \geq z$$
$$z \in \mathcal{Z}_\pi$$
$$P_G \in \mathcal{C}_G.$$

Stochastic-Feasible Set: $\mathcal{Z}_\pi := \{z \in \mathbb{R}^T : \mathbb{P}(\mathbf{L} - \mathbf{P}_r \leq z) \geq \pi\}$.
Difficulty: $\mathcal{Z}_\pi$ – handle unknown distribution, and numerical tractability. Approximate $\mathcal{Z}_\pi$ based on theoretical results about its structure

Geometry

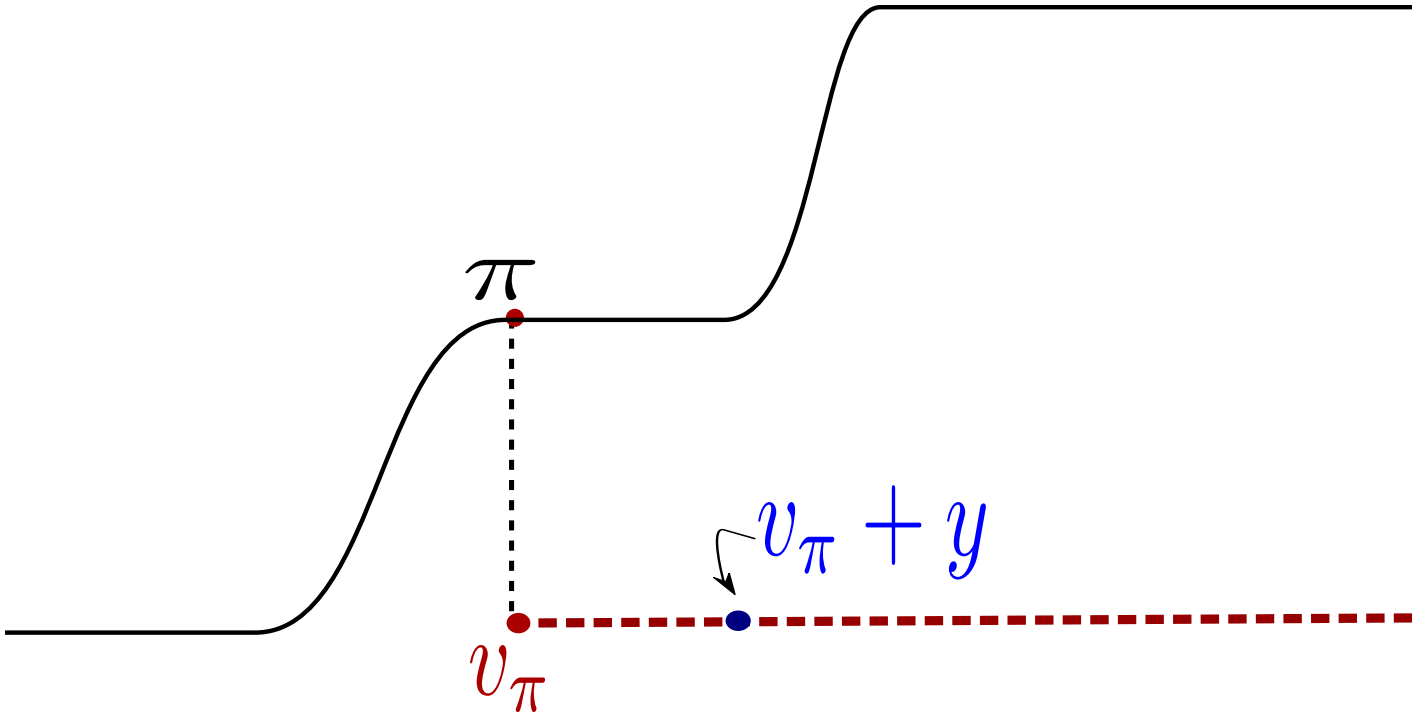
- Structure of $\mathcal{Z}_\pi$ depends on the distribution of $\mathbf{L} - \mathbf{P}_r$
- In general, it is a is non-convex set, $\mathcal{Z}_\pi = \cup\{v_j + \mathbb{R}_+^T\}$

Geometry

- Structure of $\mathcal{Z}_\pi$ depends on the distribution of $\mathbf{L} - \mathbf{P}_r$
- In general, it is a is non-convex set, $\mathcal{Z}_\pi = \cup\{v_j + \mathbb{R}_+^T\}$
  - If $T = 1$,
    $\mathcal{Z}_\pi = \{z \in \mathbb{R} : \mathbb{P}(\mathbf{L} - \mathbf{P}_r \leq z) \geq \pi\} = \{z = v_\pi + \mathbb{R}_+\}$

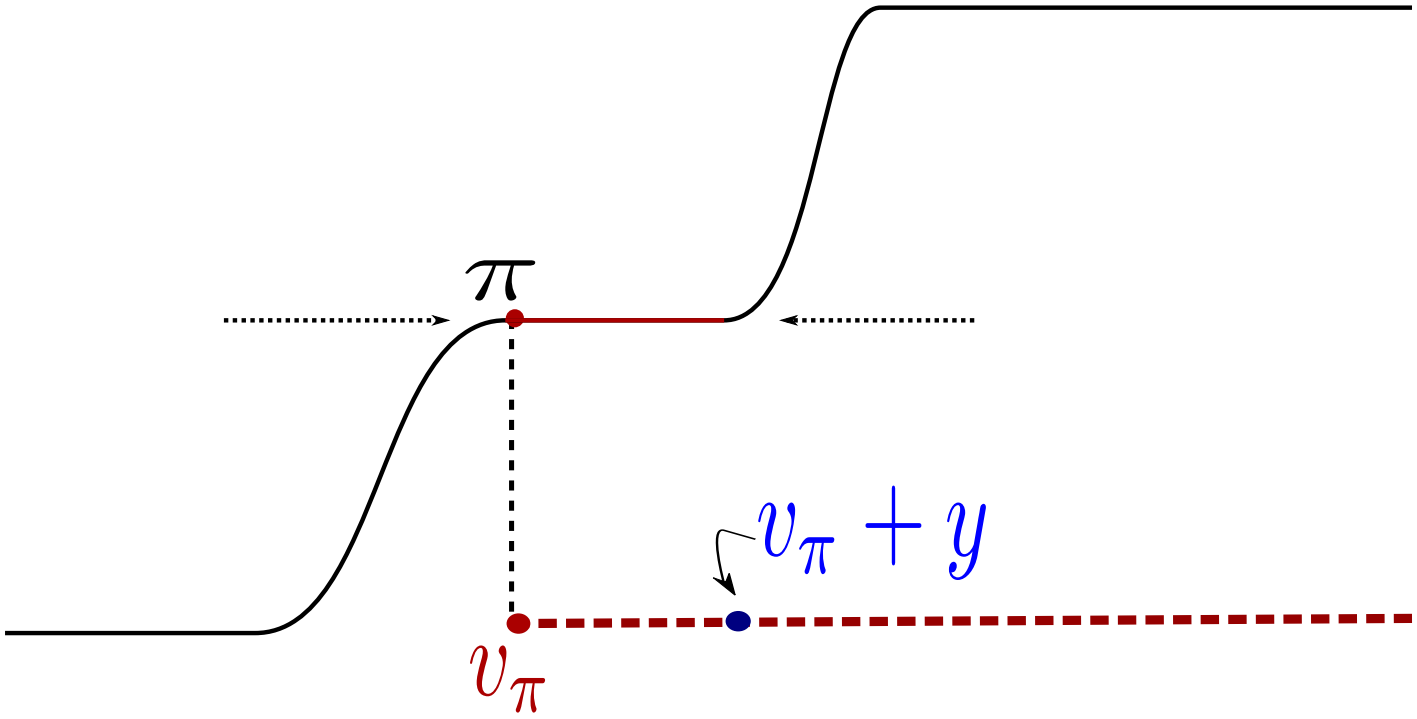
Geometry

- Structure of $\mathcal{Z}_\pi$ depends on the distribution of $\mathbf{L} - \mathbf{P}_r$
- In general, it is a is non-convex set, $\mathcal{Z}_\pi = \cup\{v_j + \mathbb{R}_+^T\}$
  - If $T = 1$,
    $\mathcal{Z}_\pi = \{z \in \mathbb{R} : \mathbb{P}(\mathbf{L} - \mathbf{P}_r \leq z) \geq \pi\} = \{z = v_\pi + \mathbb{R}_+\}$



Geometry

- Structure of $\mathcal{Z}_\pi$ depends on the distribution of $\mathbf{L} - \mathbf{P}_r$
- In general, it is a is non-convex set, $\mathcal{Z}_\pi = \cup\{v_j + \mathbb{R}_+^T\}$
  - If $T = 1$,
    $\mathcal{Z}_\pi = \{z \in \mathbb{R} : \mathbb{P}(\mathbf{L} - \mathbf{P}_r \leq z) \geq \pi\} = \{z = v_\pi + \mathbb{R}_+\}$

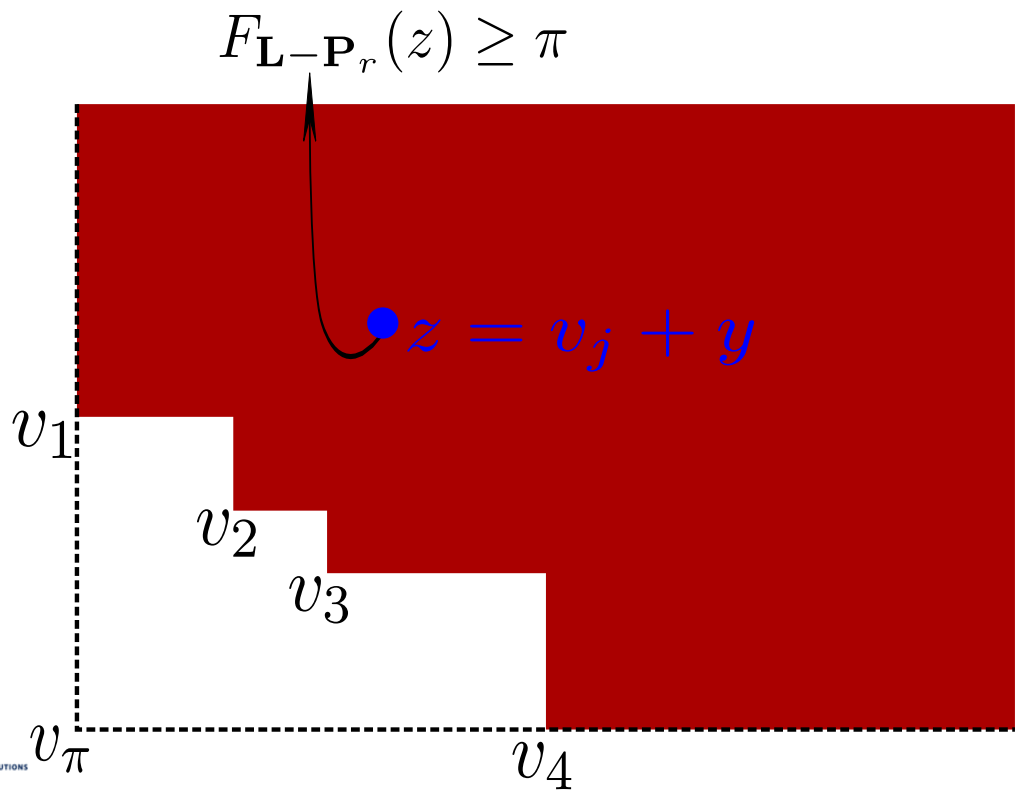


Geometry

- Structure of $\mathcal{Z}_\pi$ depends on the distribution of $\mathbf{L} - \mathbf{P}_r$
- In general, it is a is non-convex set, $\mathcal{Z}_\pi = \cup\{v_j + \mathbb{R}_+^T\}$
  - If $T > 1$, more combinations to consider to find the quantile-vectors v
    Example, $T = 2$: $\mathbf{L} - \mathbf{P}_r = (\mathbf{L}^1 - \mathbf{P}_r^1, \mathbf{L}^2 - \mathbf{P}_r^2)$

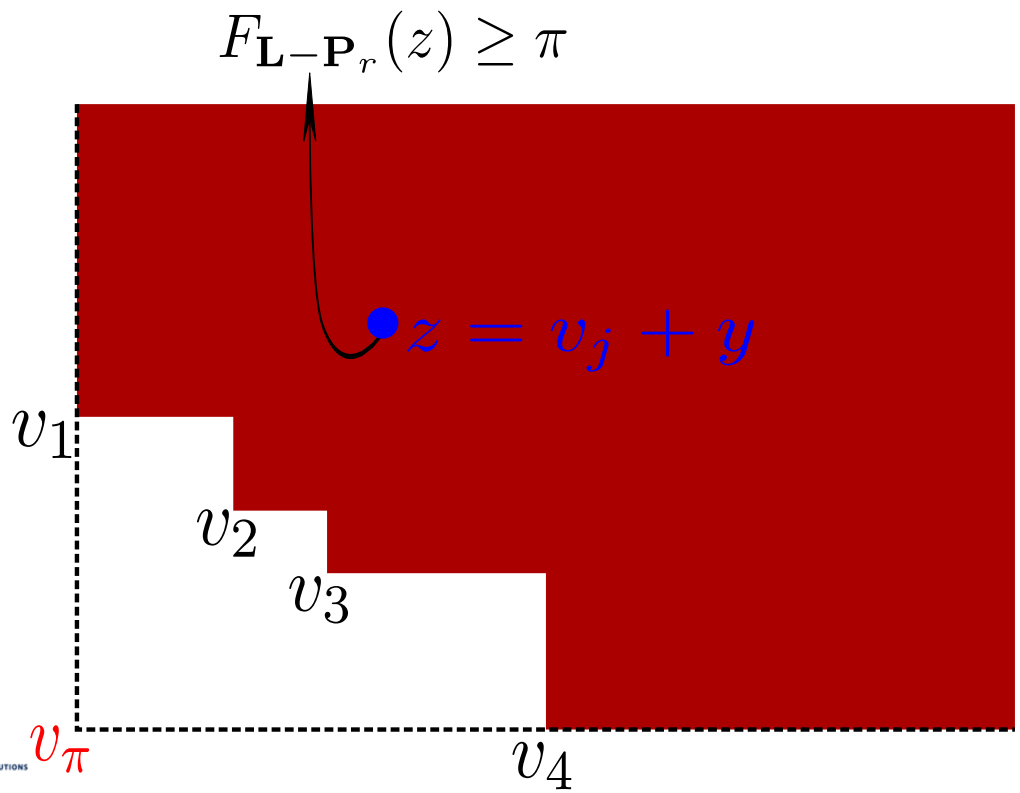
Geometry

- Structure of $\mathcal{Z}_\pi$ depends on the distribution of $\mathbf{L} - \mathbf{P}_r$
- In general, it is a is non-convex set, $\mathcal{Z}_\pi = \cup\{v_j + \mathbb{R}_+^T\}$
  - If $T > 1$, more combinations to consider to find the quantile-vectors v
    Example, $T = 2$: $\mathbf{L} - \mathbf{P}_r = (\mathbf{L}^1 - \mathbf{P}_r^1, \mathbf{L}^2 - \mathbf{P}_r^2)$



Geometry

- Structure of $\mathcal{Z}_\pi$ depends on the distribution of $\mathbf{L} - \mathbf{P}_r$
- In general, it is a is non-convex set, $\mathcal{Z}_\pi = \cup\{v_j + \mathbb{R}_+^T\}$
  - If $T > 1$, more combinations to consider to find the quantile-vectors v
    Example, $T = 2$: $\mathbf{L} - \mathbf{P}_r = (\mathbf{L}^1 - \mathbf{P}_r^1, \mathbf{L}^2 - \mathbf{P}_r^2)$

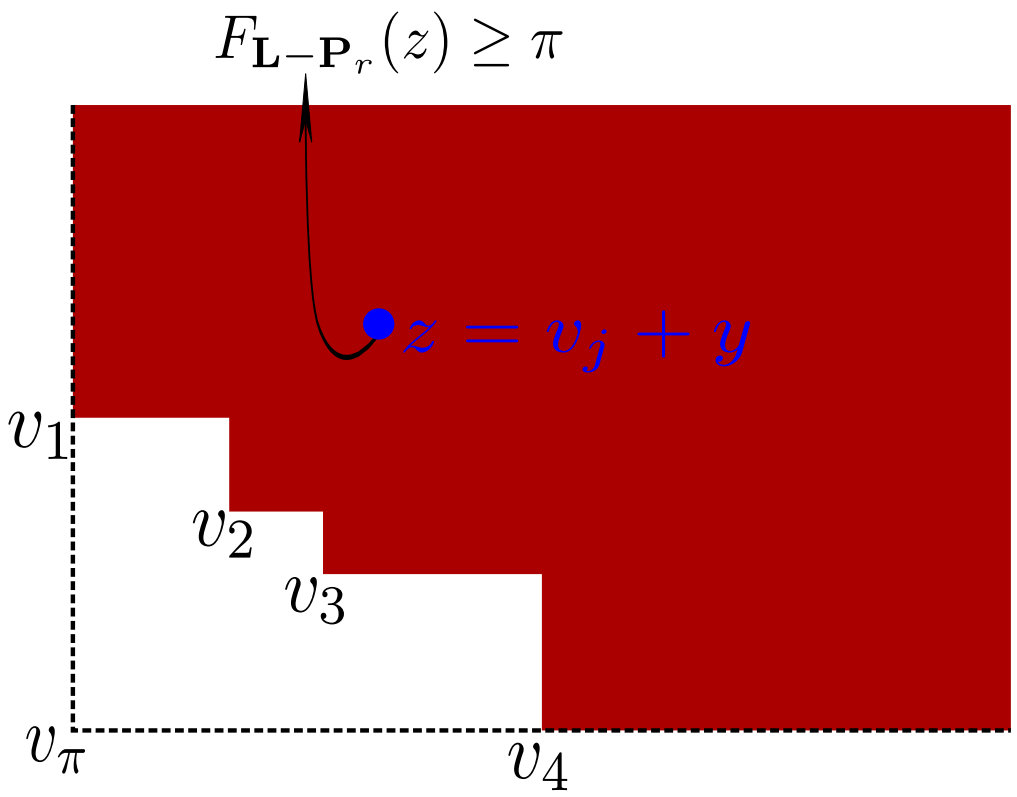


$v_\pi = (v_\pi^1, v_\pi^2)$

$F_{\mathbf{L}^1-\mathbf{P}_r^1}(v_\pi^1) \geq \pi$

$F_{\mathbf{L}^2-\mathbf{P}_r^2}(v_\pi^2) \geq \pi$

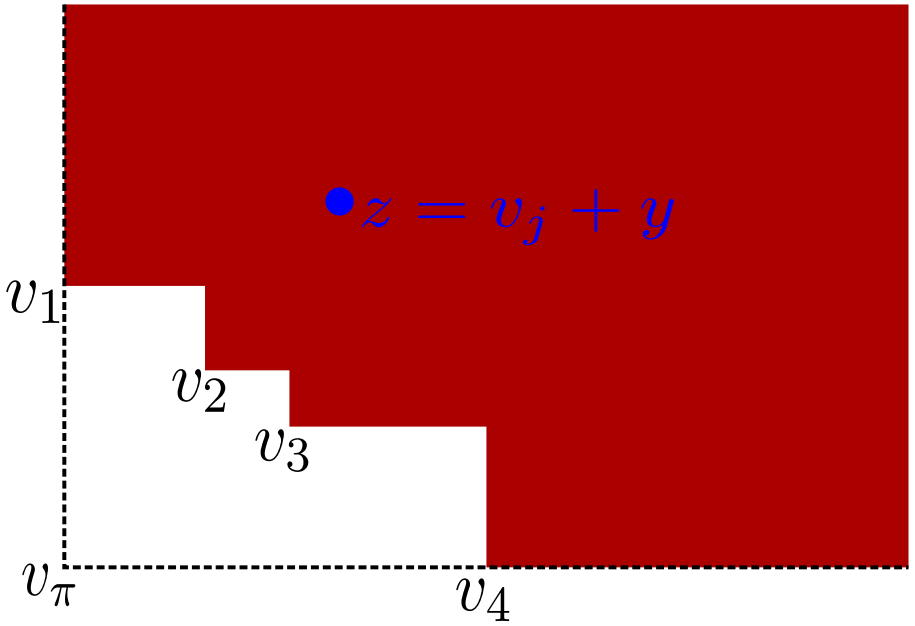
Geometry
General properties



- $v_\pi = (v_\pi^1, v_\pi^2)$
  $F_{\mathbf{L}^1-\mathbf{P}_r^1}(v_\pi^1) \geq \pi$
  $F_{\mathbf{L}^2-\mathbf{P}_r^2}(v_\pi^2) \geq \pi$
- Bounded
  $v_\pi \leq z, v_\pi \leq v_j$
- Extreme Points
  $F_{\mathbf{L}-\mathbf{P}_r}(v_j) \geq \pi$
  $v_1 = (v_1^1, v_\pi^2)$
  $v_4 = (v_\pi^1, v_4^2)$

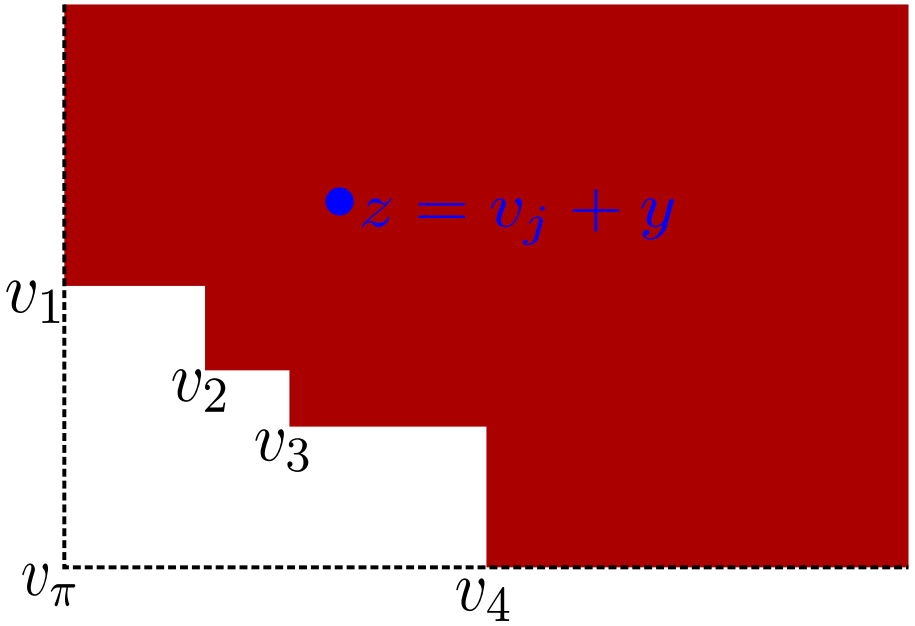
* finite realizations

Optimization problem



$$\min C_G(P_G),$$
$$\text{s.t:} P_G \geq z$$
$$z \in \mathcal{Z}_\pi$$
$$P_G \in \mathcal{C}_G.$$

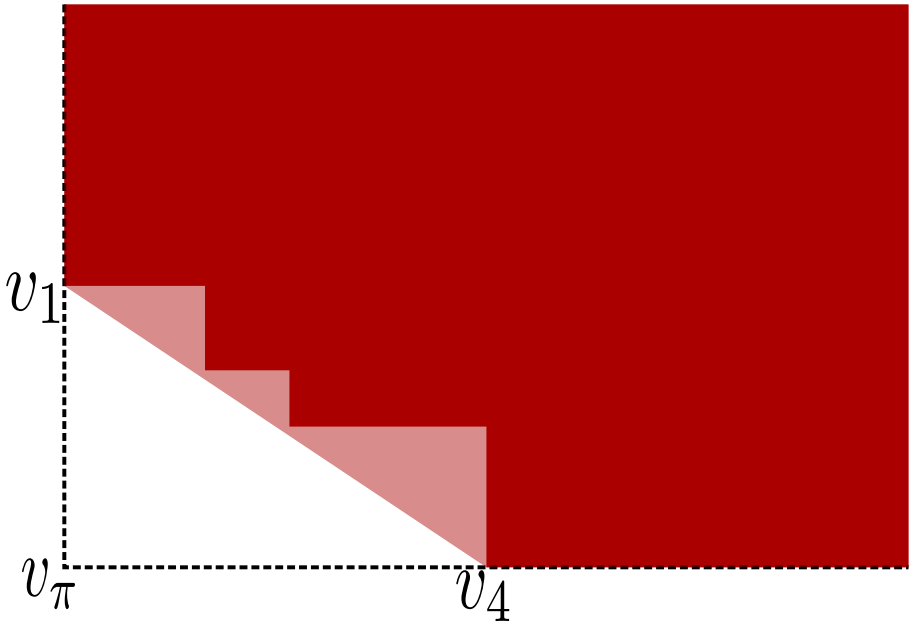
Optimization problem



$$\min C_G(P_G),$$
$$\text{s.t:} P_G \geq z$$
$$z \in \mathcal{Z}_\pi$$
$$P_G \in \mathcal{C}_G.$$

Optimal solution: $z = v_k$.
Convexification $z = \sum_{k=1}^{K} \lambda_k v_k$.

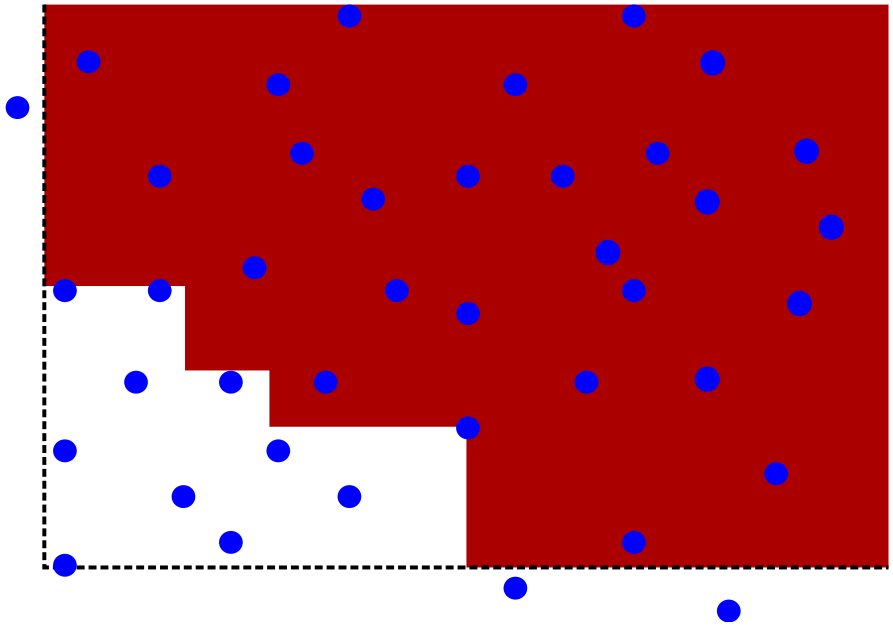
Optimization Problem



$$\min C_G(P_G),$$
$$\text{s.t:} P_G \geq z$$
$$z \in \mathcal{Z}_\pi$$
$$P_G \in \mathcal{C}_G.$$

Optimal solution: $z = v_k$.
Convexification $z = \sum_{k=1}^{K} \lambda_k v_k$.
Observation: v_j unknown

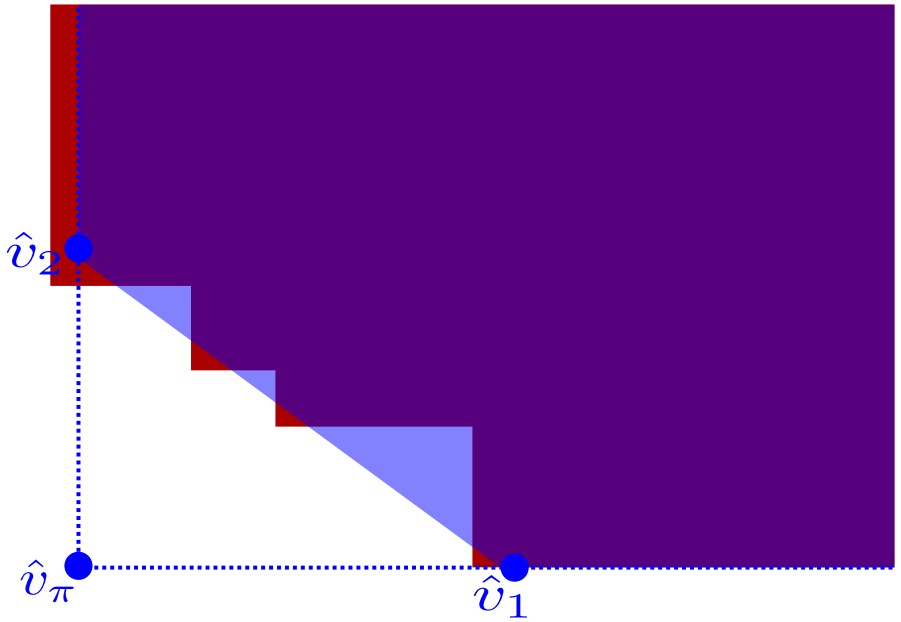
Optimization Problem



$$\min C_G(P_G),$$
$$\text{s.t:} P_G \geq z$$
$$z \in \mathcal{Z}_\pi$$
$$P_G \in \mathcal{C}_G.$$

Optimal solution: $z = v_k$.
Convexification $z = \sum_{k=1}^{K} \lambda_k v_k$.
Observation: v_j unknown
Approximate red region with a sample and estimate v_π, v_i

Optimization Problem
Approximation

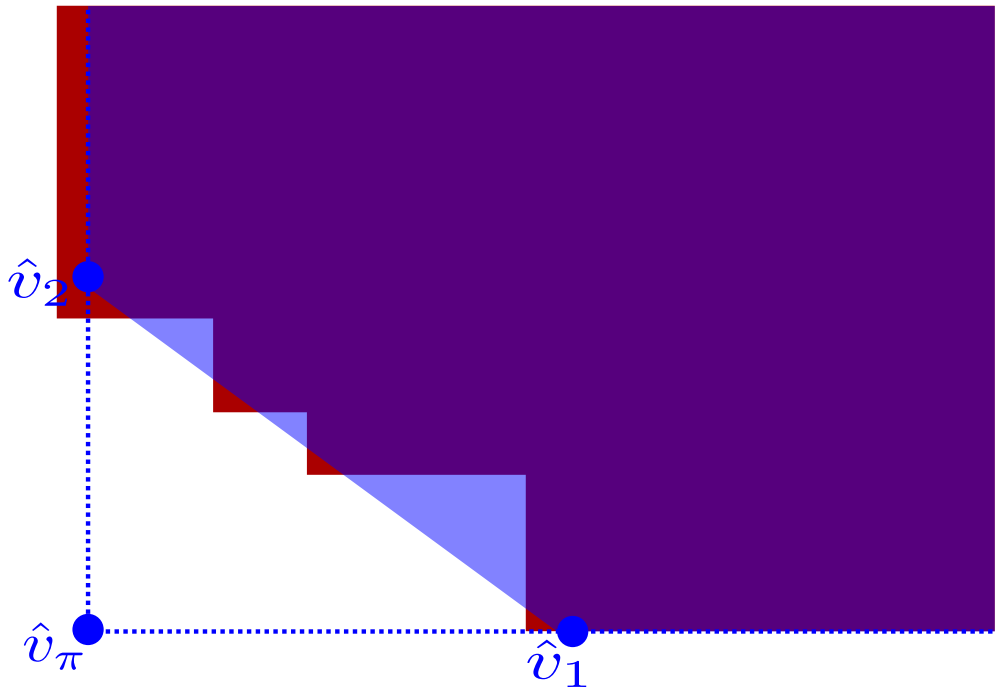


$$\begin{aligned} &\min C_G(P_G), \\ &\text{s.t:} P_G \geq z \\ &\quad z \in \mathcal{H}_\pi \\ &\quad P_G \in \mathcal{C}_G. \end{aligned}$$

$\mathcal{H}_\pi$ data-driven approximation of $\mathcal{Z}_\pi$.

Quantile-based set $\mathcal{H}_\pi$

For $t = 1, \ldots, T$, let $\mathbf{V}^t = \mathbf{L}^t - \mathbf{P}_r^t$. Let $V_1, \ldots, V_S$ be a sample of $\mathbf{V}^t$.



- Ordered sample

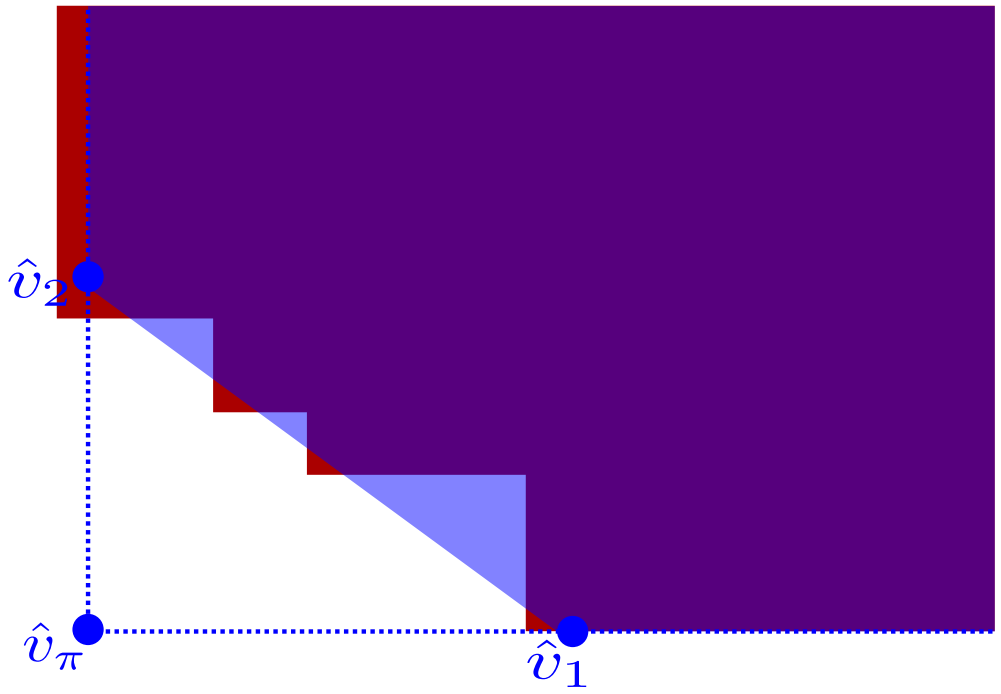
$$V_{[1]} \leq V_{[2]} \leq \cdots \leq V_{[S]}$$

- Estimate bound: $\hat{v}_\pi^t$
- Estimate coordinate: $\hat{v}_\tau^t$

CERTS
Consortium for Electric Reliability Technology Solutions

Quantile-based set $\mathcal{H}_\pi$

For $t = 1, \ldots, T$, let $\mathbf{V}^t = \mathbf{L}^t - \mathbf{P}_r^t$. Let $V_1, \ldots, V_S$ be a sample of $\mathbf{V}^t$.



- Ordered sample

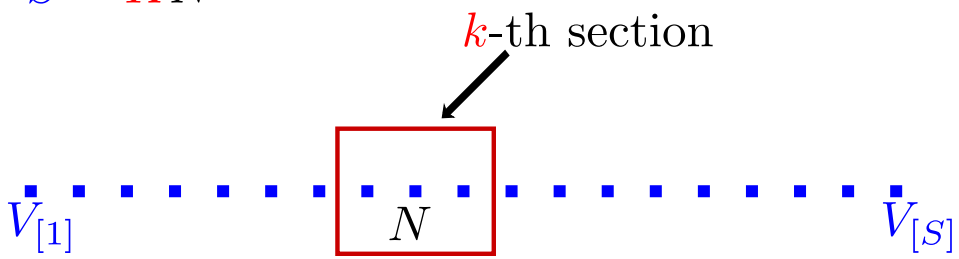
$$V_{[1]} \leq V_{[2]} \leq \cdots \leq V_{[S]}$$

- Estimate bound: $\hat{v}_\pi^t$
- Estimate coordinate: $\hat{v}_\tau^t$
- Simple estimator: $k = \lceil \pi S \rceil$, $V_{[k]}$ (problem: biased)

Quantile-based set $\mathcal{H}_\pi$

For $t = 1, \ldots, T$, let $\mathbf{V}^t = \mathbf{L}^t - \mathbf{P}_r^t$. Let $V_1, \ldots, V_S$ be a sample of $\mathbf{V}^t$.

$S = KN$



Sample: $V_{[\lceil \pi S \rceil]}$

Sample\section: $V^k_{[\lceil \pi (S-N) \rceil]}$

Jackknife: $V_{[\lceil \pi S \rceil]}$ - $\frac{K-1}{K} \sum_{k=1}^{K} V^k_{[\lceil \pi (S-N) \rceil]}$

- Ordered sample

$$V_{[1]} \leq V_{[2]} \leq \cdots \leq V_{[S]}$$

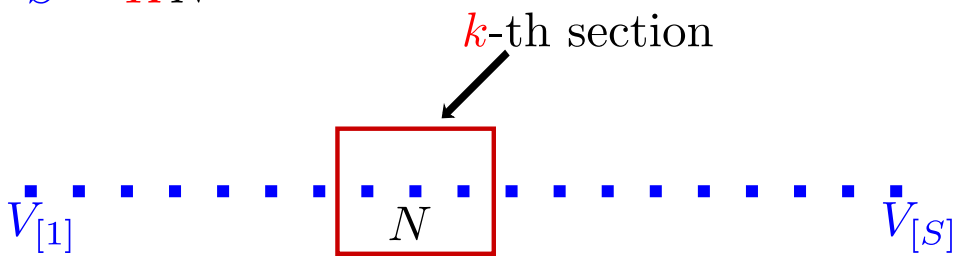
- Estimate bound: $\hat{v}_\pi^t$
- Estimate coordinate: $\hat{v}_\tau^t$
- Jackknife estimator

CERTS
Consortium for Electric Reliability Technology Solutions

Quantile-based set $\mathcal{H}_\pi$

For $t = 1, \ldots, T$, let $\mathbf{V}^t = \mathbf{L}^t - \mathbf{P}_r^t$. Let $V_1, \ldots, V_S$ be a sample of $\mathbf{V}^t$.

$S = KN$



Sample: $V_{[\lceil \pi S \rceil]}$

Sample\section: $V^k_{[\lceil \pi (S-N) \rceil]}$

Jackknife: $V_{[\lceil \pi S \rceil]}$ - $\frac{K-1}{K} \sum_{k=1}^{K} V^k_{[\lceil \pi (S-N) \rceil]}$

- Spacing $i = 1, \ldots, S-1$

$$sV_i = V_{[i+1]} - V_{[i]},$$

$k = \text{argmax}_{i \geq \lceil \pi S \rceil + 1} sV_i.$

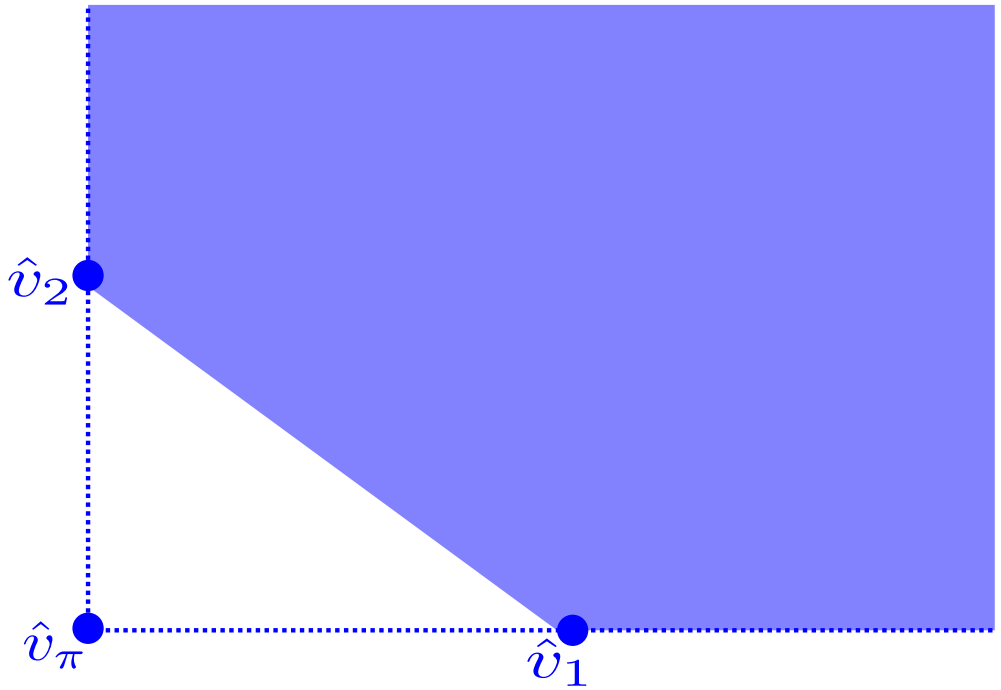
$\rho = \frac{\lceil \pi S \rceil + k + 1}{S}.$

- Estimate bound: $\hat{v}_\pi^t$
- Estimate coordinate: $\hat{v}_\rho^t$

CERTS
Consortium for Electric Reliability Technology Solutions

Quantile-based set $\mathcal{H}_\pi$

For $t = 1, \ldots, T$, let $\mathbf{V}^t = \mathbf{L}^t - \mathbf{P}_r^t$. Let $V_1, \ldots, V_S$ be a sample of $\mathbf{V}^t$.



- Define the vectors, $\tau = 1, \ldots, T$

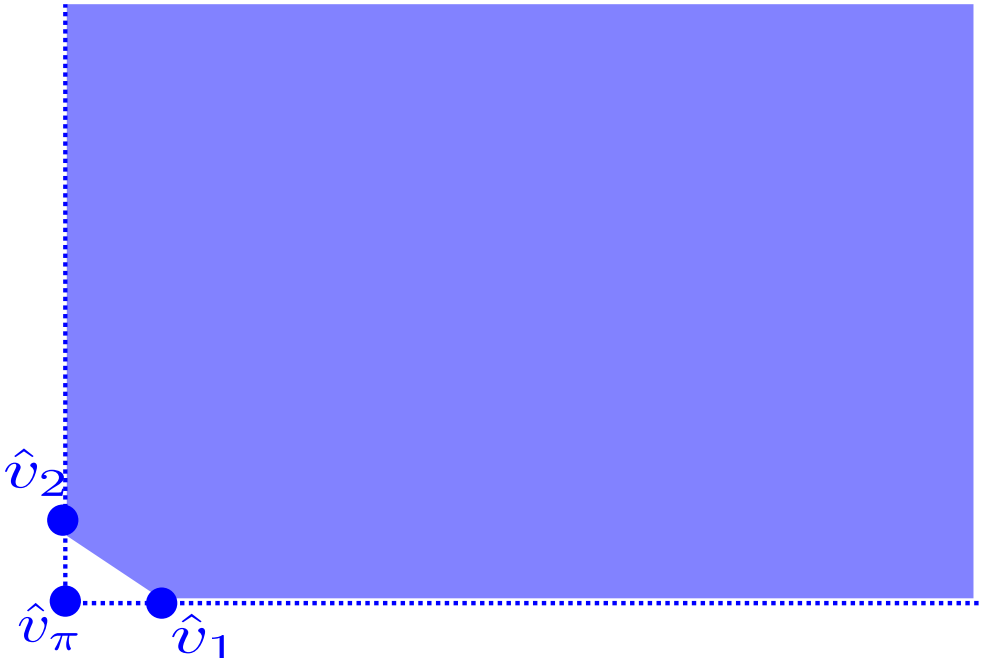
$$\hat{v}_\tau = \begin{cases} \hat{v}_\rho^t & t \neq \tau, \\ \hat{v}_\pi^t & t = \tau. \end{cases}$$

$$\mathcal{H}_\pi = \{v = \sum_{\tau=1}^{T} \beta_\tau \hat{v}_\tau, \beta_\tau \geq 0, \sum_{\tau=1}^{T} \beta_\tau = 1\} + \mathbb{R}_+^T$$

CERTS
Consortium for Electric Reliability Technology Solutions

Quantile-based set $\mathcal{H}_\pi$

For $t = 1, \ldots, T$, let $\mathbf{V}^t = \mathbf{L}^t - \mathbf{P}_r^t$. Let $V_1, \ldots, V_S$ be a sample of $\mathbf{V}^t$.



- If $\pi \approx 1$ then

$$\lceil \pi S \rceil \approx S,$$

$$\rho \approx S$$

- $\hat{v}_\rho^t \approx \hat{v}_\pi^t$

$$\mathcal{H}_\pi = \hat{v}_\pi + \mathbb{R}_+^T$$

CERTS
Consortium for Electric Reliability Technology Solutions

Numerical Method

- UC convex-relaxation model

$$\min C_G(P_G),$$
$$\text{s.t:} P_G \geq v$$
$$v \in \mathcal{H}_\pi$$
$$P_G \in \mathcal{C}_G.$$

$$\lambda \in \mathbb{R}_+^T \xrightarrow[\text{multiplier}]{\text{Lagrange}} P_G \geq v$$

- Dual decomposition

$$\mathcal{D}(\lambda) = \mathcal{D}_G(\lambda) + \mathcal{D}_v(\lambda)$$
$$\mathcal{D}_G(\lambda) = \min_{P_G \in \mathcal{C}_G} C_G - \langle \lambda, P_G \rangle$$
$$\mathcal{D}_v(\lambda) = \min_{v \in \mathcal{H}_\pi} \langle \lambda, v \rangle$$

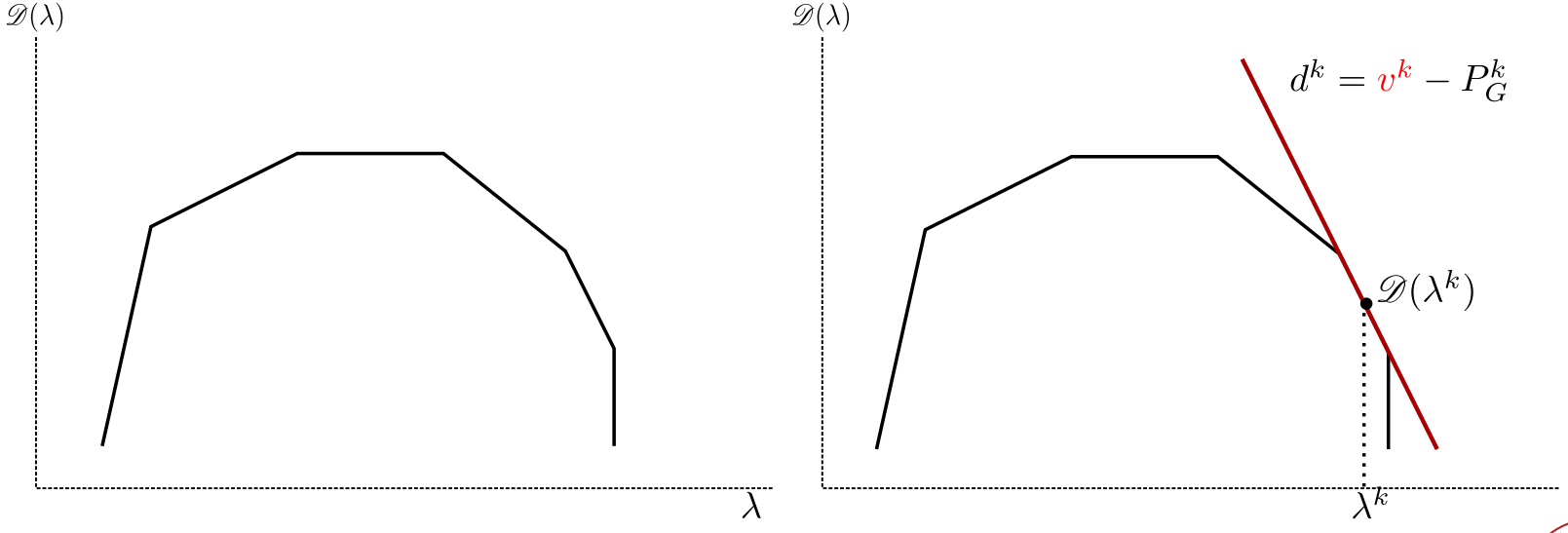
- Subgradient (inexact)

$$d = v - P_G$$

CERTS
Consortium for Electric Reliability Technology Solutions

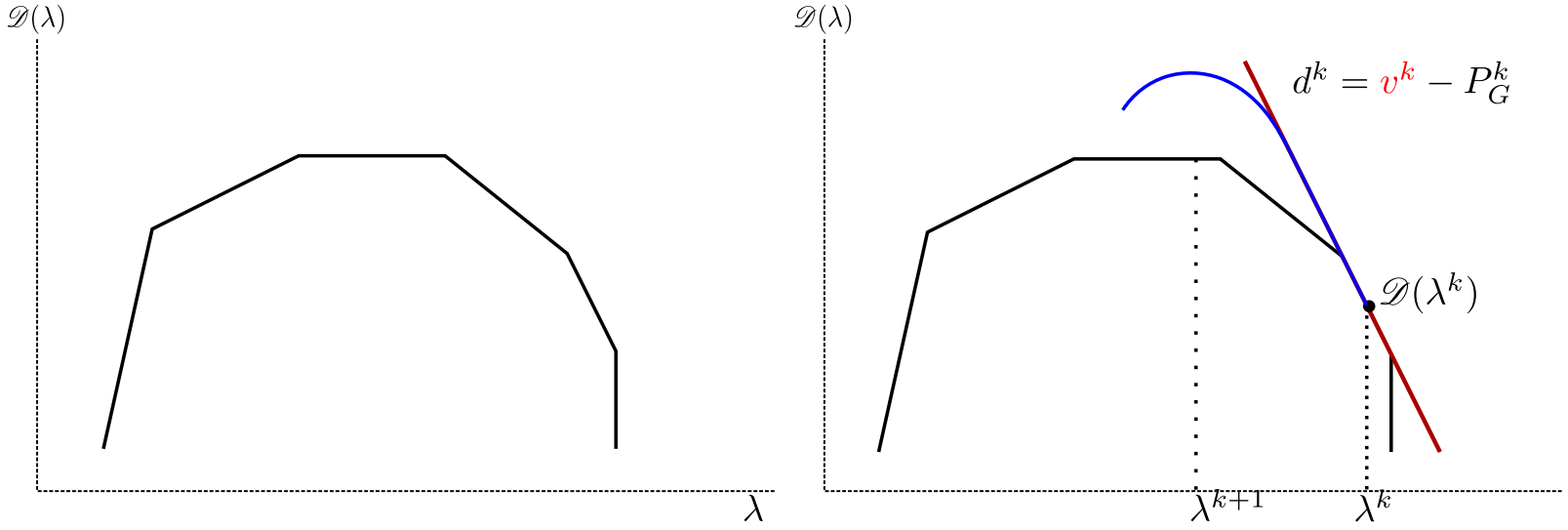
Numerical Method

- UC convex-relaxation model
- The approximation $\mathcal{H}_\pi$ introduces inexact objective values and subgradients
- Numerical method: proximate bundle method



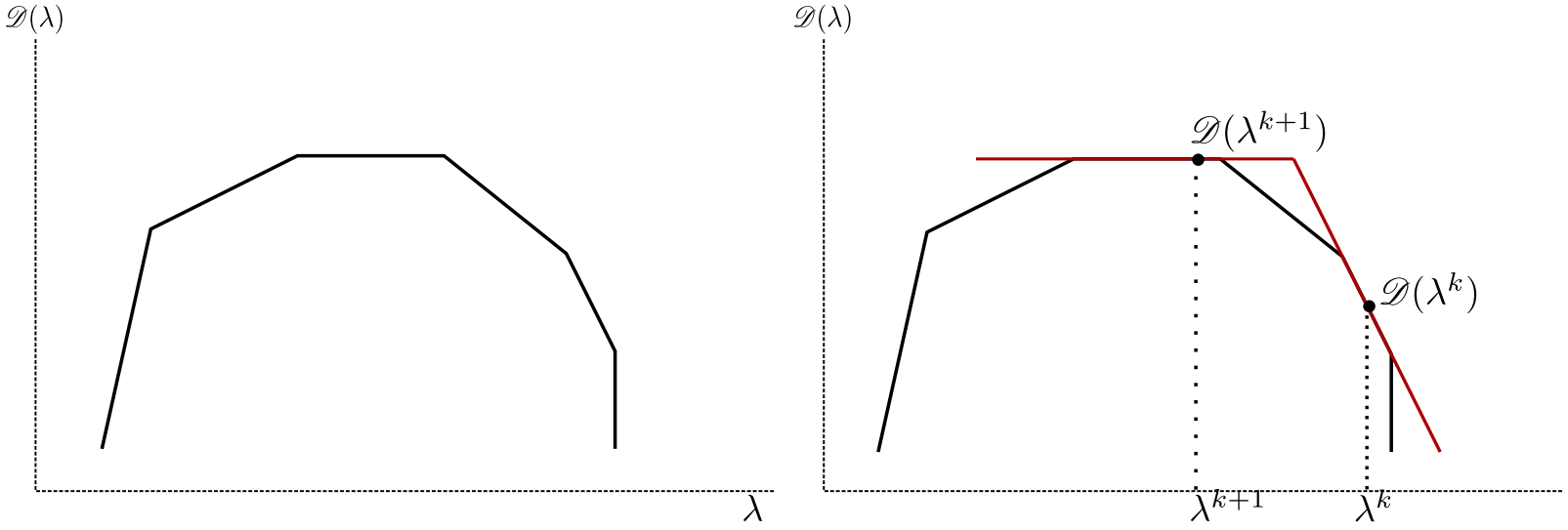
Numerical Method

- UC convex-relaxation model
- The approximation $\mathcal{H}_\pi$ introduces inexact objective values and subgradients
- Numerical method: proximate bundle method



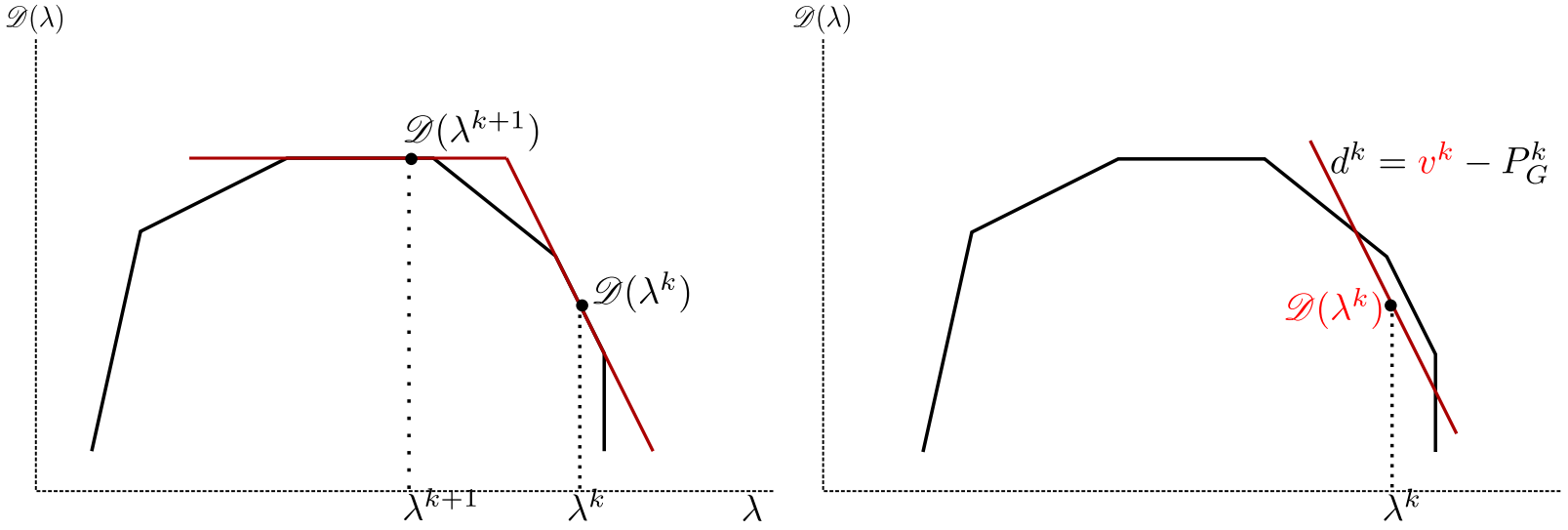
Numerical Method

- UC convex-relaxation model
- The approximation $\mathcal{H}_\pi$ introduces inexact objective values and subgradients
- Numerical method: proximate bundle method



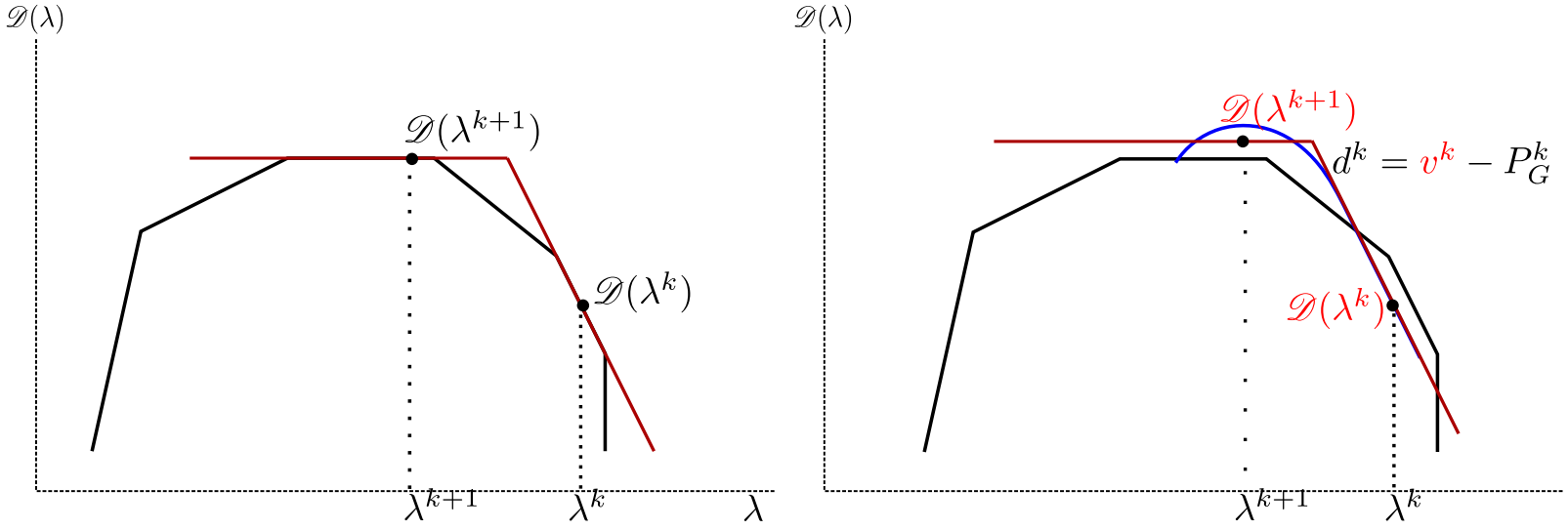
Numerical Method

- UC convex-relaxation model
- The approximation $\mathcal{H}_\pi$ introduces inexact objective values and subgradients
- Numerical method: proximate bundle method



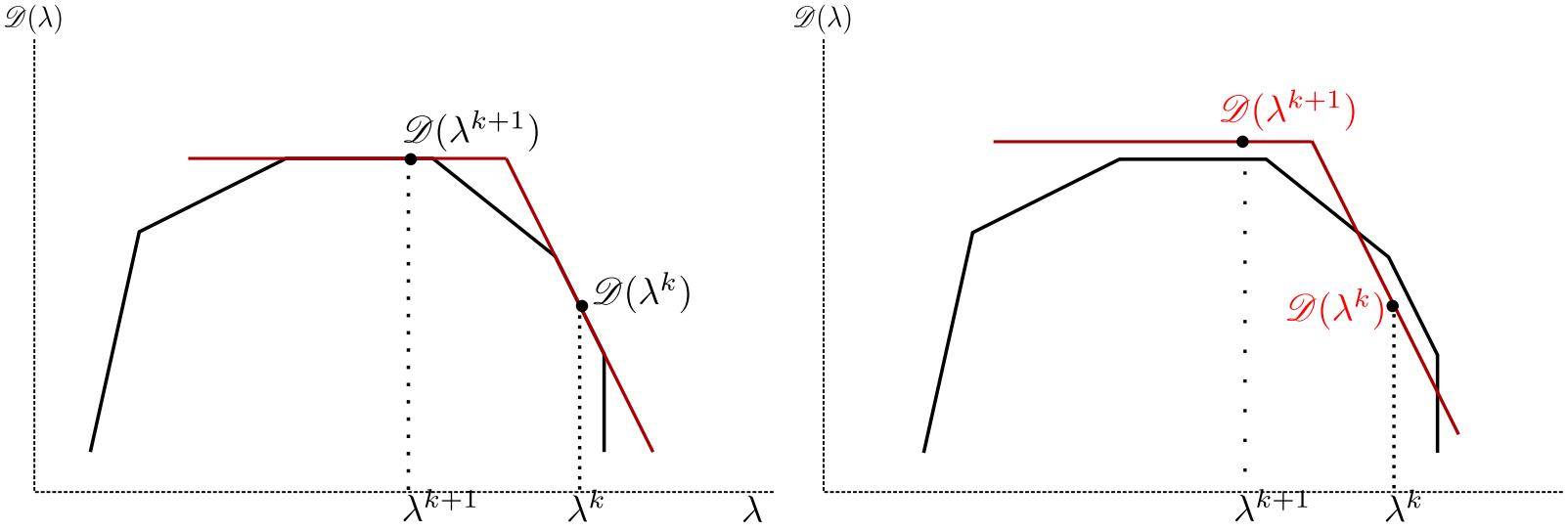
Numerical Method

- UC convex-relaxation model
- The approximation $\mathcal{H}_\pi$ introduces inexact objective values and subgradients
- Numerical method: proximate bundle method



Numerical Method

- UC convex-relaxation model
- The approximation $\mathcal{H}_\pi$ introduces inexact objective values and subgradients
- Numerical method: proximate bundle method



CERTS
Consortium for Electric Reliability Technology Solutions

Algorithm

Stochastic input: scenarios of $\mathbf{L} - \mathbf{P}_r = (\mathbf{L}^t - \mathbf{P}_r^t, t = 1, \ldots, T)$

- Construct $\mathcal{H}_\pi$ from scenarios:

$$\begin{array}{ll} \text{Bound:} & \hat{v}_\pi^t, t = 1, \ldots, T, \\ \text{Coordinates:} & \hat{v}_\rho^t, t = 1, \ldots, T. \end{array}$$

  $\hat{v}_\tau$, $\tau = 1, \ldots, T$, represent risk-level π

- Solve the dual of the UC convex-relaxation model with *Proximal Bundle*.

$$(\bar{\lambda}, \bar{P}_G, \bar{v}) \in \mathbb{R}_+^T \times \mathcal{C}_G \times \mathcal{H}_\pi,$$
$$(\bar{\lambda}, \bar{P}_G, \bar{v}) \text{ is } \epsilon\text{-optimal solution}$$

- Primal-feasible recovery: Inexact Augmented Lagrangian heuristic as in
  Belloni, A., Lima, A. D. S., Maceira, M. P., and Sagastizábal, C. A. (2003). Bundle relaxation and primal recovery in unit commitment problems. The Brazilian case. Annals of Operations Research, 120(1-4), 21-44.

Proximal Bundle Algorithm

- *Stopping criterion*: if $A_k < \texttt{tol}$ stop; otherwise continue.
- *Iteration k*: $\mathcal{J}_k \subset \{1, \ldots, k\}$, $(\mathcal{A}(\lambda^j), d^j), j \in \mathcal{J}_k$
- *Dual model*: $\phi^k(\lambda) = \min_{j \in \mathcal{J}_k} \mathcal{A}(\lambda^j) + \sum_{t \in \mathcal{T}} d^{tj}(\lambda^t - \lambda^{tj})$
- *Master problem*: $\max_{\lambda \geq 0} \phi^k(\lambda) - \frac{1}{2\mathfrak{t}_k}\|\lambda - \gamma^k\|^2$
- *Inexact measure*:

$$\begin{aligned} \delta_k &= \varphi^k - \mathcal{A}(\gamma^k), \quad & \alpha_k &= \|\gamma^k - \lambda^{k+1}\|/\mathfrak{t}_k, \\ \beta_k &= \mathfrak{t}_k \alpha_k^2 + \delta_k, & A_k &= \max\{\alpha_k, \beta_k\}. \end{aligned}$$

  If $\delta_k + \beta_k > 0$ then increase $\mathfrak{t}_k$ go to "Master"; otherwise continue.
- *Compute*: $\mathcal{D}_G(\lambda^{k+1})$, $\mathcal{D}_v(\lambda^{k+1})$, and d^{k+1}
- *Proximal point*: If $\mathcal{A}(\lambda^{k+1}) \geq \mathcal{A}(\nu^k) - \mathfrak{a}(\mathcal{A}(\gamma^k) - \phi^k(\lambda^{k+1}))$ then $\gamma^{k+1} = \lambda^{k+1}$; otherwise $\gamma^{k+1} = \nu^k$
- Set $\mathcal{J}_{k+1} = \{k+1\} \cup \{j \in \mathcal{J}_k : \mu_j^k \neq 0\}$, $k := k+1$ go to "Stopping criterion".

Results

Test-system of 100 thermal units. Generated sample for wind-generation, NREL pv-generation sample, NYC load sample. Linux machine `Intel®Core i5`, `4 GB RAM`. Python, PYOMO, CPLEX.

- $\mathcal{H}_\pi$ generation time

S \ π	0.8	0.9	0.95	0.99	0.999
$1e+3$	0.16644	0.16284	0.16283	0.16380	0.16281
$1e+4$	0.38818	0.38786	0.38473	0.38599	0.40159
$1e+5$	3.06511	3.06647	3.05644	3.03444	3.03745
$1e+6$	36.3308	36.0005	35.4912	35.5562	34.9820

- Proximal bundle solution time
- Primal-feasible schedule time

Results

Test-system of 100 thermal units. Generated sample for wind-generation, NREL pv-generation sample, NYC load sample. Linux machine `Intel®Core i5`, `4 GB RAM`. Python, PYOMO, CPLEX.

- $\mathcal{H}_\pi$ generation time
- Proximal bundle solution time

S \ π	0.8	0.9	0.95	0.99	0.999
$1e+3$	288.407	277.155	248.1538	228.030	213.338
$1e+4$	294.526	272.319	235.641	229.197	211.717
$1e+5$	290.500	262.473	231.087	222.972	210.498
$1e+6$	289.534	253.946	239.6616	218.215	207.761

- Primal-feasible schedule time

Results

Test-system of 100 thermal units. Generated sample for wind-generation, NREL pv-generation sample, NYC load sample. Linux machine `Intel®Core i5`, `4 GB RAM`. Python, PYOMO, CPLEX.

- $\mathcal{H}_\pi$ generation time
- Proximal bundle solution time
- Primal-feasible schedule time

S \ π	0.8	0.9	0.95	0.99	0.999
$1e+3$	349.615	305.905	260.871	294.778	223.165
$1e+4$	335.263	320.730	347.149	310.208	231.084
$1e+5$	303.610	344.920	318.462	264.668	227.088
$1e+6$	333.292	323.921	301.033	336.169	234.761

CERTS
Consortium for Electric Reliability Technology Solutions

UC-DC Flow

$\mathcal{Z}_\pi := \{z = (z^1, \ldots, z^T), z^t \in \mathbb{R}^K : \mathbb{P}(\mathbf{P}_D^t - \mathbf{P}_r^t \leq z) \geq \pi\}$
$\mathcal{Z}_\pi \subset \mathbb{R}^{TK}$.
$\mathcal{K} = \{1, \ldots, K\}$, $\mathcal{T} = \{1, \ldots, T\}$.

$$
\begin{array}{ll}
\min & C(P_G) \\
\text{s.t.} & \sum_{i \in I_k}^{I} P_{G_i}^t - \sum_{km} B_{km}(\delta_k - \delta_m) \geq z_k^t, k \in \mathcal{K}, t \in \mathcal{T} \\
& z \in \mathcal{Z}_\pi \\
& |B_{km}(\delta_k - \delta_m)| \leq F_{km}, km \in \mathcal{B}r \\
& P_G \in \mathcal{C}_G
\end{array}
$$

CERTS
Consortium for Electric Reliability Technology Solutions

Approximation UC-DC Flow

$\mathcal{H}_\pi \subset \mathbb{R}^{TK}$.
Estimators from data: per node, per time. Since KT is large, we have to use $\rho = 1$ for some coordinates.

$$
\begin{array}{ll}
\min & C(P_G) \\
\text{s.t.} & \sum_{i \in I_k}^{I} P_{G_i}^t - \sum_{km} B_{km}(\delta_k - \delta_m) \geq v_k^t, k \in \mathcal{K}, t \in \mathcal{T} \\
& v \in \mathcal{H}_\pi \\
& |B_{km}(\delta_k - \delta_m)| \leq F_{km}, km \in \mathcal{B}r \\
& P_G \in \mathcal{C}_G
\end{array}
$$

Dual decomposition: thermal-subproblem, network-subproblem, stochastic-subproblem.

Future work

- Sensitivity analysis
  - other percentile estimators - stability analysis
  - apply contamination techniques - sensibility perturbation distribution function.
    Dupačová, J. (2006). Stress testing via contamination. In Coping with uncertainty (pp. 29-46). Springer Berlin Heidelberg.
- Value of Information
  - probability weight selection for clusters: $\|\hat{v}_\tau - \hat{v}^c\|$
  - measure performance of representative of clusters
- Tighter MIP formulations for primal-feasible recovery

Stochastic Two-stage Models

- First-stage. Unit commitment decisions (binary variables)

$$\begin{array}{ll} \min & S_{\text{up}}(u_G) + S_{\text{dn}}(v_G) + \mathbb{E}(Q(u,\xi)) \\ \text{s.t:} & u_G \in \mathcal{S}_{\text{up}}, v_G \in \mathcal{S}_{\text{dn}}, \\ & \bar{u}_G \in \mathcal{S}_{\text{on}}, \bar{v}_G \in \mathcal{S}_{\text{off}}, \\ & u := (u_G, v_G, \bar{u}_G, \bar{v}_G) \in \mathcal{S}_{\text{lk}}. \end{array}$$

u_G start-up, v_G shutdown, $\bar{u}_G$ on, $\bar{v}_G$ off.

- Second-stage. Power dispatch. $Q(u, \xi^s)$ optimal value of

$$\begin{array}{ll} \min & C(P_G) + V(y) \\ \text{s.t.} & \sum_{i \in I_k}^{I} P_{G_i}^t - \sum_{km} B_{km}(\delta_k - \delta_m) + y_k^t = \mathbf{L}(s)_k^t - \mathbf{P}_r(s)_k^t, \\ & k \in \mathcal{K}, t \in \mathcal{T} \\ & |B_{km}(\delta_k - \delta_m)| \leq F_{km}, km \in \mathcal{B}r \\ & P_G \in \mathcal{C}_G(u) \end{array}$$

CERTS
Consortium for Electric Reliability Technology Solutions

Stochastic Two-stage Models

- First-stage. Unit commitment decisions (binary variables)

$$\begin{array}{ll} \min & S_{\text{up}}(u_G) + S_{\text{dn}}(v_G) + \mathbb{E}(Q(u,\xi)) \\ \text{s.t:} & u_G \in \mathcal{S}_{\text{up}}, v_G \in \mathcal{S}_{\text{dn}}, \\ & \bar{u}_G \in \mathcal{S}_{\text{on}}, \bar{v}_G \in \mathcal{S}_{\text{off}}, \\ & u := (u_G, v_G, \bar{u}_G, \bar{v}_G) \in \mathcal{S}_{\text{lk}}. \end{array}$$

u_G start-up, v_G shutdown, $\bar{u}_G$ on, $\bar{v}_G$ off.

- Second-stage. Power dispatch. $Q(u, \xi^s)$ optimal value of

$$\begin{array}{ll} \min & C(P_G) + V(y) \\ \text{s.t.} & \sum_{i \in I_k}^{I} P_{G_i}^t - \sum_{km} B_{km}(\delta_k - \delta_m) + y_k^t = \mathbf{L}(s)_k^t - \mathbf{P}_r(s)_k^t, \\ & k \in \mathcal{K}, t \in \mathcal{T} \\ & |B_{km}(\delta_k - \delta_m)| \leq F_{km}, km \in \mathcal{B}r \\ & P_G \in \mathcal{C}_G(u) \end{array}$$

CC. P_G fixed v some risk-level π. TS. P_G per realization s, minimize expected cost.

Future work

- Scalable numerical approach. Stochastic dual dynamic programming (SDDP approximate dynamic programming technique)
  Shapiro, A. (2011). Analysis of stochastic dual dynamic programming method. European Journal of Operational Research, 209(1), 63-72.
- Value of information
  - Value of stochastic information (VSS) - performance of expected value problem
  - VSS for cluster - measure information loss
  - Use VSS to select number of clusters

Scenario Identification and Reduction

Laurie Tupper

CERTS
Consortium for Electric Reliability Technology Solutions

Motivation

The scale and dimensionality of power system models result in problems of computational tractability, so we seek methods to effectively represent uncertainty in low cardinality sets.
This is challenging due to

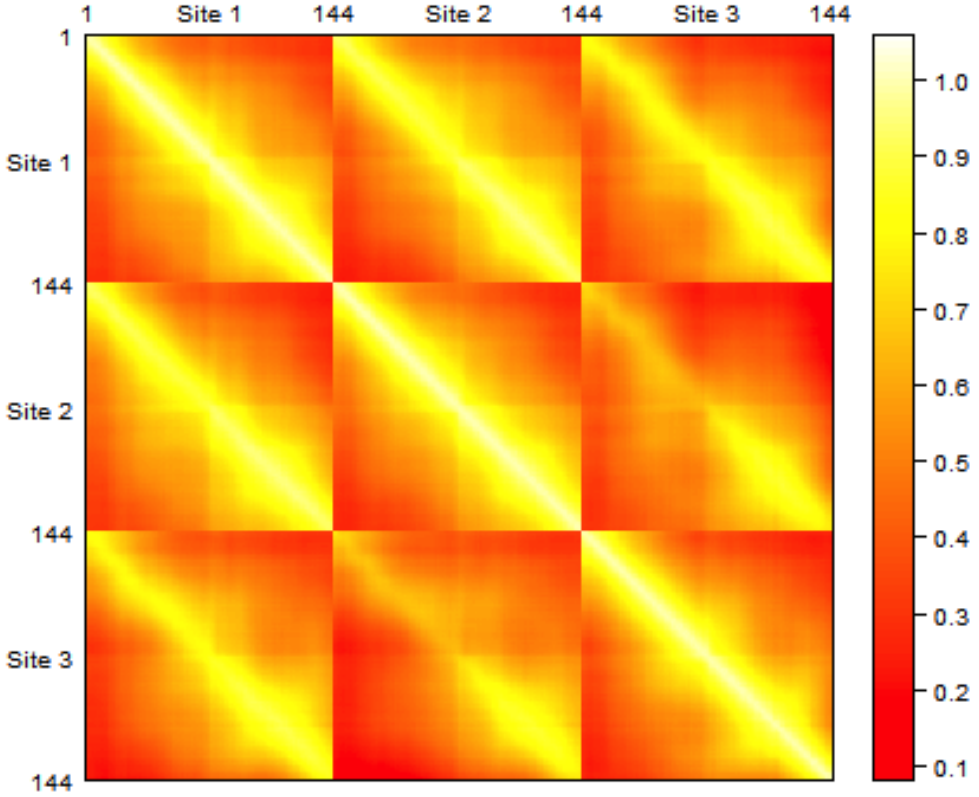
- highly non stationary time-series
- multiple locations, various levels of correlation between locations
- peaks and sudden drops are important
- no clear clusters appear in the data

Outline

- Wind data characteristics (a quick tour)
- k-means clustering
- Alternative distance metric
  - Band Depth
  - Jaccard distance
- k-medoids clustering
- Clustering results
- Conclusions and future directions

Inter-site correlation

Correlation (sometimes lagged) between sites changes over time

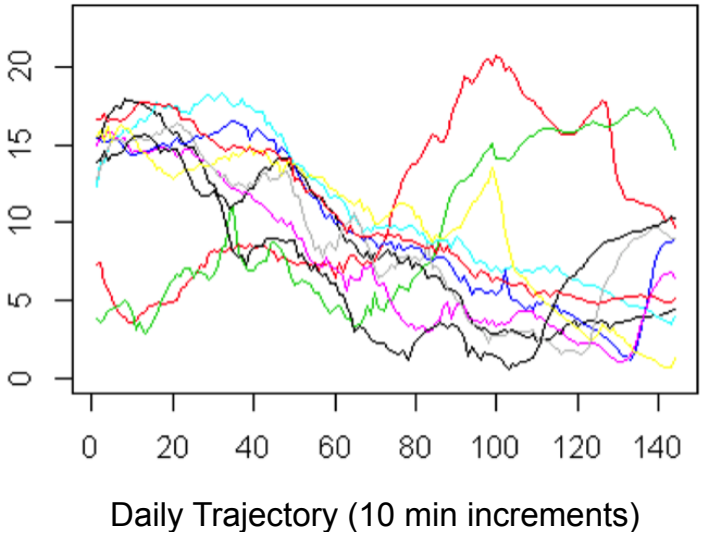


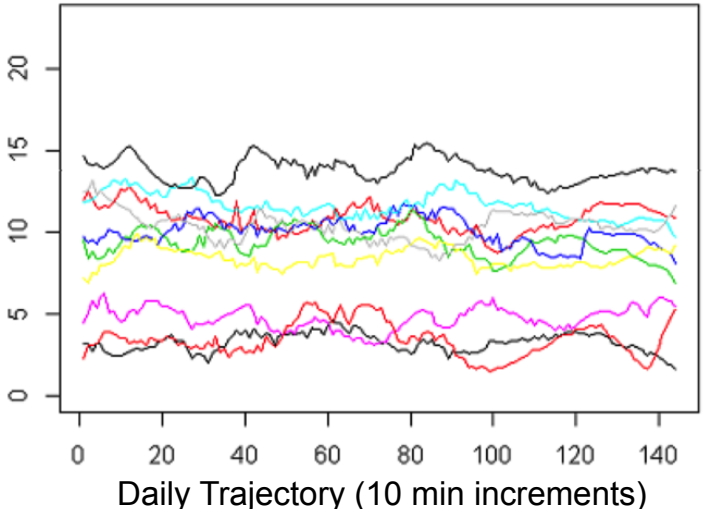
CERTS
Consortium for Electric Reliability Technology Solutions

Cornell University Founded A.D. 1865

Daily behavior

- No consistent daily shape in wind speed
- Days vary widely in mean, amount of variance, smoothness



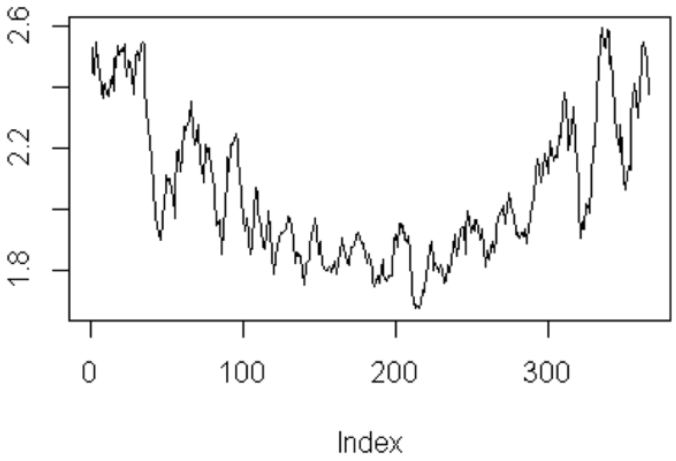


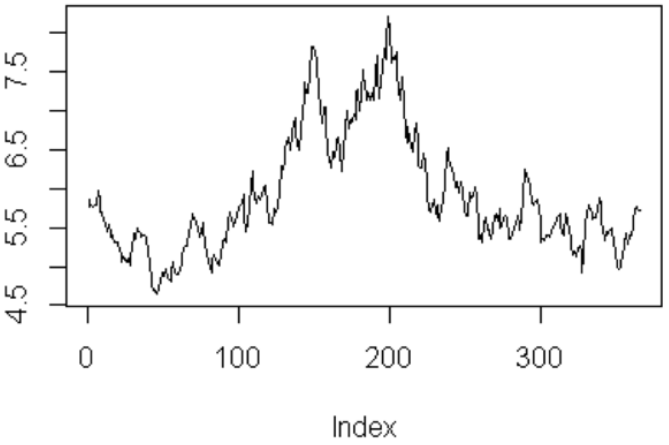
Wind speeds for high variance (left panel) and low variance (right panel) days for a single site, shows significantly different behaviors.

CERTS
Consortium for Electric Reliability Technology Solutions

Annual variation

- Daily behavior does, however, follow annual cycle
- For example:
  - Variance over course of day higher in winter (left panel)
  - More short-term erratic behavior in summer (right panel)

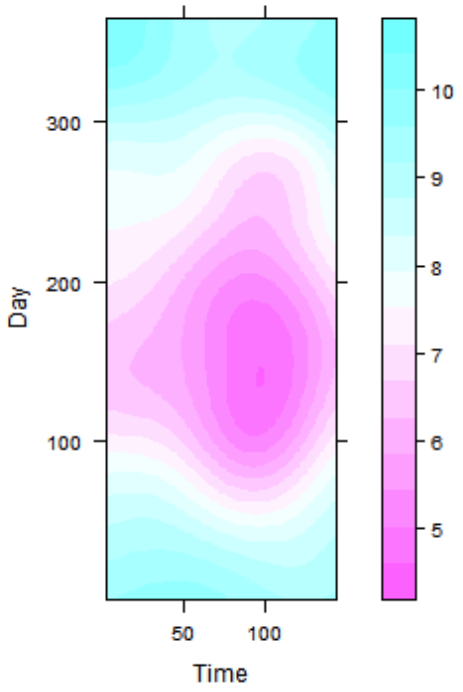




To obtain a typical daily wind speed curve, we fit a generalized additive model of the form

$$W = s(t, d)$$

where W represents wind speed, and $s(t, d)$ is some smooth function of time of day, t, and day of year, d.



CERTS
Consortium for Electric Reliability Technology Solutions

Cornell University Founded A.D. 1865

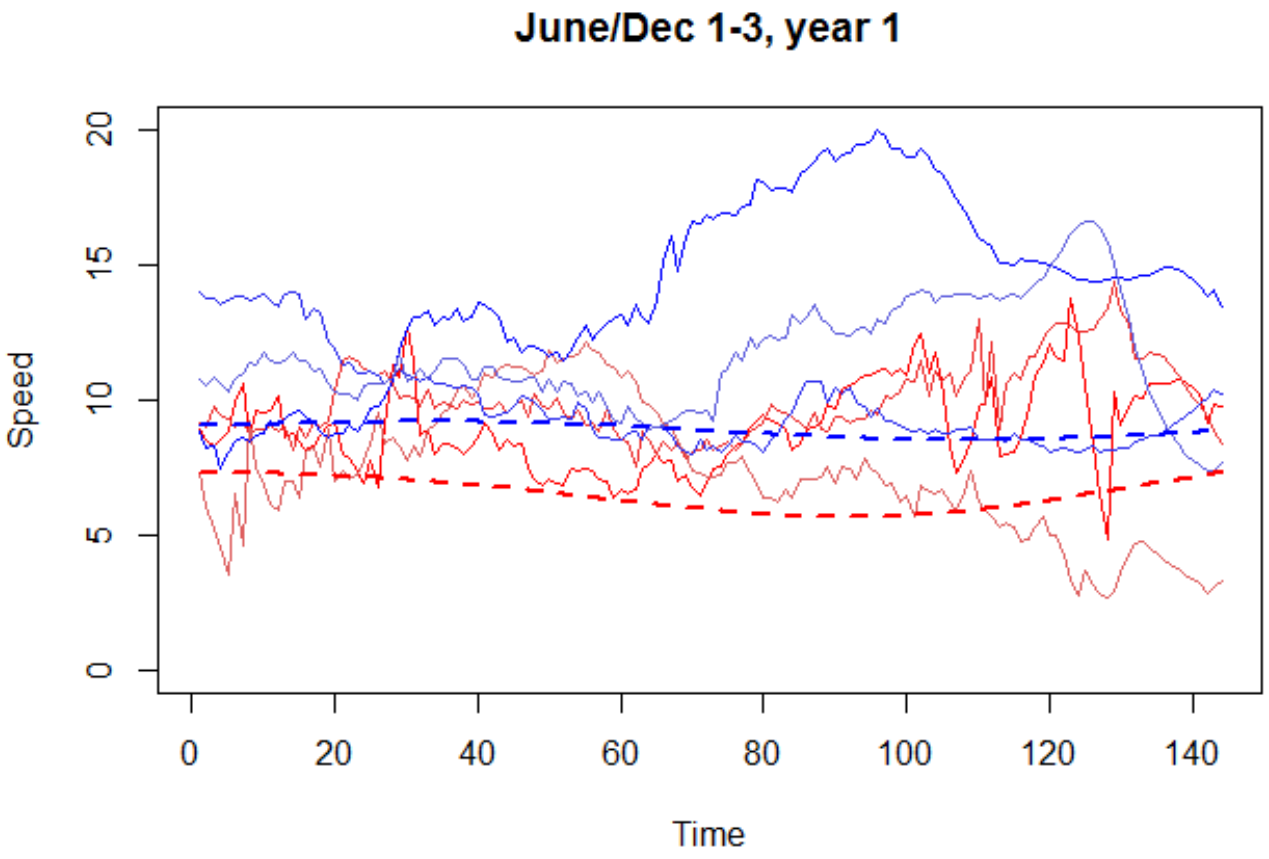


Figure : Observations from June (red) and December (blue), with the typical daily curves shown as dashed lines.

CERTS
Consortium for Electric Reliability Technology Solutions

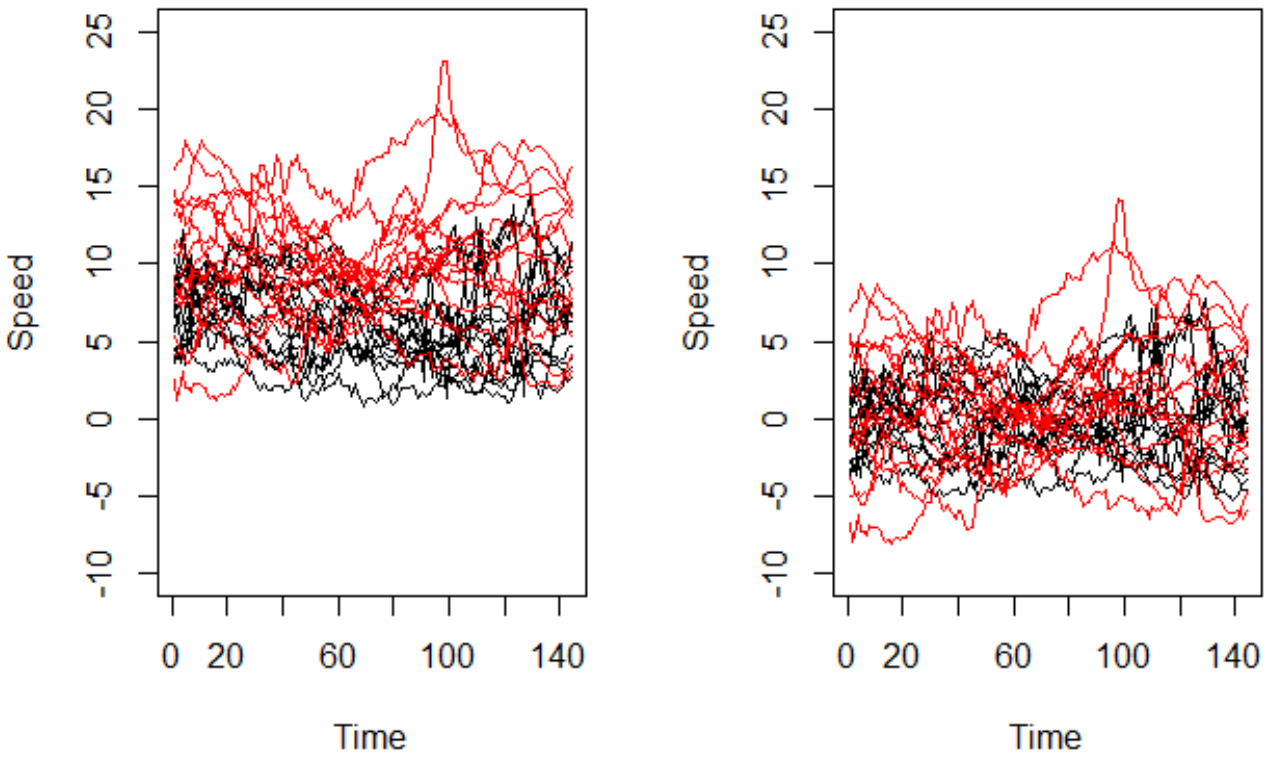
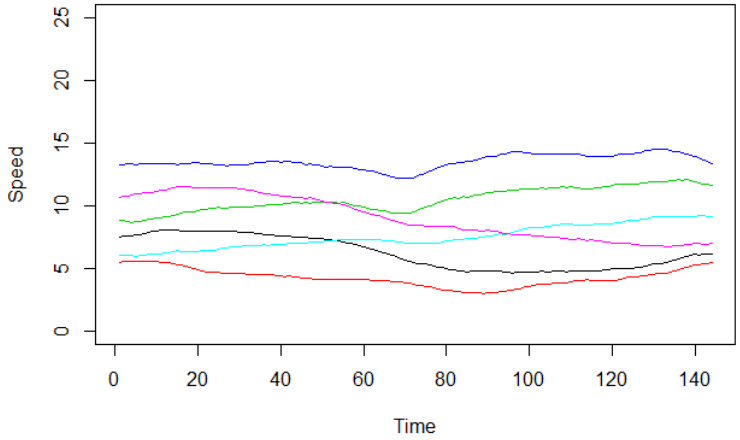


Figure : The first 15 days of June (black) and December (red) in year 1. Left, orignal observations; right, typical daily curves removed. The predominant effect is to shift the center of the observations.

K-means clustering

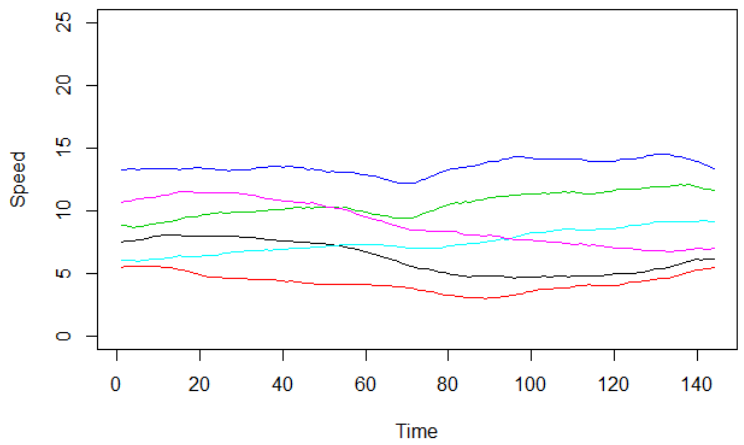
- Straightforward k-means clustering is not effective
- Mean speed of each day dominates cluster assignment
- Aggregating observations gives unrealistically smooth centers



CERTS
Consortium for Electric Reliability Technology Solutions

K-means clustering

- Straightforward k-means clustering is not effective
- Mean speed of each day dominates cluster assignment
- Aggregating observations gives unrealistically smooth centers



Unsupervised clustering is convenient for large data sets; alternatives are investigated that can provide more effective scenarios

Dissimilarity between observations

- recall that k-means clustering is based on the distance between observations in a set
- observations are assigned to sets to minimize total distance within the cluster
- distance is the euclidian distance at each point
- slight differences in mean behavior dominate major differences in shape

The band depth

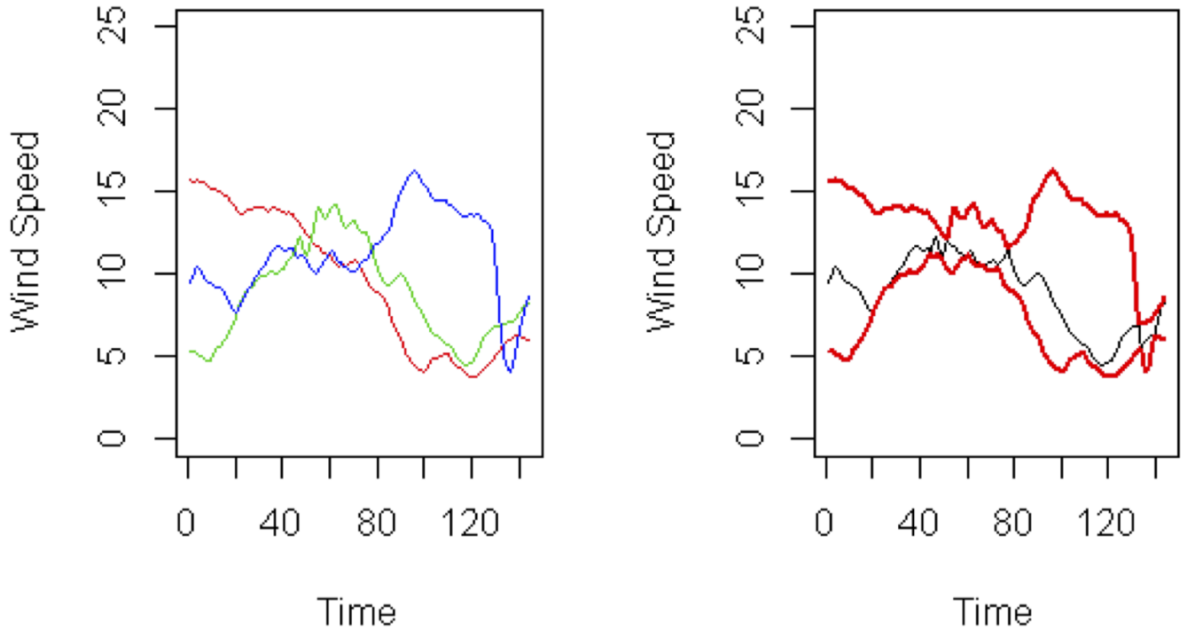
- Developed by Lopez-Pintado and Romo (2009) to judge typical or atypical shape
- Let x_i represent an observation, x_{it} represent the value at time point t, and X denote the entire set of observations
- A band b is defined by B observations $(x_1, \ldots, x_B)$ drawn from X; its upper limit is the maximum value of all the observations defining the band:

$$b_{ut} = \max_{i=1,\ldots,B} x_{it}$$

and lower limit is the minimum value:

$$b_{lt} = \min_{i=1,\ldots,B} x_{it}$$

Construction of a band: defining observations and resulting band limits



Original band depth is binary, modified band depth uses proportion of observation within band

CERTS
Consortium for Electric Reliability Technology Solutions

Band Depth to Distance

- clustering algorithms require a distance measure
- similarity scores of the band depth are not a distance metric
- Jaccard distance provides a conversion:
- Jaccard measure of distance between two sets A_1 and A_2 is

$$1 - \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|}$$

Distance for modified band depth

- Similar observations will fall into each band at similar times
- Each pair of observations x and y receives a similarity score for each band b:

$$J_{xy}^b = \frac{|T^b(x) \cap T^b(y)|}{|T^b(x) \cup T^b(y)|}$$

  where $T^b(x)$ is the set of times when x is in b
- This is converted to a distance score for each band:

$$L_{xy}^b = 1 - J_{xy}^b$$

- The overall distance is the average score over $b \in U_{xy}$, the set of all bands containing x or y at any time:

$$D_{xy} = \frac{1}{|U_{xy}|} \sum_{b \in U_{xy}} L_{xy}^b$$

K-medoids

- Can now use any clustering method that requires only a pairwise distance matrix
- One method is the k-medoids clustering algorithm:
  - Iterative but non-hierarchical
  - Similar to k-means, but cluster centers must always be observations from the dataset
  - Begin with random centers
  - Assign all observations to nearest center, then reset centers to most central observation in each cluster
  - Iterate until stable

Comparison of clustering results

- 90 observations, corresponding to the first 15 days of June and December in each year of the dataset
- Six clusters
- First, we apply the distance metric and clustering directly to the wind speed time series

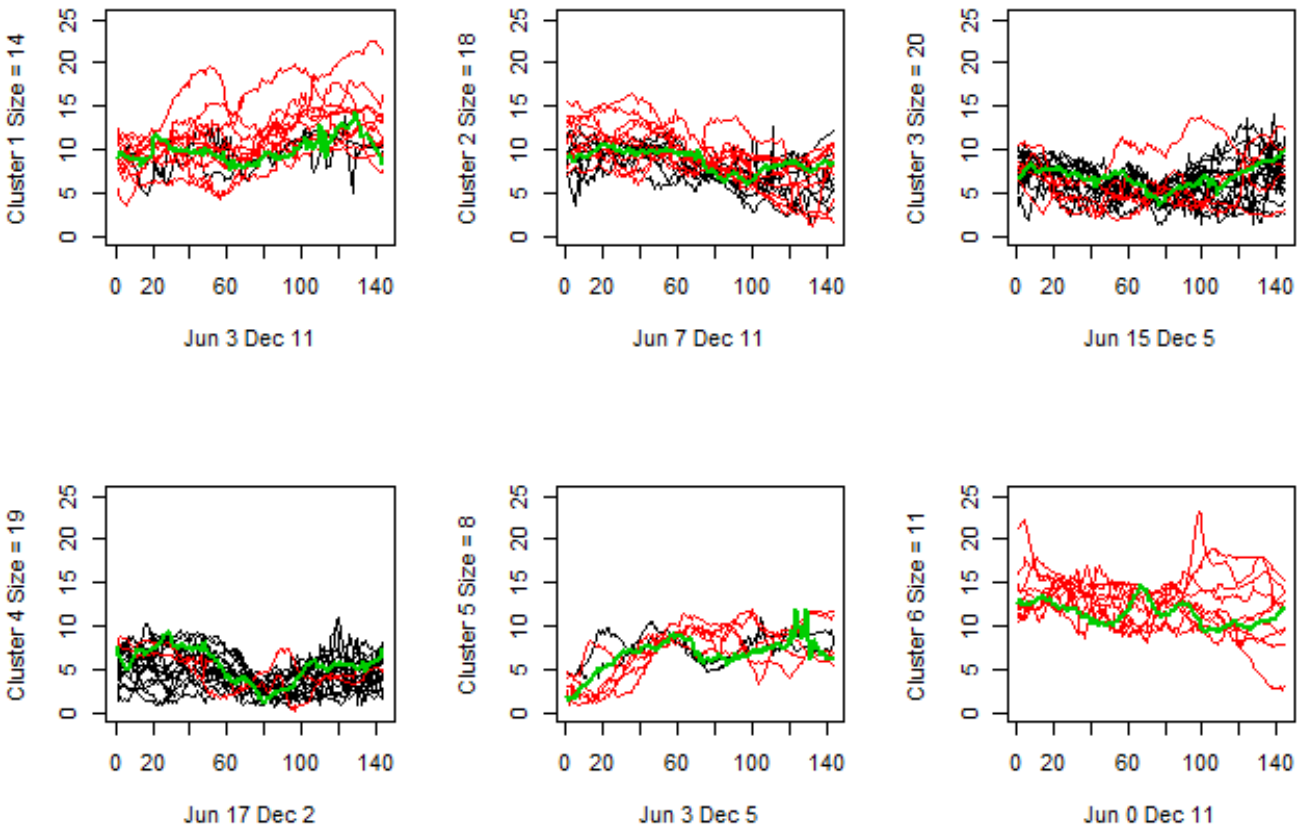


Figure : Clustering results for original wind speed data, with number of June (black) and December (red) observations in each cluster noted on axis. Cluster centers (medoids) are shown in green.

- Observations from the same month tend to be clustered together, reflecting the different average levels in winter and summer
- Removing the GAM-based daily shape from each day before clustering gives more balanced groups
- Adjusted Rand Index between the two classifications of the observations is only 0.126, indicating that classifications are dissimilar

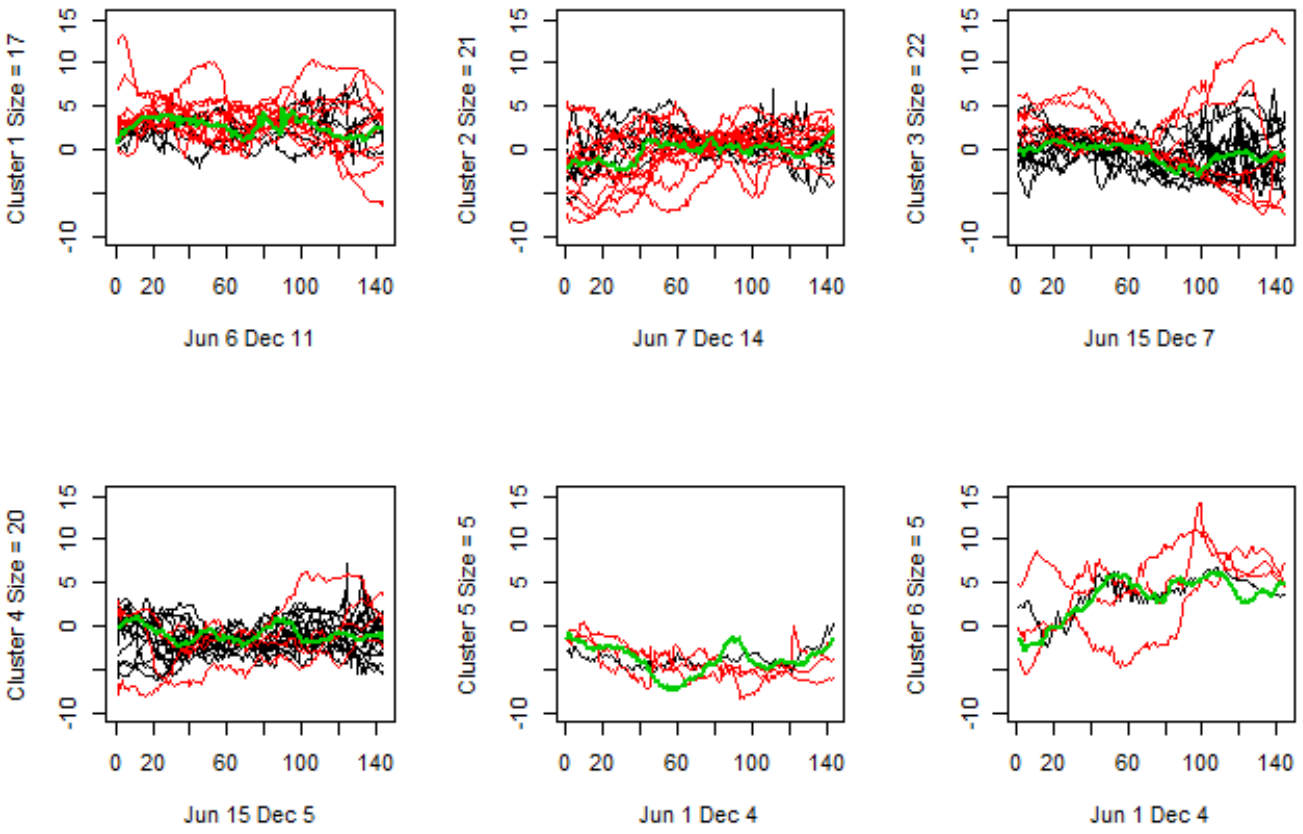


Figure : Clustering results for observations with typical daily shape removed, with number of June (black) and December (red) observations in each cluster noted on axis. Cluster centers (medoids) are shown in green.

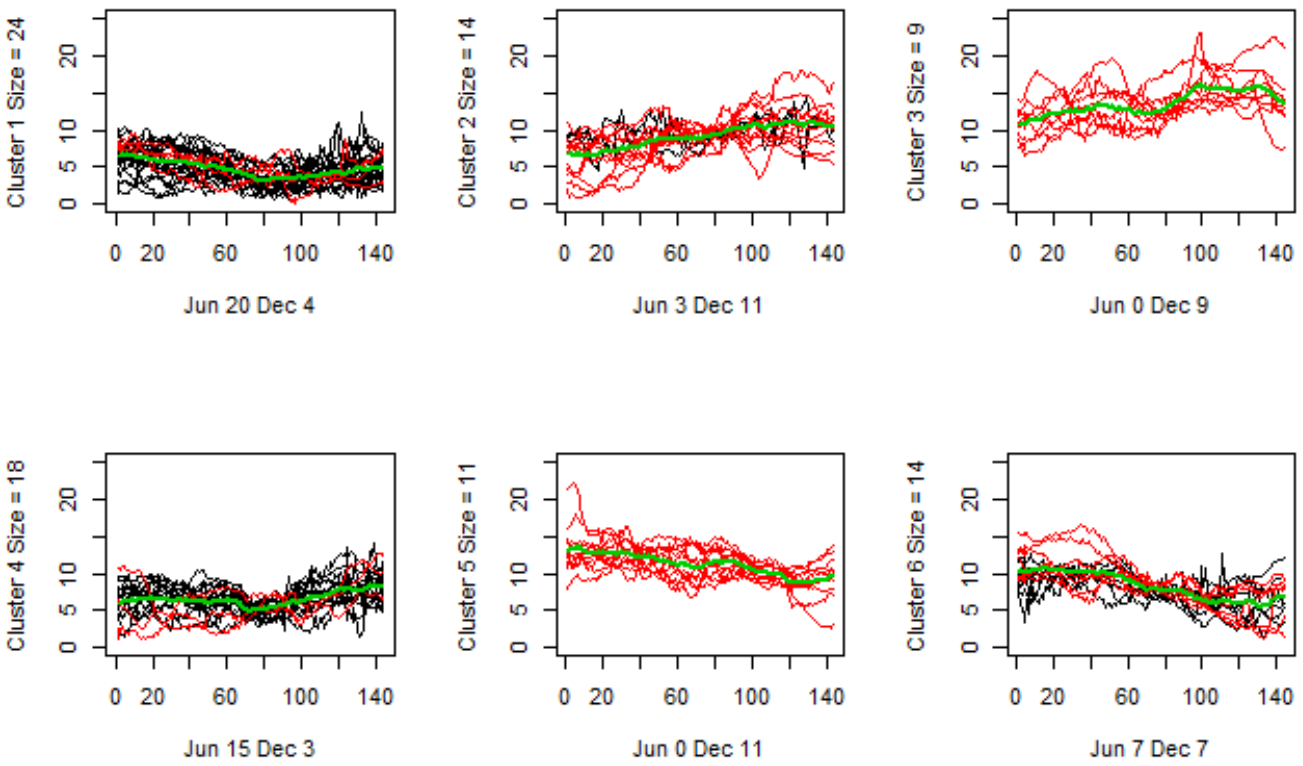


Figure : For comparison, using standard k-means with number of June (black) and December (red) observations in each cluster noted on axis. Cluster centers (medoids) are shown in green.

Theoretical questions

- How can we assess effectiveness of these methods?
  - What objective function do we use to judge optimality of cluster assignments?
  - Can we determine how much information is lost by using our chosen representatives?
  - If we do not have a fixed number of clusters, can we find the optimal number?
- Are these methods sensitive to initialization or slight changes in observations?
  - Cluster sizes and assignments
  - Representatives from each cluster

Extensions of the problem

- Spatial component
  - Extension to multiple locations does not incorporate spatial information explicitly
  - Non-constant correlations between sites, with lags depending on spatial information
  - Will increase both complexity and size of the problem
- Different lengths of observations
  - Longer outlook means taking into account correlations between days

Conclusions

The primary conclusions of recent work are as follows:

- Implementation separates the challenge of the stochastic variables from the system complexity, providing computationally scalable solutions
- The chance-constrained formulation of SCUC provides a balance of risk between expected value methods and robust methods
- Initial tests show promising computation times for simplified problems
- Unsupervised clustering can still be implemented and some shortcomings avoided
- Clustering on a new distance metric, with centroid representatives realizations show more realistic behaviors

Contributions

- Martinez, M. G., & Anderson, C. L. A Risk-averse Optimization Model for Unit Commitment Problems (pp. 19). Submitted to the 2015 48th Hawaii International Conference on System Sciences (HICSS).
- Laura L. Tupper, C. Lindsay Anderson & David Matteson (2104) Spatiotemporal Analysis and Clustering of Wind Speeds, presented at the Joint Statistical Meetings, Boston MA. August 2-7, 2014.

Next directions

The next goals for this project are summarized as follows:

- Stochastic Unit Commitment
  - Explore other percentile estimators and sensitivity of optimal solution
  - Test scalability of CCP formulation with larger networks
  - Expand formulation to include AC power flow
  - Test value of lost information in two-stage dynamic formulation
- Scenario Reduction
  - Estimation of the optimal number of clusters for capturing uncertainty
  - Sensitivity of cluster representative to initial conditions
- Test the sensitivity of solution quality to scenario selection and probability estimates.